The Use of Bone Turnover Markers in Monitoring Osteoporosis Treatment in Clinical Practice

Antonia Ugur

A thesis submitted in partial fulfilment of the requirements for the degree of Masters of Philosophy

The University of Sheffield
Faculty of Medicine, Dentistry and Health
Department of Oncology and Metabolism

May 2021

Acknowledgements

This thesis was born out of my clinical interest in metabolic bone disease and my desire to advance my knowledge in the field and immerse myself in the amazing world of the Academics who tirelessly work to advance the evidence base for the good of my patients.

The opportunity my supervisor Professor Richard Eastell gave me to study under him, has catapulted me into my career as Consultant Endocrinologist with special interest in metabolic bone disease, and continues to feed my best evidenced-based patient care. For this I will forever be grateful to him. I owe Professor Eastell particular thanks for giving me an incredible amount of time over the last year to motivate and support me to complete this thesis in the exceptional circumstances of a Pandemic.

Navigating the world of clinical academic research was daunting for a novice, but I would like to thank Dr Jennie Walsh, my second supervisor, as well as Dr Margaret Paggiosi amongst the other students and fellows of the Academic Unit of Bone Metabolism for guiding the way. The former, along with Dr Kim Naylor were significant in the generous provision of anonymised subject data from the historic Trials included within my thesis.

I would like to thank the AUBM study coordinators Simon Bowles and Selina Bratherton, with their team of students from Sheffield Hallam University who recruited 80% of the subjects included in my thesis, recording the relevant information on the clinical database.

This thesis would not have been possible without the work of Fatma Gossiel who has educated me in biochemical assays with never-ending patience, and carried out all the biochemical measurements for this work.

Every day of investigative work for this thesis was spent embedded within the Metabolic Bone Unit of Sheffield Teaching Hospitals and I could not have been made to feel any more welcome by Dr Nicola Peel and all the staff at the MBU.

Above all, I would like to thank my truly selfless husband Dr Andrew Jurd, who didn't hesitate in the sacrifices he has made and the support he has provided since 2017, without which I could not have completed this thesis.

My work and studies at the AUBM have been funded by Immunodiagnostic Systems, Boldon, UK.

Abstract

Osteoporosis is a skeletal disease of compromised bone strength resulting in fragile bones at greater risk of fracture. Clinical trials of bisphosphonate and denosumab treatments show fracture reduction, but with secondary end points of suppression of bone turnover markers (BTM). Markers of bone formation, procollagen I N-extension peptide (PINP) and osteocalcin (OC) suppress after 3 months of anti-resorptives, because of the coupled bone turnover cycle.

I explored the utility in monitoring osteoporosis treatment with a biological target of BTM named the threshold level, defined as the geometric mean of a healthy premenopausal reference interval. Using a single-centre, cross-sectional study I examined the use of the threshold of PINP and OC by two commonly used assays (IDS iSYS and Roche Cobas e411) in an unselected mixed-gender clinical sample on oral or parenteral anti-resorptives. Comparison was made with the performance in controlled subjects of a RCT in which least significant change (LSC) was also analysed as a biological target.

In practically all analyses the two assays of each BTM were equivalent in performance owing to almost perfect method agreement. The sensitivity of the BTM at threshold to detect apparent response to oral bisphosphonates, was lower in the Clinical cohort than in the Trial cohort, with a low specificity in the Clinical cohort also impacting on the significant difference in the AUC in ROC analysis. The sensitivity improved in the cohorts on IV zoledronate and SC denosumab, owing mainly to unquestionable adherence, but with a subset of individuals that did not supress. Impacting on the sensitivity were recent fracture history and renal impairment, in which the assays differed in behaviour. The specificity remained low in the parenteral cohorts likely because of the previous anti-resorptive use as well as calcium supplementation and glucocorticoid prescription. The impact of this on change in BMD was inconclusive.

Contents

List of Figures

List of Tables

Declaration

1 Declaration

I, the author, confirm that the Thesis is my own work. I am aware of the University's Guidance on the Use of Unfair Means (www.sheffield.ac.uk/ssid/unfair-means). This work has not been previously been presented for an award at this, or any other, university.

2 Abstracts Arising from this Thesis

- Comparison of the reference intervals for procollagen I N-propeptide and osteocalcin in men and women by two assays. Antonia Ugur et al. *(2018) Calcified Tissue International,* Volume 102, Issue 1 - Springer. P98
- Apparent Response Rate by PINP to Oral Bisphosphonates in Clinical Practice and Clinical Trial Settings. Antonia Ugur et al. (2018) *J Bone Miner Res* 32 (Suppl 1). Available at http://www.asbmr.org/Meetings/AnnualMeeting/AbstractDetail.
- The clinical utility of bone formation markers to monitor intravenous zoledronate treatment. Antonia Ugur et al. (2018 Dec) *Osteoporosis Int.* 29(Suppl 2): 601-649. doi: 10.1007/s00198-018-4738-8.

Chapter 1 Introduction

1 Overview

Osteoporosis as a term was first used in 1835 by French surgeon and pathologist Jean Lobstein, who derived it from the Greek for porous bone. Thirteen years earlier a British surgeon Sir Astley Paston Cooper noticed fractures related to abnormal bones (Lorentzon *et al.*, 2015). These early observations still form the basis of the accepted definition today (Peck *et al.*, 1993; NIH, 2001). Osteoporosis is a systemic skeletal disease of compromised bone strength resulting in fragile bones that are at greater risk of fracture (Eastell *et al.*, 2016; Bonewald, 2011).

The chronic disease of osteoporosis and subsequent fragility fractures can negatively impact on the diverse structural functions of the skeletal system; locomotion, providing form and support to the body, growth, self-repair and adaptation, and protection of internal organs (Carter *et al.*, 2000).

Much of the knowledge of the biology of skeletogenesis and bone homeostasis required for a healthy skeletal system, has been acquired through the study of specific bone phenotypes, especially those with altered external appearances or a predisposition to fracture. As understanding of the pathophysiology behind loss of bone homeostasis and osteoporosis has advanced, important interactions with the metabolic functions of the skeletal system have been identified; haematopoiesis (Long, 2011), reservoir of body calcium and phosphate, and as an endocrine organ (Walker and Smith, 2017; Cui *et al.*, 2016).

2 Bone Biology in Osteoporosis

2.1 Bone Macrostructure

The total human skeleton is made up of approximately 15% cancellous or trabecular bone and 85% cortical bone (Mundy, 1999), with composition varying by site. This is most markedly reversed in the vertebrae where the percentage composition has been shown by two techniques to be predominantly trabecular (Eastell *et al.*, 1990). As shown in figure 1.1,

the cancellous bone is sponge-like consisting of plate-like trabeculae that form multiple connections like bracing struts around which mineralisation occurs. The trabecular or cancellous bone is more metabolically active (influenced by the nature of the red bone marrow) leaving it more vulnerable to external stimuli. Cortical bone is the more solid outer shell and makes up much of the long bone shafts surrounding yellow bone marrow (figure 1.1). Although only 20% of total skeletal remodelling occurs in cortical bone, it has more recently been shown to be porous, a feature which appears important in the propagation of osteoporotic changes (Andreasen *et al.*, 2017).

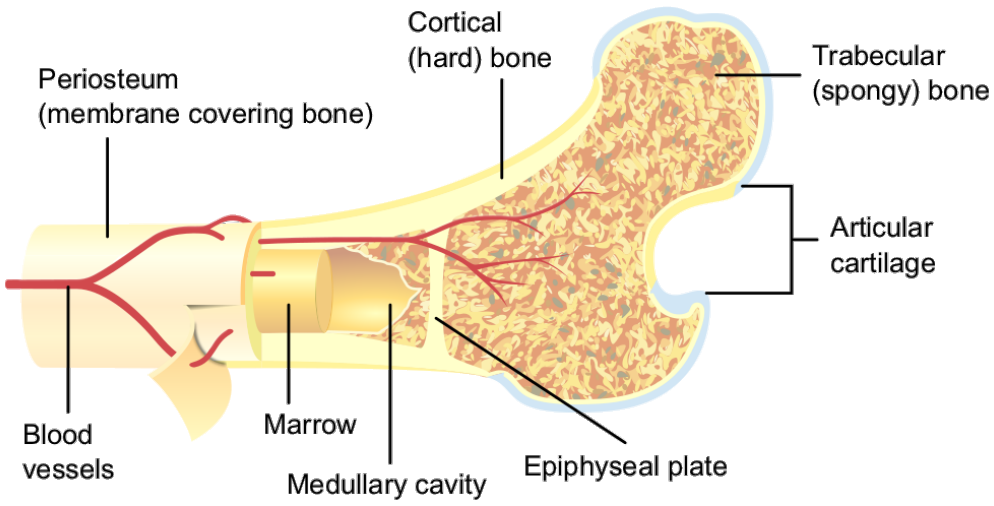


Figure 1.1. Schematic of a long bone demonstrating the distribution of cortical and trabecular (cancellous) bone.

With trabecular bone being the most vulnerable, this explains the sites of the common osteoporotic fractures. (By Pbroks13 CC BY 3.0)

Bone strength, or the force which is tolerated before failure, depends on the mechanics of the force, as well as the density of the bone and skeletal volume (radius of the bone and for the hip particularly "hip axis length") (Sharir, Barak and Shahar, 2008). The bone density is a representation of trabeculae number, spacing and thickness but also quantity of mineralisation and cortical thickness (Peck *et al.*, 1993; NIH, 2001).

Also important for the strength of bone, but not captured by calculation of bone density, is the quality of the microarchitecture and quality of mineralisation (Eastell *et al.*, 2016; Lorentzon *et al.*, 2015), and cortical porosity.

2.2 Bone Remodelling Cycle and Cells

Self-repair and adaptation of bone occurs through a constant remodelling process, a cycle of bone resorption and formation shown in figure 1.2 (Novack and Teitelbaum, 2008).

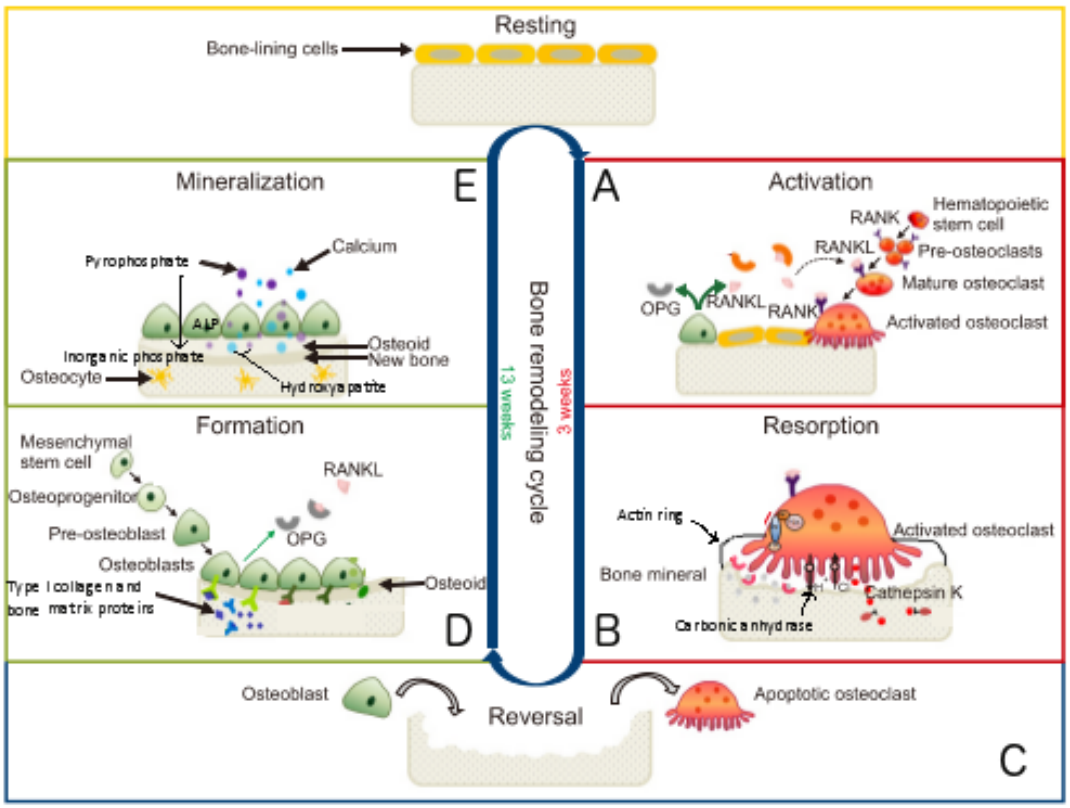


Figure 1.2. Phases of the remodelling cycle with key cells and processes

(Adapted from Kim *et al.*, 2017)

A; Osteoclasts are delivered to activated sites of bone via the systemic circulation from marrow rich bone as osteoclast precursors. The bone lining cells retract.

B; With direct opposition of the bone surface, osteoclast forms an actin ring to seal off its ruffled border with functional secretory domain (Novack and Teitelbaum, 2008). Resorption occurs by two mechanisms, the minerals dissolve in the acidic environment, then the enzymes break down the organic matrix.

C; The osteoclast migrates away after resorption and osteoblasts move in.

D; Osteoblasts secrete osteoid; osteocytes (entombed mature osteoblasts), 10% bone matrix proteins and 90% type 1 collagen.

E; Osteoblasts begin primary mineralisation over several weeks by the dephosphorylation of pyrophosphate to inorganic phosphate in surface vesicles by alkaline phosphatase. The inorganic phosphate combines with calcium to form hydroxyapatite. Osteocytes continue secondary mineralisation over 12-24 months.

Remodelling of the cortical bone and trabecular bone are considered as two distinct processes being facilitated by different structures and stimulated by different factors (Andreasen *et al.*, 2017). The main determinant of trabeculae number and thickness and cortical porosity and thickness is universally the success of coupling the activities of cells of bone remodelling, osteoclasts and osteoblasts. An uncoupling, together with an increase in frequency of activation of the cycle creating more resorption pits as seen in section B of figure 1.2, are one pathophysiological mechanism of osteoporosis (Gossiel *et al.*, 2018).

Osteocytes, which make up 95% of bone cells, are considered in more recent history to play a key role in orchestrating the bone remodelling cycle. Key features demonstrated in osteocytes are polarity, and dendritic connecting processes directed towards where mineralisation is occurring and to the nearest vascular space, and release of RANK ligand and sclerostin, key messengers of coupling as shown in figure 1.3. They appear to respond to bone strain or change in the microenvironment (Bonewald, 2011).

Secondary mineralisation occurs over 12 to 24 months. The frequency of bone turnover therefore effects the amount and quality of mineralisation. In states of high turnover, less mineralisation can occur before more bone resorption pits are created in the same loci again and the bone will become weakened (Cui *et al.*, 2016). However, if bone remodelling is completely suppressed the bone will become over mineralised producing brittle composition, also susceptible to fracture.

There are other disease states of the systemic skeleton that also have an inherent risk of fracture (different in nature to a fragility fracture from osteoporosis), but through different pathophysiology leading to lack of mineralisation. These include osteomalacia, alkaline phosphatase deficiency in hypophosphatasia (Bianchi, 2015), or hypophosphataemic rickets (Vlot *et al.*, 2018). These conditions have very different management strategies to osteoporosis and failure to recognise them could cause more morbidity. Finally, loss or disorganisation of trabecular connections can be caused by defects in the osteoid, the stroma of bone secreted by osteoblasts, before mineralisation. Faulty type I collagen in osteogenesis imperfecta increases activation of bone remodelling as well as uniquely increasing the mineral to collagen matrix ratio, resulting in more brittle bones, with an unpredictable bone mineral density (Dijk *et al.*, 2011; Bishop, 2016).

2.3 Balance in Bone Remodelling

In order for bone resorption and formation to remain in balance, osteoclast and osteoblast differentiation and function must be coupled processes and figure 1.3 details the signalling pathways that control this.

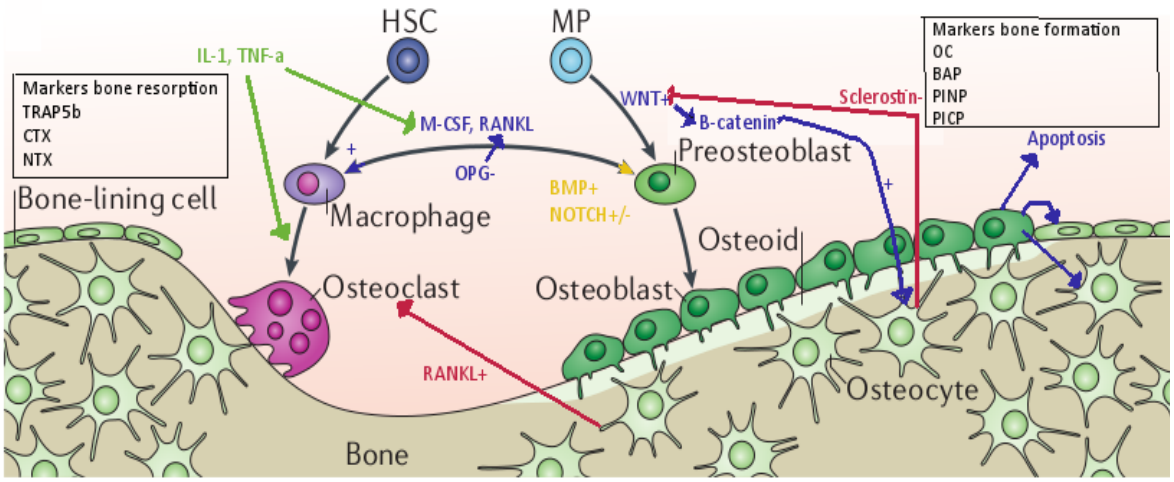


Figure 1.3. Signalling pathways promoting differentiation and action of bone cells showing biochemical markers identifying different stages

(Modified from Long, 2011)

Osteoclast differentiation from haemopoietic stem cells (HSC) is dependent on the presence of osteoblast lineage cells and their expression of RANKL and macrophage colony stimulating factor (M-CSF), as well as other osteoclastogenic pathways that are common to other parts of the immune system. The differentiation of osteoclasts is inhibited by osteoblast osteoprotegerin (OPG) (Novack and Teitelbaum, 2008). Messengers secreted from osteoclasts are shown in yellow.

Osteoblast lineage cells are formed from mesenchymal progenitor cells (MP) which may also differentiate into an adipocyte depending on the activation of the WNT pathway. Once fully matured osteoblasts either apoptose, become inactive lining cells, or are incorporated into the matrix as osteocytes (Bonewald, 2011). Messengers secreted from osteoblasts are shown in blue.

Sclerostin is produced by osteocytes and inhibits WNT pathway within osteoblast lineage cells causing B-catenin to degrade. the WNT pathway which stabilises B-catenin is vital alongside the BMP pathway and to some extent the NOTCH pathway to promote osteoblast intra-membranous differentiation (Long, 2011). The later pathways are activated by osteoclast secreted and membrane-bound factors. B-catenin is also important for osteocyte viability which acts as a stop mechanism on sclerostin inhibition. Messengers secreted from osteocytes are shown in red.

The biochemical markers of these processes are functional enzymes and breakdown products.

The cytokine signals invoke their effect with high enough concentration in the enclosed space of one remodelling compartment, as well as by cell-to-cell contact.

Self-repair is required by bone remodelling in response to microcracks from strain, with increasing evidence of the role of the osteocyte in detection and initiation (Bonewald, 2011). Initiation is facilitated through communication with the bone lining cells or through direct recruitment of osteoclasts by apoptosis of the osteocytes at the location of the micro-crack. Apoptotic bodies express RANKL. Increasingly considered important in remodelling in response to micro-cracks as well as in repair of fracture, is the ingrowth of marrow capillaries to supplement the progenitor cells within the remodelling compartment (Khosla, Westendorf and Mödder, 2010).

The balanced processes of bone remodelling continue as per the descriptions in figure 1.2 and 1.3.

2.4 Changes in Bone Remodelling

There are many situations in which the number of remodelling compartments accelerate, and resorption and remodelling can become unbalanced, with resultant fragility of bone.

The two key precipitants of this state are increased osteocyte apoptosis or necrosis, with or without the generation of multiple lacunae void of ongoing osteocyte presence.

Increased osteocyte apoptosis in the absence of strain or damage, occurs with aging, oestrogen deficiency and states of immobilisation (Crockett *et al.*, 2011). In these situations, osteoclasts are recruited in greater number with less coupled osteoblast differentiation. With subsequently fewer osteocytes, there are resultant empty lacunae in which detection of micro-cracks fails to occur (Bonewald, 2011). Immobilisation causes reduced oxygenation of the osteoid, and oestrogen deficiency is accompanied by increased pro-inflammatory markers TNF-a and IL-1. As well as causing osteocyte apoptosis these cytokines interact with osteoclast differentiation through macrophage colony stimulating factor (M-CSF) seen in figure 1.3. This is theorised to be one mechanism by which postmenopausal osteoporosis is caused (Novack and Teitelbaum, 2008; Bonewald, 2011).

Age-related and postmenopausal osteoporosis are states of exaggerated physiological processes of bone remodelling and demonstrate that the balance of osteocyte viability

versus apoptosis is crucial. Other pathological influencers of bone remodelling leading to fragility by the same mechanism are glucocorticoids and secondary osteoporosis as a result of disease states such as CKD-metabolic bone disorder, primary hyperparathyroidism, hyperthyroidism and diabetes mellitus.

3 Biology of Bone Turnover Markers

Bone turnover markers are molecules produced during the remodelling cycle by osteoclasts, osteoblasts and osteocytes as seen in figure 1.3.

<table>
<tr><th colspan="3">Bone Formation Markers</th><th colspan="3">Bone Resorption Markers</th></tr>
<tr><td>Enzyme</td><td>BAP</td><td rowspan="3">Stage of secretion pre-osteoblast to mature osteoblast</td><td>Enzyme</td><td colspan="2">TRAP</td></tr>
<tr><td>Matrix Protein</td><td>Propeptides of type I collagen- PINP
PICP</td><td>Collagen degradation products</td><td>Pyridinium cross-links of collagen- free (DPD) or membrane bound (NTX/CTX)
Or longer fragment (CTX-MMP/ICTP)</td><td>Digestion by Cathepsin K or MMP</td></tr>
<tr><td></td><td>OC</td><td></td><td></td><td></td></tr>
</table>

Figure 1.4. Categorisation of Bone Turnover Markers

Table of all proposed bone turnover markers with available assays (Vasikaran *et al.*, 2011).

Proteins or fragments released in bone formation; bone alkaline phosphatase (BAP), procollagen type 1 N terminal peptide (PINP), procollagen type 1 C terminal peptide (PICP), Osteocalcin (OC).

Proteins or fragments released in bone resorption; Tartrate resistant acid phosphatase (TRAP), deoxypyridinoline (DPD), (amino)N-terminal and (carboxy) C-terminal telopeptide of type 1 collagen (NTC, CTX), with equivalent fragment produced by matrix metalloproteinase digestion.

They may be break-down products of the processes involved, or they may have enzymatic functions in the processes, detailed by figure 1.4. Through identification of which cell line or

process is the source of the marker, they can be divided into bone formation and bone resorption markers.

3.1 Markers of Bone Formation

3.1.1 Bone-specific Alkaline Phosphatase

High levels of alkaline phosphatase (ALP) are secreted by osteoblasts to adhere to the cell surface membrane by a glycan phosphatidylinositol anchor. Bone alkaline phosphatase performs primary mineralisation on the cell surface. Membrane fragmentation to release vesicles, and a specific phospholipase enzyme confer the ability to detect bone alkaline phosphatase in the blood, but only as a proportion of the body's total alkaline phosphatase (Moss, 1997). The isoenzyme produced in bone mineralisation can be measured distinctly from that formed by hepatocytes and to a lesser extent in the kidneys (Long, 2011). Although assay development for bone-specific alkaline phosphatase (BAP) has improved reproducibility and reduced the inter-assay coefficient of variation, it is not the most sensitive marker of bone formation and biological variation within one individual (CV_I) is still 10% over a longer term. However, its advantage is that most clinical departments have an automated analyser and reagents readily available to measure tissue non-specific ALP (TNSALP) from blood samples taken at any time of day (Bergmann *et al.*, 2009). Although a higher BAP level represents greater osteoblast activity in the context of normal renal function, the converse points more specifically to the mineralisation pathology, hypophosphatasia (Bianchi, 2015).

3.1.2 Osteocalcin

The measurement of osteocalcin in systemic circulation is specific to bone metabolism as it is secreted almost exclusively by osteoblasts into osteoid during the mineralisation phase (Szulc, 2012). OC exists in two forms with different affinities for the bone matrix, carboxylated and undercarboxylated. This refers to the carboxylation of the glutamic acid residues before secretion from the osteoblast which is a vitamin K dependent process. The way these residues of the OC relate to the bone matrix is shown in figure 1.5. Some OC leaks into the circulation when osteoid is formed, and the amount is influenced by the exposure of the bone microenvironment to insulin and glucose levels (Long, 2011) and vitamin K

antagonists like warfarin (Eastell and Szulc, 2017). OC also appears in venous blood samples as a consequence of breakdown of the bone matrix (Bergmann *et al.*, 2009) but predominantly as the undercarboxylated form as decarboxylation occurs in an acidic resorptive environment (Ducy, 2011).

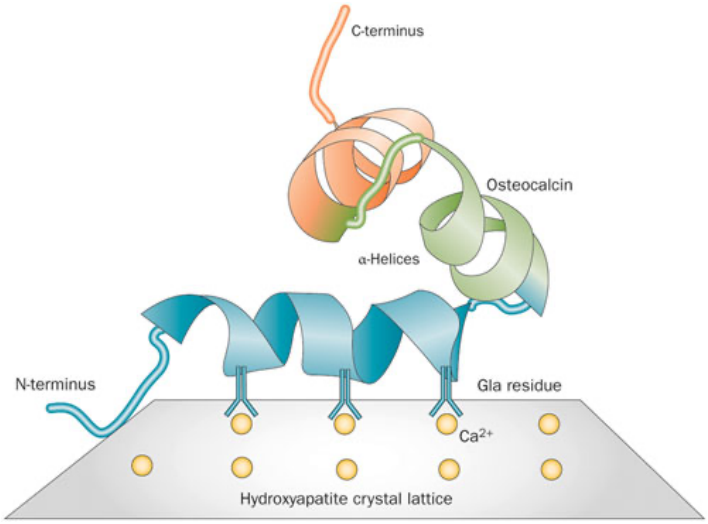


Figure 1.5. Structure of osteocalcin as determined by NMR and Xray crystallography (Booth *et al.*, 2013)

OC is a tight globular peptide made of three helices. It opposes the hydroxyapatite surface with y-carboxyglutamic acid residues that interact with calcium ions. Carboxylated OC has a greater affinity to the bone than under-carboxylated.

Another influencer of venous levels of OC is glomerular filtration rate and this greatly increases the biological variation. The inter-assay coefficient of variation (CV_A) with a newer N-mid fragment assay on IDS iSYS auto-analyser has been recently published from a European wide study at 2.3% (Cavalier *et al.*, 2020), but is effected greatly by pre-analytical factors, such as haemolysis, time to freezing, and sensitivity to the freeze-thaw cycle (Bergmann *et al.*, 2009). The CV_I of OC is 8.9% (Cavalier *et al.*, 2020). Osteocalcin does show a circadian variability with the lowest serum levels in the morning through to midday, rising to a peak just after midnight (Schlemmer and Hassager, 1999; Calvo, Eyre and Gundberg, 1996). The reduction of the circadian variability from the observed 10-30% in the older assays to use of the N-MID terminal assay , has not been fully publicised (Calvo, Eyre and Gundberg, 1996).

3.1.3 Procollagen I N-terminal Extension Peptide

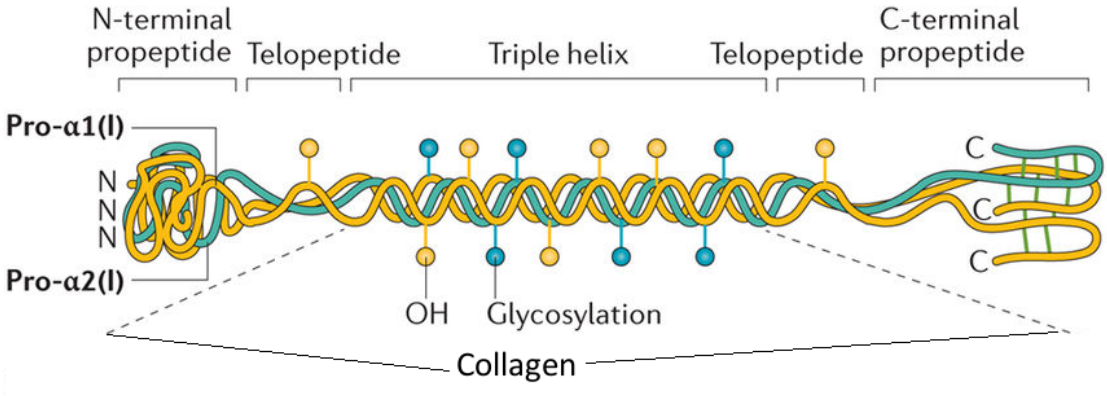


Figure 1.6. Structure of collagen with the two end terminals that are cleaved to PINP and PICP

(Marini *et al.*, 2017)

Proteases recognise a unique sequence in the two telopeptide regions to cleave at the correct site to release the carboxyl- terminal (PICP) and amino-terminal (PINP). PINP can exist in the circulation as a trimer, or additionally at body temperature, as a monomer.

There are two procollagen I extension peptides that can be measured in the systemic circulation when they are cleaved from each end (C-terminal and N-terminal) of the precursor to type I collagen as seen in figure 1.6. This occurs after the molecule is secretion into the osteoid. As neither peptide has the affinity to remain in the bone matrix, the amount detected in the blood corresponds directly to the osteoid producing activity of the osteoblast. They are secreted in good amounts as the collagen is 90% of the composition of osteoid (Chavassieux *et al.*, 2017). Any tissue that has type I collagen as a component would contribute these peptides to those measured in the circulation, but the rate of bone formation means that the proportion (90%) from bone turnover far outweighs any other source. In terms of biological variability, the peptides do show small circadian variation but are not excreted into the urine. The biological variation of PICP is high at approximately 25% owing to a shorter half-life which is influenced by both thyroid and growth hormones (Vasikaran *et al.*, 2011). PINP is more stable and unaffected by the endocrine environment and so is more clinically useful with CV_I of 8.8% and CV_A of 3.8% (Cavalier *et al.*, 2020). The time to freezing is only a problem for assays that do not identify partially degraded PINP (Bergmann *et al.*, 2009).

3.2 Markers of Bone Resorption

3.2.1 Tartrate Resistant Acid Phosphatase 5b

Tartrate resistant acid phosphatase 5b (TRAP5b) is an enzyme like Cathepsin K released from the osteoclast during bone resorption, and it enters circulation as it traffics with breakdown products into a capillary close to a resorbing osteoclast. Before an assay was developed to measure it in blood, its presence was used in the histological staining of osteoclasts (Novack and Teitelbaum, 2008). In the recent past assays have been refined to specifically measure TRAP5b rather than 5a, and even more recently to be specific to bone dependent TRAP5b rather than include that produced by alveolar macrophages. This should improve the biological variation and sensitivity to changes in bone resorption (Bergmann *et al.*, 2009; Brady *et al.*, 2014).

3.2.2 C-terminal Crosslinked Telopeptide of Type 1 Collagen

Amino-terminal and carboxy-terminal telopeptides of type 1 collagen (CTX and NTX) are peptide-bound breakdown products of resorption, and they are markers of osteoclast activity rather than number as in TRAP5b. In the osteoclast-rich form of osteopetrosis, in which the activity of the osteoclasts is defective due to failure to produce a functioning ruffled border, the levels of CTX and NTX are much reduced whereas TRAP5b is elevated.

Cathepsin K is the main enzyme secreted by the osteoclast to break down type I collagen within an acidic extracellular microenvironment as seen in figure 1.2, thereby releasing these fragments onto which crosslinks attach, into the circulation (Novack and Teitelbaum, 2008). They are then excreted by the kidney and can therefore be measured from venous blood or urine.

There are four isomers of CTX and it is ßCTX that is widely assayed from serum, but either α or ß can be measured in the urine (Bergmann *et al.*, 2009). Because of their specificity to degree of bone resorption they are very sensitive to influencers of change of bone turnover. The coefficient of variation is similar between serum and urine samples with recently reported serum CV_I of 15.1% and CV_A of 5.0% (Cavalier *et al.*, 2020). They demonstrate the most circadian variation with values much less in the early afternoon and also supressed post-prandially (Schlemmer and Hassager, 1999). Despite this pre-analytical variation sCTX is

designated the marker of resorption of choice (Vasikaran *et al.*, 2011). Looking at the ratio of alpha to beta CTX may help to distinguish how recently collagen has been formed (Gossiel, Hannon and Eastell, 2009).

If matrix metalloproteinases, such as in cancer states, break down the matrix then a circulating longer collagen crosslink and telopeptide is the result, named ICTP or CTX-MMP. This bone marker has no use in osteoporosis or it's treatments (Vasikaran *et al.*, 2011).

4 Effect of Systemic Conditions and Medication on Bone Turnover Markers

4.1 Thyroid Hormone

The effect of thyroid disorder on bone appears to be mediated through the TSH receptor rather than the nuclear thyroid hormone receptors. It is found on osteoclasts and osteoblasts and when activated has an inhibitory effect on both cell lines causing reduced bone remodelling. If it is lacking, as modelled in mouse studies, this induces global osteoporosis caused by increased bone turnover. The nuclear receptor for active thyroiodine, is found in all bone cells and can directly inhibit osteoblast function, but not significantly so (Novack and Teitelbaum, 2008).

4.2 Parathyroid Hormone

The effect of parathyroid hormone on bone is dependent on the duration of action of increased levels on the bone. Sustained increased parathyroid hormone levels activate bone resorption by osteoclasts, which couple to osteoblast increasing bone turnover with some imbalance to preference bone resorption (Bonewald, 2011). However, parathyroid hormone also reduces sclerostin levels taking the breaks off the WNT pathway, thus intermittent exposure to higher than normal parathyroid hormone has an anabolic effect on bone (Long, 2011).

4.3 Diabetes Mellitus

There is a complex interaction between the presence of altered glucose metabolism and fracture risk. It must take into account the confounder of obesity (Vestergaard, 2007), increased falls risk from hypoglycaemia or neuropathy (Oei *et al.*, 2015), as well as changes to bone turnover. The fracture risk in type 1 and type 2 diabetes mellitus (DM), is beyond that which you would expect for an individual without diabetes with the same bone mineral density (BMD). This is more evident in patients with type 2 diabetes, when perhaps from the confounder of obesity, a higher BMD than healthy controls is found (Starup-Linde *et al.*, 2016; Vestergaard, 2007). This applies to the hip fracture risk in type 1 DM, and the ankle in type 2 DM. The fracture risk of the spine has not been shown to be increased. The main hypothesis is that the glycation of the matrix and change in vasculature will be altering the bone microarchitecture (Janghorbani *et al.*, 2007) and cortical porosity. BMD does not correlate with HbA1C and there is mixed evidence of whether diabetic complications are associated with a lower BMD (Starup-Linde *et al.*, 2016; Vestergaard, 2007).

There have been very mixed reports of whether bone biopsies from patients with diabetes in small numbers reproducibly show low bone turnover. With regards to bone turnover markers, a meta-analysis in patients with diabetes does show a consensus of reduced CTX and OC (Starup-Linde, 2013). The Aarhus group went on to study a large cohort of subjects with type 1 and type 2 DM, that were unmatched in age and BMI and from which non-fasted samples were taken (Starup-Linde *et al.*, 2016). They demonstrated a statistically lower level of biochemical markers of bone resorption and bone formation in patients with type 2 compared to type 1 diabetes. In the cohort of type 1 diabetes there was a statistically significant negative correlation of HbA1C with one bone turnover marker OC, rather than the non-fasting glucose level and BTM level. This supports thinking of the effect of diabetes on bone turnover markers being mediated through amount of circulating insulin as an anabolic agent (Starup-Linde *et al.*, 2016). An association between lowering of OC and vertebral fracture risk was seen by the Aarhus group. Therefore the consistency and gravity of the effect of DM on OC suggests it would be a useful BTM to monitor in the chronic management of DM (Vlot *et al.*, 2018).

4.4 Chronic Kidney Disease- Metabolic Bone Disorder

The risk of fracture in a patient with chronic kidney disease (CKD) less than stage 3b, is 4-fold (Davina *et al.*, 2017). The fracture risk is conferred by osteoporosis associated with CKD, or by CKD- Metabolic Bone Disorder (MBD), a distinct group of bone pathophysiology (Pazianas and Miller, 2021).

CKD-MBD includes disorders of mineralisation, altered bone turnover and vascular and soft tissue calcification. The altered bone turnover may be elevated from secondary or tertiary hyperparathyroidism, or a low bone turnover state (Ketteler *et al.*, 2017).

The contributors of bone fragility in CKD should be assessed in order to optimise management strategies, but the gold standard method of identifying or excluding aspects of CKD-MBD is a trans-iliac crest biopsy. Osteoporosis becomes the diagnosis by exclusion (Pazianas and Miller, 2021). There is increasing evidence of the utility of bone-derived turnover markers rather than the traditional intact parathyroid hormone (iPTH) to determine a high or low turnover state in CKD (Salam *et al.*, 2018; Davina *et al.*, 2017). BAP and TRAP5b show low biological variation in this context, and even high intact PINP is associated with prevalent fractures (Vlot *et al.*, 2018). Most other BTM are renally excreted and rise as a consequence of reduced clearance. On the basis of evidence that BMD can predict incident fractures in CKD stages 3 and below, the KDIGO guidelines of 2017 take a more pragmatic approach to evaluation. They encourage performing bone densitometry alongside optimising markers of mineralisation in order to guide when intervention to reduce fracture risk is necessary (Ketteler *et al.*, 2017).

4.5 Sex Hormones

All cells found in bone express androgen receptors (AR) and two types of oestrogen receptors (ER*a* and ERß). Testosterone directly simulates periosteal apposition of bone during the period of longitudinal growth, but also after end-plate closure to facilitate cortical remodelling (Walsh *et al.*, 2017). Oestrogen produced from the aromatisation of testosterone, effects the ERα (Long, 2011). Activation of ER α stabilises B-catenin signalling, reducing the negative effect of sclerostin, to induce more osteoblast differentiation (Bonewald, 2011). Oestrogen causes apoptosis of osteoclasts, therefore oestrogen deficiency causes an increase in the number of osteoclasts (Novack and Teitelbaum, 2008).

With the loss of oestrogen at menopause, this period of transition is critical for increased bone turnover and bone mineral density (BMD) loss (Ebeling *et al.*, 1996). The primary effect of oestrogen loss and increased osteoclasts is suggested by the bone resorption markers elevating before those of bone formation. There follows a 3 year period of significant BMD loss, which actually commences 1 year prior to the cessation of menstrual cycle (Greendale *et al.*, 2012).

4.6 Glucocorticoids

Chronic use of systemic glucocorticoids causes a suppression of the remodelling cycle. The end result appears to be a more pronounced reduction on bone formation, despite initial disruption of osteoclastogenesis. In in-vitro studies when only the glucocorticoid receptor in the osteoclast cell line is activated, the resulting phenotype is same as systemic exposure of all bone and stromal cells to glucocorticoid (Novack and Teitelbaum, 2008).

The predominant effect on osteoblasts, is hypothesised to be from the expression of the activating enzyme 11β-hydroxysteroid dehydrogenase type 1, particularly in the osteoblast (Cooper *et al.*, 2003). Cooper et al showed a statistical correlation between baseline activity of 11βHSD and reducing PINP and OC, not CTX. The mechanism by which glucocorticoids reduce osteoclast function is through interruption of the osteoclast differentiation (Novack and Teitelbaum, 2008).

As widespread activation of the glucocorticoid receptor occurs in oral and parenteral (intramuscular depo, intra-articular injection or IV bolus) steroids, all these modes of administration may reduce bone remodelling. Even intra-articular steroid produces a rapid effect on reducing osteocalcin in the systemic circulation within 24hours, wearing off by 7 days (Weitoft *et al.*, 2005).

Cooper at al have also demonstrated the onset of effect on glucocorticoids to be rapid; 5mg prednisolone twice daily caused maximal suppression of BTMs at 4 days, but with statistically significant reduction in PINP and OC from baseline remaining at 7 days (Cooper *et al.*, 2003).

Even with reduced bone formation markers on glucocorticoids, a further reduction in PINP was seen in a study of alendronate for glucocorticoid induced osteoporosis (GCIO). Despite

the reduction exceeding LSC, positive gain in BMD on alendronate was blunted (Burshell *et al.*, 2010).

5 Bone Turnover Markers in Clinical Trials

A clinical population attending primary or secondary care with fragility fractures or osteoporotic bone mineral density, is markedly heterogeneous. An individualised approach to addressing the cause, lifestyle factors and considering treatment is therefore required. In the consideration of treatment, bone turnover markers have a part to play potentially in establishing baseline turnover and monitoring treatment (Vasikaran and Paul Chubb, 2016).

The first step in monitoring a chosen treatment is the knowledge of the anticipated effect of it on bone turnover markers.

The available treatments that have been developed to reduce fracture risk, can be divided broadly into those reducing osteoclastic bone resorption, and those increasing osteoblastic bone formation. Because of the coupled phases of bone remodelling, the treatments for osteoporosis will over time suppress or stimulate total bone turnover with consequential matched decrease or increase of markers of both bone formation and bone resorption.

When to initiate treatment has predominantly been guided by trials categorising patients by the WHO definition of osteoporosis and determining the subsequent effect of treatment on reducing fracture risk (Black *et al.*, 2000). The incidence of fracture is the primary outcome in the pivotal treatment studies (Black *et al.*, 2007; Cummings *et al.*, 2009; Cummings *et al.*, 1998). However, change in bone turnover markers and whether there is a significant difference in change between the treatment and placebo groups has often been a secondary outcome or the subject of a post-hoc analysis.

Any change in bone turnover markers with treatment in the pivotal trials is related to cohort or population change only, and usually within a controlled population with osteoporosis. It is not directly transferable to the individual.

5.1 Hormone Replacement Therapy and Selective Oestrogen Receptor Modulators

Oestrogen causes osteoclast apoptosis and therefore hormone replacement therapy supresses bone resorption, as much as bisphosphonate therapy (Greenspan, Resnick and Parker, 2005). Hannon et al looked specifically at the use of bone turnover markers to monitor HRT (Hannon *et al.*, 1998). When considering the biological variability of the biochemical markers, PINP and OC demonstrated the greatest proportion of subjects that had a significant reduction in the BTM on HRT. PINP reduction at 24 weeks was 40.5% (Hannon *et al.*, 1998). When considering its use for osteoporosis, the prolonged action of HRT on the breast and uterus are important factors to consider.

Selective oestrogen receptor modulators (SERM) such as raloxifene, avoid these considerations as well as targeting a secondary mechanism by which deficiency in oestrogen causes osteoporosis. Studies show that elevated levels of FSH in primary oestrogen deficiency stimulate TNFa production which drives osteoclasts. As raloxifene activates oestrogen receptors on the pituitary as well as bone, it reduces the FSH levels without the unwanted effects on breast and uterus (Novack and Teitelbaum, 2008). The mean reduction in PINP of 31% on SERM (Naylor *et al.*, 2010) is relatively lower than publications for other treatments, however it maintains a stable BMD over time (Eastell *et al.*, 2009) and significantly reduces vertebral fracture rate (Eastell, 2007).

5.2 Bisphosphonates

Bisphosphonates, as an analogue of pyrophosphate, are incorporated into the bone matrix in the same way as the calcium phosphate compound, hydroxyapatite. When osteoclasts resorb the matrix, the bisphosphonate molecule enters the osteoclast and alters pathways that cause the osteoclast to dysfunction and then undergo apoptosis (Novack and Teitelbaum, 2008). The result is suppression of bone resorption. There are two subsets of bisphosphonate molecules, nitrogen containing and not. The N Bisphosphonates such as alendronate, risedronate, ibandronate are much more potent in terms of anti-resorptive potential than the non-N bisphosphonates, etidronate and clodronate (Russell, Croucher and Rogers, 1999). They can be given orally with possible compliance issues that may be due to upper gastrointestinal side effects, or intravenously.

Initial trials of older oral bisphosphonates such as ibandronate and etidronate showed the main effect was to reduce vertebral fractures. The Fracture Intervention Trial Group produced multiple publications between 1996 and 2000 on the effect of alendronate on fracture risk in large cohorts of variable baseline fracture characteristics and total hip T-score (Cummings *et al.*, 1998; Black *et al.*, 1996). A summary pooled analysis in 2000 demonstrated that for patients with defined osteoporosis by hip bone mineral density or vertebral fracture, there is reduction in vertebral and non-vertebral fractures with NNT of 13 and 11 respectively. All clinical fractures saw a significant reduction within 12 months of the treatment start (Black *et al.*, 2000).

The alendronate trials did not include analysis of bone turnover markers. The Vertebral Efficacy with Risedronate Therapy Trial (VERT) showed a 35% reduction from baseline of bone alkaline phosphatase at six months that remained plateaued for the 3 years on 5mg daily risedronate (Harris *et al.*, 1999). Without many significant fracture risk reductions its main conclusion was that "risedronate is effective and well tolerated".

The Horizon trial studied once yearly intravenous zoledronate and showed the reduction of vertebral fractures and hip fractures with zoledronate over 3 years were 70% and 40% respectively. Alongside this as a secondary outcome was demonstration of the PINP and sCTX reductions by 56% and 59% respectively at 12 months, remaining plateaued for the 3 year study (Black *et al.*, 2007).

5.3 Denosumab

Denosumab is also an anti-resorptive, but rather than cause osteoclast apoptosis like the bisphosphonates, it was developed to target the RANK signalling that promotes differentiation of osteoclasts. It is a monoclonal antibody to RANK ligand, thereby causing reversible inhibition of the interaction of RANKL and the receptor RANK on the osteoclast cell line.

The FREEDOM trial was pivotal in demonstrating a relative reduction in vertebral fractures of 68% and hip fractures by 40%. This trial analysed sCTX and PINP and showed a massive early reduction in sCTX of 86% at one month. The PINP nadir was by six months, at 50% below the placebo group (Cummings *et al.*, 2009). The Freedom extension study had arms

which ceased Denosumab (Bone *et al.*, 2013). In these subjects, the markers of bone turnover, rapidly reversed to be higher than baseline.

5.4 Teriparatide

Teriparatide is the only licensed anabolic agent. It is a recombinant of parathyroid hormone consisting of the first 34 amino acids (out of 84). The anabolic effects are seen at an intermittent dosing interval of once daily or once weekly. Teriparatide only has safety and efficacy data for administration over 24 months due to fears over promotion of neoplasm in rat studies (Eastell *et al.*, 2009). The BMD gain from teriparatide occurs in an "anabolic window" with early increases of bone formation markers by 75% at one month, but no significant increase in the markers of bone resorption (Glover *et al.*, 2009). The markers of bone resorption do subsequently catch up (Vescini and Grimaldi, 2012; Yamamoto *et al.*, 2014), but there continues to be hip and lumbar spine BMD gain from 12 to 24 months of teriparatide (Eastell *et al.*, 2009).

Burshell et al have hypothesised that in glucocorticoid induced osteoporosis there is an even bigger window of predominantly bone formation (Burshell *et al.*, 2010). There was a different observed effect on BTMs, by which PINP and BAP increased and remained down-trending but significantly higher than baseline at 18 months, with sCTX falling back to baseline at 18 months after the initial peak.

On cessation of therapy, the positive effect of teriparatide on cortical bone especially is maintained in the presence of an anti-resorptive, but lost very quickly if no ongoing bone therapy is used (Eastell *et al.*, 2009).

6 Monitoring Methods as a Potential Surrogate for Fracture Risk Reduction with Treatment

Treatment for osteoporosis undoubtedly causes a change in bone turnover marker level shown in the secondary outcomes of the clinical trials. Few have reflected on whether the relative or absolute change of bone turnover markers are predictable, reproducible and associated with a predictable reduction in fracture risk.

6.1 Bone Mineral Density as a Surrogate

Bone mineral density has more frequently been included as a secondary outcome in the pivotal clinical trials. The response of bone mineral density (BMD) to anti-resorptive treatment has been shown to be predictable with 80% of patients showing a response (Vasikaran and Paul Chubb, 2016). It is because of the inclusion of BMD data in all the trials that the FNIH Bone Quality Project has performed a meta-regression of 38 RCTs, with the inclusion of 19 treatments for osteoporosis (Bouxsein *et al.*, 2019). It found a correlation between percent change in BMD on treatment compared to placebo and the log of the relative risk of vertebral and hip fractures. If the total hip BMD in the treatment cohort improves by 6%, this would confer a 66% reduction in vertebral fracture risk and a 40% reduction in hip fracture risk in the population. To this end, change in total hip BMD has now the evidence to be used a surrogate for fracture reduction in the clinical trials of new agents to treat osteoporosis.

But the more recent data applies to population groups only, it isn't the case that a BMD percent change of x will confer a y% reduction in fracture risk in the individual.

Therefore I return to older evidence to appreciate that bone mineral density with a T-score of -2.5 at the total hip has a sensitivity for hip fracture of 80% (Yang *et al.*, 2009). It is through sequential evidence that BMD is used routinely as a surrogate for monitoring reduction in fracture risk with treatment in individuals (Kanis *et al.*, 2014). The best percentage change in bone mineral density at a recommended interval of two years is very small (variable by treatment group), and there is often heterogeneity at different sites. Once one has considered technical factors such as change in scanner, change in patient weight, change in region of interest; the change in BMD certainly is very limited in predicting reduction in fracture risk in the individual patient (Delmas *et al.*, 2000). Added to this that bone mineral density is not the only determinant of bone strength and fracture risk, particularly with certain pathologies such as diabetes, and therapies such as raloxifene (Bjarnason *et al.*, 2001), that the use of bone density as a measurement of prediction is even more flawed. (Eastell *et al.*, 2003).

Despite the limitations, using BMD change as the gold standard for whether there has been a response to treatment is the more efficient way of evaluating a biochemical tool for monitoring response of the individual to treatment for osteoporosis.

6.2 Need for an Alternative Monitoring Method

Ravn thought about the necessity of another way to monitor response to treatment with the emergence of bisphosphonates as a new therapy for osteoporosis in 1999. Her drive was that the duration the clinician and patient would have to wait to see a predicted 3-5% increase in bone mineral density does not lend itself to identifying non-responders promptly. Even after 2 years the expected change is only just above the least significant change in the individual (Ravn *et al.*, 1999). But despite the clear data as presented from the clinical trials, that BTM were a useful tool in population based research studies, it was only around 2000 that the clinical utility of BTMs in individuals for different aspects of bone health started to be explored (Fink *et al.*, 2000).

6.3 Relationship of Bone Turnover Markers to Fracture Incidence

Direct analysis of bone turnover markers with fracture outcomes are few. Retrospective analysis of the VERT trial of bisphosphonates, specifically risedronate, in 2003 by Eastell et al, has equated change in the bone resorption markers, urinary CTX and NTX directly with fracture reduction (Eastell *et al.*, 2003). As well as demonstrating that more than a 55-60% reduction in uCTX conferred no additional fracture risk reduction benefit, it also described that the reduction in fracture risk was as low as possible with any uCTX levels that were lower than -0.5 standard deviation from the premenopausal mean. From results from the 3 year MORE trial, analysis of the change in osteocalcin and bone alkaline phosphatase pre and post raloxifene treatment calculated an odds ratio for new vertebral fractures for each third of the cohort divided by BTM percentage change (Bjarnason *et al.*, 2001). It appears to show that beyond a reduction of 8.7ug/l in osteocalcin, there is no further reduction in fracture risk. It also shows that a third that did not achieve a large reduction in osteocalcin on treatment for a year, did not have any benefit from the treatment in terms of fracture risk compared to placebo. It would be impossible to determine from this study if the non-response was due to non-adherence. It also doesn't refer these values of osteocalcin back to the premenopausal reference range (Chapurlat *et al.*, 2000). The post-hoc analysis of the

Horizon trial, also demonstrated no difference in fracture risk in the subgroup with PINP supressed below the 95% lower limit of the reference range to the subgroup who did not (Delmas *et al.*, 2009).

Although in individual trials there appears to be a limit in the benefit in anti-resorptive treatment that can be determined by the percent reduction in bone turnover markers, a recent meta-regression by the FNIH Bone Quality Project has analysed 11 and more trials of bisphosphonate therapy (in addition to SERM trials). Two publications present data showing a clear relationship between percent reduction in bone turnover markers and the fracture risk reduction particularly of the vertebra (Bauer *et al.*, 2018; Eastell *et al.*, 2020). The Bauer paper shows a significant correlation in the percent reduction in bone formation markers only and the odds ratio of vertebral fracture, that does not plateau down to an included percent reduction in PINP of approximately 50% (Bauer *et al.*, 2018). The Eastell paper shows that 85% and 60% of the anti-resorptive positive effect on vertebral and non-vertebral fracture risk reduction can be explained by the 6 month change in PINP, with an even higher percentage of treatment effect for sCTX (Eastell *et al.*, 2020). This evidence should lead to the validation of the use of BTM primary outcomes in clinical trials for novel agents as a surrogate for effect on fracture risk.

6.4 Relationship of Bone Turnover Markers to Bone Mineral Density

Bone turnover markers taken in isolation have shown no correlation to the level of BMD amongst research groups and this was looked at again with no correlation between the BTM levels and BMD except in OC and total hip BMD, in small study in 2000 (Fink *et al.*, 2000). The variation in the BTM levels in this study arose from the cohorts of premenopausal, early postmenopausal and osteoporosis in later postmenopausal subjects. There have been multiple studies suggesting that the baseline bone turnover status, as indicated usually by the baseline biochemical markers of bone turnover seems to predict the response in bone mineral density to a treatment. The exception to this is teriparatide for postmenopausal osteoporosis, where studies are not concordant as to baseline markers predicting degree of positive effect (Burshell *et al.*, 2010). There is more a consensus for anti-resorptive treatment that if the baseline bone turnover markers are higher, regardless of osteoporosis

severity, then more of an effect is seen in bone mineral density and also fracture risk reduction as demonstrated for alendronate in the FIT trial (Bauer *et al.*, 2006).

6.5 Relationship of Bone Turnover Markers to Histomorphometric Measures

There are a few studies that have examined whether systemic bone turnover markers, bone ALP, PINP and sCTX in postmenopausal women with established osteoporosis correlate with pre-treatment bone turnover on a microscopic level (Chavassieux *et al.*, 2017). Although cellular level bone activity at one site, in this instance the iliac crest, does not represent overall skeletal risk, the significance of the correlation supports that bone turnover markers directly show what is occurring in trabecular bone.

6.6 Limitations in Using Bone Turnover Markers for Monitoring Treatment

The main challenge that arises with treatment with oral bisphosphonates is adherence to taking the medication and in the manner that is recommended for maximal absorption. Adherence is reduced by the inconvenience of the instructions of administration and possible gastrointestinal side effects (Eastell, 2007). The idea that monitoring bone turnover markers would highlight these problems to the clinician and encourage adherence was disputed. There is evidence for the use of bone turnover markers to encourage compliance from a study that showed conveying only a good reduction in the levels encouraged patients (Delmas *et al.*, 2007). However, there is also a RCT that showed that measuring the markers did not affect compliance as compared to nurse monitoring alone (Clowes, Peel and Eastell, 2004). The first study seemingly self-selects patients with a compliant nature and the second doesn't show positive results for use of bone turnover makers for improving adherence. However, many guidelines recommend the use of monitoring bone turnover markers to highlight when a patient is non-adherent to think about a change in treatment rather than promoting compliance (Kanis *et al.*, 2014). The IOF working group that convened in 2017, specifically to provide guidance on monitoring BTMs to identify nonadherence to oral bisphosphonate therapy (Diez-Perez *et al.*, 2017), based its finding on the TRIO study outcomes (Naylor *et al.*, 2016), and recommended measuring PINP and CTX at baseline and at 3 months after commencing treatment.

The International Osteoporosis Foundation working group on BTMs monitored the gathering of evidence with respect to using bone turnover markers for different purposes in clinical practice from 2011 to 2017 (Morris *et al.*, 2017). Initially the priority was to create standardised reference ranges and encourage investigations into the variability of the various markers of bone turnover. This has been predominantly achieved for multiple geographical areas including the USA, Australia and Mid Europe (Glover *et al.*, 2008; Eastell *et al.*, 2012; Jenkins *et al.*, 2013; Michelsen *et al.*, 2013), and the pre-analytical variability has been well documented for most of the markers now (Eastell *et al.*, 2017). Since 2017 Cavalier and collaborators have been working on the two recommended BTMs, CTX and PINP to describe the direct relationship between different assays (Cavalier *et al.*, 2019). There remains awareness that in certain situations, such as renal impairment, harmonisation of the assays for total and intact PINP will not be possible (Bhattoa *et al.*, 2021).

Using bone turnover markers to monitor treatment is only useful if the evidence can be translated into clinical practice. This means relevance to a cohort of individuals that can have multiple comorbidities or be on co-prescribed medications that also influence bone turnover and therefore the biochemical markers of it. With many of the studies using controlled osteoporotic populations with many exclusion criteria, more representative evidence is required.

7 Targets of Response to Treatment

It is fair to conclude that a positive change in markers of bone turnover have just as much validity from studies in controlled populations to be used as a surrogate for reducing fracture risk as a positive change in bone mineral density. The level of or change in bone turnover marker that has the potential to represent true risk reduction has been studied over the years with different treatments for osteoporosis.

7.1 Least Significant Change

The concept of least significant change (LSC) comes from data from using bone mineral density change in the individual. In bone densitometry the multitude of technical factors

determines a substantial short-term precision error, that is variable dependent on not only the machine, but also the age and habitus of the patient (Delmas *et al.*, 2000). For biochemical tests the product of the coefficient of variation and biological variability should be highlighted when analysing change. To be significant one must demonstrate a change that does not fall within the least significant change, and this was first explored in response to hormone replacement therapy (Hannon *et al.*, 1998) as shown in figure 1.7.

We are aware from clinical studies of some treatments, including the HRT work by Hannon, that the best degree of change in bone mineral density and best percentage change in bone turnover markers that is seen in the treatment group doesn't reach the LSC for an individual in one or other of these modalities. For example, raloxifene would not be expected to induce a response in either bone mineral density or sCTX that exceeds least significant change in an individual (Naylor *et al.*, 2010).

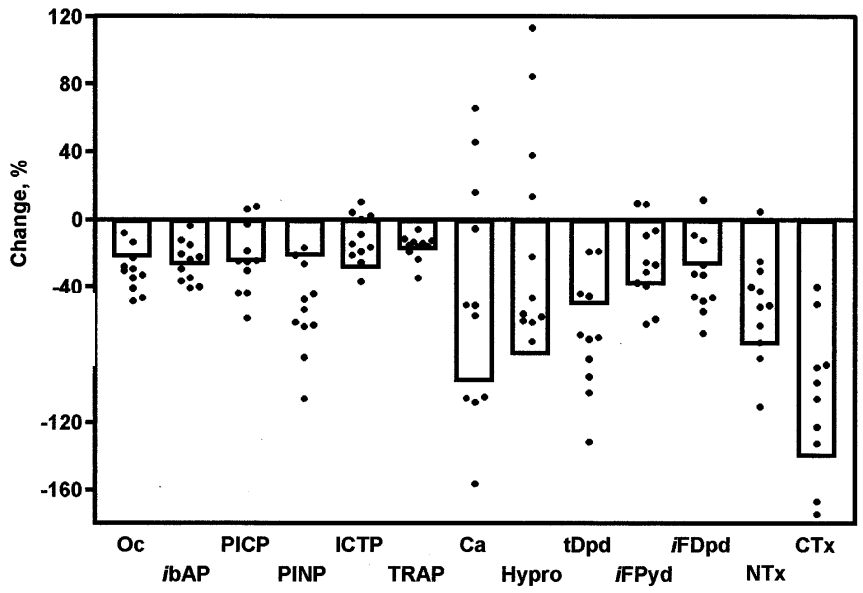


Figure 1.7. Least significant change of different bone turnover markers and individual percentage change with HRT over 24 weeks

(Hannon *et al.*, 1998)

With n=11, PINP and OC demonstrate the most sensitive result to identify responder by LSC (10, 9 respectively).

The work with raloxifene demonstrating the poor negative predictive value of response, elucidates the main concern of using least significant change as a biological target for

treatment (Kanis *et al.*, 2014). However early work with least significant change and bisphosphonates, demonstrates that it is a good way to demonstrate response to treatment with more sensitive markers of bone turnover such as sCTX, PINP and OC (Fink *et al.*, 2000). Although in this study, the least significant change was only calculated from four blood samples from each of 9 premenopausal individuals, the percentage values are very similar to the TRIO study in which two blood samples a week apart on treatment for osteoporosis were taken from 147 individuals (Naylor *et al.*, 2016). Interestingly it is CTX that demonstrates the most variability with a least significant change of 57.4% and 56% in each study compared to that of PINP of 38.3% and 38% (Fink et al and Naylor et al respectively). Despite this, sCTX remains the most sensitive in identifying subjects that achieve beyond this on alendronate treatment after a median 4 months in the Fink study with n=20 and after 12 weeks in the TRIO study with n= 51 reaching 85% and 98% achievement respectively. The response in PINP on 12 weeks of oral bisphosphonates in the TRIO study is seen in figure 1.8 with the alendronate achievement of LSC being 82%, with the equivalent in the Fink study 75% achievement (Naylor *et al.*, 2016; Fink *et al.*, 2000).

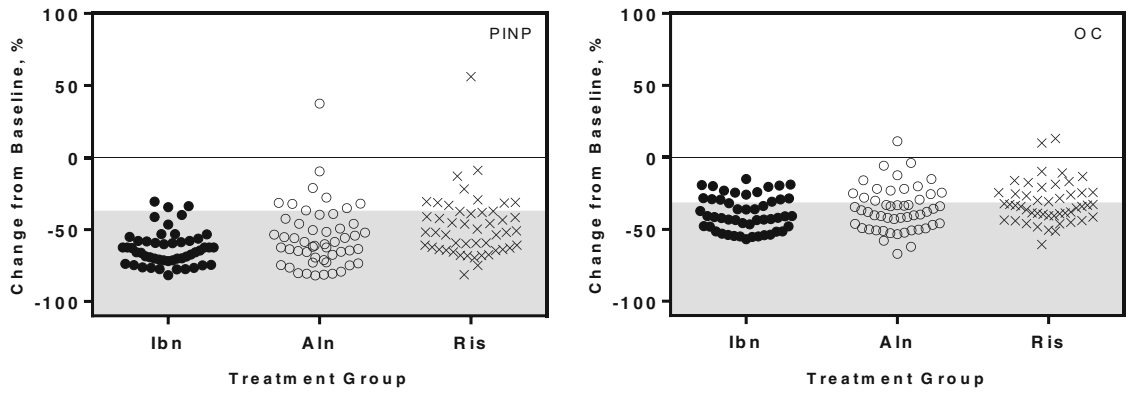


Figure 1.8. Least significant change for PINP and OC shown with the individual subject percentage change result at 12 weeks on three oral bisphosphonates

(Naylor *et al.*, 2016)

The greyed-out area is exceeding LSC. The three groups of n=57,57,58 for ibandronate, alendronate and risedronate.

It is for this reason that the International Osteoporosis Federation working group continue to support the use of sCTX in clinical practice, despite the pre-analytical requirements for an accurate measurement (Vasikaran *et al.*, 2011; Morris *et al.*, 2017). More specifically the IOF

working group on screening for adherence to oral bisphosphonates (Diez-Perez *et al.*, 2017) recommends using the achievement of LSC in PINP and CTX at 3 months to guide discussions on adherence.

7.2 Geometric Mean of the Premenopausal Reference Range

Although some studies have talked about the aim of treatment in postmenopausal women to be to render the levels of bone turnover into the premenopausal range (Garnero, Darte and Delmas, 1999), it has been shown that the range of bone turnover markers after menopause are not consistently and reliably above the premenopausal range (Kanis *et al.*, 2014). It is clear from the reviews that there are promoters of the use of targets in bone turnover markers on treatment (Eastell *et al.*, 2017) and sceptics(Kanis *et al.*, 2014). However, the analysis of the VERT results performed by Eastell et al in 2003 should not be ignored. The demonstration that there was a best expected response to risedronate in terms of fracture risk and this corresponded to bone turnover markers in the lower half of the premenopausal range. The geometric mean of the premenopausal range has since been termed the threshold value for treatment response with anti-resorptives.

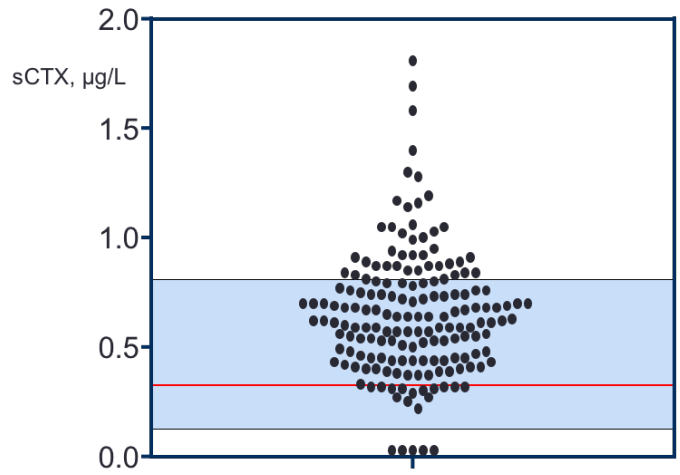


Figure 1.9. Baseline sCTX measurements in postmenopausal women (Naylor *et al.*, 2016)

The mean age of the n=149 women was approximately 67yrs old. Only 10% of the subjects had serum CTX levels below the geometric mean of the premenopausal reference interval (0.32ug/l) before any treatment.

The TRIO study (Naylor *et al.*, 2016) has clearly demonstrated that only 10% of postmenopausal women with T-score less than -2.5 have baseline CTX level under the threshold as shown in figure 1.9. A much smaller study by Fink et al also calculated the premenopausal mean but from 28 women with a mean age of 45.7 years (Fink *et al.*, 2000). From a multitude of bone formation and resorption markers, serum CTX by far demonstrated the most change in the early postmenopausal period with 73% of subjects being above the entire premenopausal range rather than just the mean. For PINP and OC this was 40% and 53% respectively.

7.3 Male Reference Range

Since 2011 more reference intervals studies have been performed demonstrating subtle differences in PINP and CTX reference intervals between men and women, in Australia (Roche Modular) and Germany (IDS ISYS) and also in both genders between different age groups (Jenkins *et al.*, 2013; Michelsen *et al.*, 2013; Chubb, Mandelt and Vasikaran, 2016; Chubb *et al.*, 2017). But despite having more established male reference intervals now, there appears to be no studies using the geometric mean of a male reference interval as a threshold for response to treatment, and how this might relate to change in bone mineral density. Equally there is no evidence of whether using the geometric mean of the female premenopausal reference interval which is far easier to obtain for each bone turnover marker with each assay can be applied as a biological target for men in clinical practice on osteoporosis treatment.

7.4 Sensitivity and Specificity

Sensitivity and specificity are terms used when evaluating a new measurement for clinical diagnosis or screening. They are widely understood by clinicians and easily reportable to patients (Hess *et al.*, 2012). The importance of having a target of treatment response set at a level for a high specificity is important to identify patients without response so as not to leave them on treatment that confers no positive effect on fracture risk. However it has generally been seen that the specificity of tests in osteoporosis is poor (Bergmann *et al.*, 2009) . Ravn in 1999 was the first to evaluate osteocalcin and urinary NTX together with bone mineral density change. Although the gold standard positive response to alendronate was a lumbar spine bone mineral density set at no change 0%, the best sensitivity and

specificity for urinary NTX to show treatment response was 86% and 48% respectively for a target change of -40%. 52% of subjects that showed no response to alendronate by BMD change would not be identified by urinary NTX. The performance of OC against hip BMD from the same study is shown in figure 1.10.

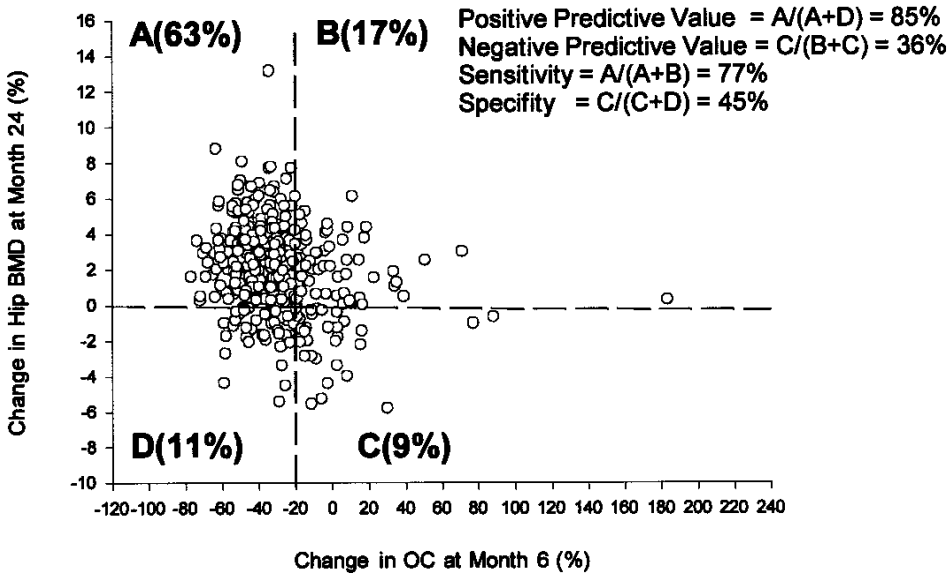


Figure 1.10. Subject response to alendronate by change in OC measurement at 6 months plotted against BMD change at the hip at 12 months

(Ravn *et al.*, 1999)

The legend shows the sensitivity and specificity at a change of 20% reduction on treatment.

One way to improve specificity without losing sensitivity by reducing the target of bone turnover markers on anti-resorptive treatment, was explored by the group of Delmas and Garnero. They looked at subjects on alendronate that were responders by a 3% change in lumbar spine bone mineral density and performed ROC curve analysis for using the change in bone alkaline phosphatase and the absolute value of BAP as predictors of this. Using both an absolute level of 9.5ug/l and a percentage change of -38.2% produced a specificity of 90% and a corresponding sensitivity of 72%. They do not describe how these values equate to the premenopausal reference range (Garnero, Darte and Delmas, 1999). The TRIO study and that of Fink et al, do describe the sensitivity of using the least significant change method (98%, and 100% respectively at 12 or 16 weeks for sCTX) and also for TRIO the threshold

method (96%) to identify responders to alendronate treatment as seen in figure 1.11 (Naylor *et al.*, 2016; Fink *et al.*, 2000).

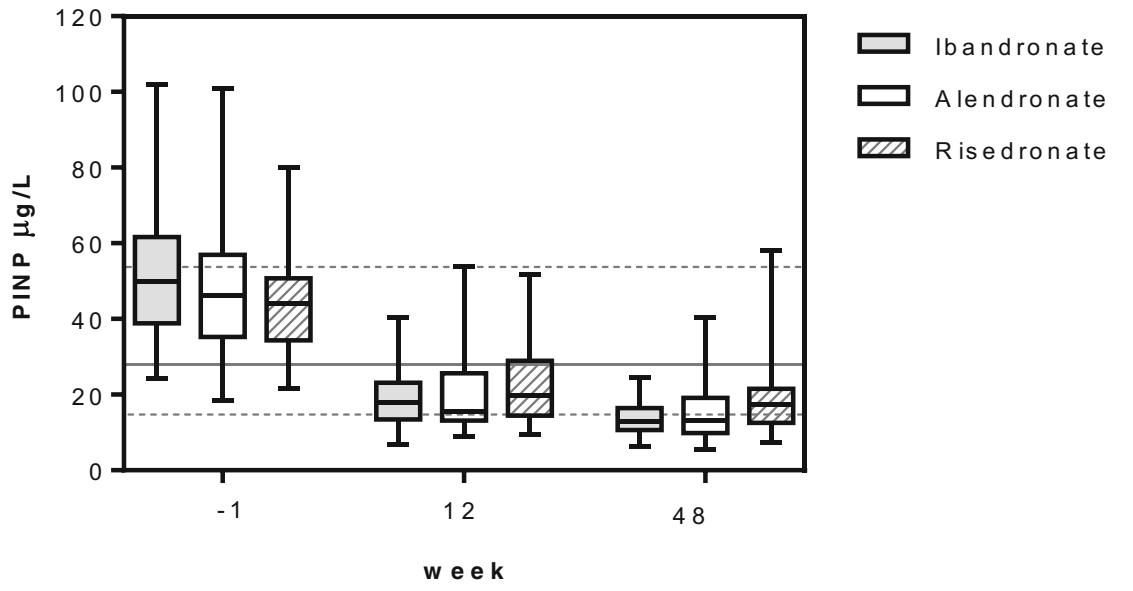


Figure 1.11. Response to three oral bisphosphonate treatments at 12 and 48 weeks by absolute PINP value.

(Naylor *et al.*, 2016)

The solid line is at the geometric mean of the premenopausal range of 28.3ug/l and at 48 weeks 100%, 94% and 82% of subjects suppress below this on ibandronate, alendronate and risedronate respectively.

However, because in TRIO a true positive was only dictated by the knowledge of whether an individual was on treatment, and there was no placebo group, specificity is not calculated. In the Fink et al study with only 16 patients, the gold standard response to alendronate was determined by change in bone mineral density (greater than LSC at 12 months). Although not clearly stated, the specificity of the least significant change method can be calculated at 75% for sCTX (3 out of 4 true negatives).

8 Summary

To answer a question about clinical utility of a tool, one first needs to understand the tool and what it is being used for, which in the case of the bone turnover markers and

osteoporosis, there is much informative evidence. The next step is to understand how to use the tool. For a biochemical test this includes understanding the assays and how they differ from each other, and a consensus statement in 2017 has stated that more evidence on this is required (Morris *et al.*, 2017) and in recent years more understanding of the relationship of different assays for PINP has been gathered from across Europe (Cavalier, 2019). However, it also includes setting the level of which the biochemical test is seen as significant. There is building evidence to support the use of least significant change and the geometric mean of the premenopausal range as biological targets (Naylor *et al.*, 2016; Paggiosi *et al.*, 2014). There has been limited studies up till now on the performance of these two biological targets in two populations; in men, and in those common in the clinical cohort with co-morbidities (Eekman *et al.*, 2011). Multiple different conditions, influence bone turnover, as do many drugs (Vlot *et al.*, 2018). Therefore, to see if the sensitivity of the tool at the biological targets remains acceptable to monitor osteoporosis treatment in the clinical population is crucial. The need also remains to evaluate whether the specificity of the tool can be improved to avoid missing non-responders to treatment.

9 Aims and Objectives

To define the geometric mean and 95% limits of the reference interval of two bone formation markers (procollagen I N-propeptide, PINP, and osteocalcin, OC) for men and premenopausal women and to calculate an absolute value of least significant change (LSC) for each bone turnover markers with each assay in women.

To evaluate the usefulness of performing two bone formation markers by two assays, in clinical practice using a single-centre, observational, retrospective, cross-sectional study.

To validate a target marker level (threshold) as a surrogate for responders to osteoporosis treatment by comparison with the least significant change (LSC) method in a trial patient cohort and by examining the rate of bone mineral density gain in the unselected patient and the trial patient cohorts.

9.1 Main research question:

- What is the sensitivity and specificity of procollagen I N-propeptide and osteocalcin at a target marker level, to identify the responsiveness of clinical patients on anti-resorptive treatments for osteoporosis? (The target is the threshold level defined as the geometric mean of a healthy premenopausal reference interval)

9.2 Secondary research questions:

- Is there a difference in the distribution of the healthy male and the healthy premenopausal female reference interval for procollagen I N-propeptide and osteocalcin?
- Is there method agreement in the combined measurements of the male and female reference groups for each individual bone formation marker, when performed with two different assays?
- What is the Least Significant Change expressed in percentage and absolute terms of PINP and OC with the two different assays, calculated from short interval paired data?
- Do two different assays for PINP and OC have an equivalent sensitivity and specificity of the biological target in identifying clinical patients on anti-resorptive treatments for osteoporosis?
- Is the proportion of female trial patients that achieve the LSC on an anti-resorptive for postmenopausal osteoporosis, in each of the bone formation markers by each assay, the same as the proportion achieving the threshold in the respective assay?
- Is there a significant difference in the rate of gain of bone mineral density between trial “responders” and “non-responders” to osteoporosis treatment defined by both the threshold and LSC biological targets for each bone formation marker assay?
- Is there a significant difference in the rate of gain of bone mineral density between the sub-groups of clinical patients identified by each bone formation marker as “responders” and “non-responders”?

Chapter 2 Materials and Methods

1 Overview

The study design consisted of three stages. The first was establishing the reference intervals for procollagen I N-propeptide (PINP) and osteocalcin (OC) in a UK population, and thereby the proposed biological targets to demonstrate treatment effect; least significant change (LSC) and threshold. The second was the evaluation of data of PINP and OC from the published research trial, TRIO, postmenopausal women on oral bisphosphonate therapy. The third stage was a cross-sectional observational study of PINP and OC in patients from a single centre in South Yorkshire who were on anti-resorptive treatment for osteoporosis. The distributions of PINP and OC at all stages were from measurements made on two different platforms, the IDS iSYS and Roche Cobas e411 auto analysers.

The measurements were made in February and March 2018. Sensitivity and specificity of the assays for each analyte at the biological target threshold to show treatment response were compared, in both the selected patients and clinical practice. The sensitivity was explored in comparison to response rate by LSC in the trial subjects and by studying bone mineral density change in the trial and clinical subjects.

2 Stage A: Reference Groups

2.1 Population

2.1.1 Participants

Three different control cohorts were selected from historic participants with archived trial samples:

- Female reference cohort; 130 healthy premenopausal women 30-45 years (with regular menstruation or any non-hormonal contraception) with a BMI less than 35kg/m^2 and 25-hydroxyvitamin D greater than 30nmol/L.
- Male reference cohort: 50 men aged 30-60 years with a BMI less than 35kg/m^2 and 25-hydroxyvitamin D greater than 30nmol/L.

- Female LSC cohort; 135 healthy postmenopausal (≥5 years) women 85 years or less, on oral bisphosphonates T-score ≤ -2.5 OR ≤ -1 with non-traumatic fracture with BMI between 18 and 35.

The historic participants of the trials were screened for influencers of bone metabolism.

2.1.2 Recruitment

Subjects were recruited within the Academic Unit of Bone Metabolism, University of Sheffield and had their samples archived in the South Yorkshire and North Derbyshire Musculoskeletal Biorepository (SYNDMB) with Oxford Rec reference 15/SC/0132 and Sheffield Teaching Hospitals Study number STH15691.

The study from which the female reference cohort originated was "TRIO" (Naylor *et al.*, 2016) for which 226 control cohort participants were recruited, sampled according to the above inclusion criteria from those with stored serum available for chosen time points.

The studies from which the male cohort originated were "Extreme CT" (XCT) (Walsh *et al.*, 2017) in which 90 men were recruited into age categories, all with a BMI less than 35kg/m^2, and "Fat and Bone" (PRP) (Walsh *et al.*, 2016) for which 126 men were recruited into BMI and age categories. These historic subjects were sampled according to the inclusion criteria above from those with stored serum available.

The female LSC cohort also originated from "TRIO" (Naylor *et al.*, 2016), from the 180 treatment cohort subjects at baseline, sampled according to inclusion criteria above from those with stored serum available for chosen paired time points of week 12 and week 13.

With one week between the two samples these subjects represented a short term intra-individual variation, with incorporation of the intra-assay variation. The value of LSC was representative of variability in a sample with osteoporosis, but on bisphosphonate treatment (with no other influencers of bone metabolism).

2.2 Interventions

There were no interventions in the male and female reference cohorts.

Ad Cal D3 was initiated at the baseline visit (week -1) of the 135 subjects of the female LSC cohort. Oral alendronate 70mg weekly or risedronate 35mg weekly or ibandronate 150mg monthly was commenced at week 0, in the same cohort.

2.3 Assessments

2.3.1 Sample collection and storage

Morning fasted samples taken in 2012-2013 (PRP) (Walsh *et al.*, 2016), 2013 (XCT) and 2007-2009 (TRIO) (Naylor *et al.*, 2016) remained in storage at -80 ^{oC} in the SYNDMB biorepository as aliquots. Upon successful application for use of these samples (SYNDMB045) they were released and remained frozen in the University of Sheffield Bone Laboratory until the time of sample measurement.

2.3.2 Baseline characteristics

I requested the following baseline characteristic data of each subject for which an archived sample was used, from the study records at recruitment. I only had access to the non-identifiable study number:

- Age
- BMI
- BMD at the proximal femur (total hip) from densitometry performed for the purpose of research. This was used to calculate the T-score for the baseline characteristics using the gender-specific NHANES III database (Shuhart, Yeap and al, 2019).

2.4 Outcome Measures

From the biochemistry measurements detailed below, 25-hydroxyvitamin D was measured to reconfirm eligibility. PINP and OC with two assays was used to calculate the reference intervals of the male and female reference cohort, described by the geometric mean (GM) and lower and upper 95% limits. The GM of the premenopausal reference interval was the outcome that represented the biological target of the threshold and the paired data results produced the other biological target of the least significant change. The methods by which this was done is described in the statistics section.

3 Stage B: Archived Trial Treated Cohort

3.1 Population

3.1.1 Participants

A cohort of 135 all female subjects named the “Trial cohort” were on oral bisphosphonate treatment for 48 weeks and acted as their own controls before starting oral bisphosphonate treatment. The subjects were sampled from historic trial participants determined by the availability of the archived serum aliquots with the originally determined inclusion criteria:

- Postmenopausal (≥5 years) women
- 85 years or less
- T-score ≤ -2.5 OR ≤ -1 with non-traumatic fracture
- BMI between 18 and 35.

The historic participants of the trials were screened for influencers of bone metabolism.

3.1.2 Recruitment

These archived trial samples had also originated from “TRIO” (Naylor *et al.*, 2016), from the 180 treatment cohort postmenopausal women at baseline and at week 48 (149 retained subjects), excluding recent fracture or condition or medication effecting bone metabolism.

3.2 Interventions

The participants were randomized to oral alendronate 70mg weekly, risedronate 35mg weekly or ibandronate 150mg monthly, as well as receiving calcium carbonate 3 g (1200 mg elemental calcium) and cholecalciferol 20 micrograms (800 IU) a day. The Ad Cal D3 was commenced at the baseline line visit (week -1) and the oral bisphosphonate at week 0.

The adherence to the medication had been monitored using medical events monitoring system (MEMS).

3.3 Assessments

3.3.1 Sample collection

Morning fasted samples taken in 2007-2008 (TRIO) (Naylor *et al.*, 2016) at baseline (week -1) and at week 48, and remained in storage at -80 ^{oC} in the SYNDMB biorepository as aliquots. Upon successful application for use of these samples (SYNDMB045) they were released and remained frozen in the University of Sheffield Bone Laboratory until the time of sample measurement.

3.3.2 Baseline characteristics

I requested the following baseline characteristic data of each subject for which an archived sample was used, from the study records at recruitment. I only had access to the non-identifiable study number:

- Age
- BMI

3.3.3 Clinical dual-energy x-ray absorptiometry review

Within the TRIO study, bone densitometry assessment was performed at the baseline visit and at week 48 by Dual-energy x-ray absorptiometry (DXA) with Discovery A densitometer (Hologic Inc). I used the proximal femur (total hip) bone mineral density (BMD g/cm^2) for the Trial cohort at both these time points for this study, but I did not include all 135 sets of data for analysis, in instances of more than 10% weight change between the two scans (Yu *et al.*, 2012), and in instances of incomplete data sets (including missing PINP and OC by both assays at either time point).

Total hip T-score for the baseline characteristic of the Trial cohort, was calculated from BMD using the young female NHANES III database (Shuhart, Yeap and al, 2019).

3.4 Outcome Measures

25-hydroxyvitamin D was measured at both time points to characterise confounders of bone turnover marker (BTM) measurements. The PINP and OC with both assays was measured at both time points with the primary outcome of the Trial cohort, to compare the

approaches of using the biological targets of LSC and threshold for identifying women that respond to oral bisphosphonate therapy. It was then assessed if a response in the BTMs by these two approaches was associated with change in bone mineral density.

4 Stage C: Clinical Patient Groups

4.1 Population

4.1.1 Participants

To represent mixed gender clinical patients on anti-resorptive treatment for osteoporosis of any type, three cohorts were identified from the recruitment of all-comers attending a Metabolic Bone Unit to that of the SYNDMB repository:

(a) Clinical cohort on oral bisphosphonates for at least 3 months (n=86).
(b) Zoledronate cohort receiving at least the 2nd annual injection (n=99).
(c) Denosumab cohort receiving at least the 2nd 6 monthly injection (n=15).

First the treated groups of each cohort were gathered, then with the baseline characteristics known at recruitment I used an excel formulated spreadsheet to identify matched patient no treatment groups, for each cohort. They were recruited to SYNDMB whilst attending for bone densitometry for investigation or monitoring of osteoporosis but had been on no osteoporosis treatment for 12 months at recruitment (except calcium or cholecalciferol). Matching as best fit, was based on gender, age (round down +/- 3 years), body mass index (+/- 2.5) and Tscore of the total hip (+/- 0.3).

To represent the breadth of clinical variation of a typical osteoporotic population, provided each patient had the capacity to consent, there were no exclusion criteria.

4.1.2 Recruitment

Subjects were recruited within the Academic Unit of Bone Metabolism, University of Sheffield to the South Yorkshire and North Derbyshire Musculoskeletal Biorepository (SYNDMB) with Oxford Rec reference 15/SC/0132 and Sheffield Teaching Hospitals Study number STH15691. From a recruitment date of February 2016 there was enough clinical

information about the participants obtained to identify subjects for the treatment cohorts and the matched non-treatment group. Recruitment continued over 24 months with varied recruitment pathways:

- Between February 2016 and April 2017 recruitment of all participants was from the population of patients attending the Sheffield Metabolic Bone Centre for fracture risk assessment (FRAS), all of whom received an information letter about the SYNDMB with their appointment letter. Patients that required blood tests as part of their usual clinical care were identified by bone densitometry technicians and asked if they would give extra blood samples for research. Some patients volunteered themselves. This pathway remained open for recruitment for the duration. It was the main source of recruitment for participants eligible for the matched non-treatment groups.
- Between October 2017 and December 2017, patients attending for a DXA scan for FRAS or clinic review were approached based on review of the referral documentation (scan packs), who were suspected to have been on particular treatment for osteoporosis. The bone densitometry technicians received an alert on the patients' scan packs to ask them if they were happy to meet the research team. This pathway increased subjects in the oral bisphosphonate Clinical cohort.
- Between April 2017 and January 2018 participants were recruited from the day unit who had attended at least 12 months after parenteral osteoporosis treatment. These patients attended through two routes: "Direct Access" for IV zoledronate treatment after a fracture without a previous DXA scan, or from primary care requiring treatment for osteoporosis but intolerant of oral therapy and outpatient clinic patients on IV zoledronate, IV ibandronate or SC denosumab. The patients that required cannulation for their treatment were approached by the day unit staff and asked if they would be happy to speak to the research team.
- From October 2017 I specifically identified 2 weeks in advance, patients that were booked to attend for parenteral treatment. We contacted them by telephone and sent the patient information sheet in the post. An alert was given to reception staff to contact the research team at the patients' attendance. The preliminary communication with the patients was conducted to maximise recruitment in the latter stage.

4.2 Interventions

There was no specific intervention for the matched no treatment groups, although 50% were likely to be on calcium or cholecalciferol and this was recorded.

The Clinical cohort on oral bisphosphonates, treatment group had received treatment for at least 3 months of oral alendronate 70mg weekly, risedronate 35mg weekly or ibandronate 150mg monthly, with a record made of calcium and cholecalciferol supplementation. Checking adherence was not seen as an intervention.

The Zoledronate cohort treatment group had received an annual zoledronate 5 mg dose intravenously or ibandronate 3mg every 3 months intravenously for at least one year.

The Denosumab cohort treatment group had received subcutaneous denosumab 60 mg 6-monthly for at least 6 months.

4.3 Assessments

4.3.1 Sample collection

Serum and plasma samples were collected from the recruited subjects at any time of the day by the BD Vacutainer® blood collection system into two SST II Advance 8.5ml bottles and one spray-coated KEDTA bottles. If all bottles were not available whilst maintaining the subject's comfort, then any volume of venous blood attained was processed. Blood samples were attained by venepuncture or from a cannula inserted for the purpose of giving parenteral treatment for osteoporosis. The process of sample collection and handling has been described in Appendix B.

The aliquots remained in storage at -80 ^{oC} in the SYNDMB biorepository. Upon successful application for use of these samples (SYNDMB045) they were released and remained frozen in the University of Sheffield Bone Laboratory until the time of sample measurement.

4.3.2 Baseline characteristics

The study recruitment log for the South Yorkshire North Derbyshire Musculoskeletal Biobank collection (STH15691), recorded the anthropometric data of age at recruitment, height and weight and gender.

4.3.3 Workbook

When a patient was recruited to SYNDMB, a workbook was completed (V1.0 16Dec15) as in Appendix A, with the following information;

- Anthropometric data; Height and weight
- Questionnaire check (see below) with any updates required to older information
- Medication check
- Supplements check

4.3.4 Metabolic bone questionnaire

All patients attending the Metabolic Bone Centre, Northern General Hospital, Sheffield for a first clinic review or subsequent DXA scan, had completed a Metabolic Bone Centre Questionnaire (PD6421-PIL2472 August 2011, March 2014, September 2016) as seen in Appendix C. This contained self-reported information about risk factors and fracture history. From this I recorded the following information:

- Osteoporosis treatment details including start date of the current therapy or end date of the last completed therapy, and previous therapy.
- Vertebral/non-vertebral fracture history above 40 years old; particularly any within 12 months prior to recruitment
- Co-morbidities including coeliac, diabetes, thyroid disease, parathyroid disease, or renal disease
- Details of co-prescribed bone supporting medication; calcium, cholecalciferol or activated vitamin D
- Details of co-prescribed medication effecting bone metabolism in the past or present; corticosteroids, aromatase inhibitor therapy, GnRH agonist therapy
- Details for female subjects of ongoing periods, hysterectomy, hormone replacement therapy or parenteral progesterone

4.3.5 ArQ database

A summary of the clinical information gathered with the two above tools could be extracted from the Metabolic Bone Centre computerised database “ArQ”. As well as creating a series of queries of the database to extract large amounts of data in CSV files, I could also use

individual SYNDMB biobank study numbers (SBM)HM2…. “HPP Biobank Ref” to clarify details.

As well as the information inputted by the research team, the ArQ database detailed the clinical contacts the patient had had with the Metabolic Bone Centre. These included nurse appointments for the Day Unit for a parenteral treatment for osteoporosis, from this information, the date of onset of the therapy and whether there was more than one month delay from its advised dosing interval at recruitment, could be ascertained. If a patient was unsure how long they have been on oral bisphosphonate, this was sometimes recorded on ArQ or within sequential bone densitometry formal reports that were saved to Sheffield Teaching Hospitals radiology (PACS) or results (ICE) server.

The co-prescribed medications of interest were those that had the potential to influence bone formation markers;

- Corticosteroids- Current dose within the last week and whether the patient had an extended oral (>3 months) or parenteral course within the past (more than 3 months prior to recruitment to five years earlier).
- Aromatase inhibitor therapy- past (more than 3 months prior to recruitment) or present and which one.
- Gonadotrophin releasing hormone analogue/agonist therapy- past (more than 3 months prior to recruitment) or present and which one.
- Calcium and either cholecalciferol or activated calcitriol/alfacalcidol.
- Hormone replacement therapy (see menopausal status) or parenteral progesterone (Depo-Provera) if the patient reports no periods.

The ArQ database was also interrogated for clinical biochemistry related to bone health. Only blood screens organised by the Metabolic Bone Unit, related to the “Fracture Risk Assessment Service” were transferred to ArQ from Sheffield Teaching Hospitals results server “ICE”. These were searched for measurement of calcium and creatinine within 6 months of recruitment as well as TSH and testosterone for men. The measurement closest to the sample at recruitment was used for creatinine, where there was multiple taken within 6 months. To complete the creatinine data set for all subjects in the ‘on treatment'

groups of the three cohorts, a further search was completed of clinical system “Open ICE” to see if out of area results were available.

4.3.6 Menopausal status

The unselected patient cohorts consisted of male and female subjects. With no specific age inclusion, some female subjects were potentially premenopausal. Their status was determined by their response on the questionnaire as to whether they continued to menstruate and if they did not, the age of their last period.

As the absence of menstrual periods could be for variety of reasons, including menopausal status, and as FSH was not measured in this study, the following were taken into account to conclude menopausal status:

- Age ≥55 years Postmenopausal in the absence of positive or negative record of menstrual cycle (with note of HRT)
- Age 45-55 Unknown status if presence of menstrual cycle was not recorded or absence of menstrual cycle in context of Hysterectomy without oophorectomy
- Age ≤45 Premenopausal in the absence of positive or negative record of menstrual cycle, or absence of menstrual cycle in context of Hysterectomy without oophorectomy or contraceptive administration
- Age ≤ 45 Premature Menopause if a note of HRT prescription or absence of menstrual cycle in context of oophorectomy/treatment for breast cancer, or no hysterectomy or contraceptive administration

4.3.7 Radiology review

The fracture history for each subject could be predominantly obtained from ArQ. I had to process the data extracted to note any large bone fractures (exclusive of hand/feet/rib) that had occurred in the 12 months prior to recruitment. The reported fracture history above 40 years old (deemed fragility fractures) was characterised as vertebral or non-vertebral.

4.3.8 Clinical dual-energy x-ray absorptiometry review

The ArQ database was interrogated for all densitometry data available regarding locally performed clinical DXA scans for the three cohorts, which were performed on one of three

Hologic scanners (2 Horizon and 1 Discovery). The total hip measurement was used as this was in keeping with the Metabolic Bone Centre clinical reporting standards and treatment recommendations. Out of all available, two or three DXA scan results were identified to fulfil the following criteria:

- A DXA result at treatment commencement (12 months before or after) for estimate of pre-treatment bone mineral density
- A DXA result on treatment (1-5 years after the treatment commencement DXA date for the Clinical cohort on oral bisphosphonates and 1-3 years for the zoledronate and denosumab cohort. The difference is based on local practice.)
- A DXA result closest to recruitment date for baseline characteristic at the point of blood sample. This was used to calculate the T-score for the baseline characteristics using the young female NHANES III database (Shuhart, Yeap and al, 2019)

In instances of incomplete data sets, use was made of clinical IT systems such as the Sheffield Teaching Hospitals PACS system (on which DXA studies back dated to 2011 were to be found).

Not all patients had qualifying scans included in the rate of change of bone mineral density analysis. Factors that disqualified the BMD change results were;

1. Inconsistency in scanner used
2. Change in the side of proximal femur scanned (right or left)
3. More than a 10% weight change between the two scans
4. Any change in medication prescription between the two scans that could have affected the overall bone health over the time interval, but not the singular BTM
5. Subjects with missing proximal femur data at either of the sequential scans on oral bisphosphonate, or missing PINP and OC by both assays on treatment.

The following identified medication changes affected point 4 above:

- Extended glucocorticoid treatment (greater than three months) that was commenced prior to the baseline DXA and ceased in-between the baseline and subsequent DXA on treatment.
- Aromatase inhibitor therapy that was commenced prior to the baseline DXA and ceased in-between the baseline and subsequent DXA on treatment.

- GnRH analogue therapy that was commenced prior to the baseline DXA and ceased in-between the baseline and subsequent DXA on treatment.

The BMD percent change for the three cohorts was expressed as an annualised percent change as the DXA scans were conducted with a one to five year interval.

4.4 Outcome measures

On each of three patient cohorts we measured PINP and OC on the IDS-iSYS and Roche Cobas e411 assays at the point of recruitment in addition to 25-hydroxyvitamin D and intact PTH in order to characterise confounders of the distributions of PINP and OC.

Through the Trial cohort I introduced and started to validate the use of a single measurement of BTM on anti-resorptive treatment and evaluated it against the biological target of threshold. The main outcome from the clinical cohorts was calculations of sensitivity and specificity of PINP or OC set at the threshold to detect treatment, to assess how threshold performed in the “real life” setting.

5 Biochemistry Measurements

5.1 Auto Analyser Method

The biochemical measurements were made from February to March 2018 with the same aliquot being sampled for measurements on the Immunodiagnostic Systems (IDS) iSYS auto analyser and the Roche Cobas e411 auto analyser.

The N-MID assay of both Roche and IDS measured the large fragment of amino acids (aa) 1 to 43 after protease cleavage of the unstable 49 aa intact osteocalcin between aa 43 and 44 as well as intact OC. The two monoclonal antibodies used in the assay oppose the N-terminal, and N-mid fragment, without targeting the cleaved C-terminal fragment. In both the IDS and OC assays the monoclonal antibody that joined with an epitope on the N-Mid fragment was labelled with biotin, that after the first incubation would combine with streptavidin coated microparticles in a second incubation. This combination facilitated the intact and N-mid fragment of OC to be magnetically captured onto an electrode. It is the second monoclonal antibody that differed between the IDS and Roche assays in their

labelling of acridinium and ruthenium complexes respectively. Once the unbound substances were washed away, only the “captured” fragments were triggered to emit chemiluminescence from these labels, that was measured as directly proportional to the concentration of OC in the serum sample.

There has been little published about the epitopes with which the two monoclonal antibodies used in the sandwich method of the electrochemiluminescence immunoassays of PINP from IDS and Roche, interact. However the biotin labelled monoclonal antibody and ruthenium labelled monoclonal antibody used in the Roche Cobas e411 autoanalyser detected both the trimer (two pro-α1 and one pro-α2 chains) and monomeric forms. Both were captured by the biotin-streptavidin solid phase complexes after 2 incubations and triggered to luminesce from the ruthenium. This represented the total PINP. Whereas the biotin labelled monoclonal antibody and acridinium labelled monoclonal antibody used in IDS iSYS autoanalyser, only opposed, captured and luminesced the intact three subunit chains (trimer), also by the interaction with streptavidin coated microparticles to magnetise the intact PINP.

5.2 Outcomes

The biochemical measurements were made in the bone biochemistry laboratory within the same calibration period. Most subjects will have a singular sample for the day of recruitment. However, the Trial cohort subjects had samples to process from both baseline (pre-treatment), and on treatment.

Serum (0.5- 1ml) for all cohorts;

- PINP (Roche Total, IDS Intact)
- OC (Both Roche and IDS N-MID)
- 25-hydroxyvitamin D (IDS iSYS only)

Plasma (0.5ml) for three clinical cohorts only;

- Intact PTH (IDS iSYS only)

PINP and OC have a better signal to noise ratio than bone alkaline phosphatase (BAP), are more stable to sample handling than Tartrate-resistant acid phosphatase 5b (TRACP5b), and do not demonstrate diurnal and post prandial variation like C-terminal telopeptide (CTX).

They were therefore chosen to be the most practical bone turnover markers in clinical practice, which was relevant to be able to answer the research questions.

A large majority of clinical laboratories in the UK have Roche auto analysers, and our bone biochemistry laboratory has the Roche Cobas e411. The IDS iSYS auto analyser uses an assay for intact rather than total PINP, so was chosen as a good comparator and was advocated by the sponsorship of the study by, Immunodiagnostic Systems, Bolton.

25-Hydroxyvitamin D (25OHD) was measured on the IDS iSYS auto analyser for the entire studied cohort. The 25OHD measurements was then directly comparable between all groups, in order to use this as a descriptor of a factor that can affect bone turnover.

The BTMs measurements have been represented in micrograms/Litre as the international standard and 25-hydroxyvitamin D as nanomol/Litre for the UK recognised standard and PTH as nanograms/Litre for SI units.

The clinical creatinine measurements were processed to stratify as CKD group (Stevens, Levin and Members, 2013), calculated by eGFR-EPI 2009 formula (Levey *et al.*, 2009). The equation took into account the age at recruitment and sex of the subjects but was not adjusted for subjects of black ethnicity due to the absence of this information. It did not represent the creatinine clearance as per the Cockcroft-Gault calculation.

Alongside the record of clinical calcium measurements was the measurement of intact PTH on plasma with the IDS auto analyser. Plasma PTH was chosen for its stability as compared to serum. It was only measured during this study period on the clinical cohorts due to its relevance for clinical co-morbidities. The Trial cohort had intact PTH confirmed by the IDS iSYS auto analyser measurement for the TRIO study in 2013.

The results for albumin adjusted calcium were filtered to any above 2.6mmol/L and compared against the iPTH measurement to determine if there was evidence of PTH dependent hypercalcaemia.

Outcome	Source of measurements	Site of measurements
Parathyroid hormone (iPTH) (ng/L)	IDS iSYS analyser	Bone Biochemistry - AUBM
25-OH vitamin D (nmol/L)	IDS iSYS analyser	Bone Biochemistry - AUBM

Procollagen I N-terminal propeptide (PINP) (µg/L)	IDS iSYS analyser	Bone Biochemistry - AUBM
Osteocalcin (OC) (µg/L)	IDS iSYS analyser	Bone Biochemistry - AUBM
Procollagen I N-terminal propeptide (PINP) (µg/L)	Roche Cobas e411 analyser	Bone Biochemistry - AUBM
Osteocalcin (OC) (µg/L)	Roche Cobas e411 analyser	Bone Biochemistry - AUBM
Creatinine (µmol/L)	Roche Cobas 702 Jaffe Assay	STH Clinical Laboratory
Adjusted Calcium (mmol/L)	Roche Cobas 702 NM-BAPTA assay	STH Clinical Laboratory

AUBM Academic Unit of Bone Metabolism; STH Sheffield Teaching Hospitals

Roche Diagnostics, North America and ImmunoDiagnostic Systems, UK

6 Statistical Analysis

6.1 Statistical Power

The gold standard for a sample size for reference interval determination, is 120 (CLSI, 2008). However the CSLI document of 2008 introduced a robust method of determining the limits of a reference interval when neither the parametric or non-parametric methods would be suitable, particularly for sample sizes less than 120 (Henny, 2009).

In order to calculate sample size required for accurate sensitivity and specificity with 90% power, we used power tables from a recommended paper by the statistics department (Hess et al. 2012). We decided on what we expected the sensitivity and specificity to be evaluated as, and we had to state performance targets of what would be acceptable. As for performance targets in the field of osteoporosis, we considered one of the best tests to be the total hip bone density for diagnosis of osteoporosis, in hip fracture patients. When we used this as a diagnostic test, then total hip bone density had a sensitivity of 80% for a specificity of 60% (Yang et al. 2009). As a higher specificity was important in the use of a target of bone turnover marker to suggest treatment response, we increased our performance target of specificity to 75%. but this logically compromised the sensitivity

which was also set at 75%. Using the tables (Hess et al. 2012) in Appendix D; we were aware that acquiring a total of 134 or 118 subjects in each cohort (50% untreated and 50% treated) would power the sensitivity and specificity outcomes respectively. We accepted that this would not be possible in the Denosumab cohort due to local pathways of administration. However, a number of 20 treated and 20 untreated would have been a sufficient sample size with the expectation of the sensitivity in particular to be above 0.9, and with a lower precision of 0.20 instead of 0.10.

6.2 Statistical Programme

I used SPSS to create three combined data sets for each cohort:

1. All the above biochemical measurements, alongside the baseline demographics
2. All of the clinical data and all available DXA attendance data described in Phase C, alongside the baseline demographics
3. The relevant sequential DXA data alongside determined response status, with two cohorts together, for direct comparison

I used a combination of SPSS (IBM, version 24-26 for MacOS) and MedCalc (MedCalc Software Inc, version 16-19.8) to perform the statistical analysis, and MedCalc and GraphPad Prism (GraphPad Software, Inc, version 8-9) to create figures.

6.3 Assessment of Pattern of Distribution

To make a decision on whether to use a parametric or non-parametric statistical analysis, I first predominantly used normal probability plots, alongside histograms, to check whether there was a linear relationship and following a gaussian curve.

Secondary to this when performing the chosen statistical test for parametric data, I checked the built-in test for normality which was predominantly the Shapiro Wilk test.

The final analysis I performed to study the pattern of distribution, was to see whether when the BTM measurements were log transformed, they then fulfilled a test for normality by Shapiro Wilk. The predominant outcome of the groups ‘on treatment’ was that they did not, and hence I chose non-parametric testing in comparison of any of the BTM distributions, except the reference range analysis.

For the reference range analysis, the raw measurements of PINP and OC in the female and male reference groups did not satisfy a normal distribution and therefore were log transformed to fulfil the properties of a normal distribution.

6.4 Calculation of Biological Targets

The biological targets of threshold and least significant change were derived from the premenopausal reference interval. Both the premenopausal reference interval and the male reference interval were described with the geometric mean and upper and lower limits of the distribution with 95% confidence.

The inverse log of the mean of the log transformed reference group data is the geometric mean. This can also be calculated directly from the raw data using the following formula;

$$geometric\ mean = \sqrt[n]{a_1 a_2 a_3 \ldots . a_n}$$

The 95% confidence limits of the reference interval represent the upper and lower values of the distribution, between which 95% of the values lie. The interval is estimated by using the z value of 1.96 (above and below the mean). Log transformation was required to standardise the distribution of each BTM to apply this calculation. Therefore the mean (χ) and standard deviation (σ) in the following formula are of the log transformed data;

$$Upper\ and\ Lower\ 95\%\ Limits = 10^{(\chi \pm (1.96(\sigma)))}$$

The least significant change (LSC) for PINP and OC with each assay was calculated as an absolute value from the mean standard deviation (σ) of the paired values and a percentage value from the mean intra-individual coefficient of variation (CV_I) of the paired values.

$$LSC\ absolute = 1.96\left(\sqrt{2}\right)\left(\sqrt{\frac{\sum \sigma^2}{n}}\right)$$

$$LSC\ \% = 1.96\left(\sqrt{2}\right)\left(\sqrt{\frac{\sum CV^2}{n}}\right)$$

This is the same method as is used to calculate the LSC for a change in bone mineral density in clinical practice. I used the ISCD template for calculation, expanding it for 135 pairs (ISCD, 2017).

6.5 Method Agreement

To examine the relationship between the measurements produced by assaying the serum with the IDS iSYS and Roche Cobas auto analysers, method agreement analysis by Bland and Altman and Passing and Bablok was performed. The Bland and Altman method determines a mean offset and prediction limits, and its regression determines whether there was a statistical bias in the relationship between the two assays for high or low values. The Passing and Bablok regression analysis calculated the slope and intercept with 95% confidence intervals to compare against that of perfect agreement between the two assays.

6.6 Comparison of Baseline Characteristics

Differences in the characteristics that can affect bone turnover were analysed with a parametric 2- tailed t test if satisfying normality, or when a normal distribution was not satisfied, a 2-tailed Mann Whitney.

The analysis was predominantly between independent samples, with the one exception being within Trial cohort group analysis, which was always paired.

For both the mean comparison and rank comparison, 95% confidence intervals were selected, determining the critical p-value of 0.05. However, for the comparison in Chapter 4, which was a 4 way comparison of intra and inter cohort groups on oral bisphosphonates, the critical p-value was adjusted to make allowances for chance findings, to p=0.0125.

6.7 Comparison of BTM Distributions

The distributions of the bone turnover markers were unlikely to satisfy normality. For the reference interval data this was managed by performing the statistical comparison of means on the log transformed data, with an independent 2-tailed t-test between men and women, but paired for the comparison between the assays of each BTM. For the comparison 95% confidence intervals were selected, but the critical p value was adjusted for multiple comparison of each mean to p=0.025.

The non-parametric data of PINP and OC in the treatment cohort groups was described by the median and interquartile range (IQR). Comparison have been made using rank comparison of a Mann-Whitney test on independent samples or a signed rank Wilcoxon test

for the paired samples of the two treatment groups of the Trial cohort. For Chapter 4 there was a four-way analysis, so the critical p-value was adjusted to 0.0125. However in Chapter 5, only the two groups within each cohort were compared, so the p-value with 95% confidence was 0.05.

6.8 Diagnostic Accuracy of BTM

6.8.1 Sensitivity and specificity of biological targets

With the biological targets established and the measurements of PINP or OC by each assay obtained for all the anti-resorptive treatment cohorts, the following table was completed.

BTM y with *assay* z		**Treatment status**	
		Subjects on osteoporosis treatment	Matched subjects not on/prior to treatment
Measurement	Reaching/beyond Threshold	A	C
	Not reaching Threshold	B	D

$$Sensitivity = \frac{A}{A + B}$$

$$Specificity = \frac{D}{D + C}$$

Table 2.1. Treatment status vs Biochemical Measurement to calculate sensitivity and specificity for each cohort by Threshold.

The sensitivity for LSC was only calculated on the proportion of the Trial cohort that achieved the assay-specific LSC or greater reduction.

6.8.2 ROC analysis

Receiver operator curve analysis was another way of demonstrating the accuracy of a tool, in this instance BTM, to differentiate between a positive and a negative, which is either “on treatment” or “no treatment” respectively.

In simple terms it showed the balance between the sensitivity and the specificity when the threshold was any given value. Two "tools", such as the two BTMs, PINP and OC, or the two assays of each BTM were compared by the statistical analysis of the difference in the area under the curve. It could also be used to look at the performance of a single tool, either PINP or OC in two settings, such as the Trial and Clinical cohort in Chapter 4, or the Zoledronate and Denosumab cohorts in Chapter 5. For clinical decision making tools, an AUC of 0.8 is considered acceptable (Hess *et al.*, 2012).

6.9 Change in Bone Mineral Density

I have expressed two sequential eligible measurements of BMD as a percentage change from the baseline. Where this percentage change was over an interval more than 1 year, I calculated the annualised BMD percentage change. The distribution of the annual BMD percentage changes were divided into a subgroup on treatment that achieved the biological target and a subgroup on treatment that did not achieve the biological target. The difference in the mean of these groups was tested by independent samples t-test with 95% confidence intervals.

Any correlation between individual subject BTM and annual BMD percent change was also explored. Linear regression was not appropriate in this context because the BTM measurements in themselves were an unknown with variability. The correlation coefficient was calculated with Spearman rank, rather than Pearson because the BTMs did not follow a normal distribution.

7 Quality control

7.1 Instruments

As part of the normal clinical pathway:

- DXA: Hologic Discovery and Horizon densitometer (Hologic Inc., Bedford, MA, USA)
- Metabolic Bone Questionnaire

Additional instruments for this study:

- IDS-iSYS Analyser (ImmunoDiagnostic Systems)
- Cobas e411 Analyser (Roche Diagnostics)

7.2 QC Data

The stability of the clinical Hologic Discovery and Horizon densitometers was monitored using established in-house standard operating procedures. These included daily measurements of the device-specific phantoms provided by the manufacturer and weekly scans of the European Spine Phantom. The data was plotted on Shewhart control charts to monitor drifts or shifts in measurements by the bone density technicians who had received training from the device manufacturer.

Quality control of research and clinical biochemical analysis for each measurement obtained was maintained by in-house standard operating procedures. In-house quality control samples (clinical STH lab every 3-5hours) and manufacturer's quality control samples (clinical STH lab every 2 weeks) were used.

In order to directly compare the two assays of PINP and OC, all measurements were performed again by the Bone laboratory, with both assays from the same aliquot, directly on its thaw from Biorepository storage. The storage at -80^{oc} in the biorepository reduced the degradation of bone formation markers over time.

The measurements of PINP and OC were made on both the archived trial samples and the clinical samples within one calibration period of the analyzer immunoassay. This reduced the effect of assay variability on the comparison of results.

Because of the commodity of the archived serum from the clinical trials, and the use of auto analyzer not manual immunoassay, duplicate measurements were not taken.

The following inter-assay CV data in table 2.2 was from 10 single measurements from the in-house control mixed serum on different days, with the exception of PTH for which

measurements were made on 4 days, on the first of which were two measurements.

	IDS PTH pg/L	IDS 25OHD nmol/L	IDS PINP µg/L	Roche PINP µg/L	IDS OC µg/L	Roche OC µg/L
n	5	10	10	10	10	10
Minimum	34.4	58.3	43.1	46.9	16.0	18.5
Maximum	35.6	81.2	48.1	52.2	18.8	19.7
Mean	34.9	68.7	45.8	50.0	17.5	19.0
SD	0.56	7.71	1.75	1.83	1.21	0.38
CV_A %	3.2	11.2	3.7	3.8	6.9	2.0

Table 2.2. QC Data of 10 unique day measurements of in-house control, mixed serum for PINP and OC by IDS iSYS and Roche Cobas e411 autoanalysers and iPTH and 25OHD by IDS iSYS autoanalyser with calculation on the CV_A.

The QC data ruled out a major recalibration of the assay in the kit sent for use in this study. A drift on the result for the in-house control could have been from the storage, but still provided information on the inter-assay coefficient of variation (CV_A). The CV_A expressed as a percent was similar to recently published values made with the IDS iSYS analyser for PINP and OC of 3.7% and 2.3% respectively (Cavalier *et al.*, 2020). The panels of trend graphs in figures 2.1 to 2.3 appeared to have a lot of variability but this was only because the scale on the Y axis was so small. The IDS OC had the largest CV outside the recent publication values, but an older publication based on manufacturers data suggested up to 6.5% CV_A might have been expected for an assay of N-mid OC (CLSI, 2004). The significance of this has been explored alongside the results in later discussions. These CV results were not beyond the accepted value from which the assay and machine investigated by the manufacturer.

The CV_A data for the clinical measurements for creatinine and calcium were obtained from STH laboratory at 7.86% and 2.48% respectively.

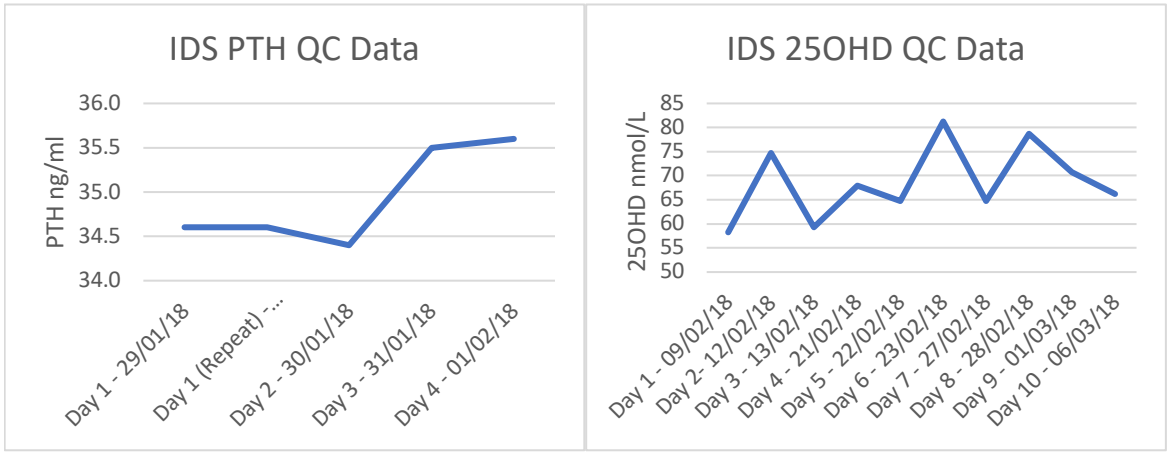


Figure 2.1. QC Plots for intact PTH on plasma and 25OHD with IDS iSYS from 4 and 10 non -consecutive days respectively January -March 2018

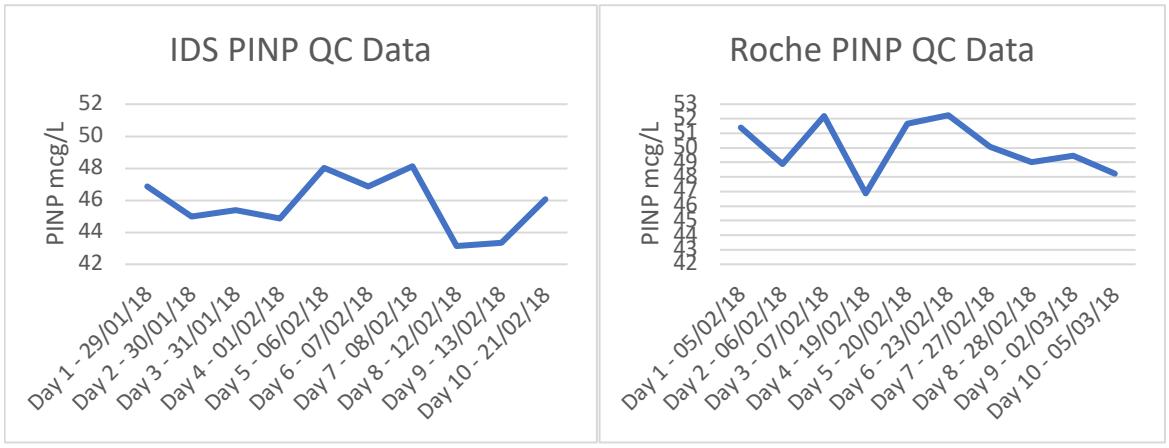


Figure 2.2. QC Plots for PINP by two methods, IDS iSYS and Roche Cobas e411 from 10 non- consecutive days over a one month duration Feb-March 2018

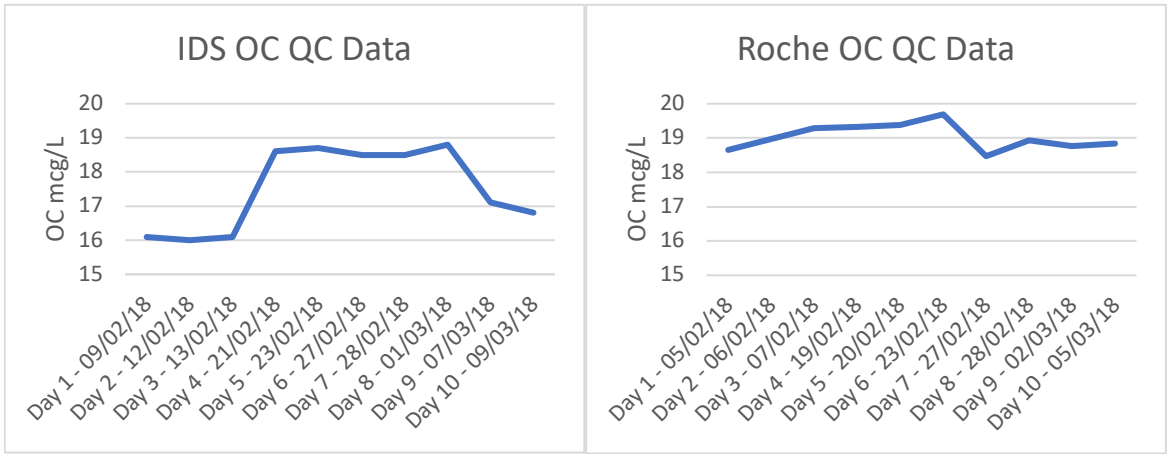


Figure 2.3. QC Plots for OC by two methods, IDS iSYS and Roche Cobas e411 from 10 non- consecutive days over a one month duration Feb-March 2018

8 Ethics

The study was conducted within the Metabolic Bone Centre, Academic Unit of Bone Metabolism, Northern General Hospital (NGH), Sheffield.

The study subjects for Stages A and B provided written informed consent for their samples to be stored after the protocol was complete in the South Yorkshire, North Derbyshire Musculoskeletal Biobank (SYNDMB). The measurements made in this report were done after seeking ethical approval for the protocol (SYNDMB045) and the samples were released from the SYNDMB for measurement.

The SYNDMB ethics approval are Oxford REC reference 15/SC/0132 and Sheffield Teaching Hospitals Study number STH15691.

The study subjects for Stage C attended the Metabolic Bone Centre at the NGH for a routine clinical appointment or investigation when they were recruited to the South Yorkshire, North Derbyshire Musculoskeletal Biobank.

They had contact with a team of staff employed specifically to undertake clinical research. Consent occurred in a private consultation room and blood samples were drawn by a trained phlebotomist in a private consultation space or purpose made phlebotomy clinic room. All other aspects of the retrospective clinical study occurred as part of the patients' normal clinical care by the NHS team at the Metabolic Bone Centre, Northern General Hospital.

All participants have given informed consent before enrolment onto the study. All participants have been given the Patient Information Sheet (PIS) V1.1 17NOV2015 in advance of the consent procedure (appendix E). This described the procedure of giving their blood sample for storage in the SYNDMB, for subsequent use in any musculoskeletal research study, and the use of their anonymised medical information.

The patients were advised that only 25mls of blood needed to be taken rather than 100mls stated in the PIS, that a urine sample was not required as mentioned in the PIS, and that no further research visits would be required. After the participant was given adequate time to

read the PIS, they had the opportunity to ask questions about the SYNDMB or this specific study on the day of consent before they signed the informed consent form.

Informed consent with ICF V1.0 31JAN2015 (appendix F) was carried out prior to venepuncture by an appropriately trained member of staff as assigned by the delegation log for SYNDMB.

The patient took a copy of their consent form and the PIS away with them and was informed where to find the research team contact numbers to ask further questions or withdraw from the study.

9 Data Collection, Handling and Record Keeping

Subjects each had a Metabolic Bone questionnaire (appendix A), and a workbook (paper source data form) (appendix C). Research clinicians and coordinators completed the workbook, after the consent procedure from subject questioning and review of the questionnaire. All data from the workbook and questionnaire was entered into designated layers of ArQ, the Metabolic Bone Centre musculoskeletal database. The data was stored according to the regulations of the Data Protection Act 1998.

Subject files were kept in locked offices in the Metabolic Bone Centre and the study database was protected by the standard University network login passwords.

A password protected subject enrolment log included participant date of birth and initials should study numbers needed to be identified to communicate significant clinical findings.

Chapter 3: Establishing the Reference Interval for Procollagen Type I N-propeptide and Osteocalcin in Men and Women by Two Assays

1 Background

1.1 Introduction

Bone turnover markers are increasingly used in clinical practice to monitor osteoporosis treatment. The longstanding use of a significant change in bone mineral density (BMD) on treatment to signify fracture risk reduction follows two steps of evidence. The first that a change in BMD in a treatment cohort, confers a percentage reduction in fracture risk (Bouxsein *et al.*, 2019), and the second is the level of BMD change that signifies a real change, as denoted by the least significant change (LSC). The same concept has been proposed for bone turnover markers (BTM) with recent evidence that a percent change in PINP in a bisphosphonate treatment cohort is correlated with vertebral fracture risk reduction (Bauer *et al.*, 2018). In the individual, the degree of change in BTM on treatment that is significant, is still in debate. Clinical trial evidence supports the use of two targets that differentiate with greater than 90% sensitivity, the response of female study subjects to treatment (Naylor *et al.*, 2016). One of these targets being LSC, the other is for treatment to suppress markers below the geometric mean of a healthy premenopausal reference interval.

A crucial step in delivering guidance on the use of BTMs to monitor treatment, is establishing the value of the biological targets that can be applicable to the majority of clinical bone health settings. Much work has been done on this between 2017 and a recent publication on the harmonisation of the assays for PINP (Cavalier *et al.*, 2019). In keeping with multiple studies in different geographic areas, the relationship between PINP measurements made by immunoassays of Roche and IDS (Morovat *et al.*, 2013) is complex and further studies are required (Vasikaran *et al.*, 2020). In addition, there has been no direct study of these assays for OC, as it is not an IFCC-IOF recommended BTM (Morris *et al.*, 2017).

This chapter explores the biological targets which will be taken forward for use in later chapters, to identify responders to treatment.

1.2 Aims

To define the geometric mean and 95% limits of the reference interval of two bone formation markers (procollagen I N-propeptide, PINP, and osteocalcin, OC) for men and premenopausal women.

To investigate whether there is a statistical difference between genders and how the two different assays, from IDS iSYS and Roche Cobas e411 autoanalysers, relate in method agreement analysis.

To calculate an absolute value of least significant change (LSC) for each bone turnover markers with each assay in women.

The following inquiries have been posed:

- Is there a difference in the distribution of the healthy male and the healthy premenopausal female reference interval for procollagen I N-propeptide and osteocalcin?
- Is there method agreement in the combined measurements of the male and female reference groups for each individual bone formation marker, when performed with two different assays?
- What is the Least Significant Change expressed in percentage and absolute terms of PINP and OC with the two different assays, calculated from short interval paired data?

2 Patients

When subjects were recruited within the Academic Unit of Bone Metabolism, University of Sheffield, consent was obtained to store serum and plasma samples within a repository, the South Yorkshire and North Derbyshire Biorepository (SYNDMB). (Oxford Rec reference 15/SC/0132 and Sheffield Teaching Hospitals Study number STH15691).

I applied to the SYNDMB (SYNDMB045) to release serum aliquots to carry out up to date measurements, to form the reference intervals for the purpose of this study.

The studies from which selected aliquots for the male reference cohort originated, were;

- "Extreme CT" (Walsh *et al.*, 2017) 90 men were recruited into age categories, all BMI less than 35kg/m^2, sampled according to inclusion criteria in table 3.1 from those with stored serum available.
- "Fat and Bone" (PRP) (Walsh *et al.*, 2016) 126 men were recruited into BMI and age categories, sampled according to inclusion criteria in table 3.1 from those with stored serum available.

The study, from which archived samples of female subjects for the reference cohort were taken, was;

- "TRIO" (Naylor *et al.*, 2016) 226 control cohort subjects at baseline, sampled according to inclusion criteria in table 3.1 from those with stored serum available for chosen time points.

All subjects recruited to these studies had been screened for any co-morbidity or medication affecting bone metabolism.

Selections were made based on the proposed subject characteristics for a reference group detailed in table 3.1.

	Male	**Female** (TRIO Controls)	**TRIO Tx Cohort** (Baseline)
Inclusion	30-60 years BMI<35 kg/m^2 Vitamin D>30nmol/l	30-45 years (with regular menstruation or any non-hormonal contraception) BMI<35 kg/m^2 Vitamin D>30nmol/l	T-score ≤ -2.5 OR ≤ -1 with non-traumatic fracture Ambulatory ≥5 years postmenopausal ≤85 years old 18≤ BMI ≤35
Exclusion	Co-morbidity or medication known to affect bone metabolism		

Table 3.1. Proposed subject characteristics by inclusion and exclusion criteria specified in choosing appropriate archived samples from the historic studies.

Further archived samples were released on application to the SYNDMB from the TRIO study subjects for paired measurements for the female LSC cohort;

- "TRIO" (Naylor *et al.*, 2016) 180 treatment cohort subjects at baseline, sampled according to inclusion criteria in table 3.1 from those with stored serum available for chosen paired time points of week 12 and 13.

At baseline neither the control or treatment cohort of the TRIO trial were on a calcium and vitamin D supplement. Ad Cal D3 was initiated at the baseline visit (week -1) and oral alendronate 70mg weekly or risedronate 35mg weekly or ibandronate 150mg monthly was commenced at week 0, in the TRIO treatment cohort.

3 Methods

3.1 Raw Data Collection

The serum samples were stored at -80 ^{oC}, until the week the biochemical measurements were made.

The biochemical measurements were made from February to March 2018 with the same aliquot being sampled for measurements on the Immunodiagnostic Systems (IDS) iSYS and the Roche Cobas e411 autoanalyser.

The following biochemical measurements were made in the bone biochemistry laboratory within the same calibration period from serum aliquots;

- PINP (Roche Total, IDS Intact)
- OC (Both Roche and IDS N-Mid)
- 25-hydroxyvitamin D (IDS iSYS)

I requested the following baseline characteristic data of each subject for which an archived sample was used, using the non-identifiable study number from the study records at recruitment:

- Age
- BMI
- BMD at the total hip from densitometry performed for the purpose of research.

3.2 Data Processing

Hip T-score was calculated using the NHANES III database (Shuhart, Yeap and al, 2019) from gender specific BMD reference data.

The raw BTM measurements were converted to the agreed units as stated in Chapter 2. The measurements of PINP and OC in the female and male reference groups were log transformed to fulfil the properties of a normal distribution.

To describe the reference intervals, I used the geometric mean and upper and lower limits of the distribution with 95% confidence. These were calculated following the equations detailed in Chapter 2.

3.3 Statistical Analysis

Differences in the characteristics that affect bone turnover have been analysed with an independent samples 2-tailed t-test with 95% confidence intervals as normality was satisfied. This was performed to compare the characteristics on the male and female reference groups only. The null hypothesis was that there was no difference in the age, BMI, hip T-score and 25 hydroxyvitamin D level between the male and female reference group, as was intended by subject selection. No comparison was required of characteristics for the study group from which the LSC was derived.

The log transformed distributions of PINP and OC by each assay were compared between men and women with an independent 2-tailed t-test and between assays with a paired sample 2-tailed t-test as described in Chapter 2. The null hypothesis was that there was no difference in the distribution of PINP and no difference in the distribution of OC between men and women when measured on each assay, in addition that there was no difference between the IDS iSYS and the Roche Cobas e411 autoanalyser assays for PINP and OC when measured in men and women.

To examine the relationship further between the measurements produced by assaying the serum with the IDS iSYS and Roche Cobas autoanalysers, method agreement analysis by Bland and Altman and Passing and Bablok has been performed as described in Chapter 2.

4 Results

4.1 Subject Characteristics

Table 3.2 demonstrates the success in the subject selection of the male (n=50) and female (n=130) group from which the reference intervals have been established. The baseline characteristics of the postmenopausal group on treatment with oral bisphosphonates in the TRIO study are also detailed in table 3.2 but were not included in comparison with the reference cohorts. The male group is a smaller cohort than the proposed statistical gold standard number of 120 (CLSI, 2008).

	Male	**Female** TRIO Controls	**TRIO COHORT** Baseline
n	50	130	135
Age at recruitment (yrs.)	38 (9.4)	38 (1.6)	67 (7.4)
BMI (kg/m^2)	26.9 (3.8) †	25.0 (3.9)	26.4 (4.0)
Total hip T-score	+0.27 (0.87)	+0.27 (0.91)	-1.35 (0.85)
Mean 25OHD (nmol/L)	57.9 (18.6)	55.5 (18.9)	68.5 (22.5)
Co-morbidities	None	None	None

Table 3.2. Reference group characteristics by cohort.

Mean and (SD) for each characteristic that could impact bone metabolism. Significant difference $p < 0.05$ by t test denoted between the Male and Female cohorts † but not compared with the TRIO cohort baseline characteristics.

There was no significant difference in age, total hip T-score or 25-hydroxyvitamin D between the male and female reference cohorts. There was a significant difference (p=0.003) in the BMI between the male and female group. However the mean BMI in the male and female reference group was at the margin of normal weight and overweight categories. The null hypothesis of no difference in baseline characteristics between the male and female reference groups was rejected for BMI, but could not be rejected for the other characteristics.

4.2 Descriptive Statistics

Uniformly, for both PINP and OC distributions, with both IDS iSYS and Roche Cobas e411 assays, the geometric mean in men was higher and the variance greater than in women. This can be seen in table 3.3 and figure 3.1. The null hypothesis that there was no difference in the distribution of PINP and OC in men and women, was rejected with strong statistical significance (p=0.000).

		PINP µg/L		**OC µg/L**	
		IDS-iSYS	Roche Cobas e411	IDS-iSYS	Roche Cobas e411
Male n=50	Geometric mean	51.8†	49.5†	19.0†*	23.0†
	95% RI UL	121.9	112.3	38.3	37.6
	95% RI LL	22.0	21.8	9.4	14.0
Female n=130	Geometric mean	37.2*	41.7	13.5*	17.8
	95% RI UL	63.9	68.5	24.3	27.9
	95% RI LL	21.6	25.4	7.5	11.3

Table 3.3. Geometric mean and Upper and Lower Reference Limits (95%) of PINP and OC by two assays for men and women.

Significant difference p <0.025 by t test on the log transformed data, denoted between the Male and Female reference cohorts † within each assay method, for each bone turnover marker; between each assay by * for PINP or OC, within gender cohort.

Looking in more detail at figure 3.1, the graphs for the distribution of OC in men and women by the two assays, clearly showed a pattern of the Roche N-Mid assay reading higher than the IDS N-Mid assay and comparison of means of the log transformed data showed a statistical difference (p=0.000). The pattern of the relationship between the two was explored by method agreement analysis.

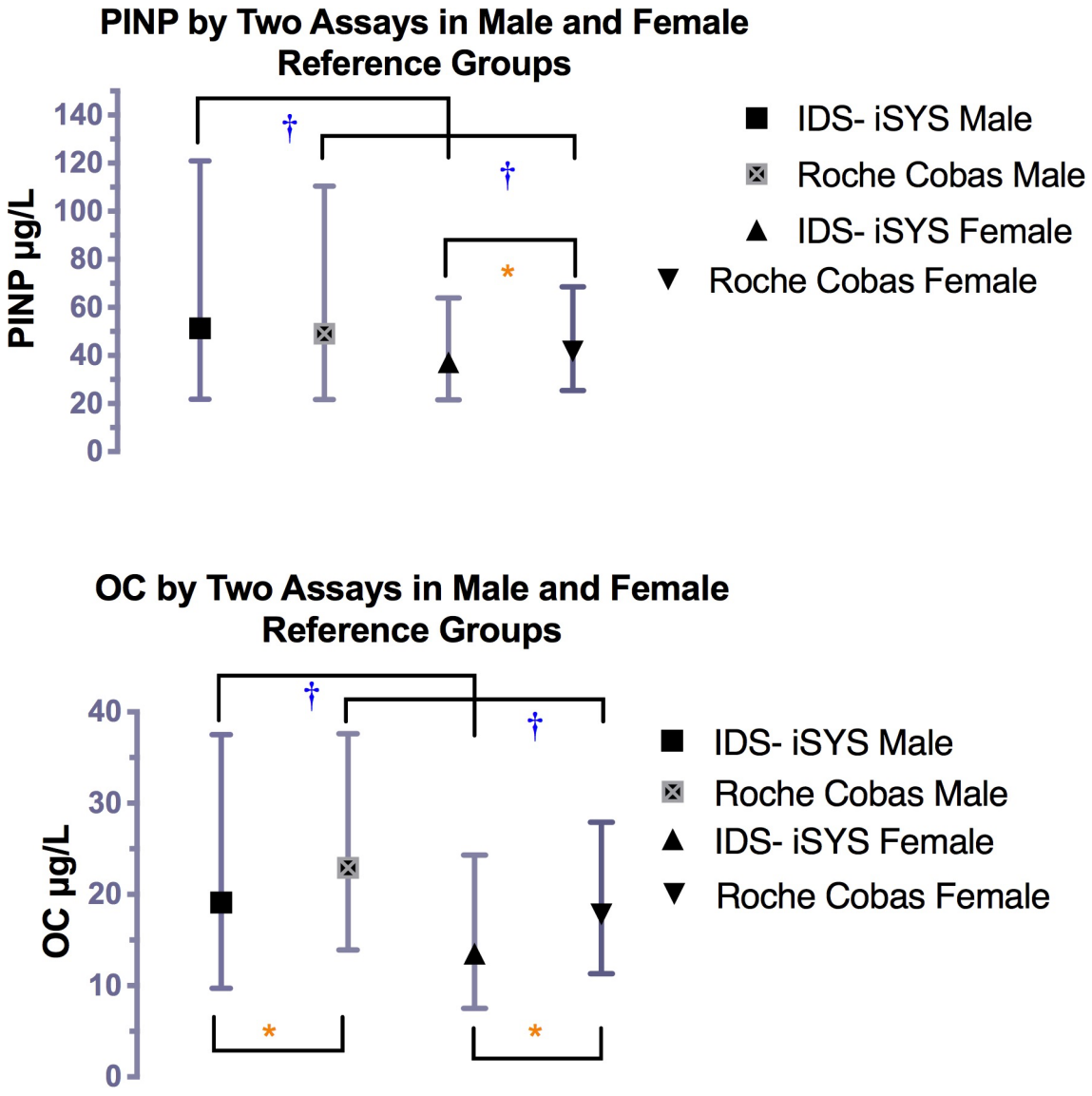


Figure 3.1. Distribution plots of the Male and Female Reference cohorts for a) PINP by IDS iSYS and Roche Cobas e411 and b) OC by IDS iSYS and Roche Cobas e411 assay.

Demonstrating the GM and 95% upper and lower reference limits. Significant difference p <0.025 by t test denoted between the Male and Female cohorts† within each assay method; and between each PINP and OC assay*, within cohort.

However, for PINP the distributions in the male and female reference cohorts of the Roche total PINP and IDS intact PINP assay varied in their relationship. In women, the geometric mean detailed in table 3.3, was significantly higher in the Roche assay (p=0.000). In men, the IDS assay produced the higher geometric mean, but not significantly so. Whether the

variation in the relationship was because of a bias in method agreement depending on the level of measurement, was explored in the method agreement analysis.

4.3 Method Agreement

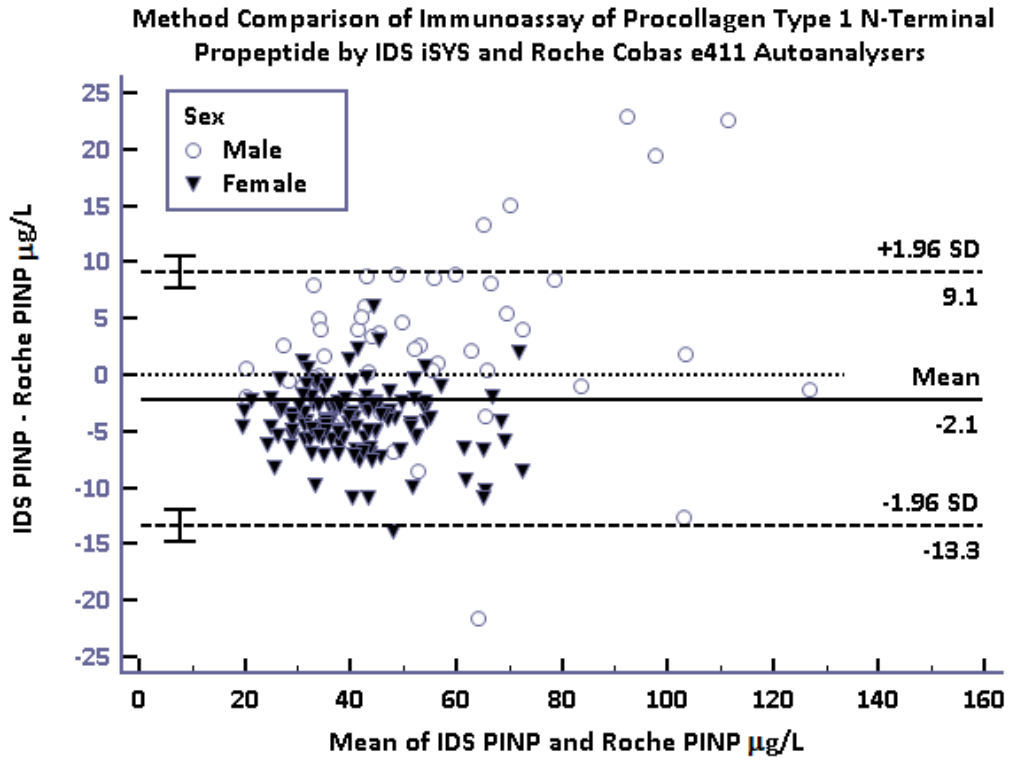
Comparison of the geometric means of the PINP and OC distributions with each assay was done separately for men and women. The method agreement and correlation analysis has been done on combined measurements, on the basis that there was no reason why serum from men and serum from women should interact differently with the immunoassay.

The predictability of the method agreement was described by the Bland Altman plots seen in figure 3.2 for PINP and OC by the two assays.

With the prediction limits of the mean difference between PINP results made by IDS and Roche assays being either side of 0 (9.1 to -13.3 µg/L), they did not significantly produce different results overall in PINP measurement. Although the mean offset is calculated at -2.1 µg/L, at higher levels of PINP, the IDS PINP would measure higher than the Roche PINP. With the subgroups of men and women shown in figure 2, this was predominantly the values arising from men that demonstrated this effect. The change in the offset, as the PINP is measured higher, resulted in the regression of the Bland Altman plot having a significant slope of 0.10 (0.06 to 0.15).

The Passing and Bablok regression shown in figure 3.3 again showed the change in the relationship by the regression line and confidence intervals that crossed the line of perfect agreement with a slope of 1.04 to 2 decimal places. Because the confidence intervals of the regression line crossed 1, the agreement of the readings by each assay were not significantly different from perfect agreement. The slope and the intercept of the regression equation to predict the result of the Roche assay from IDS assay can be seen in table 3.4.

The Bland Altman plot seen in figure 3.2 for OC showed an offset of -4.2 µg/L with prediction limits -2 to -6.3 µg/L. This demonstrated a fixed offset of the result of OC with the Roche assay compared to IDS. There was no alteration in the relationship if measurements are high or low, as there was a non-significant slope to the regression of the Bland Altman plot with 95% CI that cross zero. Unlike with PINP, the prediction limits were well away from 0 so the assays significantly produced different values.



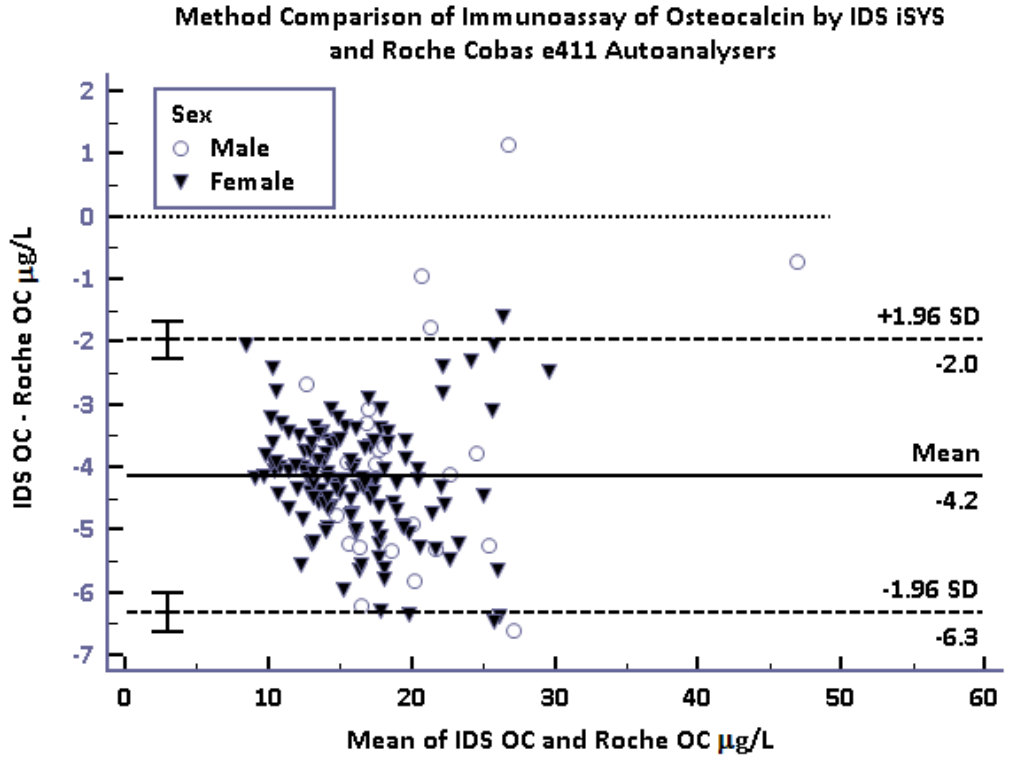
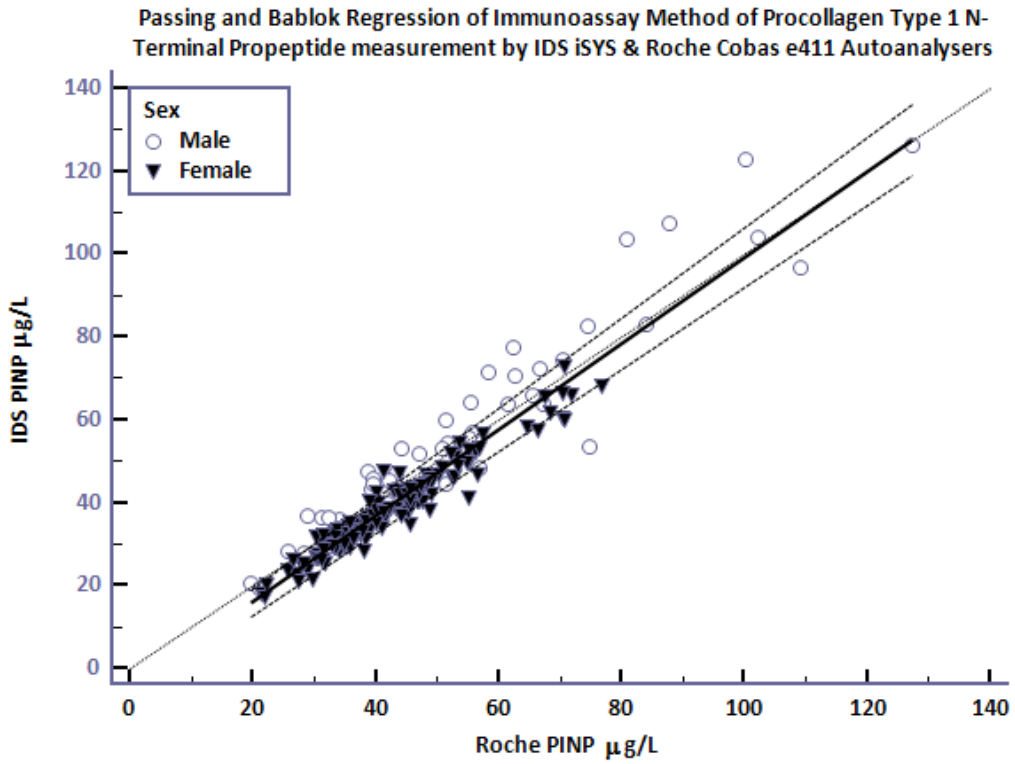


Figure 3.2. Bland Altman Plots for a) PINP and b) OC by two methods, IDS iSYS and Roche Cobas e411.

Regression equation for PINP y= -6.8048 + 0.1041* x with significant slope p<0.0001. Slope of regression for OC is NS (0.02 -0.02 to 0.06)



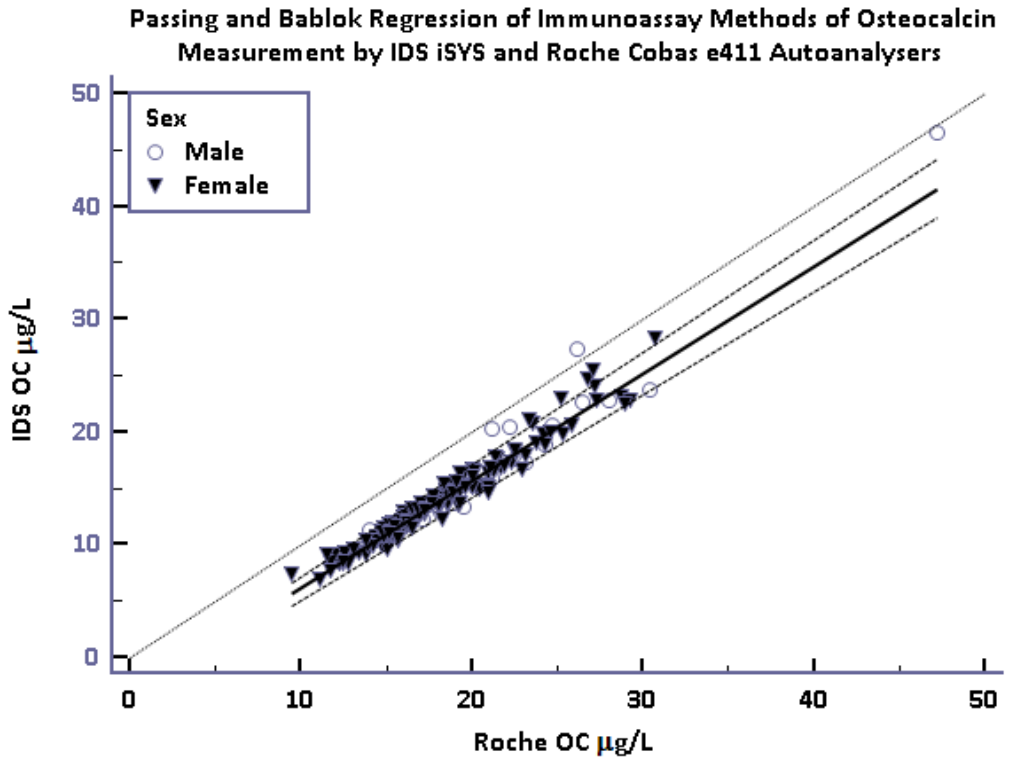


Figure 3.3. Passing and Bablok Plots for a) PINP and b) OC by two methods, IDS iSYS and Roche Cobas e411.

The slope and intercept of the regression line are detailed in table 3.4.

The Passing and Bablok regression line for the methods of OC measurement shown in figure 3.3 was parallel to the line of perfect agreement showing a consistent offset, whatever the level of measurement, with the slope of the line of 0.95 to 2 decimal places. The confidence intervals of the regression line as described in table 3.4 did not include the line of perfect agreement, but with the absence of a major deviation away from the slope, there is some agreement between the assays.

	PINP	**OC**
Slope **95% CI**	1.0389 0.9896 to 1.0912	0.9512 0.9124 to 0.9933
Intercept µg/L **95% CI**	-4.60 -7.00 to -2.70	-3.35 -4.07 to-2.70

Table 3.4. Passing and Bablok Linear Regression of PINP and OC by two methods, IDS iSYS and Roche Cobas e411, autoanalyser assays.

4.4 Least Significant Change

Although the distribution of PINP and OC did not follow a normal curve, when taking paired data, the distribution of the differences has been shown to satisfy normality by plotting histograms in figures 3.4 and 3.5. The cumulative histograms demonstrated a clear sigmoidal curve and 50% of the value below and above 0.

The LSC could therefore be calculated from non-log transformed data. The histograms in figures 3.4 and 3.5, also clearly showed outliers in the sets of paired data in terms of the measurements from week 12 and week 13 being of a greater difference. There were four data sets that produced outlying difference in the PINP values from both assays and three of these also produced outlying difference in the OC values from both assays. Because four pairs of serum samples behaved uniformly across the BTMs and assays, they contributed important variability to the data. However, one set of paired samples produced an outlying difference in the IDS OC values only and one an outlying difference in the Roche OC values only. These were subsequently removed from the LSC calculations.

The variable number of sets of paired data from which each assay specific BTM LSC was calculated, as seen in table 3.5 is due to missing raw data.

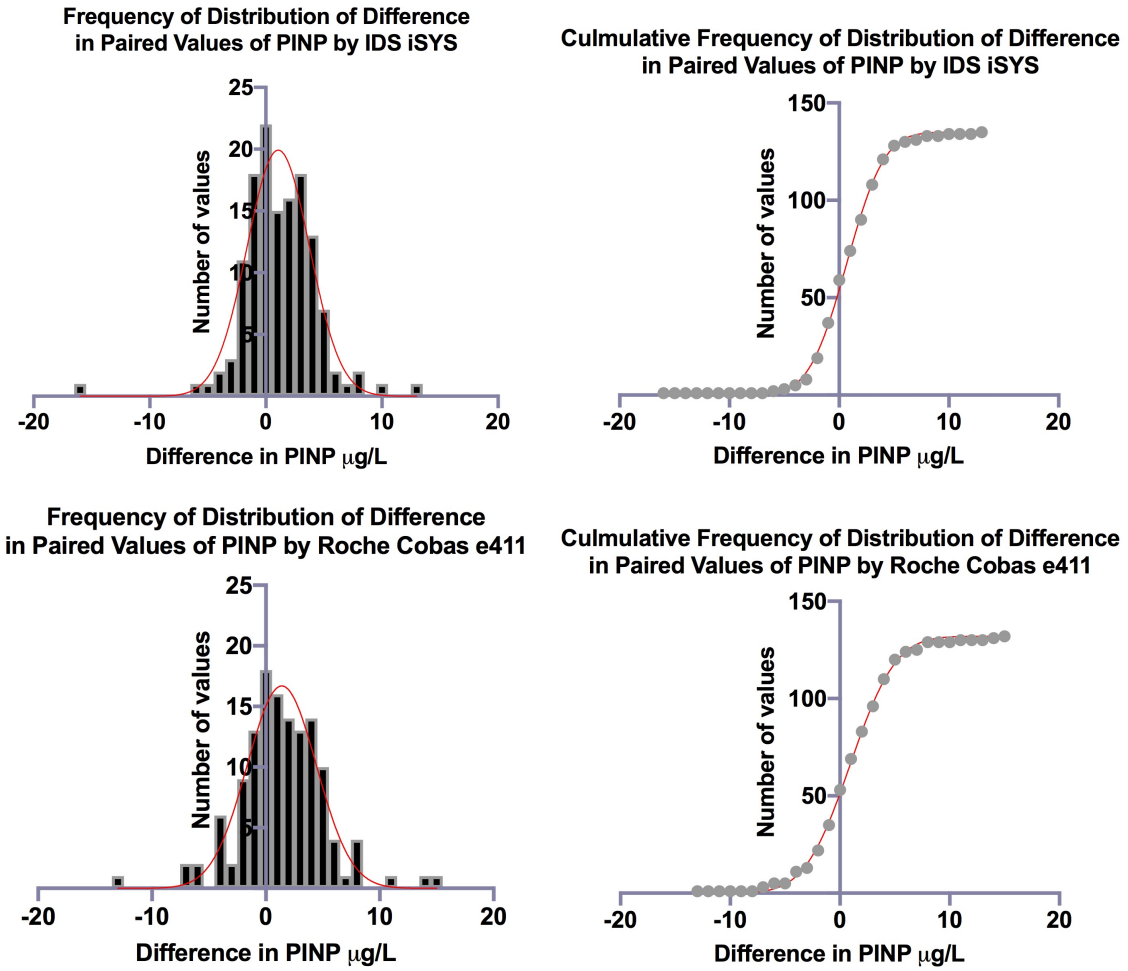


Figure 3.4. Distribution plots of the difference in Paired data for PINP by a) IDS iSYS (n=135) and b) Roche Cobas e411 (n=135) assays.

Paired data from week 12 and 13 of treated trial subjects to calculate short term LSC. Gaussian distribution lines drawn on the bar distribution nomogram and the cumulative distribution by count.

	PINP		**OC**	
	IDS-iSYS n= 135	Roche Cobas e411 n= 134	IDS-iSYS n= 132	Roche Cobas e411 n=130
LSC µg/L	6.7	7.9	4.4	4.1
CV_I %	10.0	10.2	8.4	8.6
LSC %	28	28	23	24

Table 3.5. Least Significant Change for PINP and OC by the IDS iSYS and Roche Cobas e411 assay, expressed as an absolute figure and a percentage alongside the calculated intra-individual mean CV_I

The values shown in table 3.5 showed small intra-individual variation, and they are similar between the two assays for each PINP and OC.

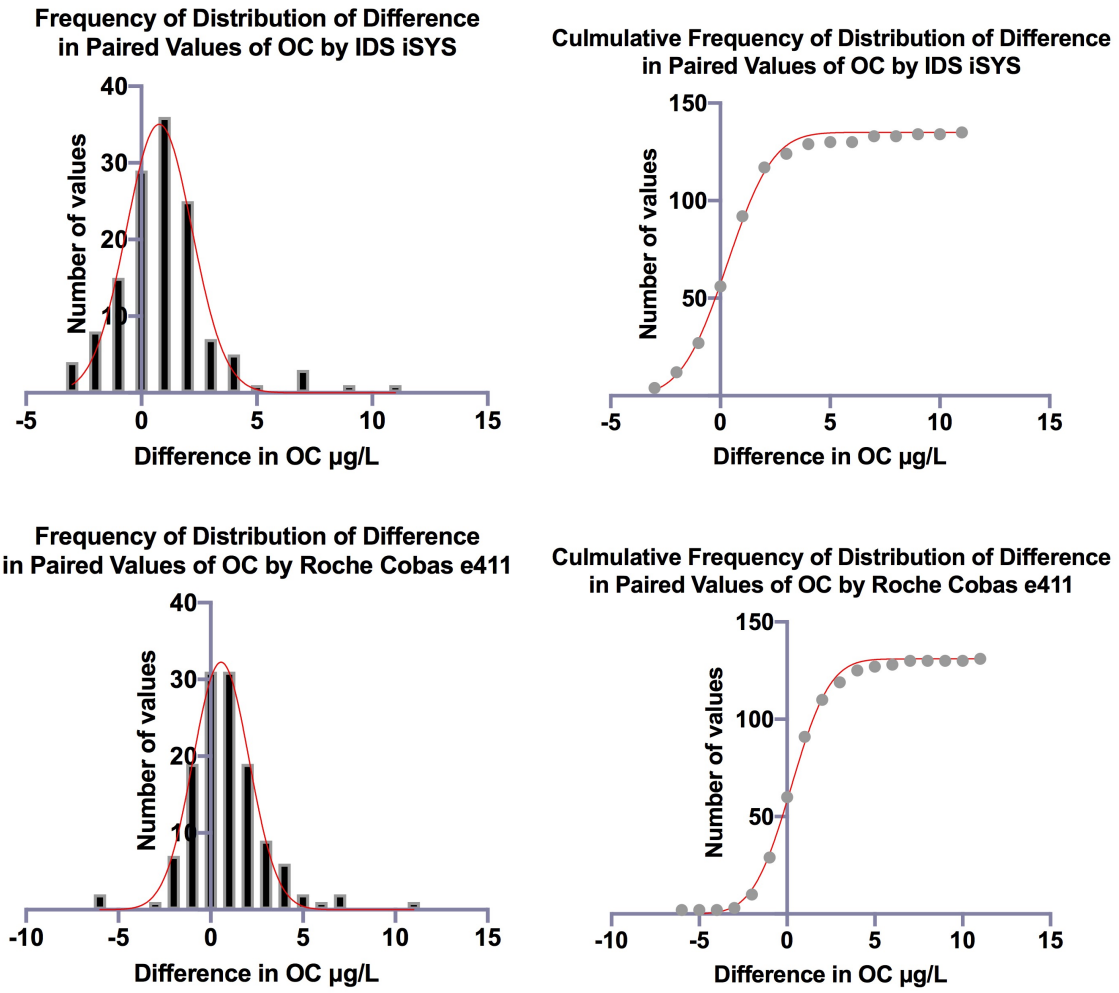


Figure 3.5. Distribution plots of the difference in Paired data for OC by a) IDS iSYS (n=132) and b) Roche Cobas e411 (n=131) assays.

Paired data from week 12 and 13 of treated trial subjects to calculate short term LSC. Gaussian distribution lines drawn on the bar distribution nomogram and the cumulative distribution by count.

5 Discussion

5.1 Reference Intervals

When considering any conclusion based on a comparison of the distribution in the male reference cohort, it must be remembered that 50 subjects does not constitute the gold standard for a reference interval of n=120 (CLSI, 2008). The sample we used fulfilled strict inclusion criteria as seen in table 3.1 and 3.2, and the historic male participants from which the sample was selected were originally recruited based on varying age and BMI criteria (Walsh *et al.*, 2016; Walsh *et al.*, 2017). Therefore the suitable archived serum samples for men were limited. However, by using the robust method of reference interval limit calculation advocated by the Clinical and Laboratory Standards Institute in 2008, the reference interval that has been produced with the smaller sample size is still a valid comparison (Henny, 2009).

The male and female reference interval subjects did satisfy criteria for stability in bone health with no discernible difference in the characteristics of the two cohorts, with the exception of the BMI as seen in table 3.2. As BMI starts to influence (reduction) BTM at the point of obesity, both cohorts just falling into the over-weight category would represent minimal impact on bone formation markers (Walsh *et al.*, 2016; Eastell and Szulc, 2017).

There was consistency in the male reference interval being significantly different and indeed higher than the female as seen in table 3.3. This was accompanied by greater variance in both the bone formation markers with each assay in the male group seen in figure 3.1. The studies by Michelsen (Michelsen *et al.*, 2013) for IDS iSYS PINP and Chubb (Chubb *et al.*, 2015) for Roche PINP consistently show male PINP reference interval being greater than the female. Although skeletal size is on average greater in men than women, so is the circulating volume and clearance of PINP by the kidneys, so it is unlikely to be a contributory factor in this finding. Work done by the Australian group with the Roche OC assay, shows the same pattern for men and women (Chubb *et al.*, 2015). A contributing factor could be that men stop growing later than women, with the later closure of growth plates and the lower age limit in both cohorts as seen in table 3.1 was 30 years. The androgens and oestrogen are

likely to play a part with periosteal growth stimulated by testosterone (Walsh *et al.*, 2017) and oestradiol suppressing the BTM (Hannon *et al.*, 1998).

The results shown in this chapter differ to the published reference intervals, in that the PINP reference intervals established here in table 3.3 have a much greater range with a much higher upper 95th reference limit. The age group for the men and women included in the other studies were similar. In particular, the female upper age limit is as recommended by an older study by Glover (Glover *et al.*, 2008). The Michelsen study included multiple different age groups (Michelsen *et al.*, 2013). Results described in this chapter from men with a mean age of 38 years are most consistent with 70 year old men in the Michelsen study. Distribution of the age range in the male reference group was 41 subjects between 30 and 45 years, akin to the female reference interval, and 9 older subjects, who were between 55 and 60 years. There was no correlation when performed for Pearson's correlation coefficient of age in the male reference group and OC by either assay. There was only a significant correlation in PINP measurements by Roche Cobas e411 with male age (r -0.31 95% CI -0.54 to -0.03), not by IDS iSYS. This limited correlation supports the idea that the lower age limit on the reference group is still influencing the upper limit of the reference interval by later closure of growth plates than in females.

A recent study by Cavalier et al in 2019 (Cavalier *et al.*, 2019) updated the understanding (Morovat *et al.*, 2013) in respect of the relationship between IDS and Roche assays of PINP. There is no previous direct comparison of these two osteocalcin assays in the literature, therefore it cannot be said that our results replicate previous findings. Studies in two different populations at geographically different centres suggest the Roche assay produces higher OC reference intervals than the IDS (Hannemann *et al.*, 2013; Chubb *et al.*, 2015). This is consistent with the results presented in this chapter, where there was a fixed offset with Roche greater than IDS measurement of OC seen in figure 3.2. This chapter provides additional useful information such as the offset showing no change dependent on lower or higher measurements and the regression is almost parallel to the line of perfect agreement described in table 3.4. Because the assays are measuring the same fragments of OC, this offset should be straight forward to amend by calibrating or harmonising the two assays.

The varied relationship in the measurements of PINP by the two assays appears to be more in the serum of men. This demonstrates how the relationship alters in the two assays as the

PINP measurements get higher given that the male cohort have on the whole higher PINP levels seen in figure 3.2. Because of the trend that the difference between the measurements from the two assays become smaller as the values increase, the net outcome was that the method agreement results of IDS and Roche immunoassays for PINP actually show no difference from perfect agreement in its confidence intervals in table 3.4. The Morovat study found a very strong correlation between the Roche E170 and IDS iSYS PINP but it was a nonlinear relationship (Morovat *et al.*, 2013). The difference between the measurements from each method could exceed 20% although the overall difference was not significant.

When considering the cause of the inconsistency in the relationship of the PINP assays, the target of the immunoassays contributes to the explanation. IDS measures intact PINP and Roche total PINP, they are therefore measuring different fragments. For this reason also, any harmonisation would not include subjects with Chronic Kidney Disease (CKD) (Bhattoa *et al.*, 2021) and plans have recently been developed to tackle this.

The 2017 International Osteoporosis Federation (IOF) meeting highlights that assay harmonisation is lacking (Morris *et al.*, 2017), and whilst OC is not on the main agenda, the information regarding the osteocalcin assay is important. The last aim of the IOF working group for harmonisation comparison studies, is to identify the bias in calculating correction factors for data from different assays (Bhattoa *et al.*, 2021). Although this is based around the IOF recommended BTM of PINP and CTX, with the N-MID assays commonly in use by several biochemical companies, this looks entirely possible for Osteocalcin now, when it didn't in 2011 (Vasikaran *et al.*, 2011).

5.2 Serum vs Plasma Samples

The measurements in this study are on serum samples. The same is the case for the past reference interval studies cited. However, Etienne Cavalier has done studies comparing measurements from serum and plasma on the Roche Cobas and IDS iSYS, in order to establish generalisability of the reference intervals. Passing and Bablok regression analysis on plasma vs serum showed a slope close to 1 for both IDS and Roche assays (Cavalier *et al.*, 2019).

5.3 CV Data

Looking individually at the use of PINP and OC, a desirable characteristic of a bone turnover marker assay in terms of its precision, is that the CV_A should be less than half of the CV_I (CLSI, 2004). Looking back on the CV data in table 2.2 of Chapter 2 and the CV_I in table 3.5 in the results above, this has been shown for all but the IDS OC data (CV_I 8.4% with CV_A 6.9%). Previous studies of intra-individual CV for N-mid OC quoted 10.2% (n=20 premenopausal women), which our CV_I falls within (IDS OC 8.4%, Roche OC 8.6%). PINP intra-individual CV on the same 20 premenopausal women was quoted as 10.6%, again in the same region as the CV_I from the paired data presented in this chapter (IDS PINP 10.0%, Roche PINP 10.2%). The CV_I from our paired data, is specifically in keeping with IDS iSYS measurement of PINP and OC CV_I in plasma in a recent multi-centre collaboration the EuBIVAS study (Cavalier *et al.*, 2020). In this study, 91 men and women age 21-69 years gave weekly samples from 7-10 weeks total, controlled for time, and the CVI was 8.8% for PINP and 8.9% for OC with no difference between men and women, but higher intra-individual CV in premenopausal than postmenopausal subjects.

Other desirable features of a bone turnover maker that could be clinically useful are;

- the specificity to bone
- the assay to be widely available on automated platforms
- demonstration of resistance in variability to sample handling
- stability in storage.

The serum samples used in this research had been stored from previous archived studies at -80 degrees Celsius. The similarity of the geometric mean calculated for this chapter seen in table 3.3, compared to that in the TRIO publication (Naylor *et al.*, 2016), with samples taken 2007-2008, is reassuring in that with time the PINP has not increased, and OC decreased. In addition, the archived aliquots used had never previously undergone a freeze thaw cycle, important for the stability of OC, with PINP being more robust in this regard.

5.4 LSC

There appears to be much debate in the literature regarding the LSC, or reference change value (RCV) (as it is referred to by biochemists), of PINP. The Roche OC LSC has not been

previously studied. The TRIO study publication (Naylor *et al.*, 2016) and work based on it in IOF guidelines (Diez-Perez *et al.*, 2017) quoted a LSC based on the same pairs as used in this study of 38% (although in as many as 180 pairs compared to the 135 I used). This is a lot higher than the 28% I calculated as seen in table 3.5 and could be because they were calculated from log-transformed data in TRIO. There have been plenty of additional studies to compare the process of calculating LSC. Fink, Cormier et al calculated the intra-individual coefficient of variation and LSC for OC and PINP in 9 healthy premenopausal women with a mean age of 32 (Fink *et al.*, 2000). The calculation was performed on samples from 4 non-consecutive days over a mean of 14 days. The cohort was different to ours in that they were not on treatment and premenopausal, but the method of calculating the LSC is similar to this study with p=0.05. The CV_I of was reported as PINP 12.4% and OC 7.2%, but with the LSC higher; PINP 38.3% and OC 20.4%. The only difference identified in the method is that the LSC presented in this chapter did not include the analytical intra-assay CV (CV_A), because serial measurements in time would encompass both the intra-individual and the intra-assay CV. This potentially explains the higher percentage LSC found in this study, as the intra-assay CV in PINP in the study is 6%, quite large compared to that of OC 1.6%. It does not explain the difference in the TRIO study results (Naylor *et al.*, 2016) in which CV_A was also not added as extra and the LSC for the OC also is higher than encountered in our study. A further study calculating LSC (Eekman *et al.*, 2011) used an in house intra-assay CV of 8% for PINP but based the intra-individual variation on a study looking specifically at the "critical difference between serial measurements" by Scariano (Scariano *et al.*, 2001). The CV_I was calculated from 7 serial values in both pre and postmenopausal females, including n=9 on anti-resorptive treatment as in this study. The CV% calculated for PINP was between 11.5% and 13.2%, but again the LSC was high at 36% for PINP. The LSC calculated in both the Fink and Scariano study were based on small number of subjects (n=9 and n= 9) and used a different assay method, the manual RAI by Orion. As a manual assay, this accounts for higher intra-assay CV, but not for the biological variation.

Eastell et al actually recalculated the data from the TRIO study to publish CV_I and LSC% for a review article in 2018 (Eastell *et al.*, 2018) and found it to be more in keeping with the results in table 3.5, at 29% and 23% for measurements by IDS iSYS and Roche Cobas e411.

This is also in keeping with the results from the EuBIVAS data which shows a very low LSC result of 19.9% for PINP by IDS iSYS on plasma (Cavalier *et al.*, 2020).

An absolute value rather than percentage is more practical to use in clinical practice but may be falsely reassuring in subjects with a very high baseline measurement of BTM. There is very little published literature giving the LSC in absolute terms as I have done in table 3.5 and take forward into Chapter 4.

One conclusion of the Scariano study (Scariano *et al.*, 2001), that biological variance is not significantly different between the premenopausal or postmenopausal cohort or on or off anti-resorptive treatment, is useful for the applicability of the LSC calculated in this study. However, this differs to a conclusion in the recent and larger EuBIVAS study, which showed a difference between premenopausal and postmenopausal subjects (Cavalier *et al.*, 2020). Using samples 1 week apart represents short term intra-individual variation. The presented LSC has the advantage that it is made from postmenopausal subjects, to which it would be applied in clinical practice, and also those on treatment, which is the purpose for which it is intended.

Chapter 4 Results: Response Rate to Oral Bisphosphonates by Procollagen Type I N-propeptide and Osteocalcin measurement by Two Assays in Two settings

1 Background

1.1 Introduction

Clinical guidelines recommend the use of bone turnover markers to monitor anti-resorptive treatments in order to identify those not responding or not adherent to oral bisphosphonates (Diez-Perez *et al.*, 2017). In clinical practice, serum procollagen I N-propeptide (PINP) is favoured as it is relatively unaffected by food intake or time of day (Schlemmer and Hassager, 1999). IDS or Roche autoanalysers are both available for use in clinical practice. The level of PINP is impacted on by glucocorticoid use, recent fracture and, chronic kidney disease (specific to Roche Total PINP assay) (Vlot *et al.*, 2018). The TRIO study demonstrated significant results of more than 90% response in postmenopausal study patients on oral bisphosphonates, when the geometric mean of the premenopausal reference interval (threshold) was used as the biological target (Naylor, 2016).

The following results explore the performance of PINP and osteocalcin (OC) by two assays, to identify apparent response (sensitivity) to oral bisphosphonates in clinical practice where the above factors are present.

Our aim was to compare apparent response rates in serum PINP and OC, and associated changes in total hip BMD, between women recruited to a trial of oral bisphosphonates (the TRIO study) and an unselected cohort of patients receiving oral bisphosphonates in our clinic.

1.2 Aims

To evaluate the usefulness of performing two bone formation markers by two assays, in clinical practice using a single-centre, observational, retrospective, cross-sectional study.

To describe the performance of the same bone formation markers in trial patient cohorts controlled for influences of bone metabolism.

To validate a target marker level (threshold) as a surrogate for responders to osteoporosis treatment by comparison with the least significant change (LSC) method in a trial patient cohort and by examining the rate of bone mineral density gain in the unselected patient and the trial patient cohorts.

1.2.1 Main research question:

- What is the sensitivity and specificity of procollagen I N-propeptide and osteocalcin at a target marker level, to identify the responsiveness of clinical patients on oral anti-resorptive treatments for osteoporosis? (The target is the threshold level defined as the geometric mean of a healthy premenopausal reference interval)

1.2.2 Secondary research questions:

- Do two different assays for PINP and OC have an equivalent sensitivity and specificity of the biological target in identifying clinical patients on oral anti-resorptive treatments for osteoporosis?
- Is the proportion of female trial patients that achieve the LSC on an anti-resorptive for postmenopausal osteoporosis, in each of the bone formation markers by each assay, the same as the proportion achieving the threshold in the respective assay?
- Is there a significant difference in the rate of gain of bone mineral density between trial “responders” and “non-responders” to osteoporosis treatment defined by both the threshold and LSC biological targets for each bone formation marker assay?
- Is there a significant difference in the rate of gain of bone mineral density between the sub-groups of clinical patients identified by each bone formation marker as “responders” and “non-responders”?

2 Patients

Subjects were recruited from 2016 to 2018 within the Academic Unit of Bone Metabolism, University of Sheffield for the South Yorkshire and North Derbyshire Musculoskeletal Biorepository (SYNDMB) with Oxford Rec reference 15/SC/0132 and Sheffield Teaching Hospitals Study number STH15691.

Study subjects recruited within the Metabolic Bone Centre had consented for storage of serum and plasma samples within the biorepository for use in future research.

Ethical application to the SYNDMB was approved (SYNDMB045) to create fresh cohorts sampled from the following historic participant groups:

- Oral bisphosphonates. 100 clinical patients treated with oral alendronate, risedronate or ibandronate.
- Matched no treatment group. 300 clinical patients who attended for bone densitometry on no osteoporosis treatment (approximately 50% were on calcium or cholecalciferol). They were matched by gender, age (round down +/- 3 years), body mass index (+/- 2.5) and Tscore of the total hip (+/- 0.3) to the above group.
- TRIO study (Naylor *et al.*, 2016), treatment cohort of 180 postmenopausal women who received one of three oral bisphosphonate treatments and were sampled at baseline and at 48 weeks (retention of 149 participants). The participants had no other co-morbidities that could affect bone turnover.

The subjects in the TRIO study were randomized to oral alendronate 70mg weekly, risedronate 35mg weekly or ibandronate 150mg monthly, as well as receiving calcium carbonate 3 g (1200 mg elemental calcium) and cholecalciferol 20 micrograms (800 IU) a day. The adherence to the medication had been monitored using medical events monitoring system (MEMS). The subjects had been screened for any co-morbidity affecting bone metabolism and therefore an informed assumption was made that there was no relevant clinical information to record. The calcium and vitamin D supplement was initiated at the same visit from which the sample for baseline biochemical measurement was taken.

For the purpose of this phase of my work, the “Trial cohort” were an all-female cohort from the TRIO study who acted as their own controls before starting oral bisphosphonate treatment.

The “Clinical cohort” were a mixed gender group treated for at least 3 months on oral alendronate 70mg weekly, risedronate 35mg weekly or ibandronate 150mg monthly, and a clinical group not on treatment, that were matched to them as above.

Table 4.1 shows the proposed subject characteristics.

	TRIO COHORT	CLINICAL COHORT
Inclusion	Tscore ≤ -2.5 OR ≤ -1 with non-traumatic fracture Ambulatory ≥5 years postmenopausal ≤85 years old 18≤ BMI ≤35	Patient on oral bisphosphonate > 3m OR On no treatment > 12m Capacity to consent to a blood test for research
Exclusion	Co-morbidity affecting bone metabolism	None

4.

Table 4.1. Proposed subject characteristics by Inclusion and Exclusion Criteria for the Trial and Clinical cohort

3 Methods

3.1 Raw Data Collection

3.1.1 Sample collection

The process of sample collection and handling is described in Appendix B and in Chapter 2. The samples were taken at any time of day and not under fasting conditions for the Clinical cohort and taken first thing in the morning from fasted subjects for the Trial cohort.

3.1.2 Clinical information collection

Clinical information was assimilated for the Clinical cohort only as any factors effecting bone metabolism was excluded during the screening process for the Trial cohort.

The full list of factors assimilated in the clinical summary can be referred to in the materials and methods and in subsequent results and the tools for data collection are in Appendix A and C. If data was not extracted from the ArQ database for a subject, the assumed and documented response was negative rather than of missing data.

3.1.3 Dual-energy x-ray absorptiometry data

The contrasting bone densitometry data collected for each cohort has been described in Chapter 2.

3.1.4 Biochemical measurements

- PINP (Roche Total, IDS Intact)
- OC (Both Roche and IDS N-Mid)
- 25-hydroxyvitamin D (IDS iSYS only)
- Plasma Intact PTH (IDS iSYS only)

3.2 Data Processing

The raw BTM measurements were converted to the agreed units as stated in Chapter 2 and the clinical creatinine results were processed into the corresponding Chronic Kidney Disease (CKD) group (Levey *et al.*, 2009; Stevens, Levin and Members, 2013).

The Clinical cohort which contained a male subset was taken as one set of measurements, evaluated with the geometric mean of the female reference interval (threshold).

Total hip Tscore for the baseline characteristic of each cohort, was calculated from BMD using the gender specific NHANES III database (Shuhart, Yeap and al, 2019) for both the clinical densitometry data and the TRIO data.

Out of the bone density scans that fulfilled timing criteria for the Clinical cohort, further exclusions were made before proceeding to rate of change, of bone mineral density analysis, as per the details in Chapter 2. This excluded n=10 from Trial cohort for lack of paired results and n=5 from Trial cohort for weight change.

The BMD percent change for the Clinical cohort was expressed as an annualised percent change, as the DXA scans were conducted with a one to five year interval. The mean BMD percent change for the Trial cohort was over 48 weeks on the oral bisphosphonate therapy.

3.3 Statistical Analysis

Differences in the characteristics that can affect bone turnover were analysed with either a 2-tailed t-test with 95% confidence intervals (independent samples for the Clinical cohort or

between the cohorts and paired for the Trial cohort) or a 2-tailed Mann Whitney with 95% confidence intervals, depending on the pattern of distribution. The critical P- value was adjusted for the multiple testing on each group of data to 0.0125.

The distribution of the PINP and OC measurements were described by the median and interquartile range (IQR). It was deemed fair to conduct a statistical comparison between the two groups within each cohort, matching Clinical cohort groups, together with Trial subjects that were their own control. The distributions were subsequently compared using a rank comparison as described in Chapter 2.

ROC curve analysis for PINP and OC by each assay was performed with "on treatment" as a positive result and "no treatment" as the negative result. The statistical analysis of the difference in the area under the curve was performed with a null hypothesis that there was no difference in the AUC for Trial and Clinical cohort on oral bisphosphonates. The outcome of the previous comparisons of PINP and OC in the "no treatment" and "on treatment" groups from each cohort, helped to explain where any difference in the performance on ROC analysis came from.

The sensitivity or specificity of the biological target threshold with each assay was calculated.

Appropriateness of categorising groups as "responder" to treatment by achievement of the threshold, was explored, by comparison of the achievement of the two biological targets, in the Trial cohort. This was evaluated by comparison of annual BMD percent change when either target was achieved. Reflecting back to the Clinical cohort, the null hypothesis showed that there was no statistical difference in change of bone mineral density in subgroup "on treatment", that achieved the biological target and a subgroup "on treatment" that did not achieve the biological target, as described in Chapter 2.

Finally, the correlation coefficient by Spearman Rank was calculated, with a null hypothesis that there was no correlation between PINP or OC by both IDS and Roche assays and the annualised percent change in BMD.

4 Results

4.1 Subject Characteristics

4.1.1 Baseline demographics

There were 133 women in the Clinical Cohort and only 34 men (n=17 with n=17 matched). The Trial cohort consisted of 135 women (table 4.2). The Clinical cohort "on treatment" were predominantly on weekly alendronate (72%), compared to a roughly equal three-way split in oral bisphosphonate in the Trial cohort, to which the subjects were randomised. As per the design of the studies, the Trial cohort "on treatment" group which I will refer to as T+, had been on oral bisphosphonates for 48 weeks, compared to the inclusion criteria in the Clinical cohort "on treatment" group, which I will refer to as C+, of being on oral bisphosphonate for at least 12 weeks. There was in fact an unclear start date in the ArQ database for 27 subjects of the C+ group, but all study recruiters were aware of the inclusion criteria.

Table 4.2 details the characteristics of both the Trial and Clinical cohort, broken down into the group not on oral bisphosphonates ("no treatment" which I will refer to as T- and C- respectively) and the group "on treatment", which for the Trial cohort represents the same participants 48 weeks on, and for the Clinical cohort, matched participants. Table 4.2 demonstrates the success in the subject selection and matching of the Clinical cohort.

The Trial cohort demonstrated a normal distribution for all characteristics at baseline and at 48 weeks of treatment with oral bisphosphonates, with the exception of BMI, which had a positive skew. This was similar in the Clinical cohort.

	TRIO COHORT		CLINICAL COHORT	
	No treatment (Baseline) n= 135	Oral Bisphosphonate (48 weeks) n= 135	No treatment n= 81	Oral Bisphosphonate n= 86
Sex	135 female		133 female; 34 male	
Age at recruitment (yrs)	67.0 (7.4) *	68.0 (7.4)	69.8 (9.7)	70.8 (9.3)
BMI (kg/m^2)	25.8 (23.7 to 29.0) #	25.6 (23.3 to 28.6) #	25.8 (22.9 to 29.4) #	25.1 (22.4 to 28.7) #
Total hip Tscore	-1.35 (0.85) *	-1.23 (0.83) †	-1.62 (0.82)	-1.68 (0.92)
25OHD (nmol/L)	68.1 * (49.4 to 84.6) #	85.1 (71.4 to 96.8) #	75.1 * (55.8 to 96.2) #	88.9 (70.5 to 105.6) #
PTH (ng/L)	36.7 † (28.6, 44.4) #	34.8 † (27.2, 47.4) #	29.3 (21.9, 43.4) #	28.2 (19.2, 40.2) #
Alendronate n	0	47	0	64
Risedronate n	0	43	0	22
Ibandronate n	0	45	0	2

Table 4.2. Treatment group characteristics by cohort.

Mean and (SD) for each characteristic that could impact bone metabolism with Median and IQR (y to z) # given when distribution doesn't satisfy normal distribution. Significant difference p <0.0125 by t test or rank comparison #, denoted within cohort treatment groups * and denoted between treatment groups of the two cohorts †.

All recruited females in the Clinical cohort were post menopause, as were the Trial subjects, by definition of the inclusion criteria. There was no statistical difference in the mean age of T- to C- and T+ to C+, despite the Clinical cohort groups mean being 2-3 years older than the Trial cohort groups. The mean age of the two groups in the Clinical cohort was well matched with no difference. As the Trial subjects acted as their own controls, there was by design 48 weeks difference in the mean age of T- and T+.

There was no difference in the mean BMI in any of the 4 comparisons.

The Tscore is the last of the matched characteristics in the Clinical cohort and there was no difference between C- and C+. Because the trial subjects had baseline measurements before commencing treatment, the Tscore mean was significantly different as one would predict

(p<0.0001). The difference in the mean Tscore of T- group to C- group was not significant. The mean Tscore was significantly different (p=0.0003) between T+ and C+ namely because of the improvement in the Trial cohort 48 weeks into treatment, with the Clinical cohort remaining matched by design.

4.1.2 Subject biochemical descriptors

Comparison of the biochemical data of characteristics of bone metabolism was performed by rank comparison, because 25-hydroxyvitamin D (25OHD) and intact parathyroid hormone (iPTH) have a negative and a positive skew respectively. This was stronger in the Clinical cohort.

There was a significant difference of 25-hydroxyvitamin D, within each cohort between the "no treatment" and "on treatment" subjects (p=0.007 Clinical cohort and p<0.0001 Trial cohort). The 25OHD measurement in T+ and C+ subjects was higher as noted in table 4.2. The Clinical cohort patients were recommended cholecalciferol supplementation for bone health as a first line to bisphosphonates and table 4.3 shows the number of subjects this was recorded in. The Trial cohort had supplementation with 800 Units daily cholecalciferol, between the "no treatment" baseline and 48 weeks "on treatment" visit. There was no significant difference between T- and C- or T+ and C+.

The median of the plasma iPTH in both groups of the Clinical cohort was within the normal clinical population reference range, recorded in table 4.2, and was not significantly different between C+ and C-. As a determinant of bone turnover, disorders of parathyroid hormone were screened for at recruitment in the Trial cohort. Although the T+ median value of PTH was lower than the T- median value, there was no significant difference between these groups. The baseline measurements in the Trial cohort were made before calcium was supplemented with 1200mg elemental calcium, whereas the week 48 measurements were all on calcium supplements, equivalent to the level of supplementation observed in the C+ group as described in table 4.3. There was a significantly lower PTH level in the C- group to T- (p=0.0008). The proportion of recorded calcium supplementation in this group of the Clinical cohort is described in table 4.3. There remained a significantly lower PTH level in the C+ group to T+ (p=0.0032).

4.1.3 Clinical cohort co-morbidities affecting bone metabolism

The available clinical information representing factors that can affect bone turnover was extracted from the ArQ database for every Clinical cohort subject and recorded as descriptive frequencies in table 4.3.

		NO TREATMENT n out of 81 (%)	ORAL BISPHOSPHONATE n out of 86 (%)
Calcium Supplement		35 (43)	77 (90)
Cholecalciferol Supplement		38 (47)	77 (90)
Activated Vitamin D		2 (3)	7 (8)
Fracture within recent 12m		13 (16)	5 (6)
CKD	1	17 (23)	19 (22)
	2	44 (59)	50 (58)
	3a	10 (13)	14 (16)
	3b	3 (4)	3 (4)
	4	1 (1)	0
	Missing	6	-
Coeliac disease		2 (3)	0
Hypo-, Hyper -thyroid (both)		11,1 (15)	12,1 (15)
Treated Diabetes Mellitus		3 (4)	8 (9)
Biochemistry of PHPTH		2 (3)	3 (4)
Glucocorticoid within 1 week		15 (19)	22 (25)

Table 4.3. Clinical cohort medical characteristics and relevant supplementation by treatment group.

Number of the treatment group in which each medical co-morbidity or supplement was recorded at source (Integer percentage calculated from the total group number).

Renal function was available for the majority of the subjects (n=86 C+ group, n=75 C- group). No subject in the treatment group had an eGFR less than 30mL/min/1.73m^2. The distribution across the Chronic Kidney Disease (CKD) groups appeared similar in both the "no treatment" and "on treatment" groups.

In the absence of subjects with CKD 4 or lower in the treatment group, the prescription of activated vitamin D was more likely to have been for hypoparathyroidism or hypophosphataemic rickets.

The incidence of other endocrine conditions such as treated diabetes mellitus, hyperthyroidism and primary hyperparathyroidism were low, as was the noted presence of coeliac disease. Hypothyroidism was a common co-morbidity at roughly 12% of the subjects in both treatment groups.

The percentage of patients with Diabetes mellitus requiring treatment in C+ was greater than in C- group.

Up to 25% of both C+ and C- had recorded a recent glucocorticoid prescription, excepting inhaled corticosteroids. The detail of delivery and dose was not included in table 4.3, except within the week prior to recruitment and measurement of PINP and OC. As the incidence was similar in both groups, the relevance of this will be considered further, when evaluating the sensitivity and specificity.

The factors with the biggest percentage difference between the C+ and C- groups in the Clinical cohort were the fracture history and supplement prescription. 16% of the "no treatment" group had experienced a fragility fracture within the previous 12 months. The DXA from which they were recruited would have been part of the fracture risk assessment, in advance of treatment recommendation. Finally, 90% of the C+ group were reported to take calcium supplements compared to 43% of the C- group.

This information was important in evaluating the usefulness of interpretating an isolated level of PINP or OC in a clinical patient.

4.2 Bone Turnover Marker Outcomes

4.2.1 Distribution of PINP and OC measurements

Biochemical measurements for PINP and OC by two assays were successfully acquired on the IDS iSYS and Roche Cobas e411 autoanalysers from almost all of the defrosted serum samples. Amongst the samples of the C- cohort, there were 80 or more biochemical measurements out of 81 possible included samples. Amongst the T- (baseline) samples, successful measurements were made in at least 130 of the possible 135 included samples, with the Roche OC measurement appearing most susceptible to a null outcome. All 86 of the C+ yielded all biochemical measurements, with only 4 samples from the T+ cohort included samples, lacking PINP and OC measurements on the Roche Cobas e411

autoanalyser. To maximise the measurements from which the distributions are formed, table 4.4 gives the number (n) for all included samples, with the distribution for each assay taking into account only available measurements.

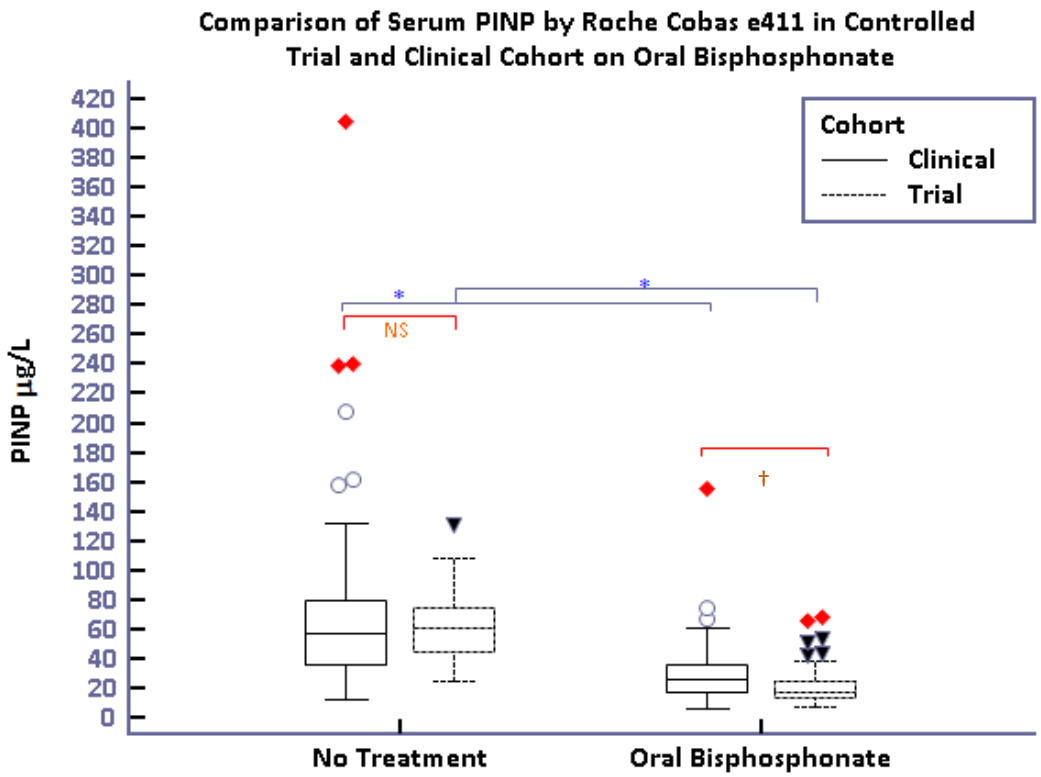
		NO TREATMENT		**ORAL BISPHOSPHONATE TREATMENT**	
		CLINICAL COHORT No treatment n= 81	TRIAL COHORT Baseline n= 135	CLINICAL COHORT Oral >3 Months n= 86	TRIAL COHORT Oral for 48 wks n= 135
PINP µg/L	IDS iSYS	52.5 * (34.9 to 78.4)	52.0 * (40.8 to 66.1)	24.6 † (14.3 to 34.5)	14.5 (11.6 to 21.3)
	Roche Cobas e411	57.6 * (35.9 to 80.0)	60.6 * (44.9 to 74.5)	25.8 † (17.9 to 36.4)	17.5 (13.8 to 24.6)
OC µg/L	IDS iSYS	19.8 † * (11.3 to 29.1)	25.7 * (21.1 to 31.7)	9.4 (6.2 to 14.2)	10.8 (8.9 to 13.2)
	Roche Cobas e411	19.3 † * (12.8 to 27.2)	24.3 * (20.5 to 30.5)	10.5 † (7.4 to 14.9)	12.0 (9.9 to 14.6)

Table 4.4. Median and IQR for the two bone turnover markers by two assays for the no treatment and on treatment group with oral bisphosphonates for each cohort.

Significant difference p <0.0125 by rank comparison, denoted within cohort treatment groups * and denoted between treatment groups of the two cohorts †.

The distribution of the PINP and OC measurements are described by the median and interquartile range (IQR). With the exception of the T- measurements, the groups did not satisfy a normal distribution. The measurements in each group demonstrated a positive skew, which was exaggerated in the “on treatment” groups. With both BTM by each assay, the T- measurements demonstrated a smaller IQR than the C- group measurements as seen by the boxplots of figure 4.1 and 4.2 (approx. 30 µg/L vs 45 µg/L for PINP and 10 µg/L vs 19 µg/L for OC). The “on treatment” groups of both cohorts also demonstrated a much smaller IQR, relative to a lower median value. Variability in the Clinical cohort relative to the Trial cohort was still increased, with a broader IQR for the PINP distribution (approx. 20 µg/L vs 10 µg/L), and the OC distributions (approx. 8 µg/L vs 5 µg/L).

a)



b)

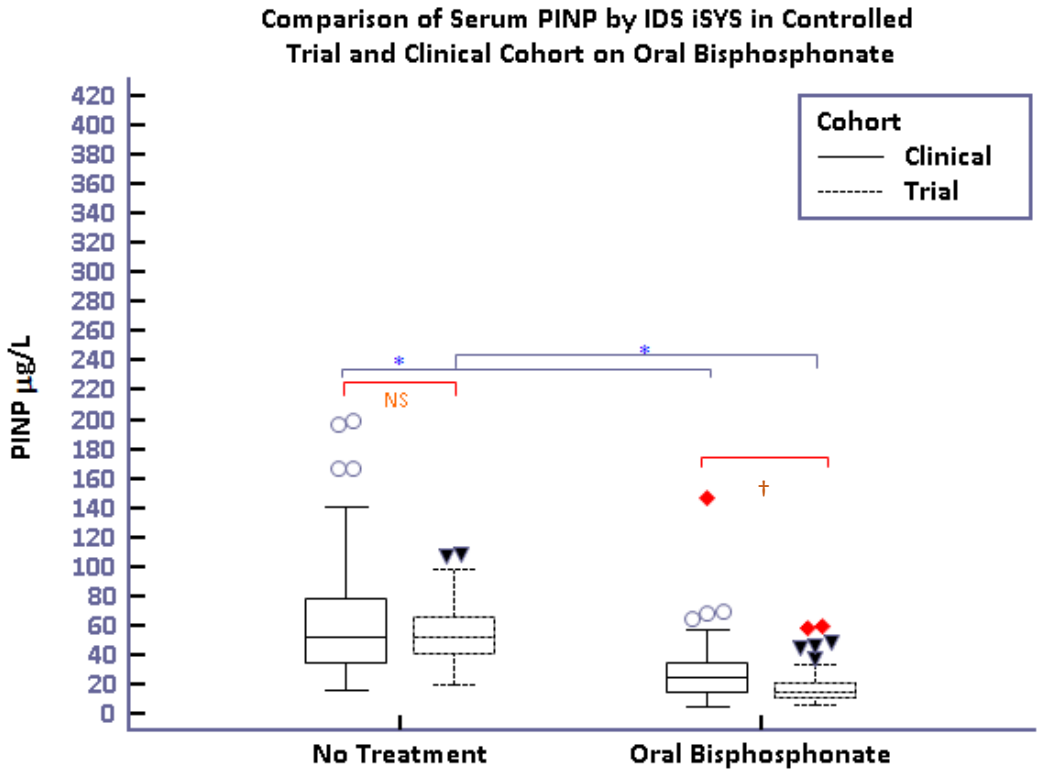
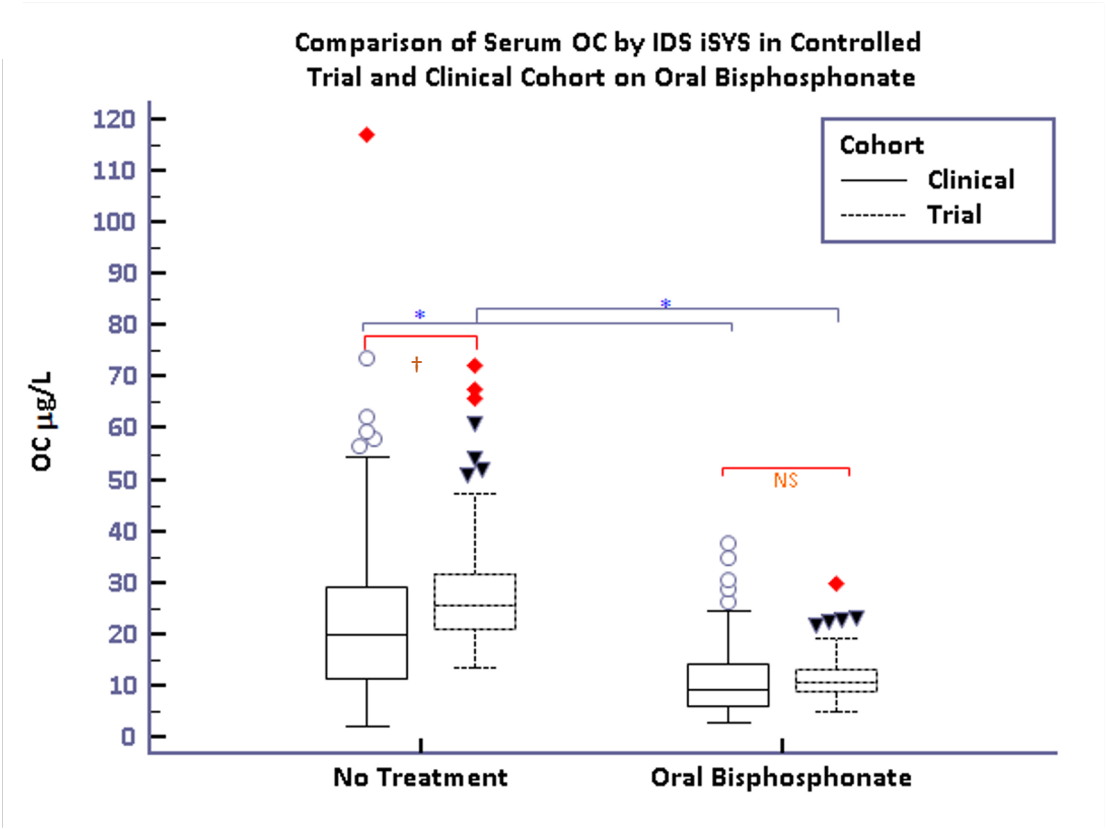


Figure 4.1. Comparison of PINP boxplots measured by a) IDS iSYS b) Roche Cobas e411 in the No treatment and On treatment with oral bisphosphonate subjects of the Trial and Clinical Cohorts.

Rank comparison of the distributions with Sig difference p <0.0125 denoted within cohort groups * and denoted between groups of the two cohorts†.

a)



b)

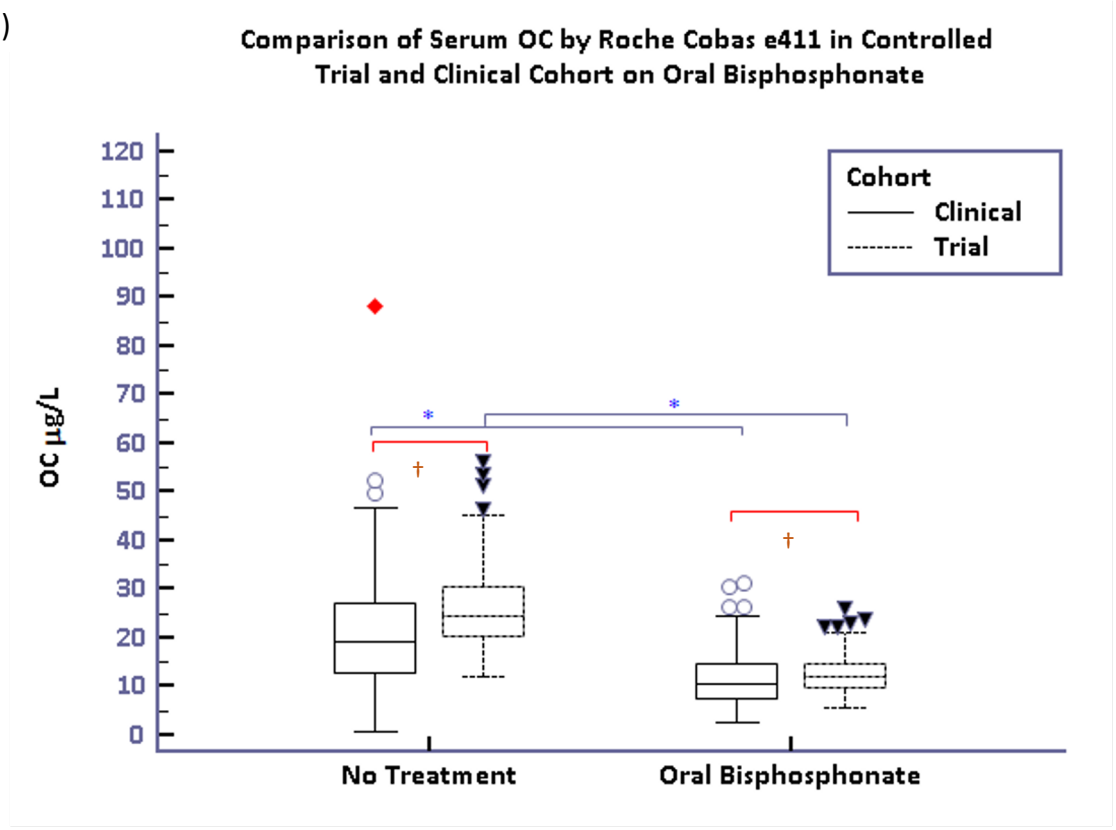


Figure 4.2. Comparison of OC boxplots measured by a) IDS iSYS b) Roche Cobas e411 in the No treatment and On treatment with oral bisphosphonate subjects of the Trial and Clinical Cohorts.

Rank comparison of the distributions with Sig difference p <0.0125 denoted within cohort groups* and denoted between groups of the two cohorts†.

4.2.2 Rank comparison of between group and between cohort PINP and OC distributions

The boxplots of the BTM measurements of C- and T- and C+ and T+ groups are shown in figures 4.1 (PINP) and 4.2 (OC) with a graph for each assay. The statistical difference between each group is demonstrated by the connector lines, with symbols described in the figure legend.

PINP and OC by both IDS iSYS and Roche Cobas e411 assays, were significantly different between the within each cohort groups (C- to C+ $p<0.0001$ and T- to T+ $p<0.0001$).

A lower median OC was found in the C- group than the T- group and this was significantly different with each OC assay ($p<0.001$). The median OC by each assay was also lower in the C+ group than the T+ group. This difference was significant for the OC distribution by Roche Cobas e411 ($p= 0.0058$) but not by IDS iSYS ($p= 0.0172$).

There was no significant difference in the distribution of PINP of the C- and T- groups by either assay, but on oral bisphosphonate treatment the median of the PINP measurements of the T+ group was lower than those of the C+ group, with a significant difference (IDS iSYS and Roche Cobas e411 $p<0.0001$).

4.2.3 Specificity and sensitivity of BTM assay by threshold

The specificity and sensitivity of the PINP and OC with the two assays to detect treatment effect at threshold has been calculated in table 4.5. The threshold value is the assay specific GM of the premenopausal reference interval.

On the graphs in figures 4.3 and 4.4 the percentage of PINP and OC measurements in the "No Treatment" groups above the marked GM represents the specificity and the percentage of measurement in the "Oral Bisphosphonate" groups supressed below the same threshold line is the sensitivity, with one assay per graph.

The performance of the two BTMs in terms of the sensitivity and specificity was consistent across each individual group with the exception of the IDS OC sensitivity and specificity by threshold in both cohorts as seen in table 4.5. Again, with the exception of the IDS OC, the sensitivity of the assays at threshold to detect treatment was better than the specificity.

Specificity to detect treatment was particularly poor by both BTMs in the Clinical cohort, compared to the Trial cohort. Sensitivity to detect treatment was also less in the Clinical cohort than the Trial cohort, but when calculated from both assays of PINP, was still above 80%.

The specificity calculated in the Trial cohort of controlled baseline measurements before treatment was also above 80%.

The exceptional results from calculation with the IDS OC assay were consistent across both Clinical and Trial cohort and how this exception impacts on the AUC as compared to the Roche OC assay has been discussed in the next section.

		SPECIFICITY		**SENSITIVITY**	
		CLINICAL COHORT n= 81	TRIAL COHORT n= 135	CLINICAL COHORT n= 86	TRIAL COHORT n= 135
% PINP VALUES	IDS iSYS	70	85	80	96
	Roche Cobas e411	66	82	83	95
% OC VALUES	IDS iSYS	68	96	71	76
	Roche Cobas e411	59	85	87	86

Table 4.5. Percentage of the two bone turnover markers by two assays relative to the Geometric Mean (GM) of the assay specific reference interval.

Representing Specificity (% of the no treatment group that ≥ GM) and Sensitivity (% of on treatment group with oral bisphosphonates that ≤ GM), for each cohort. GM for IDS PINP is 37.2 µg/L, Roche PINP is 41.7 µg/L, IDS OC is 13.5 µg/L and Roche OC 17.8 µg/L as shown in figures 4.3 & 4.4

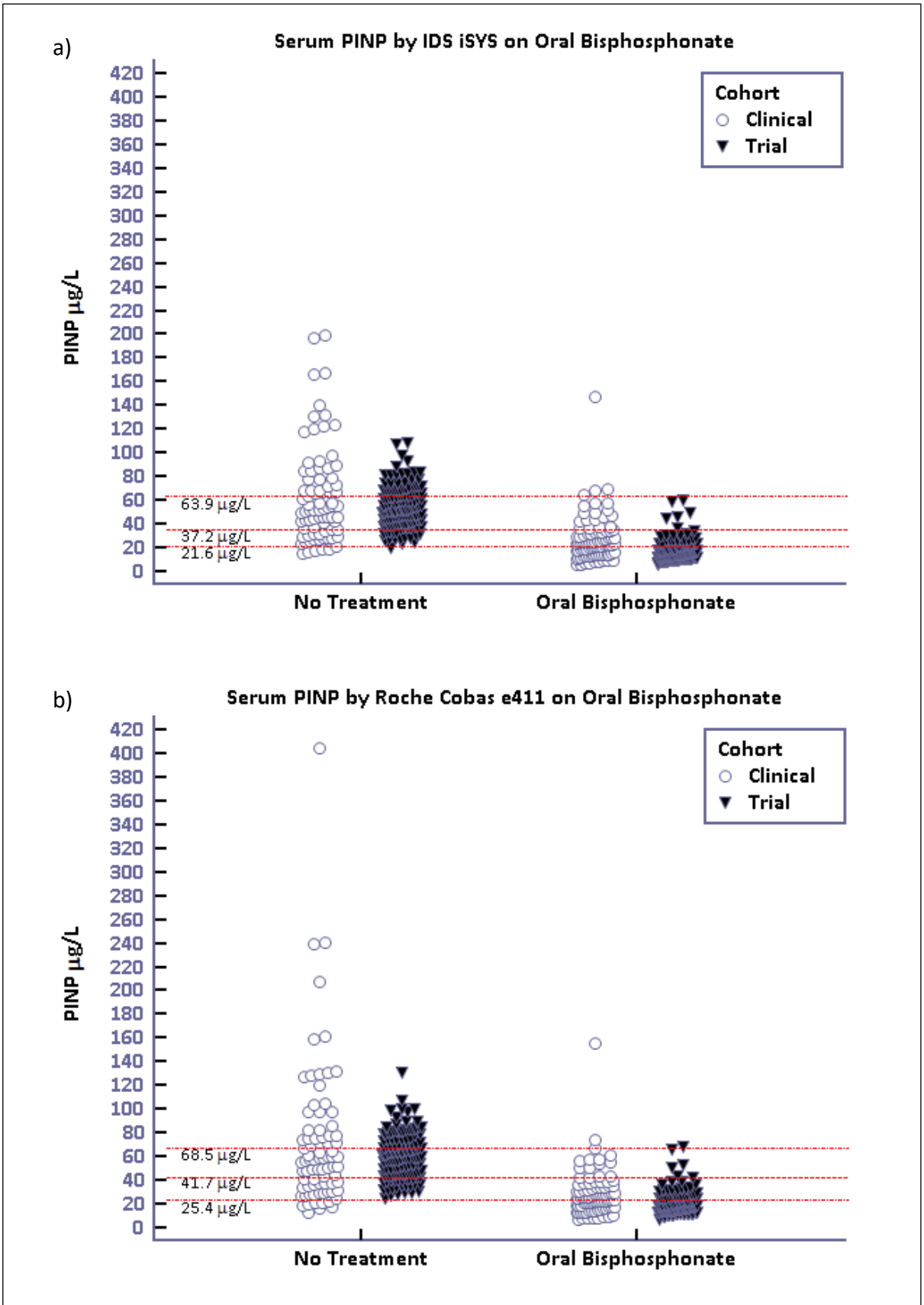
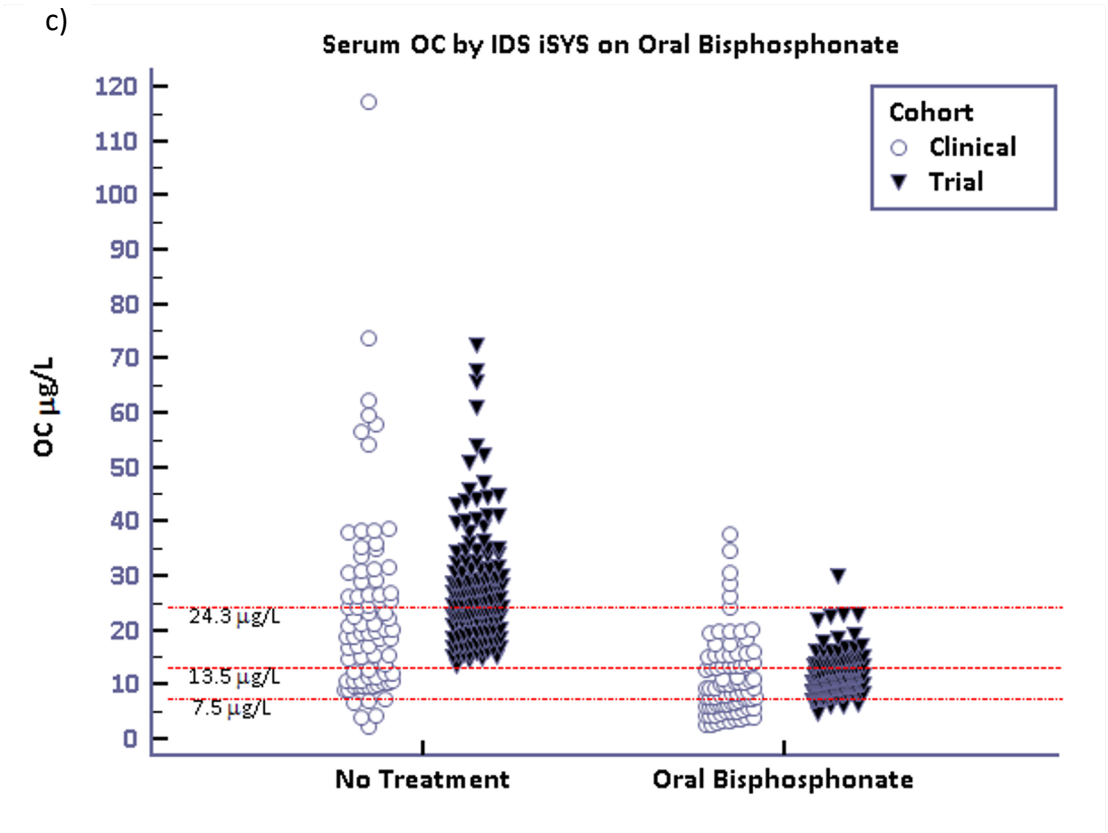


Figure 4.3. Comparison of a) PINP by IDS iSYS; b) PINP by Roche Cobas in the No treatment and On treatment with oral bisphosphonate subjects of the Trial and Clinical Cohorts.

The assay specific geometric mean (37.2 µg/L, 41.7 µg/L) and upper (63.9 µg/L, 68.5 µg/L) and lower (21.6 µg/L, 25.4 µg/L) reference limit of PINP.



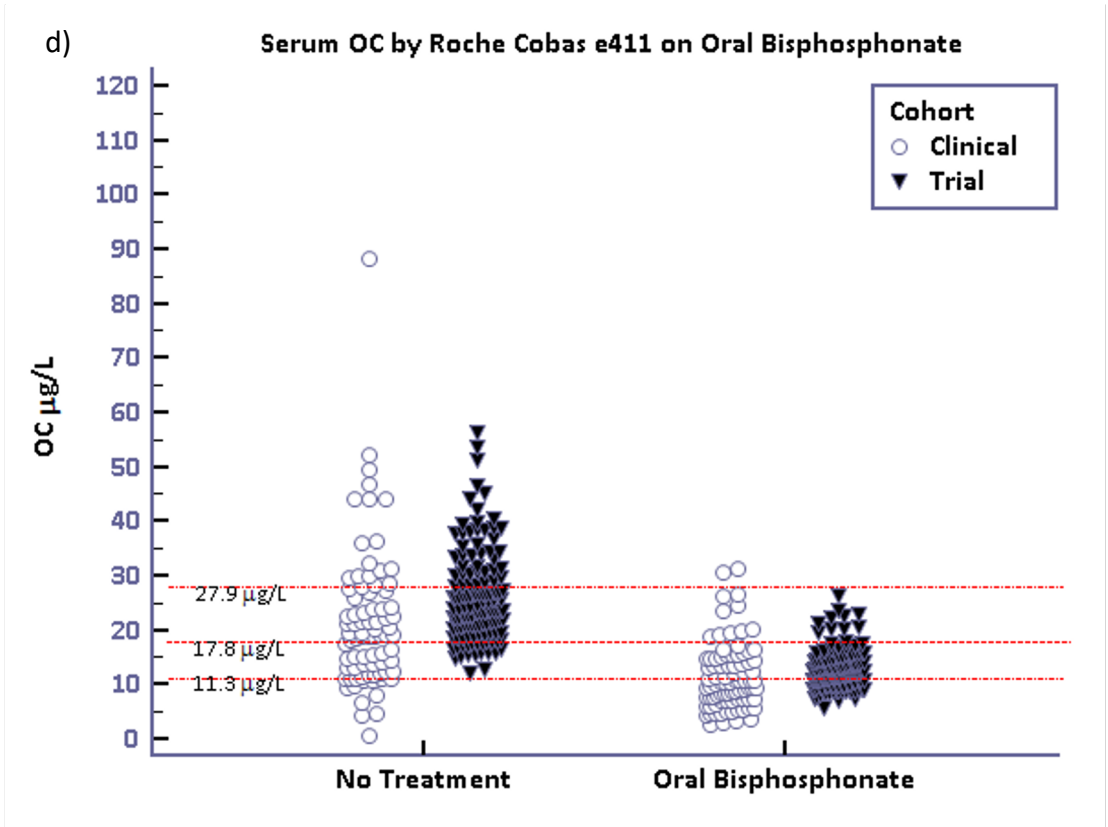


Figure 4.4. Comparison of c) OC by IDS iSYS; d) OC by Roche Cobas e411 in the No treatment and On treatment with oral bisphosphonate subjects of the Trial and Clinical Cohorts.

The assay specific geometric mean (13.5 µg/L, 17.8 µg/L) and upper (24.3 µg/L, 27.9 µg/L) and lower (7.5 µg/L, 11.2 µg/L) reference limit of OC.

Measurements in figures 4.3 and 4.4 below the GM in the "no treatment" groups were considered 'false positives'. Due to the low specificity of the BTM to detect treatment in the Clinical cohort, there were between 30-42% 'false positives'. Further descriptive analysis has been made of clinical factors impacting to reduce the bone turnover in this group. Table 4.6 uses the PINP measurements by IDS iSYS as an example. This assay has the best specificity at threshold of 70% and the best AUC for discrimination of whether a subject is "on treatment", as shown in table 4.7. The frequency and percentage of occurrence of clinical factors in the subgroup of C- with a PINP at or below the GM of the reference interval, was compared to the subgroup of C- with PINP above threshold.

The main factors that were more prevalent in the subgroup with PINP by IDS iSYS below the threshold, were; higher percentage of calcium and cholecalciferol supplementation, and exposure to previous anti-resorptive, particularly within 5 years. Glucocorticoid prescription was similar between the two subgroups.

		NO TREATMENT ≤ IDS PINP GM n out of 24 (%)	**NO TREATMENT > IDS PINP GM n out of 56 (%)**
Calcium Supplement		12 (50)	22 (39)
Cholecalciferol Supplement		14 (58)	23 (41)
Previous anti-resorptive		8 (33)	11 (20)
Anti-resorptive ceased	< 5 yrs	6 out of 7 known	4 out of 6 known
	Unknown	1	5
CKD	1	5 (21)	11 (20)
	2	12 (50)	32 (57)
	3a	2 (8)	8 (14)
	3b	1 (4)	2 (4)
	4	0 (0)	1 (2)
	Unknown	4 (17)	2 (4)
Glucocorticoid within 1 week		5 (21)	10 (18)

Table 4.6. Breakdown of clinical factors in the Clinical cohort no treatment group.

Descriptive comparison in the number of subjects recorded at source to be exposed to these factors, in the subgroup with PINP by IDS iSYS below the GM and subgroup with PINP by IDS iSYS above the GM. N (Integer percentage calculated from the subgroup number).

4.2.4 ROC Curve analysis

Figure 4.5 shows the BTMs performed better in a trial than in clinical practice as shown by the area under the curve (AUC) in a ROC curve analysis which indicates the performance of a tool to differentiate between the subjects “on treatment” and “not on treatment”. Table 4.7 demonstrates that the difference in performance between the two cohorts with each assay was significant (all comparisons p<0.0001).

		CLINICAL COHORT n= 167		**TRIAL COHORT n= 270**	
		AUC	Youden Criterion µg/L	AUC	Youden Criterion µg/L
PINP	IDS iSYS	0.834 † 0.769 to 0.887	35.2	0.975 0.948 to 0.990	26.4
	Roche Cobas e411	0.822 † 0.755 to 0.877	43.4	0.973 0.946 to 0.989	34.2
OC	IDS iSYS	0.789 † 0.719 to 0.848	16.1	0.976 0.949 to 0.990	15.1
	Roche Cobas e411	0.795 † 0.725 to 0.853	17.2	0.961 0.929 to 0.981	16.9

Table 4.7. ROC curve analysis expressed for the Clinical and Trial cohort as AUC with 95% confidence intervals for the two bone turnover markers by two assays.

Significant difference p <0.05 by comparison of independent ROC curves, denoted between the two cohorts †. Youden criteria can be viewed relative to the GM of the reference interval; IDS PINP is 37.2 µg/L, Roche PINP is 41.7 µg/L, IDS OC is 13.5 µg/L and Roche OC 17.8 µg/L.

In table 4.7 it was seen, with the 95% confidence intervals of the area under the curve, that these overlap within each cohort for PINP and OC by both assays. Therefore, there was no significant difference in the performance of PINP to OC or between the Roche Cobas e411 and IDS iSYS assays. This was a consistent finding regardless of whether the AUC was greater for the Trial cohort, or less for the Clinical cohort.

The Youden criterion given in table 4.7, at which both sensitivity and specificity of the tool could be maximised, was consistently lower for the Trial cohort. However, in the Clinical cohort, the Youden criterion is very similar to the GM of the premenopausal reference interval.

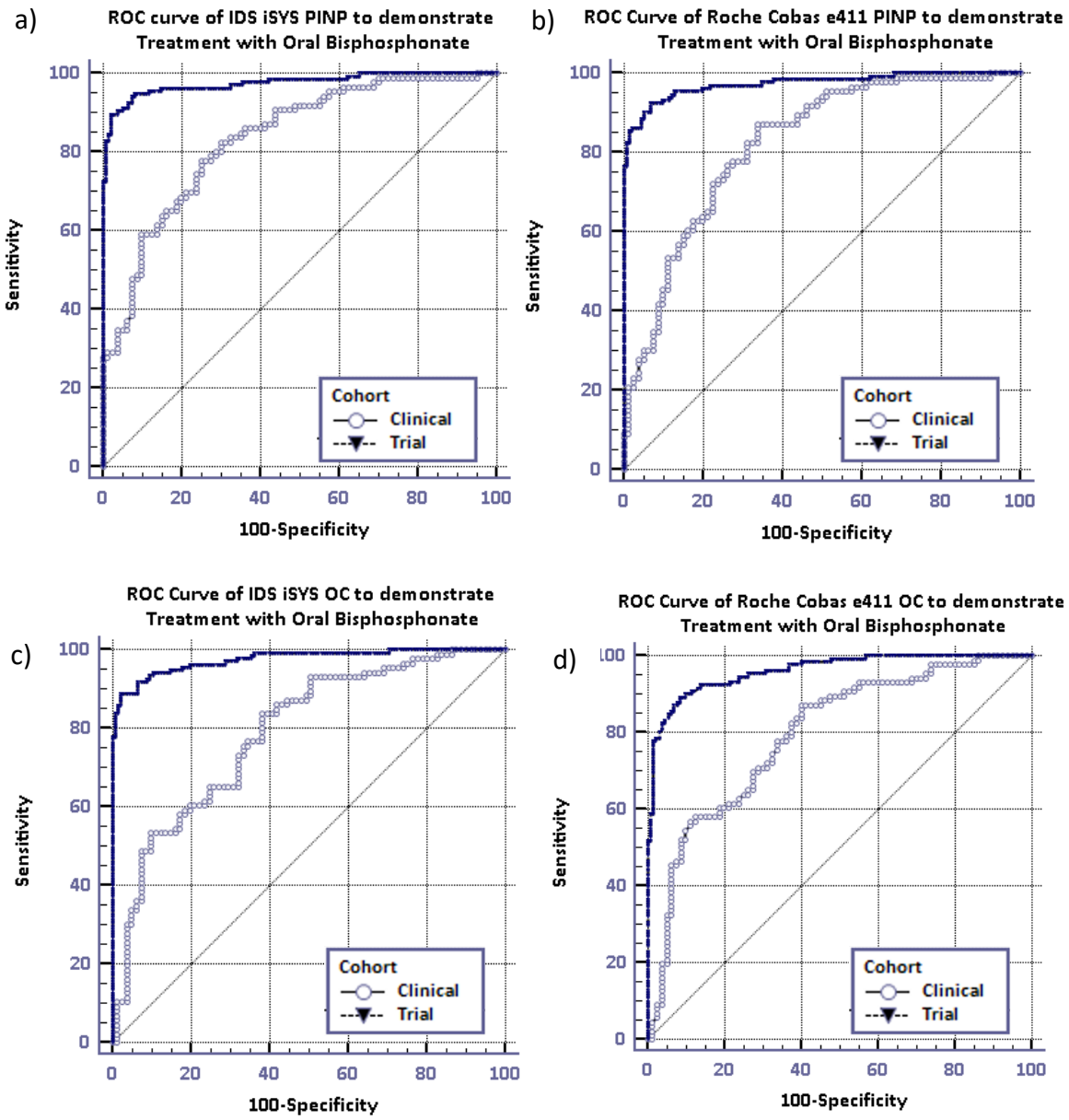


Figure 4.5. ROC Curve of measurements of a) PINP by IDS iSYS, b) PINP by Roche Cobas e411, c) OC by IDS iSYS; d) OC by Roche Cobas e411 in both Clinical and Trial cohort.

With the AUC demonstrating the performance of the tool in detecting treatment with oral bisphosphonate.

4.3 Validation

4.3.1 Comparison of responder by LSC and threshold in the Trial cohort

The measurements of BTM in the T- and T+ were from the same 135 subjects, at baseline on "no treatment" and at 48 weeks on the oral bisphosphonates, which enabled the change in PINP and OC by each assay over the 48 weeks to be calculated.

The percentage of subjects achieving the least significant change (LSC) as an assay-specific absolute figure of PINP and OC was descriptively compared in table 4.8 against the percentage of subjects achieving less than the GM (sensitivity at threshold).

		Achieving LSC	**Achieving Threshold**
% PINP VALUES	IDS iSYS	98	96
	Roche Cobas e411	98	95
% OC VALUES	IDS iSYS	97	76
	Roche Cobas e411	94	86

Table 4.8. Percentage of the subjects in the Trial cohort achieving LSC or threshold on oral bisphosphonate treatment with two bone turnover markers by two assays.

The assay specific absolute LSC and GM of the reference interval are for IDS PINP; 6.69 µg/L & 37.2 µg/L, for Roche PINP; 7.88 µg/L & 41.7 µg/L, for IDS OC; 4.41 µg/L & 13.5 µg/L, for Roche OC; 4.11 µg/L & 17.8 µg/L.

The percentage achieving the LSC was very high across all assays of PINP and OC. This was equivalent to the percentage achieving the threshold for PINP. Figure 4.6 represents this in more detail to show that there were very few subjects that achieved the biological target by one measure and not the other. The lines of achieving LSC, mostly fell below the marked line of the GM. Whereas in the panels representing the OC, because the percentage achieving LSC was higher than the percentage achieving threshold, many of the lines show LSC achieved but not falling below the line of the GM.

Two subjects can be seen in figure 4.6 to have rising PINP and OC over the 48 weeks, labelled therefore as subgroup LSC negative. This was also reported on in the TRIO publication (Naylor *et al.*, 2016) where the explanation was given to be poor compliance (less than 50%) or for one subject, no available compliance data.

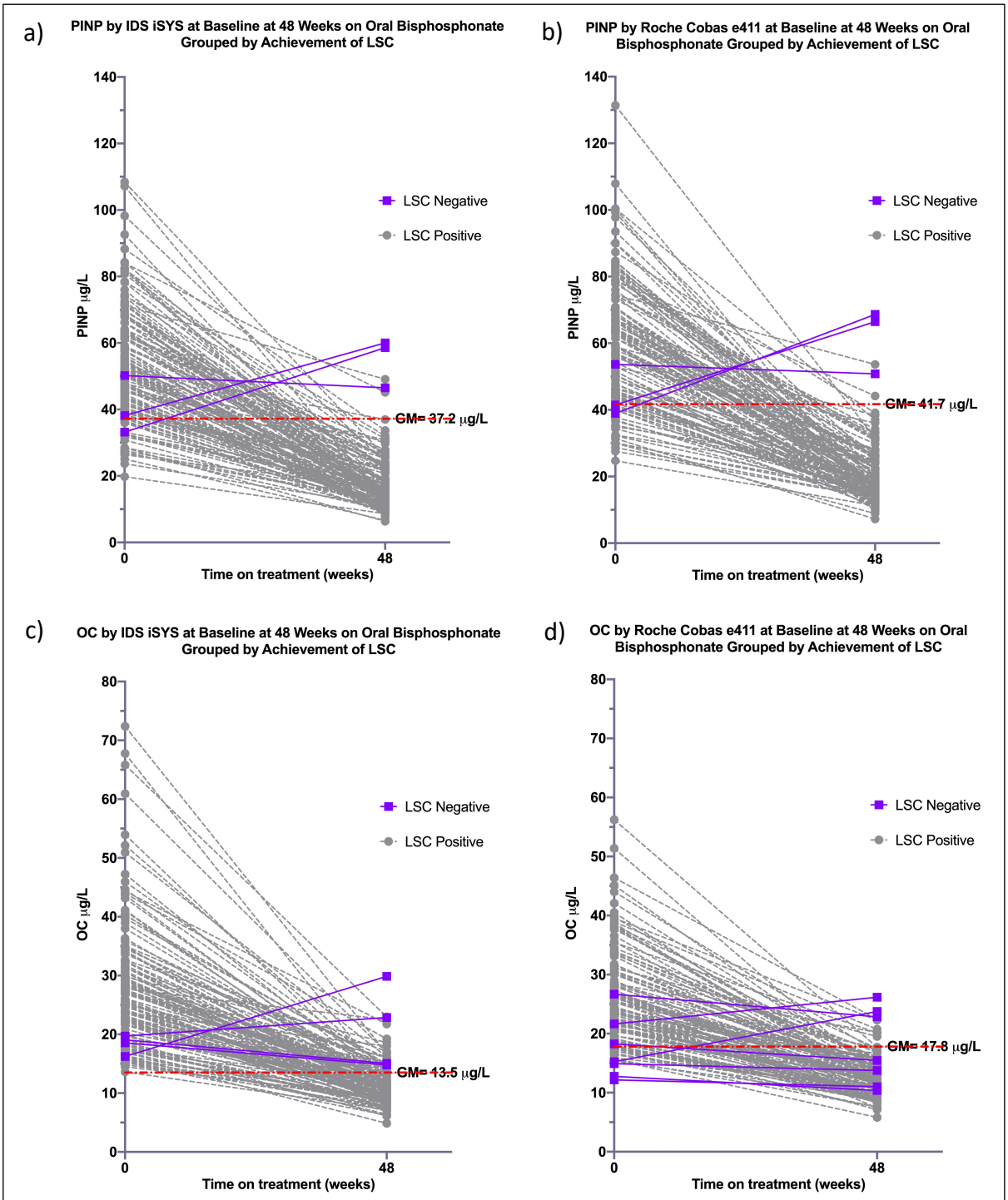


Figure 4.6. Baseline and week 48 measurements of a) PINP by IDS iSYS n=135, b) PINP by Roche Cobas e411 n=128, c) OC by IDS iSYS n=135; d) OC by Roche Cobas e411 n=126 in the Trial cohort individual subjects.

Subgroups that achieve assay specific LSC of the BTM and those that did not with both the baseline and week 48 measurements shown relative to the assay specific GM; IDS PINP is 37.2 µg/L, Roche PINP is 41.7 µg/L, IDS OC is 13.5 µg/L and Roche OC 17.8 µg/L.

4.3.2 Bone mineral density outcomes

The BMD percent change for the subjects in the Trial cohort, between baseline and week 48, followed a normal distribution. Comparison of BMD percent change when BTM LSC was achieved as the marker of treatment effect, to when the BTM threshold (GM) was achieved as the marker of treatment effect, can be seen in table 4.9.

There was a statistical likelihood that there was a gain in BMD, with a biochemical response to treatment, by either biological target. This can be seen in table 4.9 by the confidence interval of the mean BMD percent change in the subjects classified as “responders” by these targets, which did not cross zero.

		Mean BMD % change grouped by LSC	**Mean BMD % change grouped by GM (threshold)**
IDS iSYS PINP	Non-responder	1.05 (-21.12 to 23.21) n=2	0.61 (-2.29 to 3.50) n=4
	Responder	2.01 (1.44 to 2.58) n=118	2.04 (1.46 to 2.62) n=116
Roche Cobas e411 PINP	Non-responder	1.05 (-21.12 to 23.21) n=2	0.61 (-2.29 to 3.50) n=4
	Responder	2.01 (1.44 to 2.58) n=118	2.04 (1.46 to 2.62) n=116
IDS iSYS OC	Non-responder	-0.71 (-8.96 to 7.54) n=3	1.76 (0.19 to 3.34) n=28
	Responder	2.06 (1.50 to 2.63) n=117	2.06 (1.49 to 2.64) n=92
Roche Cobas e411 OC	Non-responder	-0.70 (-3.07 to 1.68)* n=6	1.95 (-0.36 to 4.26) n=8
	Responder	2.13 (1.56 to 2.71) n=114	1.99 (1.41 to 2.58) n=112

Table 4.9. Trial cohort Mean BMD Percent Change (and 95% CI) over 48 weeks in Responder and Non-responder subgroups for each assay of PINP and OC with least significant change (LSC) and threshold (GM) as the target.

N= 120 subjects included with all BTM and BMD measurements available at baseline and 48 weeks on oral bisphosphonate. Significant difference p <0.05 by independent samples t test, denoted within target groups * and denoted between target groups †.

The confidence intervals in the non-responder subgroups in table 4.9 were wide and cross zero, because the number of subjects not achieving LSC and threshold by BTM in the trial cohort, was between 2 to 8 subjects. The exception is the number of subjects that failed to meet threshold with OC on IDS iSYS assay (n=28, sensitivity= 76%). The mean BMD percent in this instance was in fact positive, adding evidence that the sensitivity result was an outlier.

When the response to treatment was measured with PINP by both assays, it can be seen in figure 4.6, that the group not achieving threshold (GM), were also predominantly those not achieving LSC. It therefore follows that the mean BMD percent change in the non-responders by the two targets, when measured with PINP were similar, as seen in table 4.9. Whereas it is clear from figure 4.6 that many of the subjects whose week 48 OC by either assay, was not below the GM, were responders by LSC, and for Roche OC, there were 4 subjects that achieved threshold but not LSC. The non-responder group of LSC demonstrated a more negative mean BMD percent change than the non-responders by threshold of OC.

There was no significant difference in how positive the mean BMD percent change was, when the responder was classed by achieving LSC or by threshold with each assay measurement and this is demonstrated figuratively in figure 4.7.

There was no significance difference in mean BMD percent change from baseline to week 48, between the responder and non-responders by each target with each BTM, with each assay, with one exception. The difference in mean BMD percent change in the 2 subgroups of using LSC as the target, measured by Roche Cobas e411 OC is 2.82, was significant with p= 0.029, as seen in table 4.9.

Figure 4.7 looked specifically at the mean BMD percent change, with standard error of the mean (SEM), in the responder subgroups only, when measured by both BTM with each assay, comparing the Trial cohort by both targets and the Clinical cohort by threshold target. The non-responder subgroups were not compared owing to the small subject numbers and wide 95% CI of the mean.

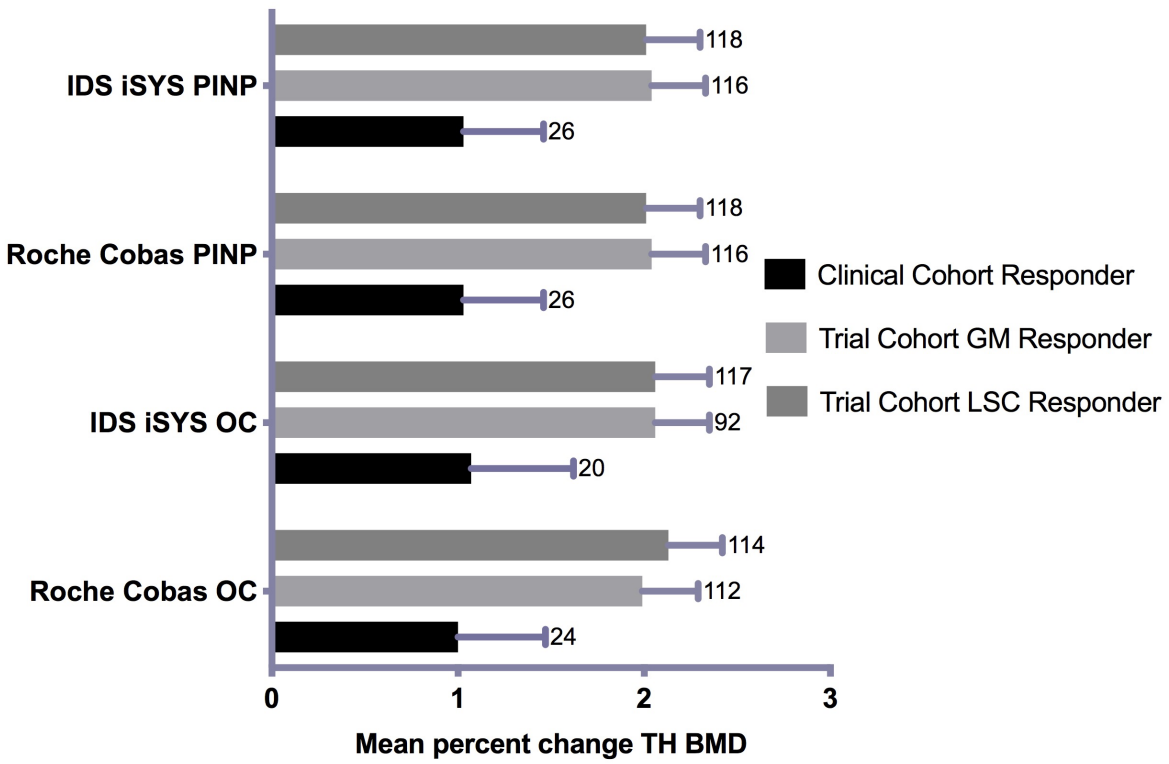


Figure 4.7. Percent Change in Total Hip Bone Mineral Density with Response in PINP and OC by two Biological Targets in the Trial and Clinical Cohorts.

The mean BMD percent change and SEM when selected postmenopausal women (n=120 over 48 weeks), and unselected clinical patients (n=30 annualised over 1-5 years) on three oral bisphosphonates reach the PINP or OC LSC and threshold on the IDS-iSYS autoanalyser and Roche Cobas e411 autoanalyser. No significant difference detected.

The data portrayed in figure 4.7 comparing the mean BMD percent change when the Trial cohort subjects and the Clinical cohort reached threshold is shown in more detail, alongside the non-responder groups, in table 4.10. Only 30 subjects in the Clinical cohort had available sequential comparable DXA scans “on treatment”. This was in part due to a lack of the exact date of commencement of the oral bisphosphonate in 26 out of 86 treated subjects.

As with the Trial cohort, if a response by threshold was identified in the Clinical cohort, in all except measurements with IDS iSYS OC, this conferred a positive mean BMD percent change, with 95% confidence interval (CI) that did not cross zero. These 95% CI of the mean BMD percent changes all overlap each other, determining no difference between BTM by

either assay. Despite appearances in figure 4.7 of the mean BMD percent change in the responders by threshold being more positive in the Trial than Clinical cohort, there was no significant difference between them.

		Mean BMD % change grouped by GM (threshold)	
		Clinical Cohort	Trial Cohort
IDS iSYS PINP	Non-responder	-0.44 (-2.38 to 1.5) n=4	0.61 (-2.29 to 3.50) n=4
	Responder	1.03 (0.13 to1.92) n=26	2.04 (1.46 to 2.62) n=116
Roche Cobas e411 PINP	Non-responder	-0.44 (-2.38 to 1.5) n=4	0.61 (-2.29 to 3.50) n=4
	Responder	1.03 (0.13 to 1.92) n=26	2.04 (1.46 to 2.62) n=116
IDS iSYS OC	Non-responder	0.34 (-0.54 to 1.23) n=10	1.76 (0.19 to 3.34) n=28
	Responder	1.07 (-0.09 to 2.23) n= 20	2.06 (1.49 to 2.64) n=92
Roche Cobas e411 OC	Non-responder	0.16 (-1.4 to 1.71) n=6	1.95 (-0.36 to 4.26) n=8
	Responder	1.00 (0.03 to 1.96) n=24	1.99 (1.41 to 2.58) n=112

Table 4.10. Trial cohort over 48 weeks compared to Clinical cohort annualised Mean BMD Percent Change (and 95% CI) in Responder and Non-responder subgroups with threshold (GM) as the target for each assay of PINP and OC.

BTM and sequential BMD measurements available on oral bisphosphonate for n=120 Trial cohort and n=30 Clinical cohort. No significant difference detected within or between subgroups.

The 95% CI of the mean BMD percent change in the non-responder subgroups of the Clinical cohort all crossed zero with small subject numbers. Although the mean BMD percent changes were lower in these subgroups than in the responders, there was no significant difference within the Clinical cohort between the mean BMD percent change in the two subgroups (responder and non-responder), when measured by either of the BTMs with each assay.

This was evaluated further by figure 4.8, in which the annualised BMD percent change on treatment with oral bisphosphonates for each subject was graphed against the BTM "on treatment".

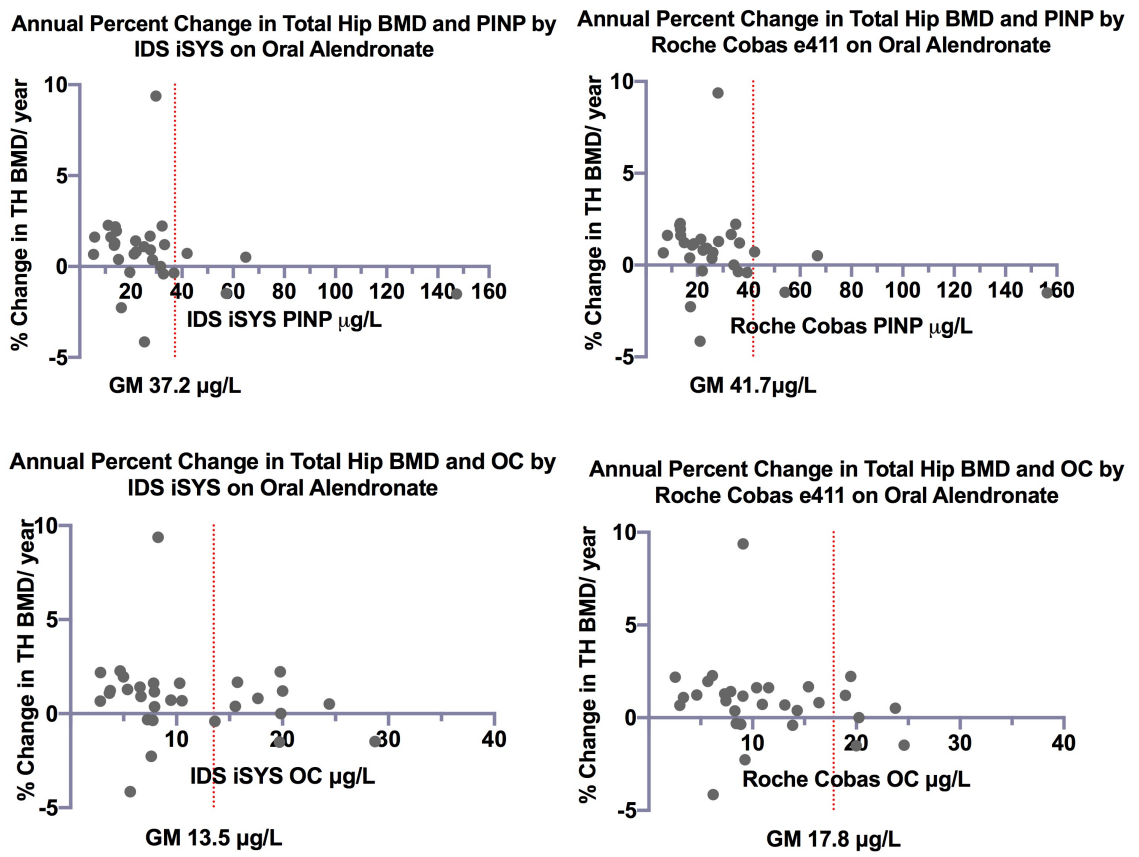


Figure 4.8. Percent Change in Total Hip Bone Mineral Density by PINP and OC by two assays.

The annualised BMD percent change in unselected clinical patients (n=30 over 1-5 years) on oral bisphosphonates against the PINP and OC measurements on treatment, with the IDS-iSYS autoanalyser and Roche Cobas e411 autoanalyser. "Responders" BTM are lower than the assay specific GM; IDS PINP is 37.2 µg/L, Roche PINP is 41.7 µg/L, IDS OC is 13.5 µg/L and Roche OC 17.8 µg/L

Spearman rank correlation analysis of the PINP measurements and BMD percent change, was significant (r= -0.41, p= 0.02 and r= -0.37, p=0.05 for IDS iSYS and Roche Cobas e411 measurements respectively). The lower the PINP measurement, the more gain in BMD "on treatment". The correlation co-efficient for the association of OC by IDS or Roche and BMD percent change, was not significant.

There were 2 subjects that consistently throughout figure 4.8 showed BTM supressed below threshold regardless of the assay but had a negative annualised BMD percent change. This

will have impacted on the mean BMD percent change in the Clinical cohort responder subgroup. Reflection of the clinical information for these 2 subjects revealed that one with supressed PINP and OC and an annual change in BMD of -2.27%, had the BMD change analysed over 4 years, whilst also on calcium and cholecalciferol, with CKD stage 3a, but with prescription of 7.5mg oral prednisolone within the week prior to recruitment. The second subject with an annual change in BMD of -4.15% did not have any clinical features that could have supressed the PINP and OC other than calcium and cholecalciferol prescription, and actually had a vertebral fracture identified 6 months before recruitment. The BMD analysis captured 2 ½ years, including a year prior to the bisphosphonate start. By the point of blood sampling for PINP and OC measurement, a further scan after more time on a bisphosphonate showed stability of BMD.

5 Discussion

5.1 Subject Characteristics

The number of men recruited in the Clinical cohort was low (n=17 with n=17 matched), because the prevalence of osteoporosis in men is a third of that in women in the UK (6.8% vs 21.2% in 2015 (Borgström *et al.*, 2020)). The Clinical cohort "on treatment" were predominantly on alendronate (72%) because of the common practice of alendronate being first line treatment for osteoporosis. Alendronate is the most cost effective choice given the outcome of a 10 year hip fracture risk of as low as 1%. ('NICE, Bisphosphonates for treating osteoporosis | Guidance,' 2019).

Any difference in the distributions of PINP and OC cannot be attributable to baseline differences of the groups within the cohorts, due to the matching strategy used in the Clinical and the Trial cohort being the same at baseline and "on treatment". The exception to this seen in table 4.2 is the significantly different 25-hydroxyvitamin D (25OHD) between the T- and T+ and C- and C+. The reduction in PINP and OC in the "on treatment" groups is in part due to higher 25OHD level, as it is known to be inversely associated with bone turnover marker level (Nair *et al.*, 2020). The PTH in table 4.2 is significantly higher in the TRIO cohort at baseline and week 48. The higher level cannot be attributed solely to the calcium

replacement occurring after the baseline sampling in the TRIO cohort, as the week 48 PTH level was on adcal D3 twice daily (100% compared to 90% in C+). This finding is likely due to the samples in the TRIO methodology being taking first thing in the morning in the fasted state (Naylor *et al.*, 2016; CALVO *et al.*, 1991). There should be no effect overall on the bone turnover marker outcomes from these differences in PTH, with all median values being in the lower half of the normal clinical population reference range.

5.2 Bone Turnover Marker Distribution

Reflecting on the distributions of PINP and OC in table 4.4, the wide IQR of the Clinical cohort is not surprising in a group of subjects in which no effort had been made to control or exclude any factors that affect bone turnover. The IQR is impacted on by the presence of outlier measurements that can be seen in figures 4.1 and 4.2 and it has been a useful exercise to reflect on the clinical data collected on these subjects. Nearly all the outliers have a clinical explanation relating to a co-morbidity effecting bone turnover or the metabolism and excretion of the BTMs. The discrepancy between the presence of an outlier of PINP measurement with one assay and not the other, is explained by the Roche assay measuring both intact and monomer products of PINP metabolism (total PINP) and the IDS assay only measuring intact PINP. The monomer accumulates in renal impairment (Vlot *et al.*, 2018). The two assays of OC are both N-MID measurements and also effected by renal impairment, due to the detection of the large fragment as well as intact (Calvo, Eyre and Gundberg, 1996). A recent fracture often in the absence of renal impairment has a greater impact on the PINP than the OC. This is in keeping with a handful of studies looking at BTMs after various fractures, from wrist to ankle (Ivaska *et al.*, 2007; Ingle *et al.*, 1999). The bone resorption markers have also been reported to be less affected after fracture, but the N-MID osteocalcin assays require further evaluation as to the circadian rhythm and impact of haemolysis and lipaemia, as to whether it could be a more practical marker in a frail secondary prevention group (Calvo, Eyre and Gundberg, 1996). Although the magnitude of change of the BTMs after fractures has been studies, there is no publication of how this compares to the magnitude of change on anti-resorptive in order to guide on the best BTM to in monitoring treatment effect after a fracture.

This study made no attempt to ascertain a self-reported adherence level to the oral bisphosphonate in the Clinical Cohort, but with known adherence rates between 50 and 70% for weekly bisphosphonates after one year (Cramer *et al.*, 2007) any outlying measurements across both assays of PINP and OC could be contributing factors to this scenario. The published TRIO data (Naylor *et al.*, 2016) reports very good adherence (administering over 80%) in 77% of the participants but indicates that the two highest PINP and OC measurements at week 48 that can be seen in figures 4.1 and 4.2 and figure 4.6 have particularly poor adherence from the MEMS bottle cap analysis.

5.3 Timing of the Sample Measurements

A publication by Fink et al in 2000 is a useful benchmark (Fink *et al.*, 2000) when considering the timing of the median PINP and OC levels observed in this study in the Clinical cohort as compared to the Trial cohort. That publication studied 20 osteoporotic women with mean age of 65.3 years, who were commenced on 10mg/day alendronate, measuring PINP and OC at mean time point of 4.1 months. Although the study recruited from clinical practice, any secondary cause for osteoporosis was excluded and no fracture within 1 year was permitted, so the selection of subjects was more controlled. Biochemical measurements were the same N-terminal-Mid OC on an autoanalyser, but PINP was measured by Orion manual RAI. The mean PINP and OC at the mean 4.1 month mark was 18.7 µg/L and 16.6 µg/L respectively from a baseline of 39.2 µg/L and 23.7 µg/L. The 52% reduction in PINP with only 30% reduction in OC suggests that by this time point, the OC may not have reached its maximal suppression.

In the TRIO publication (Naylor *et al.*, 2016), the trend in the BTMs on alendronate showed clearly a significant reduction by 12 weeks in PINP by 56% with minor further reduction to -66% from baseline at 48 weeks. The OC reduced by 36% at 12 weeks, with further substantial reduction from 12 weeks to -53% from baseline at 48 weeks. The Clinical samples had the potential to be taken from only 12 weeks “on treatment” and were compared with the Trial cohort of samples at 48 weeks “on treatment”. The OC had the potential to be significantly lower in the T+ group than the C+ group, because of the difference in timing of the samples. Actually the OC measured by Roche Cobas e411 was significantly lower in the C+ group than the T+ group as seen in table 4.4, with both median

values being much lower than the 16.6 µg/L reported in the Fink et al study (Fink *et al.*, 2000).

The IOF position statement recommends the 3 month sample in addition to the baseline for PINP and CTX specifically, for monitoring the LSC to oral bisphosphonates (Diez-Perez *et al.*, 2017). Our study cannot support the use of an OC measurement after just 3 months "on treatment" in the context of the unknown mean duration of treatment for our Clinical cohort.

5.4 Median PINP and OC Level

The higher PINP median value in the Clinical cohort "on treatment" seen in table 4.4 cannot be explained by reduced adherence compared to the trial, as this study does not provide details on adherence, and it isn't consistent as I have just discussed with the OC results. Neither is it the case that the PINP median is higher in the C- group than in the Trial cohort baseline.

Calcium supplementation may be the cause when considering the significantly lower OC median by both assays in the C- group as compared to the T- group which is clearly seen in figure 4.2. None of the Trial cohort were on it at baseline and 50% of the Clinical cohort were. Again, this explanation may be partly rejected because of the inconsistency with PINP in C-, which is not significantly different to T-. Calcium should not cause more suppression in OC than PINP. Literature shows that supplementation has a similar effect on different markers of bone turnover, although with the literature being from earlier times, BALP was the most routine bone formation marker to be included alongside bone resorption markers, which reduces within 7 days of a calcium load (Ferrar *et al.*, 2011).

33% of the subgroup of the C- subjects with supressed PINP had prior antiresorptive exposure (not within 12 months of recruitment) compared with only 20% of those that did not have supressed PINP as seen in table 4.6. The explanation for the lower OC in the C- vs T- group could be a differential effect on the offset of bisphosphonates. The TRIO offset study of 2018 (Naylor *et al.*, 2018), showed up to 2 years post completion of 2 years of oral bisphosphonate therapy, that the PINP and OC by IDS iSYS autoanalyser from 49 women, remained significantly lower, at negative 37% from the pre-treatment baseline. However another study, presented data from a post hoc analysis of subjects in a placebo n=88 and

alendronate arm n=87, who had previously been on more than 3-4 years (Saag *et al.*, 2021) of alendronate, including in the 12 months before screening and randomisation. The offset data from this study showed a similar pattern as in the TRIO offset (Naylor *et al.*, 2018), but presented differently. The PINP in the placebo group was compared to the median PINP level prior to discontinuation of the alendronate of 25.9 µg/L and to the group continuing alendronate which actually fell on commencing the study, from a median of 28 µg/L to approximately 21 µg/L despite self-reported prior adherence of at least 80%. The much lower median PINP "on treatment" in the TRIO study, may account for the difference of how significant the rise in BTMs are at 12 months after discontinuation. The TRIO offset data concurs that most of the rise in BTMs occurs within 6 months to a year, and the PINP and OC follow a similar pattern and so this could not be the reason why the OC level of the Clinical cohort "no treatment" group is lower compared to the treatment naïve Trial cohort baseline, where the PINP is not.

5.5 Sample Representativeness of Clinical Practice

We managed to acquire renal function results from clinical tests on all "on treatment" subjects of the Clinical cohort and most not on treatment as seen in table 4.3. This contributes greatly to the understanding of outlier measurements of total PINP measured by the Roche Cobas e411 autoanalyser and N-mid OC measurements measured by both Roche Cobas e411 and IDS iSYS autoanalysers.

We recorded conditions that can affect bone metabolism such as thyroid disorders and coeliac disease, at prevalence rates equivalent to the general population. Surrogates were used for more in depth clinical information like adherence to a gluten free diet, represented by 25OHD level, and mineralisation disorders for which a subject may have been prescribed activated vitamin D, represented by PTH. PTH elevation can occur as an early response to commencing bisphosphonate therapy because of the early reduction in bone resorption (Vasikaran, 2001), but as recruitment of the Clinical cohort was from 3 months or more on oral bisphosphonates, this should not be confounding the PTH levels in C+ group seen in table 4.2.

The prevalence of 25% of recent glucocorticoid (GC) prescription in table 4.3 is unsurprising with glucocorticoid induced osteoporosis being the second leading cause (Buckley *et al.*,

2017) of fragility fractures, particularly vertebral fractures. Recent GC use has the potential to lower both bone formation marker serum levels (Eastell *et al.*, 2010).

The diabetes mellitus (DM) recorded for the subjects in table 4.3 was only if they were on treatment. This probably does capture most of those in which the DM has become a risk factor for poor bone health (Starup-Linde, 2013), and those in which the OC in particular could be supressed from the effects of the DM (Starup-Linde *et al.*, 2016).

A high percentage of the “non-treatment” group experienced a fragility fracture within the previous 12 months owing in part to recruitment bias from a population of patients attending for bone densitometry. The subjects would have been referred for a DXA as part of their fracture risk assessment and decision on receiving bone protective treatment going forwards ('NICE guideline CG146, Surveillance of osteoporosis: assessing the risk of fragility fracture,' 2019). The fracture was found in many instances to still contribute to higher serum levels of PINP and OC at recruitment and the evidence for the differential impact of this has been discussed.

HRT treatment happened not to be recorded in any subject, which is likely due to the mean age of the Clinical cohort being 70 years.

Enough clinical information was gathered to confirm a representation of the clinical population of bone health management, but the incidence of the individual factors that affect bone health was not large enough to allow subgroup analysis of the PINP and OC distributions.

5.6 Bone Turnover Marker Threshold

Even with access to limited medical records, it could still be identified why the results of the bone turnover marker were not as expected or the explanation for an outlier in the distribution. Utility of the clinical information gathered, is suggestive that Bone turnover Marker (BTM) levels should always be considered in context of patient factors, as well as analytical factors. In order to establish whether the biological target of threshold (GM of the premenopausal reference interval) has merit in use as a screening tool on every patient on oral bisphosphonate treatment, we have been able to answer two important questions in this study.

5.6.1 How often would a practitioner have to explore reasons for PINP or OC level "on treatment" above Threshold?

As the sensitivity shown for these results in table 4.5 was good for the Clinical cohort, it means it is a practical biological target to use. Only up to 20% of patients "on treatment" (100-sensitivity), tested against the threshold value of PINP, would require further evaluation. In the situation that all comers on oral bisphosphonates had BTMs tested, rather than excluding those with certain clinical comorbidities such as fracture within 6 months or renal impairment (which of course should not be applicable with limits of use of oral bisphosphonates to eGFR 30).

In this study, the premenopausal geometric mean was used as a biological target for male patients. This is the same concept as advocated by the ISCD and WHO for T-score calculation from BMD reference intervals (Shuhart, Yeap and al, 2019) and has also been used for evaluation of a mixed gender group on bisphosphonate therapy in a previous study (Eekman *et al.*, 2011) where 21% of the subjects were men compared to 20% in this study. The results in Chapter 3 show that the male reference interval across both PINP and OC by both assays had a wider variance, with a higher geometric mean and 95% UL. However, having reflected specifically on the subjects that had higher bone turnover maker level on oral bisphosphonate in the Clinical cohort, none were men.

The Eekman (Eekman *et al.*, 2011) study, was very similar in design to our study, in that it was clinical subjects who had been on antiresorptive medication for at least 3 months and had 95 subjects in the group. Although they were clinical patients, those that had experienced a fracture within 6 months were excluded. A patient report of adherence to the medication was recorded, which documented only 3 to be not adherent. The baseline characteristics means were very similar to the Clinical cohort with the exception of a slightly younger age (mean 67 vs 71 years). They were a group of mixed treatments with n=71 on oral bisphosphonates and the results of these and the IV bisphosphonates were presented together. The median PINP calculated from a premenopausal reference group (n=34) was on the manual RAI Orion of 45µg/L. The higher level of suppression on treatment found in that study (95% sensitivity) couldn't be from the higher threshold level, as it should be relative to the distribution of PINP from the Orion assay. A likely contribution to the better sensitivity is the IV bisphosphonate subjects included and very few reported not to be adherent. It is

reassuring evidence that having to reflect on 20% of clinical subjects on oral bisphosphonates with a PINP above threshold, is an outside estimate.

I had expected, because an assay specific GM was used, that the sensitivity and specificity when the same group of subjects were tested by PINP and OC on both autoanalysers, would be similar. This was true of PINP but not with OC as seen in table 4.5. The OC measured on the IDS iSYS autoanalyser has a lower GM but doesn't seem to match up with the distribution both without and on oral bisphosphonates. As this occurred for both the Clinical and Trial cohorts, it isn't the case that the IDS OC should be uniquely affected by the lack of selection in the Clinical cohort compared to a selected premenopausal group. In Chapter 2 when examining the QC data from the time of the measurements, it was apparent that the IDS OC had an inter-assay CV at the outside of what one would expect, and this has an important impact on the results not being as expected. Another hypothesis is that the serum of younger patients from which threshold was calculated interacts differently with the IDS iSYS assay than older patients. The IDS iSYS and Roche Cobas e411 autoanalyser assays for OC could be harmonised more than for PINP, because they both measure any fragments containing the N-mid portion OC. Table 4.7 demonstrates that there was no statistical difference in the AUC of the ROC analysis, with the implication that although the sensitivity of the two assays appears to be very different, the performance in identifying patients on treatment is not statistically different.

There was consistent lower sensitivity in the Clinical cohort compared to the Trial cohort in table 4.5 and the ROC curve analysis showed consistent significant lower AUC in the Clinical cohort to the Trial cohort in table 4.7. I do believe the stronger performance of PINP and OC to detect treatment effect by threshold on oral bisphosphonates in Trial subjects, is because of better adherence.

5.6.2 Can a practitioner be certain that a supressed BTM below Threshold is direct effect of the current oral bisphosphonate treatment?

The uniquely low specificity in the Clinical cohort as compared to the Trial Cohort, has been shown to be due to a combination of previous anti-resorptive use, calcium and vitamin D supplementation and glucocorticoid use, all of which were controlled for in the Trial cohort.

It is the poor specificity in the Clinical cohort which likely contributes to significant difference to the Trial Cohort performance in the AUC in ROC analysis.

The Eekman study also showed that a third of the small group of clinical patients (n=31) had a baseline PINP measured below the median of the premenopausal reference interval. (Eekman *et al.*, 2011) The low specificity found in this study is in keeping with other published data.

The poor specificity of the OC in particular could be from the high percentage of patients in the "non treatment" group on glucocorticoids (Cooper *et al.*, 2003). Previous studies looking at the effect of calcium loading on BTMs, support the impact of approximately 50% of the "non treatment" group being on a calcium supplement. Ferrer et al studied the effect of dietary calcium on BTMs (Ferrar *et al.*, 2011). The impact of this would actually be more in a treated group of patients in which 90% of subjects on treatment with oral bisphosphonates were also on calcium supplementation.

The effect of previous anti-resorptive use on the specificity in the Clinical cohort, could be considered not so relevant to the clinical population starting oral bisphosphonates, many more of which would be bisphosphonate naïve, with the exception of those being managed after a treatment break. Therefore if the use of a singular sample on treatment, compared to threshold as the biological target, is limited to treatment naïve patients, the specificity could potentially be much improved. I kept to the original study design, and chose to not perform a specificity calculation with subjects with any prior antiresorptive use excluded. I discussed this limitation in Chapter 6.

In clinical practice on a patient on oral bisphosphonates, if PINP or OC were measured to be below threshold, this could be directly due to treatment or a combination of treatment and previous treatment and supplements, all of which would still confer a positive treatment effect. It is however proposed as important to keep a record of patient adherence even in the context of suppressed bone formation markers, because of the potentially mis-leading effect of glucocorticoids on PINP and OC.

5.7 Significance of Achieving LSC

This study has evaluated the use of threshold as a biological target as compared to a more established biological target, the least significant change (LSC) of a bone turnover marker, to

demonstrate treatment effect. Having explored a supressed BTM level on treatment not necessarily being the impact of the treatment (poor specificity), the advantage of using LSC is that with a before treatment and on treatment value in the same subject, the clinician would know if the only variable contributing to the change, is the introduction of the anti-resorptive. There remain only specific scenarios that a conclusion is difficult to make from achieving the LSC. These scenarios have clear other variables, such as recovery from fracture, recent glucocorticoid use, or introduction of supplements simultaneously with oral bisphosphonate.

The sensitivities with LSC as the biological target to detect treatment effect, reported in this chapter in table 4.8, are excellent at 94% and above. The comparison with sensitivities of the threshold as the biological target was only done on the Trial cohort, in which the sensitivity at threshold was at least 86%. The exception in the comparative performance was with the OC measurements by IDS iSYS. As already discussed, the sensitivity of 76% by threshold was inexplicably low. The sensitivity by LSC with IDS OC, did not demonstrate the same anomaly with sensitivity of 97%.

In Chapter 3 I reflected that the assay specific LSC percentage for PINP and OC used in this study is lower than in the TRIO study. This has led to a slightly better percentage of the Trial cohort achieving LSC than was quoted in the TRIO publication (Naylor *et al.*, 2016) for both PINP and OC at 48 weeks of treatment (IDS iSYS PINP 82 vs 100%, OC 82 vs 91%). It is not possible to directly compare the results for reproducibility because of differences in the study cohort, with fewer subjects reported on than in the TRIO study.

Where subjects did not achieve LSC, with the exception of OC measurements by Roche Cobas e411, they also did not achieve the threshold as seen in figure 6. The fact that the mean BMD change in the subgroup not achieving LSC with Roche OC was significantly different from the subgroup achieving it (table 4.9) determines that having an OC below the threshold before commencing treatment and not decreasing further (4 out of 7 non-responders by LSC), is important to identify.

This supports the proposal that even with the good specificity of Roche OC in the Trial cohort by threshold of 85% in table 4.8, it is important for the 15% of individuals that have a low Roche OC before treatment starts, to be identified. This can only be done with a baseline sample.

5.8 Bone Mineral Density Change in Biochemical Response

The choice to use the total hip (TH) bone mineral density (BMD) data in the analysis of BMD change in the subgroups of biochemical responders and non-responders was because it represents the most stable result over time as compared to the lumbar spine and femoral neck results (De castro machado, Hannon and Eastell, 1999). In this early paper the change in different bone turnover markers and bone mineral density data on alendronate was analysed in a system of signal to noise ratio. TH BMD had a higher ratio of mean percent difference to short term CV compared to the lumbar spine at 1.73 vs 1.32. This was much lower than the OC signal to noise ratio in the same analysis of 2.18, because it was conducted over only 6 month period. This does highlight, that using BMD change over approximately 1 year, has not maximised the chance of demonstrating a significant difference between subjects responding by either biological target, or between the effect of oral bisphosphonates between the Clinical and Trial cohort.

The BMD data from the Trial cohort suggests that it is more detrimental on BMD outcome if the LSC in OC measured by Roche Cobas e411 is not met compared to if the threshold is not met. However, the numbers of non-responders by either biological target in the Trial cohort are so small that this was not seen as a pattern throughout the PINP and OC with both assays.

Even in the Trial BMD data, with most subjects included, the extremely wide 95% confidence intervals (95% CI) of the mean added uncertainty as to whether in the group classified as biochemical non-responders by either biological target, there is less gain in BMD than the responders or indeed loss of BMD. The analysis of the BMD data in the TRIO study (Naylor *et al.*, 2016) also showed limited significant difference in the mean BMD percent change between biochemical responders and non-responders. It reports a significant difference only when CTX and PINP threshold were used for grouping and the BMD change in lumbar spine at 96 weeks evaluated. There was only a significant difference in the mean BMD percent change at the TH at 48 weeks when response was defined by achieving LSC in CTX measurement (p=0.048 responder +3.2% vs non-responder +1.0%).

Achieving or not achieving LSC is the more reliable surrogate for change in BMD conferred. The subjects that don't achieve GM but do achieve LSC still had a positive effect overall on

the BMD change as seen in table 4.9. If the OC measurements by Roche Cobas e411 achieve LSC there is a significant difference in the mean BMD percent change compared to those not achieving LSC (some of which have achieved threshold). Success of achieving the LSC with oral bisphosphonates, to demonstrate positive effect on bone mineral density was examined in more detail from the TRIO study in specific publication on bone mineral density (Paggiosi *et al.*, 2014). When the IDS PINP was supressed below the LSC, stated as 36%, there was significant gain in BMD at the total hip.

The baseline DXA results for the Clinical cohort were taken from the time of commencing the oral bisphosphonate therapy (within 12 months). As oral bisphosphonate therapy generally confers a gain in BMD over the first 2 years of its initiation (Bauer *et al.*, 2006) the data was able to capture this for both the Trial and the Clinical cohort. However the annual data for the two cohorts do not represent exactly the equivalent proportional change over the first year. The trial subjects' BMD is exactly the change in the first year, with an interval period between two measurements of 48 weeks. The clinical subjects BMD change is annualised from an interval between measurements of 2-5 years. The representation of what happens over a year is likely an underestimate of the BMD change over the very first year of being on treatment. This is the hypothesis for why the BMD change although not significantly different in the C+ biochemical responders to the T+ biochemical responder, appears less.

A further contributor to why the mean BMD percent change in the C+ biochemical responders appears less is that the BTM measurement was not from the same time period as the BMD change, and therefore more likely to be dis-concordant. Had the BTM been measured within the interval of the DXA scans, potentially the subject may not have been classed as a biochemical responder. One of the outlier subjects shown in figure 4.8 is potentially caused by both of these methodology limitations with the BMD change over the years closest to the blood sampling with supressed PINP and OC showing more stability than the years immediately surrounding the oral bisphosphonate start.

With the analysis from the Trial cohort BMD data of achieving LSC being more discriminating that achieving threshold, in the context of the low specificity of the biological target of threshold, more of the C+ group theoretically may have had a low BTM prior to commencing the oral bisphosphonate on which they were recruited and potentially not have achieved

LSC. Clinical factors supressing the BTM in a less adherent patient or impacting negatively on the BMD despite adherence and suppression of the BTM has been exemplified by the second outlier subject in figure 4.8. The glucocorticoid prescription of 7.5mg prednisolone would have supressed in particular the OC (Cooper *et al.*, 2003), and had it been a long term prescription will have contributed to BMD loss (Buckley *et al.*, 2017). This demonstrates the need to take the specificity into account when evaluating the use of BTMs to monitor osteoporosis treatment.

Given all these reasons why the biochemical status may not represent the BMD period of analysis in the Clinical cohort, it is perhaps surprising that the PINP measurements by both assays did actually correlate with the BMD percent change. One other small study (Fink *et al.*, 2000) presented results of lack of correlation in behaviour of n=16 BTMs on treatment with BMD on treatment. The analysis however was of change in BTM from baseline to mean 4.1 months with BMD change at 12 months on treatment but represented by absolute BMD change, not percent.

This study also showed no correlation in baseline BTM (over different groups of age) predicting BMD status without treatment. Other studies have evaluated different correlations of BTM and BMD, such as baseline BTM predicting response to treatment (Burshell *et al.*, 2010; Bauer *et al.*, 2006), which has been shown to be significant for Alendronate.

The BMD data, even in the Trial cohort provided limited conclusions on what biological target by which BTM is a better discriminatory measure of treatment effect.

Whether the biological target of threshold (GM of the premenopausal reference interval) has merit in use as a screening tool on every patient on oral bisphosphonate treatment, to highlight impact particularly on BMD stability and fracture risk reduction, cannot be answered.

Chapter 5 Results: Response Rate to Two Parenteral Anti-resorptives by Procollagen Type I N-propeptide and Osteocalcin measurement by Two Assays

1 Background

1.1 Introduction

Evidence supports the use of bone turnover markers (BTMs), in particular the geometric mean (GM) of a healthy premenopausal reference interval (threshold), to monitor response to osteoporosis treatment. It has been seen that serum procollagen I N-propeptide (PINP) has a greater than 90% sensitivity at threshold, in showing response of women to oral bisphosphonate treatment, in a trial setting. This is equivalent to the sensitivity of using the least significant change (LSC) in PINP as the indicator of effect. Although osteocalcin (OC) has shown a lower sensitivity in Chapter 4 to treatment effect of oral bisphosphonates, it has shown different strengths in an uncontrolled osteoporosis population. Co-morbidities and other supplements impact on the interpretation of the BTM measurement, but it is important to explore a PINP or OC measurement that fails to supress below threshold on oral bisphosphonates in up to 20% of subjects in order to elicit non-adherence.

Adherence with intravenous Zoledronate therapy is never in question, therefore is there any point in measuring bone turnover markers to identify non-responders?

These results have explored the diagnostic accuracy of PINP and OC set at threshold with two different assays, to identify apparent response of clinical patients to Zoledronate therapy and Denosumab therapy, in the context of the clinical comorbidities encountered.

1.2 Aims

To evaluate the usefulness of performing two bone formation markers by two assays in clinical practice.

To monitor parenteral anti-resorptive treatments using a single-centre, observational, retrospective, cross-sectional study.

To explore the importance of using the sensitivity of the bone formation markers set at the biological target of threshold to identify the non-responders to parenteral bisphosphonate, through examining the rate of bone mineral density gain in the unselected patients on zoledronate.

1.2.1 Main research question:

- What is the sensitivity and specificity of procollagen I N-propeptide and osteocalcin at a target marker level, to identify the responsiveness of clinical patients on parenteral anti-resorptive treatments for osteoporosis? (The target is the threshold level defined as the geometric mean of a healthy premenopausal reference interval)

1.2.2 Secondary research questions:

- Do two different assays for PINP and OC have an equivalent sensitivity and specificity of the biological target in identifying clinical patients on parenteral anti-resorptive treatments for osteoporosis?
- Is there a significant difference in the rate of bone mineral density gain between the subgroups of clinical patients, identified by each bone formation marker as “responders” and “non-responders”, on a parenteral anti-resorptive?

2 Patients

Subjects were recruited between 2016 and 2018 within the Academic Unit of Bone Metabolism, University of Sheffield for the South Yorkshire and North Derbyshire Musculoskeletal Biorepository (SYNDMB) with Oxford Rec reference 15/SC/0132 and Sheffield Teaching Hospitals Study number STH15691.

Ethical application to the SYNDMB was approved (SYNDMB045) to create clinical patient cohorts from these recruited populations:

- Intravenous bisphosphonates. 100 clinical patients treated with intravenous zoledronate (5 mg annually) or ibandronate (3mg every 3 months) for at least one year.

- Subcutaneous denosumab. 17 clinical patients treated with subcutaneous denosumab (60 mg 6 monthly) for at least 6 months.
- Matched “no treatment” group for each of the treatment cohorts. 300 clinical patients attending for bone densitometry on no osteoporosis treatment (approximately 50% will be on calcium or cholecalciferol). They were matched by gender, age (round down +/- 3 years), body mass index (+/- 2.5) and T-score of the total hip (+/- 0.3) to the above groups.

The Zoledronate cohort was designed as a mixed gender group on intravenous bisphosphonate treatment, and a clinical sample not on treatment that were matched to them. The Denosumab cohort was not initially restricted to a female group on subcutaneous Denosumab, and a clinical sample not on treatment that were matched to them.

Table 5.1 shows the proposed subject characteristics.

	ZOLEDRONATE COHORT	**DENOSUMAB COHORT**
Inclusion	Patient on IV bisphosphonate ≥ 12m (At least second infusion) OR On no treatment > 12m Capacity to consent to a blood test for research	Patient on SC denosumab ≥ 6m (At least second injection) OR On no treatment > 12m Capacity to consent to a blood test for research
Exclusion	None	

Table 5.1. Proposed subject characteristics by Inclusion and Exclusion Criteria for the Zoledronate and Denosumab cohort

3 Methods

Only the methodology that applies to clinical cohorts detailed in Chapter 2 “Materials and Methods” and summarised in Chapter 4, was used for this chapter. The only deviation to note was the bone densitometry data was obtained from sequential clinical DXA scans with a reduced interval of up to 3 years.

3.1 Statistical Analysis

The statistical analysis in this chapter of the performance of PINP and OC in the Zoledronate and Denosumab cohorts, follows the same system as for the analysis of the Clinical cohort in Chapter 4. The Zoledronate and Denosumab cohort were not directly compared in their subject characteristics or the distribution of the PINP and OC.

Therefore, with the statistical testing being done to compare two within cohort groups only, the critical p-value remains at 0.05 throughout.

Reflections on comparison between the performance of the PINP and OC in the two cohorts was restricted to descriptive comparison of any overlapping of the 95% confidence intervals (CI) of the median BTMs, and the 95% CI of the area under the curve in the ROC analysis.

Analysis of the BMD percent change was limited to data from the Zoledronate cohort only, with the null hypothesis that there was no statistical difference in annualised BMD percent change in subgroup “on treatment” that achieved the biological target and a subgroup “on treatment” that did not achieve the biological target, as described in Chapter 2. The correlation coefficient by Spearman Rank was also calculated for the BTM and BMD percent change for each individual on zoledronate, with a null hypothesis that there was no correlation between PINP or OC by both IDS and Roche assays and the annualised percent change in BMD.

4 Results

4.1 Subject Characteristics

4.1.1 Baseline demographics

The subjects in the cohorts were predominantly female with 23 men in the IV bisphosphonate cohort, with 22 matched men. It was not the intention to only recruit women in the Denosumab cohort, but all 15 subjects in the “on-treatment” group were women, and therefore so were the matched “no treatment” group. The Zoledronate cohort evolved from an intended mixed Intravenous cohort, as the “on-treatment” group consisted of only zoledronate administration.

The Metabolic Bone Unit attendance records confirmed eligibility of the "on-treatment" group subjects of both cohorts, with confirmation of attendance for at least their second injection. In a small number of patients, the recruitment was within a one month overdue window of the treatment, due to the timing of the appointment in a clinical setting.

Table 5.2 details the characteristics of both the Zoledronate and Denosumab cohort, broken down into the "no treatment" group which I will refer to as Z- and D-, and the group" on treatment", which I will refer to as Z+ and D+. It demonstrates the success of the matching within each cohort.

	ZOLEDRONATE COHORT		**DENOSUMAB COHORT**	
	No treatment n= 92	IV Zoledronate n= 99	No treatment n= 14	SC Denosumab n= 15
Sex	146 female; 45 male		29 female	
Age at recruitment (yrs)	70.5 (10.4)	69.7 (11.8)	77.0 (9.1)	76.9 (9.4)
BMI (kg/m^2)	25.6 (22.9 to 30.0) #	25.7 (22.6 to 29.4) #	25.2 (5.2)	26.0 (5.7)
Total hip T-score	-1.76 (0.98)	-1.74 (1.06)	-2.24 (0.91)	-2.34 (1.15)
25OHD (nmol/L)	76.3 (30.3) *	93.6 (28.3)	81.3 (39.2)	90.1 (28.2)
PTH (ng/L)	28.1 (19.1 to 43.6) #	29.2 (20.8 to 41.1) #	39.0 (32.6 to 47.7) #	46.7 (26.3 to 78.4) #

Table 5.2. Treatment group characteristics by cohort.

Mean and SD (x) for each characteristic that could impact bone metabolism with Median and IQR (y to z) # given when distribution doesn't satisfy normal distribution. Significant difference p <0.05 by t test or rank comparison #, denoted within cohort treatment groups * only.

The Zoledronate cohort demonstrated a normal distribution for all characteristics in both groups with the exception of iPTH and BMI which have a positive skew. The Denosumab cohort, had an older mean age, but the distribution of D- and D+ age still satisfied testing for normality. All other distributions, again with the exception of the iPTH, have a normal distribution. The characteristic distributions that did not satisfy normality have been represented in table 5.2 by median and IQR and marked with #.

All of the women in the Denosumab cohort, with the older mean age which was well matched between D- and D+, were postmenopausal. There was also no statistical difference in the mean age of Z- and Z+. However, in the Z+ group, there were a potential 3 subjects, between the age of 50 and 55 years, in which incomplete patient-recorded information regarding menstrual status conferred uncertainty as to whether they could have been premenopausal. The same was true of one subject in Z-.

There was no difference in the median BMI of Z- and Z+, or the median BMI of D- and D+.

The total hip T-score was the last of the matched characteristics and in this there was no difference within each cohort. The Denosumab cohort mean total hip T-score appears lower than the Zoledronate cohort.

4.1.2 Subject biochemical descriptors

There was a significant difference within the Zoledronate cohort between Z- and Z+ 25-hydroxyvitamin D (25OHD) measurements ($p<0.0001$), shown in table 5.2. The 25OHD mean in Z+ subjects was higher, however, both mean levels were 25OHD sufficient. There was no difference in the 25OHD mean in D- to D+, and again both mean levels were sufficient. As first line treatment for bone health is cholecalciferol supplementation and table 5.3 shows it was recorded in at least 60% of subjects, this underlies the finding of sufficient 25OHD levels in all four groups.

The median of the plasma iPTH in all four groups was within the normal clinical population reference range and was not significantly different between Z- and Z+ or D- and D+. However, the mean PTH in D+ appeared higher than that of the D- group, and percentage of calcium supplementation and pattern of PTH-dependent hypercalcaemia described in table 5.3 could be useful in the interpretation of this.

4.1.3 Clinical co-morbidities affecting bone metabolism

Available information on factors that can affect bone turnover was extracted from ArQ for every subject and recorded as descriptive frequencies in table 5.3.

		ZOLEDRONATE COHORT		DENOSUMAB COHORT	
		No Treatment n from 92 (%)	IV Zoledronate n from 99 (%)	No Treatment n from 14 (%)	SC Denosumab n from 15 (%)
Calcium Supplement		50 (54)	78 (79)	5 (36)	11 (73)
Cholecalciferol Supplement		55 (60)	74 (75)	9 (64)	12 (80)
Activated Vitamin D		5 (5)	2 (2)	0	0
Fracture within recent 12m		15 (16)	5 (5)	1 (7)	0
CKD	1	25 (29)	21 (21)	2 (14)	2 (13)
	2	42 (48)	58 (59)	9 (65)	7 (47)
	3a	14 (16)	15 (15)	2 (14)	2 (13)
	3b	4 (5)	5 (5)	1 (7)	2 (13)
	4	1 (1)	0	0	2 (13)
	5	1 (1)	0	0	0
Coeliac disease		1 (1)	3 (3)	0	0
Hypo-, Hyper - thyroid (both)		11,2 (14)	12,2 (14)	2,1 (21)	1 (7)
Treated Diabetes Mellitus		3 (3)	8 (8)	2 (14)	1 (7)
Biochemistry of PHPTH		1 (1)	2 (2)	0	1 (7)
Glucocorticoid within 1 week		16 (17)	12 (12)	3 (21)	2 (13)

Table 5.3. Medical characteristics and relevant supplementation by treatment group of both parenteral cohorts, Zoledronate and Denosumab.

Number of the group in which each medical co-morbidity or supplement was recorded at source. (Integer percentage calculated from the total group number).

Renal function was available for the majority of the subjects in the Zoledronate cohort (n=99 Z+ group, n=87 Z- group) and all in the Denosumab cohort. No subject in the Z+ group had an eGFR less than 30mL/min/1.73m^2. There were 2 subjects in the Z- group with CKD stage 4 and 5, but otherwise the distribution across the CKD groups appeared similar in both the “no treatment” and “on treatment” cohort. The distribution of renal function across D- and D+ also appeared comparative, but with 2 subjects in D+ having an eGFR between 10 and 30mL/min/1.73m^2.

A prescription of activated vitamin D was minimal and likely related to the Chronic Kidney Disease (CKD) reported.

The incidence of endocrine conditions such as “on treatment” diabetes mellitus, hyperthyroidism and primary hyperparathyroidism were low, as was the noted presence of coeliac disease. However, within the small Denosumab cohort, just one subject with one of these conditions, represented a higher percentage, which created a greater spotlight on the difference between D- and D+, such as with thyroid disorders (21% vs 7 % was 3 subjects vs 1 subject).

Only 17% and 12% of the Z- and Z+ groups respectively, had had a recent (within 1 week of recruitment) glucocorticoid prescription (excepting inhaled corticosteroids), and only 1 more subject in the D- group, compared to D+ group (3 vs 2).

The factors with the biggest percentage difference between Z- and Z+, were the fracture history and supplement prescription. 16% of the Z- group had suffered a recent fracture compared to only 5% of the Z+ group. Although the calcium supplementation in the D- group was lower, there was a marked difference in calcium and cholecalciferol prescription within each cohort between the” no treatment” and “on treatment” groups with between 70 and 80% supplementation recorded in the Z+ and D+ groups.

These confounding effects on the BTM distributions were useful to explore when establishing the performance of the BTMs set at the biological target of threshold to show apparent response in the group, but also when evaluating an outlier measurement of PINP or OC in a single clinical patient.

4.2 Bone Turnover Marker Outcomes

4.2.1 Distribution of PINP and OC measurements

Biochemical measurements for PINP and OC by two assays were successfully acquired on the IDS iSYS and Roche Cobas e411 autoanalysers from almost all of the defrosted serum samples. Amongst the samples of the Z- group, only 1 out of 92 possible included samples, produced a null outcome on the Roche Cobas e411 auto analyser for PINP and OC. There were results for all 99 of the Z+ group. Amongst the D- samples, successful measurements were made in all 14 included samples. 1 sample out of 15 of the D+ yielded no IDS OC

measurement, yet it produced the other 3 measurements. The n numbers in table 5.4 are the maximum possible of measurements.

The distribution of the PINP and OC measurements are described by the median and interquartile range (IQR). The groups did not satisfy a normal distribution. The measurements in each group demonstrated a positive skew.

The lowest median PINP and OC by both assays were in the D+ group (table 5.4). The D- group also had low median BTMs and this will be looked at further in relation to specificity.

The Denosumab cohort median values had broad 95% confidence intervals (CI) seen in the brackets next to the median in the table. With all assays, the median 95% CI of the D- group encompass the estimate of the median of the Z- group.

		ZOLEDRONATE COHORT		**DENOSUMAB COHORT**	
		NO TREATMENT n= 92	IV ZOLEDRONATE 5mg ANNUAL n= 99	NO TREATMENT n= 14	SC DENOSUMAB 60mg 6 monthly n= 15
PINP µg/L	IDS iSYS	52.7 * (46.0;57.3) 32.8 to 84.0	22.5 (20.9;27.2) 18.0 to 30.7	36.3 * (33.0;54.7) 33.0 to 54.7	15.2 (10.9;21.5) 10.9 to 21.5
	Roche Cobas e411	52.4 * (41.4;62.4) 33.5 to 86.0	21.1 (20.2;24.3) 16.8 to 29.9	40.6 * (33.6;55.4) 34.0 to 55.3	15.2 (12.2;21.2) 12.2 to 21.2
OC µg/L	IDS iSYS	16.0 * (12.8;22.0) 8.9 to 32.1	6.1 (5.2;6.6) 4.4 to 8.0	14.9 * (10.2;22.0) 10.4 to 22.0	4.7 (3.8;6.0) 3.8 to 6.0
	Roche Cobas e411	17.7 * (13.8;22.2) 10.2 to 27.8	7.0 (6.6;7.7) 5.4 to 9.2	15.8 * (10.7;21.5) 10.9 to 21.4	5.9 (5.1;6.9) 5.1 to 7.0

Table 5.4. Median (95% CI of the Median) and IQR for the two bone turnover markers by two assays for the no treatment and on treatment group with IV Zoledronate or SC Denosumab for each cohort.

Significant difference p <0.05 by rank comparison, denoted within cohort treatment groups * only.

The variance in the group distributions is to be expected given the factors detailed in table 5.4 that affect bone turnover. The IQR of the Z- group PINP was approximately 50 µg/L and

OC between 17 µg/L and 23 µg/L, with PINP IQR in the D- group smaller at 20 µg/L and approximately 10 µg/L for the OC IQR.

Z+ and D+ demonstrated a much smaller IQR, although this was more obvious relative to a lower median value in the Z+ group, with more variability in the Denosumab cohort.

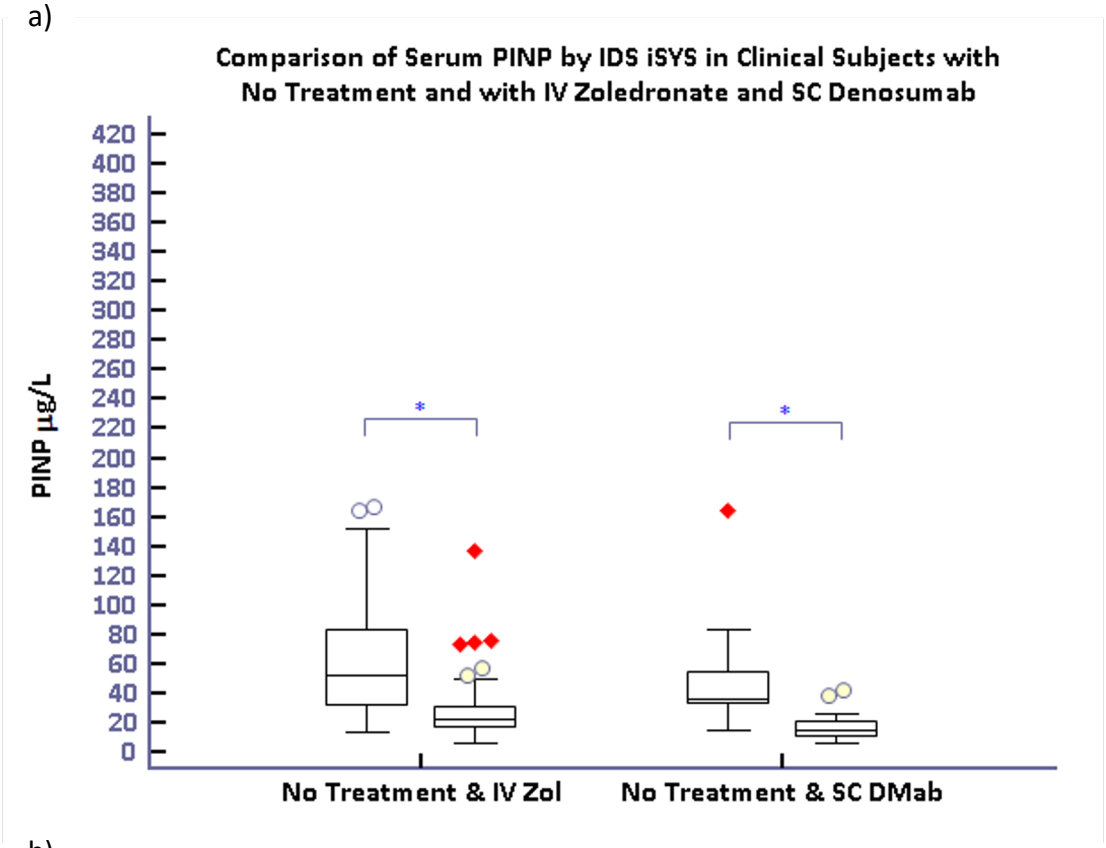
Most of the outliers of the PINP and OC distributions in each cohort, as seen in figure 5.1 and 5.2 by the diamond plotted points were not wildly off the 95% centile and therefore singularly not effecting the IQR.

The outlier in the D+ in all BTMs except intact PINP was on cholecalciferol but had CKD stage 4 and with an elevated PTH (177ng/L) and calcium corrected (2.77mmol/L), either primary or tertiary hyperparathyroidism.

The theme amongst the outliers in the Z- group was fracture history within 12 months or renal impairment. When in combination there was more impact on the total PINP than intact PINP and both OC assays (404 μ g/L compared to 89 μ g/L and 58 μ g/L, 43 μ g/L). Fracture history alone gave rise to high PINP but not OC, and significant renal impairment alone also had an elevated PTH (189ng/L) with very high OC (approx. 200 μ g/L) and mildly elevated Total PINP, again, sparing the intact PINP.

There were also significant outliers in the Z+ group. In these subjects, any information obtained on a delay in the Zoledronate dose was also important, especially the second dose. Most of these outliers were recruited only on their 2nd dose 5mg zoledronate and delays were between 2 and 4 months, with higher PINP than OC levels. One of these subjects had an elevated corrected calcium of 2.70mmol/L with PTH mid-range at 35ng/L. Three outliers had apparently no impacting clinical factors, but there was no record of when the zoledronate commenced for one. One outlier in OC and not in PINP, was a 52 year old patient with menopausal status not recorded. The subject had been on 5 years of zoledronate.

a)



b)

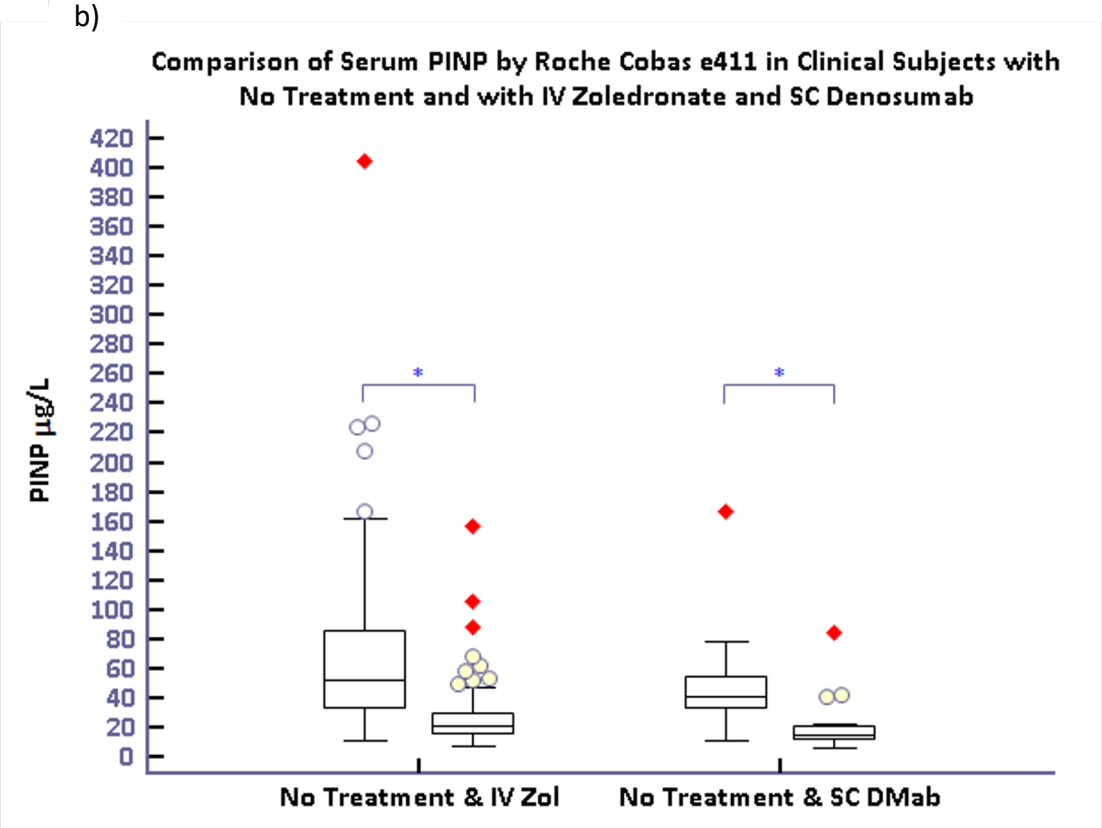
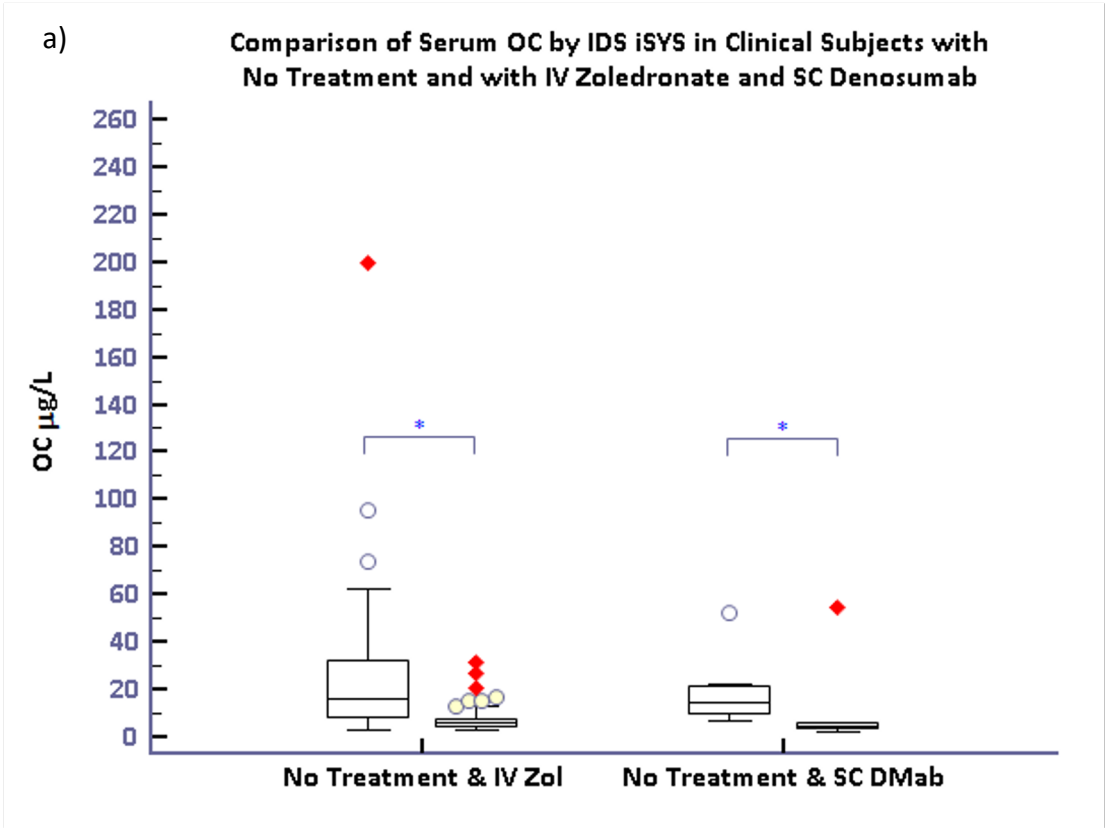


Figure 5.1. Comparison of PINP boxplots measured by a) IDS iSYS b) Roche Cobas e411 in the No treatment and On treatment subjects with 5mg annual IV Zoledronate and 60mg SC Denosumab Cohorts.

Rank comparison of the within cohort group distributions with Sig difference p <0.05 denoted by *.



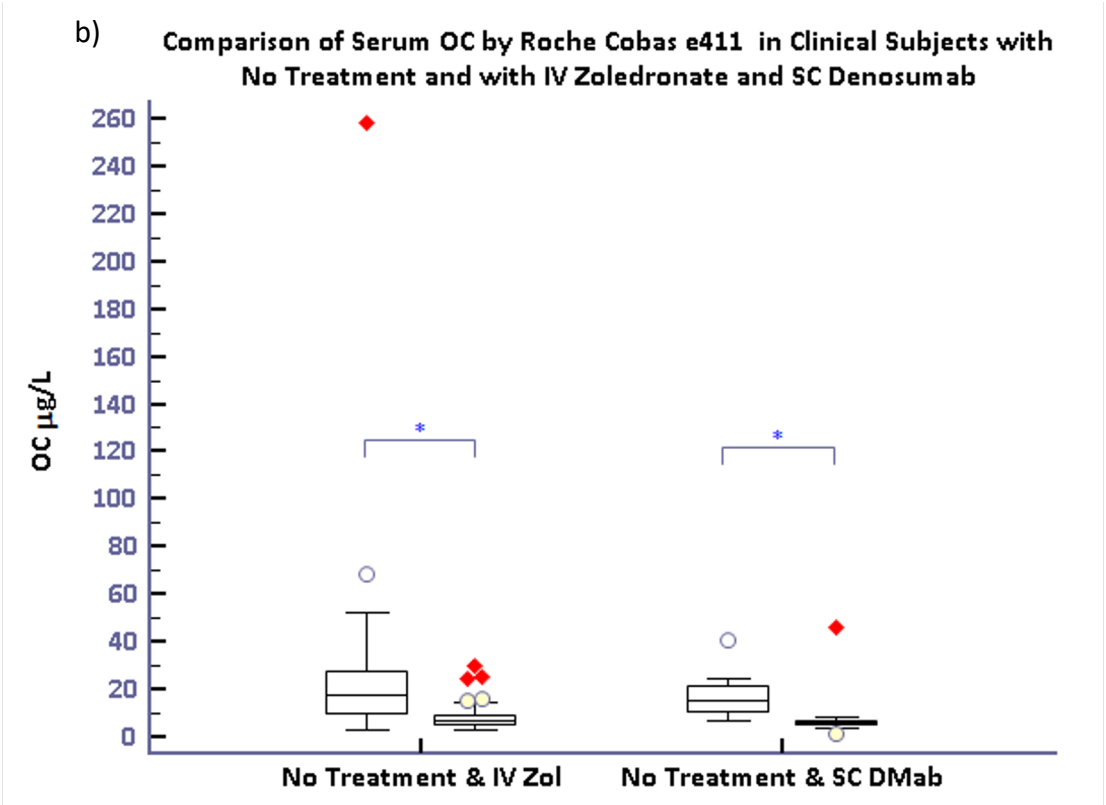


Figure 5.2. Comparison of OC boxplots measured by a) IDS iSYS b) Roche Cobas e411 in the No treatment and On treatment subjects with 5mg annual IV Zoledronate and 60mg SC Denosumab.

Rank comparison of the within cohort group distributions with Sig difference p <0.05 denoted by *.

4.2.2 Comparison of within cohort PINP and OC distributions

The boxplots of the BTM measurements of Z- and Z+ and D- and D+ groups are shown in figures 5.1 (PINP) and 5.2 (OC) with a graph for each assay. The statistical difference within each cohort is demonstrated by the connector lines, with symbols described in the figure legend.

PINP and OC by both IDS iSYS and Roche Cobas e411 assays, were significantly different in the "no treatment" groups to the "on-treatment" groups within each cohort. Each statistical difference conferred a p value of <0.0001, with the exception of the difference in PINP by both assays between D- and D+ (IDS PINP p<0.0004, Roche PINP<0.00056).

Statistical testing between cohorts was not appropriate with different subjects and different treatments. However the 95% CI of the median in the D+ group, does not include the estimate of the median of the Z+ group and vice versa with the exception of OC by the Roche Cobas e411 assay. Therefore on treatment, the median values of PINP by IDS iSYS and Roche Cobas e411 and OC by the Roche Cobas e411 assay, with denosumab, are significantly lower than on zoledronate.

4.2.3 Sensitivity and Specificity by threshold

		ZOLEDRONATE		**DENOSUMAB**	
		SPECIFICITY n= 92	SENSITIVITY n= 99	SPECIFICITY n= 14	SENSITIVITY n= 15
% PINP VALUES	IDS iSYS	67	82	36	87
	Roche Cobas e411	58	89	29	87
% OC VALUES	IDS iSYS	58	94	57	94
	Roche Cobas e411	50	97	36	93

Table 5.5. Percentage of the two bone turnover markers by two assays relative to the Geometric Mean (GM) of the assay specific reference interval.

Representing Specificity (% of the no treatment group that ≥ GM) and Sensitivity (% of on treatment group with parenteral antiresorptive that ≤ GM), for each cohort. The GM for IDS PINP is 37.2 µg/L, Roche PINP is 41.7 µg/L, IDS OC is 13.5 µg/L, Roche OC 17.8 µg/L.

The specificity and sensitivity of the PINP and OC with the two assays to detect treatment effect of zoledronate and denosumab at threshold was calculated in table 5.5. The GM of the premenopausal reference interval, threshold value, was assay specific. On the graphs in figures 5.3 and 5.4 the percentage of PINP and OC measurements in the “No Treatment” groups above the marked GM represents the specificity and the percentage of measurement in the “IV Zol” and “SC DMab” groups supressed below the same threshold line is the sensitivity, with one assay per graph.

The performance of the two assays of the two BTMs in terms of the sensitivity and specificity as seen in table 5.5 was consistent within both cohorts.

The sensitivity of the assays at threshold to detect treatment was better than the specificity and a very similar range of sensitivities was seen in the Zoledronate and Denosumab cohort, for PINP (86-89%) and OC (93-97%).

Specificity to detect treatment was particularly poor by both BTMs. The specificity of the D-group in particular was extremely poor at less than 50% in all but the IDS OC assay.

Measurements in figures 5.3 and 5.4 below the GM in the “No Treatment” groups were considered ‘false positives’. Due to the low specificity of the BTM to detect treatment, there were between 30-70% ‘false positives’. Further descriptive analysis has been made of clinical factors impacting to reduce the bone turnover in both “non treatment” groups. Table 5.6 uses the PINP measurements by IDS iSYS as an example. This assay has the best specificity at threshold of 67% and a representative AUC of all the assays for discrimination of whether a subject is on treatment, as shown in table 5.7.

The frequency and percentage of occurrence of clinical factors in the subgroup of Z- and D- with a PINP at or below the GM of the reference interval, were compared to the subgroup of Z- and D- with PINP above threshold.

The main factors that were more prevalent by percentage of subjects of Z-, in the subgroup with PINP by IDS iSYS below the threshold, were; higher percentage of calcium and cholecalciferol supplementation, and exposure to previous anti-resorptive, particularly within 5 years. Also, glucocorticoid prescription was similar in subject numbers between the two subgroups, but this represented a much greater percentage in the smaller subgroup with “no treatment” but supressed bone turnover. With the D- group being much smaller

overall (n=14), and with a lower percentage occurrence of calcium and glucocorticoid prescription throughout both subgroups, it was only the prior use of anti-resorptive therapy that differed between those with suppressed IDS PINP and those with IDS PINP measurements above the threshold. Only 3 subjects in whom exposure in the last 5 years before recruitment to an anti-resorptive could have contributed to the suppression (not within 12 months) whilst the other 6 subjects in this groups had no apparent accountable clinical features.

		ZOLEDRONATE No Treatment Group		**DENOSUMAB No Treatment Group**	
		≤ IDS PINP GM n from 30 (%)	> IDS PINP GM n from 62 (%)	≤ IDS PINP GM n from 9 (%)	> IDS PINP GM n from 5 (%)
Calcium Supplement		19 (63)	31 (50)	3 (33)	2 (40)
Cholecalciferol Supplement		21 (70)	34 (55)	6 (67)	3 (60)
Previous anti-resorptive		13 (43)	10 (16)	4 (44)	0
Anti-resorptive ceased	< 5 yrs	7 out of 11	5 out of 7	3 out of 4	0
	Unknown	2	3	-	-
CKD	1	5 (17)	20 (32)	2 (22)	0
	2	15 (50)	27 (43)	6 (67)	3 (60)
	3a	5 (17)	9 (14)	0	2 (40)
	3b	3 (10)	1 (2)	1 (11)	0
	4	0	1 (2)	0	0
	5	0	1 (2)	0	0
	Unknown	2 (6)	3 (5)	-	-
Glucocorticoid within 1 week		9 (30)	7 (11)	1 (11)	2 (40)

Table 5.6. Breakdown of medical characteristics known to suppress bone turnover in the No Treatment group of both Zoledronate and Denosumab cohort.

Descriptive comparison in the number and percentage of subjects recorded at source to be exposed to these factors, in the subgroup with PINP by IDS iSYS below the GM and subgroup with PINP by IDS iSYS above the GM. N (Integer percentage calculated from the subgroup number).

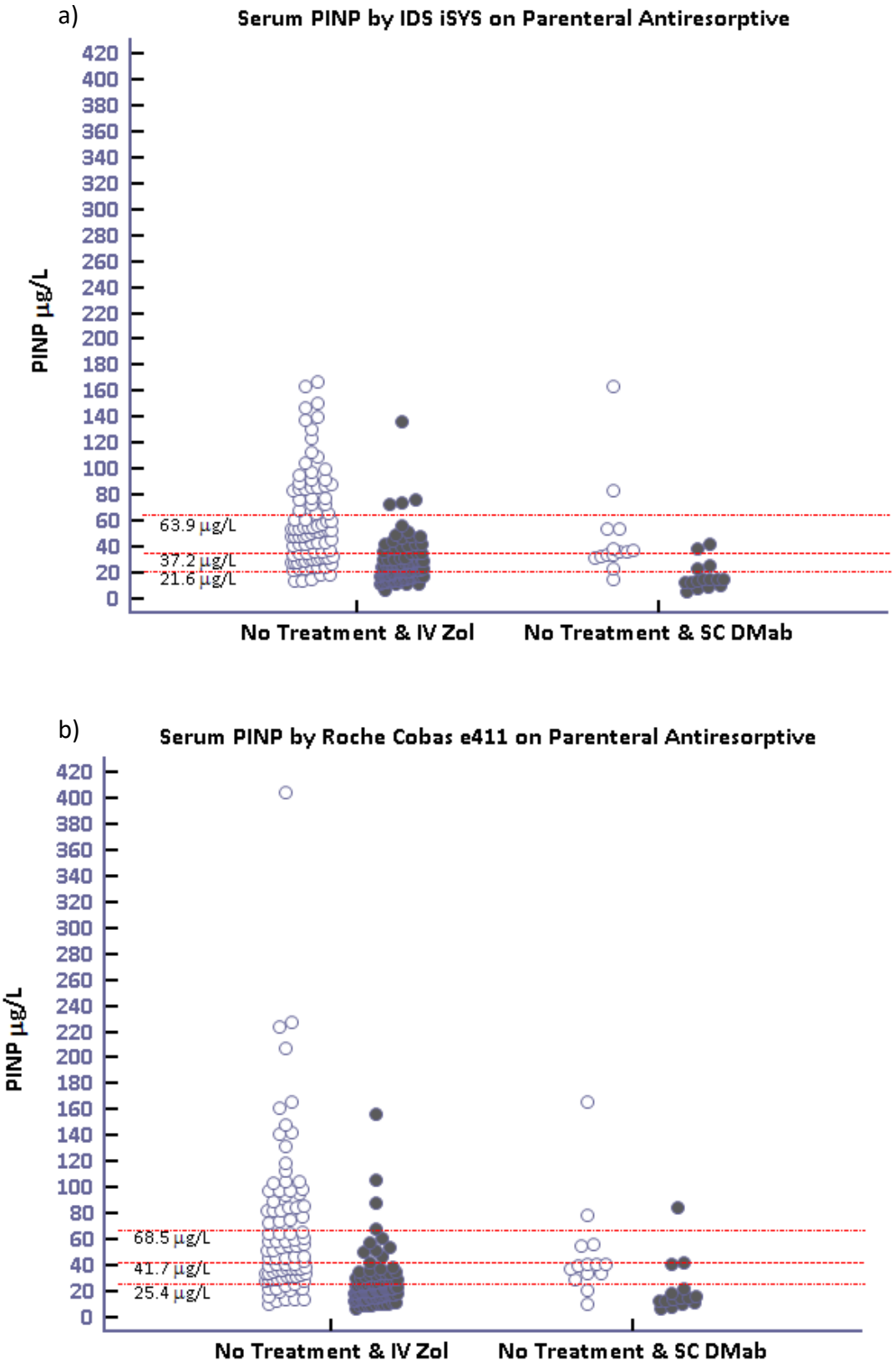
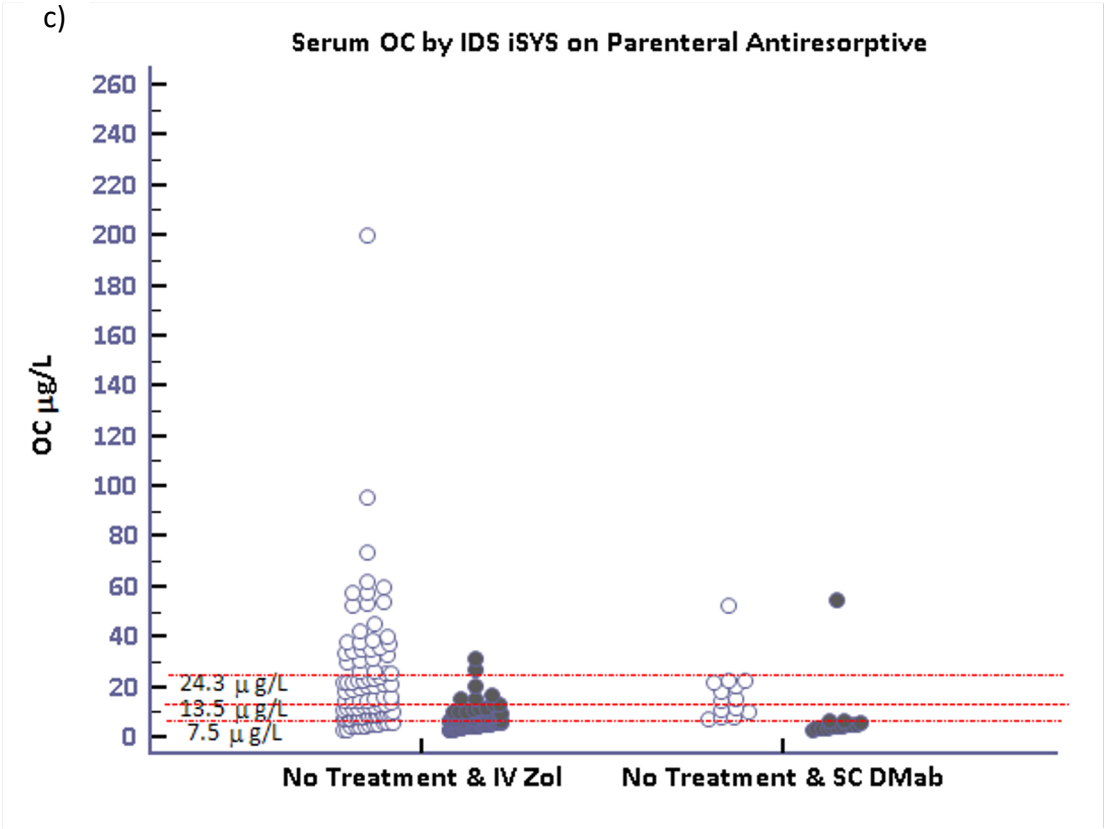


Figure 5.3. Comparison of a) PINP by IDS iSYS; b) PINP by Roche Cobas in the No treatment and On treatment with 5mg annual IV Zoledronate and 60mg SC Denosumab. The assay specific geometric mean (37.2 µg/L, 41.7 µg/L) and upper (63.9 µg/L, 68.5 µg/L) and lower (21.6 µg/L, 25.4 µg/L) reference limit of PINP.



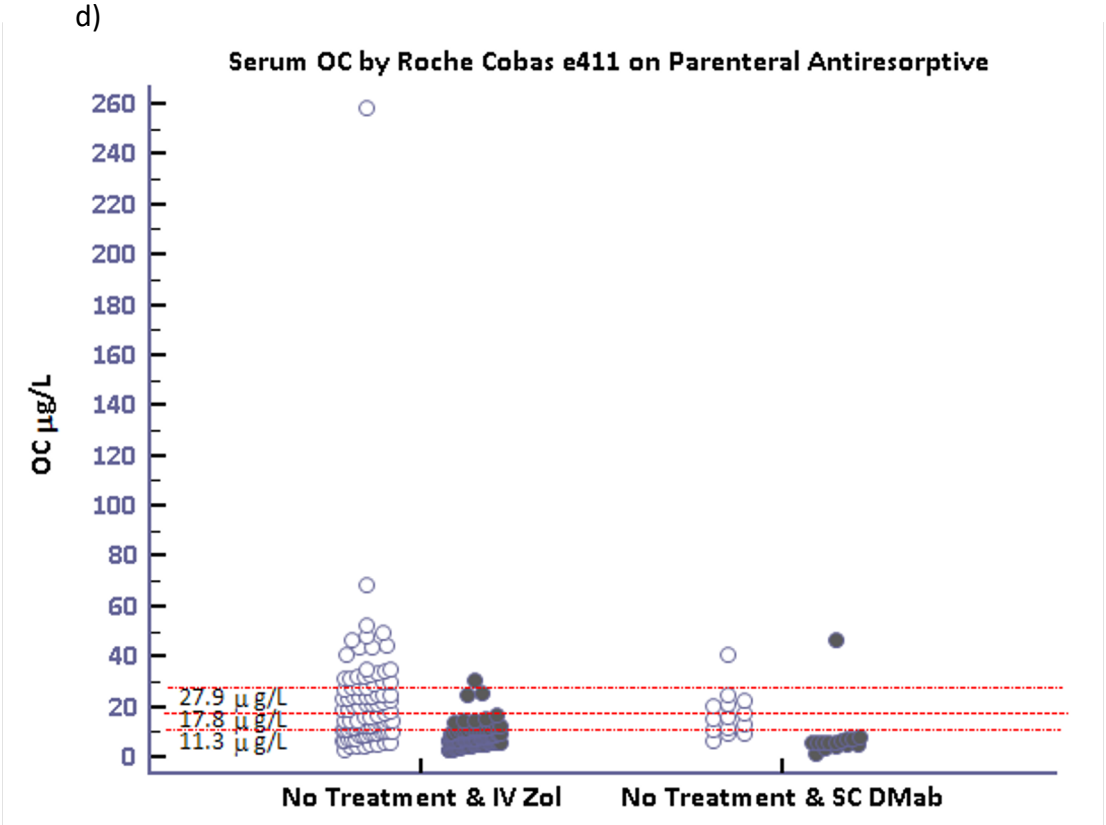


Figure 5.4. Comparison of c) OC by IDS iSYS; d) OC by Roche Cobas e411 in the No treatment and On treatment with 5mg annual IV Zoledronate and 60mg SC Denosumab.

The assay specific geometric mean (13.5 µg/L , 17.8 µg/L) and upper (24.3 µg/L, 27.9 µg/L) and lower (7.5 µg/L, 11.2 µg/L) reference limit of OC.

4.2.4 ROC curve analysis

Where the area under the curve (AUC) in a ROC curve analysis indicates the performance of a tool to differentiate between the subjects “on treatment” and “not on treatment”, figure 5.5 shows the BTMs performed comparatively in the Zoledronate and Denosumab cohorts with values between 0.84 in the former and 0.80 to 0.93 in the latter.

		ZOLEDRONATE n= 191		**DENOSUMAB n= 29**	
		AUC	Youden Criterion µg/L	AUC	Youden Criterion µg/L
PINP	IDS iSYS	0.845 0.786 to 0.893	29.59	0.882 0.704 to 0.972	23.65
	Roche Cobas e411	0.838 0.778 to 0.887	29.51	0.803 0.609 to 0.928	18.57
OC	IDS iSYS	0.848 0.789 to 0.895	7.77	0.929 0.761 to 0.992	6.49
	Roche Cobas e411	0.845 0.786 to 0.894	9.38	0.918 0.751 to 0.988	8.35

Table 5.7. ROC curve analysis expressed for the Zoledronate and Denosumab cohort as AUC with 95% Confidence intervals for the two bone turnover markers by two assays.

Youden criteria can be viewed relative to the GM of the reference interval; IDS PINP is 37.2 µg/L, Roche PINP is 41.7 µg/L, IDS OC is 13.5 µg/L and Roche OC 17.8 µg/L.

Because of the much smaller Denosumab cohort, the 95% confidence intervals of the area under the curve (95% CI of AUC) given in table 5.7 were much wider than in the Zoledronate cohort measurements by each assay of each BTM. The 95% CI of AUC for each assay all overlapped within the Denosumab, as well as within the Zoledronate cohort and the performance of the BTMs showed no discernible difference for determining treatment status with either.

Following on from the low median PINP and OC in all the distributions “on treatment”, lower in the Denosumab cohort; the Youden Criterions in both cohorts were also low, and close to the lower limit of the normal reference interval.

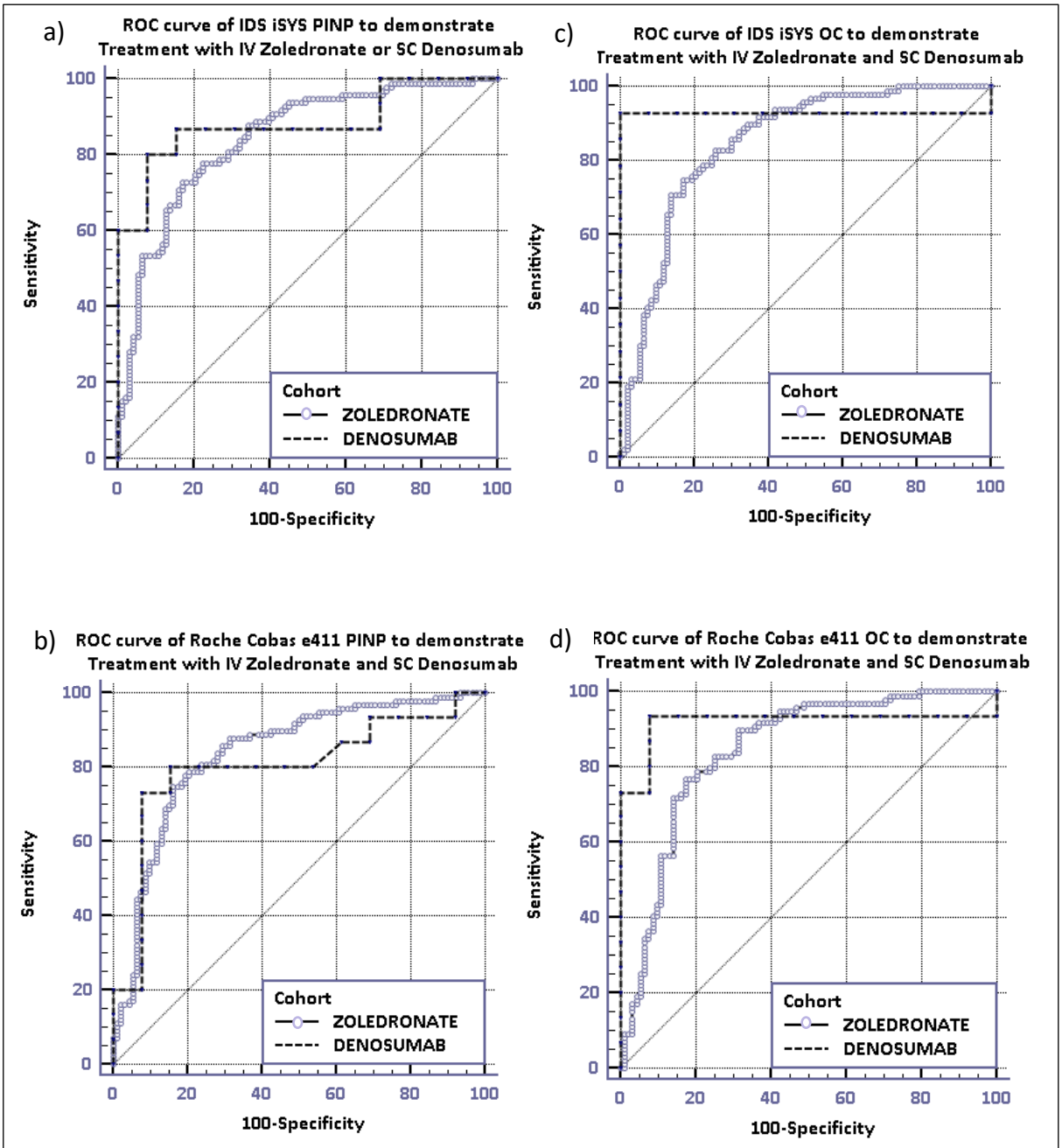


Figure 5.5. ROC Curve of measurements of a) PINP by IDS iSYS, b) PINP by Roche Cobas e411, c) OC by IDS iSYS; d) OC by Roche Cobas e411 in both Zoledronate and Denosumab cohort.

With the AUC demonstrating the performance of the tool in detecting treatment with each parenteral treatment.

4.3 Bone Mineral Density Outcomes

The mean BMD percent change "on treatment" was analysed to contribute to our understanding of whether the "non-responders" are important to identify. We have

established that between 3% and 18% of patients on parenteral treatment do not show treatment effect by supressing PINP or OC below the assay specific threshold.

Only 29 subjects in the Zoledronate cohort had available, sequential between 1 and 3 years, comparable DXA scans on treatment. This was in part due to a system in the Metabolic Bone Unit of "Direct Access" to receiving Zoledronate in the event of contra-indication to or intolerance if oral bisphosphonate. If the subject had satisfied criteria, such as a fragility fracture in over 75 year old, they may not have undergone a baseline DXA at the time of commencing Zoledronate. 33 out of 99 subjects were recruited at their 2nd or 3rd Zoledronate infusion and had therefore not yet undergone the follow up DXA scan. The results from the 29 subjects have been shown in table 5.8.

		Mean Annual BMD % change grouped by GM (threshold) ZOLEDRONATE
IDS iSYS PINP	Non-responder	0.79 (-1.19 to 2.76) n= 6
	Responder	0.33 (-0.8 to 1.47) n= 23
Roche Cobas e411 PINP	Non-responder	0.86 (-1.71 to 3.44) n= 4
	Responder	0.36 (-0.71 to 1.43) n= 25
IDS iSYS OC	Non-responder	-2.15 (-10.23 to 5.93) n= 3
	Responder	0.73 (-0.19 to 1.64) n= 26
Roche Cobas e411 OC	Non-responder	-0.28 (-2.39 to 1.83) n= 2
	Responder	0.48 (-0.53 to 1.49) n= 27

Table 5.8. Zoledronate cohort Annualised Mean BMD Percent Change (and 95% CI) in Responder and Non-responder subgroups with threshold (GM) as the target for each assay of PINP and OC.

Only 7 out of the 15 subjects in the Denosumab cohort "on treatment" had possible DXA data to analyse, as most were recruited on the second dose before being referred back to

primary care on a “shared care” pathway. They had not undergone follow up scans at the time of recruitment for data analysis. Further to this, some of the baseline scans fell out of 12 months proximity of commencing denosumab. As the very small amount of data would have been minimised by subgroup analysis, it has not been included.

As shown in table 5.8 the 95% CI of the mean BMD percent change in both the non-responder and responder subgroups of Z+ all crossed zero. The seemingly lowest mean BMD percent change (most negative) conferred by non-response when measured by OC on IDS iSYS assay has the widest CI. This can be seen in figure 5.6 which show the mean and standard error of the mean (SEM).

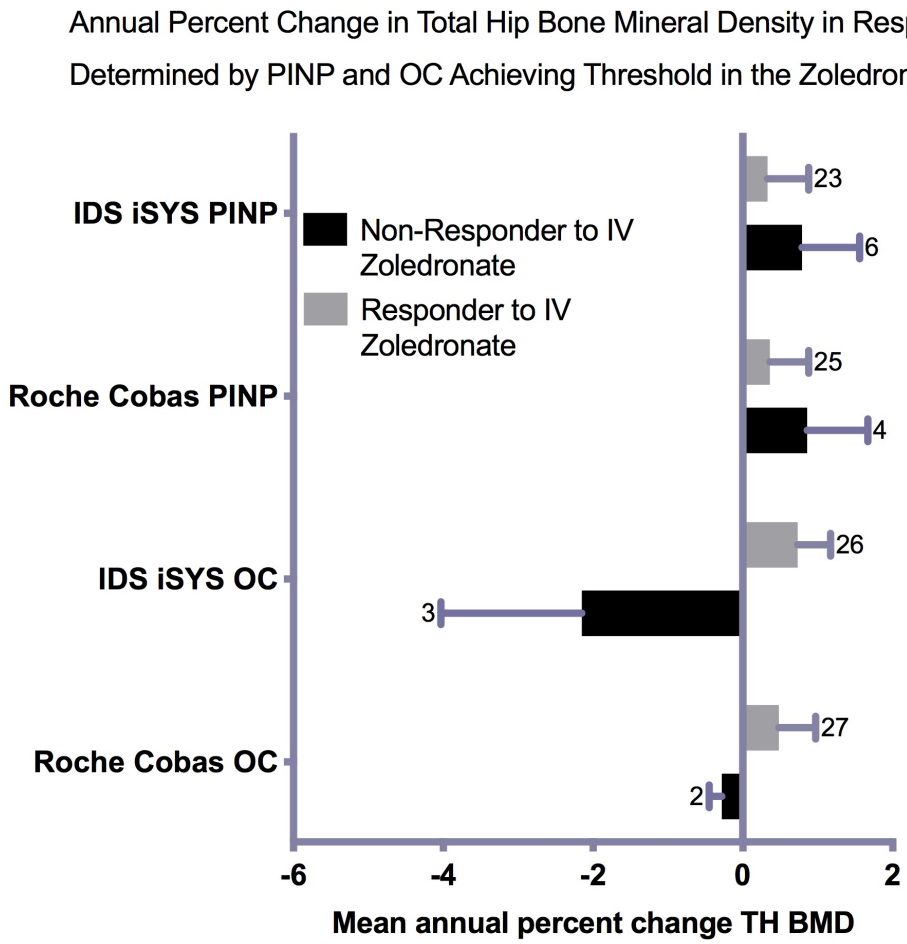


Figure 5.6. Percent Change in Total Hip Bone Mineral Density with Response in PINP and OC by two assays.

The annualised mean BMD percent change and SEM when unselected clinical patients (n=29 over 1-3 years) on zoledronate reach the PINP or OC GM (threshold) on the IDS-iSYS autoanalyser and Roche Cobas e411 autoanalyser. No significant difference detected.

Although the mean BMD percent changes were lower in the responder subgroups by PINP both assay, there was no significant difference between the mean BMD percent change in the two subgroups (responder and non-responder), when measured by either of the BTMs with each assay.

Correlation of the annualised BMD percent change on zoledronate and the BTM "on treatment" was analysed to explore the unexpected findings of BMD change. Figure 5.7 as a panel of 4 shows PINP and OC by each assay.

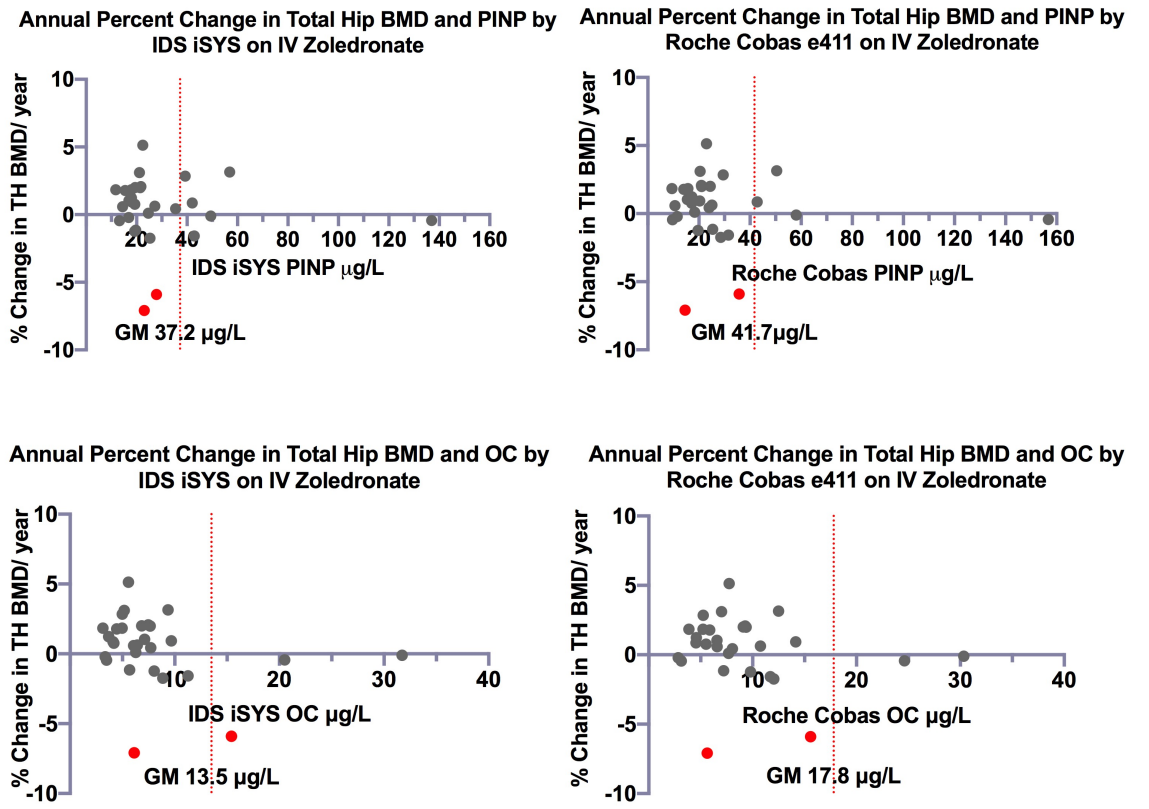


Figure 5.7. Measurement of PINP and OC by two assays in the Zoledronate Cohort with the annual percent change in BMD as Zoledronate commenced (n=29).

No R values of a Spearman Test of Correlation are significant. Geometric Mean (GM) demonstrated to show the predominant suppression of the group; IDS PINP is 37.2 µg/L, Roche PINP is 41.7 µg/L, IDS OC is 13.5 µg/L and Roche OC 17.8 µg/L

Figure 5.7 shows 2 outliers in BMD change that are marked in red. The scans were checked a second time to ensure there was no change in scanner used, positioning of the femur, and no presence of artefact that could have led to the sequential data being unreliable. The

subject that had a change of -11.8% BMD over 2 years (annualised 5.9%), had a supressed PINP by both assays below threshold and also OC by Roche Cobas e411, impacting on the mean BMD percent change in these responder subgroups. The OC by IDS iSYS did not supress below threshold and hence the absence of this impact on the responder subgroup by IDS OC as shown in figure 5.6. The subject was recruited at dose 4 of 5mg Zoledronate and clinical data review shows the PINP by Roche autoanalyser was not supressed historically, between dose 1 and 2 (the period of BMD change). There were no other apparent clinical co-morbidities that may have contributed to the initial lack of suppression. Data from subjects with over 10% weight change between the sequential scans was not used and the subject with 14% loss in BMD over 2 years (annualised 7%) experienced a 6% weight loss. There were no other clinical comorbidities that could have accounted for the BMD change out of keeping with the BTM levels.

Figure 5.7 shows that if the BTMs were recorded to be much higher above threshold there was no change in total hip BMD, however, there was no significant correlation by Spearman Test between the percent change in TH BMD on zoledronate and the BTM level at recruitment.

5 Discussion

5.1 Subject Characteristics

5.1.1 Gender and age

The mixed gender Zoledronate cohort was representative of the population with osteoporosis (6.8% men vs 21.2% women in 2015 (Borgström *et al.*, 2020)). With the small Denosumab cohort it proved more achievable to keep to an all-female cohort. Denosumab has been licensed in men with approval from the European Medicines Authority since 2014. However minimal prescribing for men in 2017 may be a consequence of the NICE technology appraisal of 2010 not being updated in its recommendations from use only in men with androgen deprivation therapy ('NICE, Denosumab for the prevention of osteoporotic fractures in postmenopausal women | Guidance,' 2010).

Because of the common practice of zoledronate being the most common parenteral treatment for osteoporosis, ibandronate was not studied. The preferential use of zoledronate is not NICE guidance, which discusses Zoledronate and Ibandronate as equal consideration and similar cost effectiveness at a MOF risk of 10%. ('NICE, Bisphosphonates for treating osteoporosis | Guidance,' 2019). However this cost analysis refers to the drug alone and not the practical costs of delivery every 3 months rather than annually.

Zoledronate is also more commonly prescribed than denosumab in the context of primary prevention, unless the T-score is significantly low to be eligible for Denosumab ('NICE, Denosumab for the prevention of osteoporotic fractures in postmenopausal women | Guidance,' 2010). Zoledronate is also chosen more in younger patients even with severe osteoporosis or secondary prevention because of the concern regarding offset of Denosumab at the end of treatment course (Cummings, 2017). For this reason a high number of subject recruitment in the Denosumab cohort was never the aim. The tertiary Metabolic Bone Centre from which the subjects were recruited has a denosumab pathway that also impacted on Denosumab cohort recruitment. Only the 1st two denosumab injections are given at the centre, before referral to primary care, out of which only attenders for the 2nd injection were eligible for recruitment (n= 15).

The Denosumab cohort mean age recorded in table 5.2 was higher than that of the Zoledronate cohort for the reasons just explored although not directly compared. It is important to note the 4 patients in the Zoledronate cohort that were potentially in the early menopausal period, 3 in the Z+ group and 1 in the Z-. According to a small study in 2000 of 28 women with a mean age of 45 years and 15 with a mean age of 53 years, the bone formation markers rise significantly, with OC most likely to be above the normal reference interval, then they actually settle slightly in a group of 20 women with a mean age of 65 years (Fink *et al.*, 2000). Equally these 4 subjects may have still been premenopausal and therefore 50% likely to have a PINP and OC level below the threshold level. In a more recent very large study (n= 657) on the balance of bone resorption and bone formation over the menopausal period (Gossiel *et al.*, 2018), the PINP is elevated compared to the young premenopausal mean (T-score method) for 20 years post menopause. This study provides reassurance that even though there was a potential difference in the age of the Zoledronate and Denosumab cohorts (they were not directly compared in table 5.2), as the mean age of

all the groups are more than 20 years post menopause, this characteristic difference alone would not impact on the BTM distributions. In addition to the impact of the menopausal years on BTM levels, is the impact on change in bone mineral density (BMD). If these subjects have been included in the 29 sets of BMD data, one would need to appreciate that without treatment, from 1 year prior to and 2 years after the final menstrual period, between 5.8 and 7.4% BMD loss can occur (Greendale *et al.*, 2012).

5.1.2 Differences in the 25-hydroxyvitamin D and PTH measurements

As 25-hydroxyvitamin D level (25OHD) is inversely associated with bone turnover marker level (Nair *et al.*, 2020) it is relevant whether the 25OHD level is significantly different between the "no treatment" and "on treatment" groups. The mean 25OHD levels in both groups of the Denosumab cohort were higher than the Zoledronate cohort and not statistically different between D-and D+, as seen in table 5.2. This finding is in the context of table 5.3 showing the percentage of the Z- and D- group reported to be prescribed cholecalciferol was 60% vs 64% and Z+ and D+ groups 75% vs 80%. The dose of cholecalciferol prescribed, was a detail not recorded, that could have explained this finding.

The PTH median level seen in table 5.2 in the D- and D+ group are higher than in the Zoledronate cohort, but they have very wide 95% confidence intervals (95% CI), due in part to the small number for analysis. There is a higher PTH level in D+ than D- (47 ng/L vs 39 ng/L) as expected, but not a statistical difference because of the wide 95% CI. Denosumab has a significant PTH stimulatory effect at a month after delivery, lasting the 6 months, even in the context of additional calcium administration (Makras *et al.*, 2013). A study with n=47 controlled for secondary osteoporosis and renal impairment, showed that the rise in PTH by 1 month could be blunted by using short term additional calcium and vit D supplementation (2g/1600 Units vs 1g/800Units). This supports the theory that the rise in PTH is partly compensatory for the denosumab reducing calcium availability from bone. In both groups of low and higher calcium supplementation, the PINP measured by Roche ECLIA supressed by month 6 to a mean of 16.6 µg/L and 17.8 µg/L, a significant difference to the baseline and no difference between groups.

5.2 Bone Turnover Marker Distribution

5.2.1 Distribution differences between parenteral cohorts

The subject characteristics of the Zoledronate and Denosumab cohort were not directly compared and the same is true of the distribution of the PINP and OC.

I have reflected on the reasons why in clinical practice, more so over the last 5 years, the group of patients suitable to receive each parenteral therapy are different, including an important difference in the renal function in which Zoledronate and Denosumab can be safely administered (Black *et al.*, 2007; Cummings *et al.*, 2009). No co-morbidities have been controlled for in this clinical study, so with different suitable patient populations, the comorbidities were expected to differ between the groups, impacting on the PINP and OC distributions.

Individual and group percent change in the BTMs is not relevant because the baseline measurements were matched subjects rather than the same pre- treatment. The fairest group wise descriptive comparison was whether the 95% confidence intervals (95% CI) of the median BTMs on treatment, overlapped with the estimate of the median on the other treatment. As table 5.4 shows, this found that predominately the D+ group PINP and OC median was lower and distinct from the 95% CI of the Z+ group, with the converse true of the Z+ group median.

5.2.2 Reflection on outlier measurements

The reasons for outliers of the PINP and OC distributions seen in figures 5.1 and 5.2 have been explored in the presentation of the results. The themes that are consistent to the oral bisphosphonate Clinical Cohort, are renal impairment with or without evidence of CKD-MBD (Ketteler *et al.*, 2017) and, or fracture within 12 months. The evidence behind the pattern of differential impact from renal impairment and fracture on the PINP and OC with the IDS and Roche assays has been fully explored in the discussion of Chapter 4. Additional factors identified for the outlier measurements included a delay in 2nd dose of zoledronate by more than 2 months. Ordinarily, a clinician would not be concerned by a delay in zoledronate dosing as the second Horizon extension trial and publication with n=95 in the placebo group, demonstrated after 6 years of annual 5mg IV zoledronate, in the following 3 years

the BTMs did not increase by a significant amount. As to whether the PINP and OC levels are more likely to rise after 12 months after just one injection, Grey at al gave just one injection to 25 osteopenic women (Grey *et al.*, 2010), with subsequent reduction of PINP to 50% of baseline up to 3 years after. This evidence tells us the outlier measurements should not be attributed to a delay in the zoledronate administration.

The inclusion of younger subjects in the Z+ group, as discussed when reflecting on the subject characteristics, shouldn't have been the cause of an outlier measurement in only OC and not PINP, and also after 5 years of zoledronate. Potentially the younger age of subject is a suggestion of another metabolic cause of the poor bone health for which zoledronate was required, that hasn't been captured with the clinical data collection.

5.2.3 Median PINP and OC level

The early clinical trials into the then new therapeutic agents, provide a framework on which to feel that co-morbidities allowing, the results of this study do follow what is recognised for zoledronate and denosumab. The pivotal Horizon Trial (Black *et al.*, 2007) reports a median 56% reduction after 3 annual infusions of 5mg Zoledronate. Only the percent change in the PINP measurements was given, as fracture was the primary outcome. The post-hoc analysis of Horizon (Delmas *et al.*, 2009) provides further breakdown of the PINP distribution at baseline (20 out of 619 3% had level below 95% LL). The median PINP measurement by the Roche Elecsys 2010 immunoassay system at baseline is equivalent to the Z- group in our study seen in table 5.4, at 49.95 µg/L reducing to 18.12 µg/L at 12 months, compared to 21-22 µg/L seen in table 5.4.

In comparison with the Horizon Trial primary and post hoc analysis is the Freedom Trial (Cummings *et al.*, 2009) and Freedom Extension data (Bone *et al.*, 2013). As expected of a pivotal clinical trial for a new therapeutic agent, denosumab, the primary outcome of Freedom was vertebral fracture. It was reported that at 6 months on denosumab there was a 50% reduction from the placebo group in PINP, which represented approximately 70% of baseline. Although not directly comparable this is an equivalent, if not greater reduction in PINP to the Horizon data for zoledronate (Black *et al.*, 2007). When the denosumab data was presented in more detail in the publication from Freedom Extension (Bone *et al.*, 2013), the median value of PINP at month 6 after injection 7 was 13 µg/L (n=69) and was not

statistically different to the 31 subjects that were cross-over over 36 months of placebo and receiving their first injection. These median values “on treatment” are similar to the 15 µg/L reported in the results of this study in table 5.4. This data also supports that it does not matter at which injection the subjects in the Denosumab cohort were recruited, provided it was at least 6 months after the first.

There was a significant difference in the median PINP and OC within both the Zoledronate and Denosumab cohort between the “no treatment” and “on treatment” groups marked on figure 5.1 and 5.2. There are clinical confounders described in table 5.3 that have contributed to this result. Recent fracture prevalence was greater in the Z- and D- group compared to the on treatment counterparts and the use of calcium supplementation in the Z+ and D+ groups was 70-80%.

5.3 Sample Representation of Clinical Practice

With the data collected on comorbidities, it has been possible to show the themes, across all the treatment groups, including the oral bisphosphonate Clinical cohort in Chapter 4, as to the importance of reviewing the PINP and OC level in the context of the individual.

As in the Clinical cohort “non treatment” group, the percentage of Z- having experienced a fracture within 12 months was high (seen in table 5.3) because of the recruitment from a population of patients attending for bone densitometry. This does highlight the difficulty in clinical practice in interpreting baseline and follow up BTM on treatment in secondary prevention. One would hope the occurrence of fractures in the Z+ and D+ group was low, as it was shown to be, but because these subjects represent very high fracture risk patients, it was not necessarily going to be the case. The fact that only one subject in the D- group was recorded to have had a fracture within 12 months, is testimony to the difficulty in showing sample representation of the population when the sample is small.

The clinical data in table 5.3 shows that the fracture history and the identification of any secondary osteoporosis was much greater in the Zoledronate cohort than the Denosumab cohort. It is of little surprise that CKD 4 and a reasonable representation of diabetes mellitus (DM) was identified in the small group of D+ in view of DM being a leading cause of CKD.

As with the Clinical cohort from Chapter 4, the recent prescription of glucocorticoids was again high. The percentage was similar across both groups of both cohorts but less than for the Clinical cohort on treatment with oral bisphosphonate, in which many of the subjects may have been on prophylactic treatment (Buckley *et al.*, 2017).

Table 5.3 shows that the percentage of cholecalciferol and calcium supplementation in the no treatment and treatment groups was similar across the two cohorts with only the record of calcium prescription in the D- being low. The recorded 70-80% seems lower than expected compared with the supplementation in the Clinical cohort "on treatment" in Chapter 4 of 90%. The only patient group on parenteral treatment that ought not to be on calcium supplementation would be those with primary hyperparathyroidism. This either shows the problems with accuracy in recording the clinical data or indicates the poor adherence that is often seen to calcium containing products. This is generally due to the unpalatable nature, or reflux (GORD) which is problematic for individuals receiving parenteral treatment because of similar side effects with oral bisphosphonates.

5.4 Bone Turnover Marker Performance by Threshold

5.4.1 Sensitivity

In Chapter 4 I explored various reasons why the sensitivities of the PINP and OC at the biological target of threshold, were lower than in the Trial cohort. Reduced adherence in the Clinical cohort, most likely contributed. Adherence in this parenteral section of the study is not in question, and the median PINP and OC levels on zoledronate and denosumab, are within previously reported expected ranges. Despite this, the sensitivities seen in table 5.5 for PINP in particular to detect treatment at threshold, are only between 82-89% for both zoledronate and denosumab. OC sensitivity by threshold, with both assays in these two cohorts is consistent at 93% and above. As fracture incidence in the on treatment groups was not high, and CKD less than stage 3b did not feature in the Z+ group in table 5.3, these 2 clinical factors are not contributors to the 18% of individuals on zoledronate that do not supress below the GM of the premenopausal reference interval.

Because the Denosumab cohort was small, just the single outlier as a consequence of a creatinine clearance less than 30, in the "on treatment" group could be affecting the

sensitivity results. Because of the nature of clinical patients receiving denosumab having renal impairment, the clinical utility of the Roche PINP and both OC assays in patient on denosumab is very limited. There is no post-hoc analysis of the Freedom data, showing the proportion of treated subjects achieving PINP levels below any target representative of the premenopausal mean.

There are also very few equivocal studies to compare the sensitivity results of the Zoledronate cohort to. The post hoc analysis of Horizon, studies the PINP measurement relative to the 95% LL of the premenopausal reference interval, with only 19% achieving a supressed level below this at each annual time point (Delmas *et al.*, 2009). However, there was no significant difference in fracture incidence in the group with very supressed PINP to those on active treatment without suppression below the 95% LL. Having a biological target of the lower end of premenopausal reference interval confers no additional benefit, but similar publications with respect to the median of the premenopausal reference interval are lacking.

The clinical study of BTM suppression on bisphosphonates by Eekman (Eekman *et al.*, 2011), included 24 out of 95 subjects on intravenous bisphosphonates (zoledronate and pamidronate), 7 of which had baseline and follow up PINP measured at a median of 4 months. However, the publication does not did not look separately at the percentage of this small subgroup with a PINP in the lower half of the premenopausal reference interval, only to mention that the sensitivity of 95% overall may have been so high as adherence to IV bisphosphonates is not in question. None of the 7 subjects on IV bisphosphonates were highlighted as failure to achieve LSC. The assay dependent threshold value in our study was lower than that in the Eekman study (37µg/L IDS iSYS vs 42µg/L Roche Cobas e411 vs 45µg/L Manual RAI Orion). The high sensitivity of this trial is interesting given the early time point of the sample. Although the Reid group had shown that PINP does reach a nadir after a single zoledronate 5mg dose at 3 months (Grey *et al.*, 2010).

There is no reliably similar study to contradict the suggestion that up to 18% of patients treated with zoledronate with a PINP measured by IDS iSYS autoanalyser, would require further exploration as to whether the lack of suppression below the biological target of threshold, could indicate lack of treatment effect.

5.4.2 Specificity

It has already been reflected on in Chapter 4 why the specificity of the biological target of threshold performed poorly, with calcium supplementation and previous treatment effect being profound, despite the converse impact of recent fracture occurrence.

The specificity in the Zoledronate and even more so the Denosumab cohort, was low as seen in table 5.5. It is suggestive that up to around 70% of patients could have a PINP or OC level supressed below the GM of the premenopausal reference interval that was not related to the current treatment.

Table 5.6 again demonstrates the impact of previous anti-resorptive treatment within 5 years on the specificity, particularly in the Denosumab cohort. The matched median T-score of -2.24, lower than that of the Zoledronate and the Clinical cohort from Chapter 4 (-1.76 and -1.62 respectively) is the probable explanation. The no treatment group were more likely to have been "on treatment" before (29% D- vs 25% Z- vs 24% C-).

The TRIO offset study of 2018 (Naylor *et al.*, 2018), showed up to 2 years post completing 2 years of oral bisphosphonate therapy, that the PINP and OC by IDS iSYS autoanalyser distribution from 49 women that had been in the TRIO study, remained significantly lower, at negative 37% from the pre-treatment baseline. A usual treatment course is longer than 2 years, and with the mechanism of incorporation into the osteoid, there is the potential for a total dose effect on time to offset. The offset of zoledronate has been extensively studied by the Reid group. In particular, after a singular 5mg dose in 25 healthy postmenopausal women with osteopenia, compared to 25 matched controls receiving placebo, the effect on PINP was ongoing at 3 years to approximately 50% of baseline, controlled for calcium and vitamin D intake (Grey *et al.*, 2010). The subjects in our study that were "no treatment" matches were included providing they had been off oral bisphosphonates for longer than 1 year and off zoledronate for more than 2 years. Although this has likely impacted on the specificity, it is representative of clinical patients escalating osteoporosis treatment to parenteral administration, because of failure of oral bisphosphonates, intolerance or restarting zoledronate after a period of time without it.

Despite an unexpected low sensitivity of PINP at threshold to demonstrate apparent treatment effect, and a very low specificity, the area under the curve on ROC analysis shown

in table 5.7, was 0.80 and above. This is an acceptable performance of a tool to differentiate between being on treatment or not (Hess *et al.*, 2012).

5.5 Bone Mineral Density Change in Biochemical Response

It has been seen and discussed in Chapter 4 that although there was some correlation in the PINP measurement by both assays and the percent change in bone mineral density (BMD) in the Clinical cohort on oral bisphosphonates, there was no significant difference in the mean percent change in BMD between the biochemical responders and non-responders.

The Zoledronate cohort mean BMD percent change when grouped by biochemical response positive and negative, have 95% CI that cross zero with all assays of PINP and OC, shown in table 5.8. This is due to the small number of sequential DXA from which the data could be gathered (n=29). Consequently, there is no difference in the mean BMD percent change in "responders" vs "non-responders".

One potential factor seen in the Clinical cohort of Chapter 4 and playing a part in the interpretation of the BMD results in the Zoledronate cohort is the underestimate of the annualised mean BMD percent change. The method of collecting sequential DXA data for this group was to obtain the first DXA scan no more than 12 months before or after the zoledronate start and a follow up scan with the interval within 3 years. The Horizon extension data shows the pattern of BMD gain to be more pronounced in the first 2 years and then plateau (Black *et al.*, 2015). The underestimate of the annualised change should have been less than for the cohort on oral bisphosphonates. However, with a possible 12-month period of BMD loss before starting zoledronate being captured, this would underestimate the BMD change that occurred on the zoledronate.

The lack of synchronicity of the timing of the BTM measurement and the change in BMD interval should not be so important where adherence to zoledronate cannot fluctuate. In fact, in analysing one particular instance of BMD loss over the first couple of years of being on zoledronate, seen in figure 5.7 with a supressed PINP and OC at recruitment, the only anomaly found was that during the interval of the BMD loss, the clinical PINP measurement by Roche Cobas autoanalyser was above the upper limit of the premenopausal reference interval. There could have been any number of temporary reasons for this (fracture, temporary acute kidney injury) that should not have impacted on systemic bone turnover.

Figure 5.7, in which no significant correlation was found between the PINP or OC level and the mean BMD percent change, does show how the majority of points sit below the premenopausal GM, and it is within this bunch that there is no visible correlation. This is in keeping with the suggestion in the post-hoc analysis of Horizon (Black *et al.*, 2007; Delmas *et al.*, 2009) that there is no discernible benefit in fracture reduction of supressing the BTM level below the lower limit of the premenopausal reference interval.

I have been unable to demonstrate that the 18% of the Z+ group that did not achieve the biological target of threshold, had a worse outcome in terms of change in BMD. Because not all of the cohort had DXA data, it leaves the question rather open as to whether there is clinical utility in testing BTM in clinical patients on zoledronate.

Chapter 6: Summary and Conclusions

1 Summary of Results

This study on the use of bone turnover markers in monitoring osteoporosis treatment in clinical practice was performed with measurements of two bone formation markers procollagen I N-propeptide (PINP) and osteocalcin (OC). The study was performed in three stages from establishing two biological targets to monitor osteoporosis treatment by, to testing the performance of these targets in both a trial and clinical setting to monitor oral bisphosphonate therapy and finally studying whether the use of the biological target of threshold adds monitoring information in a clinical patient group treated with parenteral therapy.

1.1 Establishing LSC and Threshold as Biological Targets

A UK premenopausal reference interval of PINP and OC from a group of 130 subjects was significantly lower than the UK reference interval from 50 men age 30 to 60 years. The geometric mean (GM) for the female cohort of PINP of 37.2 µg/L (IDS) and 41.7 µg/L (Roche) and male GM of PINP 51.8 µg/L (IDS) and 49.5 µg/L (Roche); the female GM of OC 13.5µg/L (IDS) and 17.8µg/L (Roche) and male GM of OC 19.1µg/L (IDS) and 22.9µg/L (Roche). This difference is in keeping with reference ranges published from other geographical areas. Men have wider reference intervals than the women.

In each pair of numbers, the lower geometric mean was calculated from measurements by the IDS iSYS autoanalyser and was significantly lower than that of the Roche Cobas e411 autoanalyser with the exception of the male PINP reference intervals in which there was no difference. The Roche N-MID OC assay has previously shown a higher reference interval than the IDS N-MID assay (Chubb *et al.*, 2015; Hannemann *et al.*, 2013)and the Roche total PINP assay a higher reference interval than the IDS intact PINP (Morovat *et al.*, 2013).

When method agreement was analysed with Bland and Altman and Passing and Bablok there was almost complete agreement in PINP or OC when measured by the two assays. The slope was close to or incorporated 1 through 95% confidence intervals. The relationship between the PINP by IDS iSYS and Roche Cobas e411 altered with slightly higher values of PINP for Roche Cobas e411.

The assay specific GM of the premenopausal reference interval was termed the “threshold” and has been taken forward as one of the biological targets by which to monitor osteoporosis treatment.

The within-subject CV was calculated from paired measurements were made on samples 1 week apart. These were obtained from 135 postmenopausal subjects on oral bisphosphonate treatment, controlled for other effects of bone metabolism. The least significant change (LSC) for PINP of 6.7 µg/L and 28% by IDS iSYS and 7.9 µg/L and 28% by Roche Cobas e411, were lower than previously reported. The method I used to calculate the LSC made the assumption that the within subject CV includes a proportion of inter-assay variation. The same method of calculation did not impact on the LSC for OC of 4.4 µg/L and 23% by IDS iSYS and 4.1 µg/L and 24% by Roche Cobas e411 compared to previously reported values (Eekman *et al.*, 2011; Scariano *et al.*, 2001; Naylor *et al.*, 2016). The absolute figure of LSC by both assays was taken forward as the more established biological target by which to monitor osteoporosis treatment.

1.2 Evaluating Performance of the Biological Targets to Detect Treatment with Oral Bisphosphonates in a Trial and Clinical Cohort

A direct comparison was made between the PINP and OC distributions from a controlled Trial cohort of n=135, the subjects of which acted as both the control and the treated group and a Clinical cohort of subjects on treatment (n=86) with matched controls (n=82) that were not. The cohorts were comparable in most respects except gender, with a mean age of 68 and 71 years, and total hip T-score of -1.35 and -1.62 respectively, with equivalent median 25OHD levels and median PTH levels higher in the Trial cohort. Comorbidities that affect bone metabolism were recorded in 51% of the treated clinical subjects, including fracture within 12 months (6%), significant glucocorticoid use (25%), thyroid disorder (15%), PHPTH (4%) and diabetes mellitus (9%). Adherence was high in the TRIO cohort but unrecorded in the clinical cohort.

A descriptive comparison and ROC analysis was made of how the biological target of threshold performed to differentiate treated subjects with three oral bisphosphonates and those not on treatment in each cohort. The apparent response rate (sensitivity) by threshold in the Clinical cohort was high at over 80% but lower than in the Trial cohort (86% for Roche

OC and 96% for PINP with both IDS and Roche). The significantly lower AUC in the Clinical cohort of 0.79-0.83 (assay and BTM dependent) compared to that of the Trial cohort of 0.96 and above was likely from the much lower specificity of the threshold to detect treatment in the Clinical cohort (59% with OC to 70% with PINP). PINP and OC, each by the two different assays showed no difference in the AUC with a Youden Criteria similar to the threshold value.

Recent fracture is likely to have impacted on the sensitivity in the Clinical cohort, and previous anti-resorptive exposure (not within 12 months but within 5 years of recruitment) alongside calcium supplementation greatly reduced the specificity.

The threshold was compared against the absolute LSC as the biological target in the Trial cohort only. There were equivalent sensitivities by PINP measurement of 95% and above, and less concordance in the sensitivity by OC measurement with lower response rates by threshold of 76 and 86% compared to LSC of 97 and 94% (IDS OC and Roche OC respectively). There were 6 subjects that achieved LSC and did not supress below the threshold of OC measured by Roche Cobas e411 and 4 subjects that had measurements below the threshold at baseline and on treatment without achieving LSC.

Serial total hip (TH) bone mineral density (BMD) data was available for 120 subjects in the Trial cohort at baseline and week 48. This showed that the “responder” subgroups by either biological target had a positive change in BMD. Although there was a significant difference in the change in BMD between the “non-responders” and “responders” to treatment by OC LSC (Roche Cobas e411), there was no difference in the change in BMD between the “non-responders” and “responders” to treatment by OC threshold.

The change in BMD results for the Clinical cohort were limited by number and by an underestimate in the way the initial annual BMD change could be calculated from an interval of 2-5 years from treatment start. TH BMD did improve in the “responders” by threshold in the Clinical cohort and there was no difference on comparison of the mean change in BMD between “responders” in the Clinical and Trial cohort. There was however, no difference in the mean BMD change between the “non-responders” and “responders” to treatment measured by threshold in the Clinical cohort.

1.3 Does the Biological Target of Threshold aid Monitoring of Clinical Parenteral Treatment Cohorts?

A Zoledronate cohort of 99 mixed gender subjects receiving annual IV zoledronate, matched with 92 subjects on no treatment and a much smaller female Denosumab cohort of 15 subjects on 6 monthly SC denosumab with 14 matches, had measurements of PINP and OC by both the assays to ascertain the sensitivity and specificity.

There were expected differences between the Zoledronate and Denosumab cohorts with the former being younger (70 vs 77 years), having a higher TH T-score (-1.74 vs -2.34) and a lower PTH (29.2 ng/L vs 46.7 ng/L), with CKD 4 also recorded in 2 subjects on denosumab. The other clinical co-morbidities for the Zoledronate cohort were equivalent to the oral bisphosphonate Clinical cohort earlier described, but with less calcium supplementation (79% vs 90%) and less glucocorticoid prescription (12% vs 25%). The clinical comorbidities in the small Denosumab cohort were representative of a typical population.

The distributions of PINP and OC were not directly compared between the Zoledronate and Denosumab cohort because of all the differences detailed above. However, the median PINP and OC levels on these parenteral treatments were similar to how they have been previously reported in the literature (Delmas *et al.*, 2009; Cummings *et al.*, 2009; Eekman *et al.*, 2011), despite the co-morbid conditions effecting bone metabolism. The clinical information was useful for identifying a cause of the outlier measurements.

The sensitivity of PINP at threshold to demonstrate treatment with IV zoledronate and SC denosumab was high (82% and above and 87%) and with OC at threshold even higher, 93% and above with both treatments. This is an interesting result given that incomplete adherence does not feature as a reason for treatment not supressing the BTMs like it does in the cohort on oral bisphosphonates. The small Denosumab cohort is being affected heavily by more cases of CKD 4, in which situation the biological target becomes futile.

Due to the previous exposure to anti-resorptive, although the sensitivity of the OC was higher, the specificity was extremely poor with both, lower with OC and in the Denosumab cohort matches. The Denosumab cohort had the lowest mean T-score of all the cohorts. The subjects therefore recruited on no treatment at that time or for the previous 12 months, were more likely to have been exposed to treatment within the previous 5 years. This is

relevant to clinical practice where treatment is often escalated from oral to a parenteral therapy. The AUC for both IDS and Roche assays of PINP and OC demonstrated no difference.

2 Strengths and Limitations

2.1 Recruitment from a Tertiary Metabolic Bone Centre (MBC)

The method protocol enabled maximum recruitment, good representation of a typical clinical population, and ensured low impact on the patient's usual care pathway. Subjects recruited from the Fracture Risk Assessment Service (FRAS), or attendance for IV Zoledronate didn't even require additional venepuncture.

Clinical review attendances are minimised in the MBC by efficient systems such as virtual assessment and treatment recommendations to primary care through the FRAS DXA and biochemistry reporting, shared care with primary care for continued denosumab administration and direct access for zoledronate administration in cases of contra-indication or intolerance to oral bisphosphonates. This meant that a method protocol designed around the recruitment of clinical subjects prior to commencing treatment and obtaining samples at both baseline and on treatment, would have required input from the subjects additional to their normal clinical care and led to diminished recruitment.

Recruitment of the well-matched no treatment groups was from attendance for bone densitometry. A large proportion of patients undergo DXA scans for monitoring to see whether bone protective treatment requires re-starting after a break of approximately 2 years. The 25% to 30% of the no treatment groups having had prior anti-resorptive exposure was a contributory factor to the poor specificity of threshold to detect treatment. Having a method design of the subjects being their own controls at baseline may reduce some of this effect. Recruitment would be even more severely restricted from the Metabolic Bone Centre using exposure to anti-resorptives within 5 years as an exclusion criterion, regardless of the design. It would be almost impossible for the parenteral cohorts where treatment is often escalated from an oral therapy due to ineffectiveness or intolerance. In conclusion therefore the poor specificity is indeed representative of real life.

Eligibility for recruitment in terms of length of time on or without treatment, was only checked with the patient at the point of their attendance to the MBC. This was difficult to ratify for oral bisphosphonates which would have required access to primary care prescribing records. It could be confirmed for the parenteral treatment from the record of administration on the ARQ clinical database. This led to exclusion of subjects in the cohort when data extraction from the database provided confirmation that they did not meet inclusion criteria. For many of the Clinical cohort, there was not a formal check of eligibility.

Despite the challenges in recruitment we did achieve an appropriate sample size to power the sensitivity and specificity calculations with n=270 in the Trial cohort, n=167 in the Clinical cohort and n=191 in the Zoledronate cohort (Hess *et al.*, 2012). The Denosumab cohort as expected was under powered at n=29.

2.2 Data Collection from a Clinical Environment

The only successful acquisition of biochemistry from the clinical systems within 6 months of the recruitment and sampling date was renal function and this was useful in interpreting outlier measurements in the no treatment groups and Denosumab cohort. Availability from the clinical biochemistry systems of contemporaneous calcium, testosterone, and TSH was poor. Adjusted calcium measurements were useful alongside the study PTH measurements if they were elevated, but where there was an absence of a clinical calcium reading and a high normal PTH, the individual would not have been noted to have PHPTH.

Although serum CTX is a recommended BTM to use in clinical practice (Morris *et al.*, 2017), it would not have been possible to compare measurements of CTX in the clinical cohorts. The patients were recruited at all times of the day and were not required to be fasted, which would have increased the pre-analytical factors in the measurements (Szulc, 2012). The treatments I have studied are anti-resorptive in mechanism of action, but because of the coupled bone turnover cycle, the bone formation markers of PINP and OC are valid to have assessed for utility of monitoring (Vasikaran *et al.*, 2011). Recent meta-regression of 28,000 subjects from 11 bisphosphonate trials (and 3 SERM), demonstrated a significant correlation in the percentage change in PINP from placebo to treatment group and the odds ratio of vertebral fracture. This was also found in BAP as another bone formation marker, but the correlation was not significant in the markers of bone resorption (Bauer *et al.*,

2018). This indeed validates the practical and scientific decision making of measuring bone formation markers only in this study.

Only positive information on factors that can affect bone metabolism were recorded. If no input was present, this could mean either a negative response or no response. Extraction of the information would have under-estimated, rather than over-estimated the proportion of co-morbidities or supplementation in the cohorts and it was not possible to be specific regarding missing data. Information on the proportion of the subjects with diabetes mellitus (DM), was obtained primarily from the prescriptions of diabetes medication extracted from the database, because it was not a question on the study workbook. Therefore, the proportion of the cohort with diabetes mellitus actually was not inclusive of diet-controlled DM. Other more specific details that would have had an impact on the bone turnover and change in BMD, was adherence to a gluten free diet, over-treatment with thyroid hormones or the reason for activated vitamin D prescription.

Using a baseline BMD reading within 12 months of commencing the osteoporosis medication guaranteed capture of the initial gain in BMD with bisphosphonate therapy. Although as I have previously discussed, this had the potential to be underestimated due to the interval including time periods of the plateau phase or even BMD loss prior to treatment commencing. This method of data collection also led to diminished number of sequential clinical scans for use in the analysis. In the Clinical cohort on oral bisphosphonates, the inability to ratify the start date of treatment, meant there was no confirmation of which scan results to use. In the Zoledronate cohort the direct access to the treatment, particularly after a fragility fracture above the age of 75 years old, meant a baseline DXA would not have been necessary to perform. The absence of the BMD analysis for the Denosumab cohort is the biggest missing data analysis of this study. The two limitations impacting on the numbers for inclusion in BMD analysis of the Clinical and Zoledronate cohort, are not relevant for the Denosumab cohort, but two further factors were. Denosumab is passed to primary care administration after the first two doses, and therefore for 50% of the denosumab treated group, recruitment occurred only 6 months into treatment start. There could have been a baseline DXA result, but no follow up scan. A further 25% of this group demonstrated that the decision for denosumab and preparation and administration take time, and the baseline DXA scans were more than 12 months before the treatment was

actually started. The decision was made that data from 4 subjects sequential scans would not offer any additional insight into the use of BTMs to monitor osteoporosis therapy.

The follow up DXA scan in the clinical patients was at variable time points between 1 to 5 years after the baseline scan. This diluted the initial gain phase in BMD seen in bisphosphonate treatment to different degrees.

Comparing the Clinical cohort on oral bisphosphonates to the Trial cohort, was a very useful exercise to determine how much “real life” interplays with the utility of the BTMs to monitor anti-resorptive treatment. It would have been beneficial to go through a similar comparison for the parenteral therapies, but we did not have access to trial data on these treatments.

3 Conclusions

This study shows that bone turnover markers would provide the clinician useful information in monitoring osteoporosis treatment even in the context of secondary osteoporosis. The utility is not limited to oral bisphosphonate treatment but requires further evaluation as to what the consequence of a non-response by threshold means on parenteral treatment.

Having studied two different bone formation markers, there was no difference in diagnostic accuracy of PINP and OC as analysed with ROC. However, the differences in the two different assay reference intervals, underline the importance of assay standardisation or harmonization (Cavalier *et al.*, 2019; Bhattoa *et al.*, 2021) and meanwhile of being aware of the assay used in clinical decision making. As the first direct comparison of the IDS iSYS and Roche Cobas e411 OC assays, this study concludes that because the two chemiluminescent immunoassays of OC being similar in method, there is good method agreement, but with a fixed offset that could be recorded as a reliable adjustment.

The differences in the reference interval by gender raise the question of whether the biological targets for treatment response should be gender specific, but using the premenopausal geometric mean for evaluating mixed gender cohorts did not cause a lower response rate and this should continue to be an accepted practice.

Apparent response rates to oral bisphosphonates are lower in clinic patients than in trial participants with lower adherence and secondary osteoporosis being possible contributory factors. Fracture history as distant as 12 months impacted on the usefulness of measuring the BTM with OC displaying an apparent advantage over PINP in terms of its more subtle rise in the face of a fracture as seen in other literature. PINP is more consistent in detecting response and intact PINP has an advantage over total PINP and N-MID OC in that it is less effected by CKD stage 4 or lower. The maximal suppression of PINP can be detected at 3 months, but measurement of OC would be optimised at a longer time period.

Detection of an apparent positive biochemical response by threshold to oral bisphosphonates is associated with beneficial changes in BMD comparable to that seen in trial settings but no conclusions can be made about the BMD change in the group that do not respond. The trial data in which both biological targets to oral bisphosphonates could be evaluated, indicated that LSC was a more discriminatory biological target for demonstrating change in BMD.

The sensitivity of PINP and OC at threshold to demonstrate parenteral treatment is higher with a diagnostic accuracy (AUC in ROC analysis) over the clinically accepted 0.80 (Hess *et al.*, 2012). As many as 18% of patients on zoledronate may be shown to have lack of biochemical response in PINP by threshold. It is not possible to say if for these patients IV zoledronate is not effective enough to treat their osteoporosis with BMD stability.

From reflecting on the sensitivity and specificity data, a single measurement of a BTM that supresses below the biological target of threshold, may only be usefully reassuring of treatment effect in treatment naïve subjects. Non the less, important review and discussions are stimulated if the single measurement on treatment does not supress below the biological target of threshold.

Our data does not aid in a conclusion regarding whether a patient who does not supress below the biological target of threshold with zoledronate, would be guaranteed to do so on denosumab.

4 Application to clinical work

Many services that treat osteoporosis do not routinely use bone turnover markers (BTMs) for monitoring anti-resorptive treatment despite extensive evidence of the changes that occur; perhaps because of the selective population the studies have been performed in, with uncertainty of how applicable this is to a clinical population. Additionally, some clinical units may find it impractical to perform one of the International Osteoporosis Foundation recommended bone resorption marker, C-terminal cross-linked telopeptide (CTX), which should be a fasted morning sample to reduce the biological variation.

The conclusions of this study which is on a non-selected clinical population will add to the directly applicable data, with transparency in the description of all influencers of bone turnover in the study population, for use of BTMs routinely in clinical practice.

Another benefit is the principal aim of the study to calculate the sensitivity and specificity of PINP and OC at a biological threshold, to identify treatment effect. These are terms that are now commonly used to evaluate and report new diagnostic tests and are widely understood by clinicians. They are relevant terms to use for any clinical patient that has only had the opportunity of one measurement of a BTM on treatment, rather than a baseline and on treatment measurement by which to assess response by the least significant change value.

This simple, cost-effective method (Burch *et al.*, 2014) could prevent the patient remaining with a higher risk of fracture for the two year standard interval before checking response to a treatment with bone densitometry.

A guideline on patient and analytical factors that should be considered when interpreting a singular BTM level in the primary care environment would be useful.

5 Strategy for taking the work forward

This research has produced values of threshold, sensitivity and specificity of two bone formation markers for osteoporosis treatment monitoring in clinical practice using two different assays. However, bone resorption markers could be looked at next if we can navigate how to obtain fasting morning blood samples in the clinical groups.

The sensitivity and specificity in this study was determined by ability to detect treatment. The Fink paper (Fink *et al.*, 2000) calculates sensitivity and specificity as per the achievement of BMD LSC at 1 year being true positive and negative. Using BMD as the gold standard would highlight if subjects with BTM supressed below threshold with comorbidities that may confound the result, can still be classed as responders that see a positive effect from treatment. This is a similar method to a study on HRT and oral alendronate (Greenspan, Resnick and Parker, 2005) and RAVN et al studying alendronate by BTM percentage change (Ravn *et al.*, 1999). But the BTM would have to be in the same time period at the BMD change. If obtaining a pre-treatment and 3 month measurement is added into this method, it could highlight whether there is a need to supress below threshold, if the LSC has been met. Separate cohorts could be recruited for treatment naïve subjects and post treatment break or escalation of treatment subjects.

Recent meta-regression work on multiple anti-resorptive clinical trials showing that change in BTM (BAP, PINP and CTX) in trial treatment cohorts accounts for approximately 80% vertebral fracture risk reduction on treatment and 60% non-vertebral fracture risk reduction (NS) (Eastell *et al.*, 2020), should enable BTM to be used as a surrogate for fracture risk reduction in novel agent trials. It would be ideal to reach a point where an individual could be informed that with a certain percent reduction in BTM monitoring on treatment, by how much in theory their fracture risk had been reduced.

It would be beneficial to recruit big enough cohorts to allow subset analysis with sufficient power in different comorbid conditions or co-prescribed medications to analyse the proportion with which they may affect sensitivity and specificity.

A large aspect of osteoporosis care is secondary prevention in the context of fracture liaison services. More work is required to identify which BTM has the best response to treatment in a ratio with the response to a fracture. This could guide a specific recommendation for using BTM in FLS services.

A longer longitudinal study is required looking at the incorporation of BTM in clinical practice and how that effects adherence to oral bisphosphonates and treatment decisions of escalation. BMD data over a defined 5 year period would allow evaluation of the precision of monitoring based on the LSC and threshold biological targets to compare with the cost.

Appendices

1 Appendix A: SYNDMB Workbook (V1.0 16Dec15)

HPP Module 2 Biobank Collection

STH15691 South Yorkshire and North Derbyshire Musculoskeletal Biobank

1. Appendix A: SYNDMB Workbook (V1.0 16Dec15)

Subject Initials

Study ID H M 2

Date of birth d d m m y y y y

Gender M/F

Age years

Consent

Consent obtained? ☐ Yes ☐ Wants more time →

Details

Date of informed consent d d m m y y y y

Time of informed consent h h m m

Version and date of PIS v . d d m m y y y y

Version and date of ICF v . d d m m y y y y

Consent form copies:
- ☐ Original to ISF
- ☐ PIS to subject
- ☐ Copy to subject
- ☐ Copy to source notes
- ☐ Copy to FRAS pack

Anthropometric data

Height . cm

Weight . kg

BMI . kg/m²

HPP Module 2 Biobank Collection

STH15691 South Yorkshire and North Derbyshire Musculoskeletal Biobank

Subject Initials | | | |

Study ID | H | M | 2 | | | | |

Questionnaire check

- [ ] Questionnaire completed
- [ ] Check medication history
- [ ] Check supplements/OTC
- [ ] Check coeliac disease, underactive thyroid, B12 deficiency/pernicious anaemia, Wilson's disease
- [ ] Copy of questionnaire for source notes

Blood samples

- [ ] SST 8.5 ml
- [ ] SST 8.5 ml
- [ ] EDTA 4.0 ml
- [ ] EDTA 4.0 ml
- [ ] Sample handling form completed
- [ ] Copy of sample handling form to source notes
- [ ] Samples sent to CRF

Study log

- [ ] Subject details entered on study log

Comments

- [ ] Happy to be contacted about future research

Completed by:

Name

Signature

2 Appendix B: Sample Collection Standard Operating Procedure

Blood samples will be collected from the recruited subjects at any time of the day within the Clinical Research Facility opening hours.

Use the BD Vacutainer® blood collection system into 1st two SST II Advance 8.5ml bottles and 2nd two spray-coated KEDTA bottles.

The venepuncture may be performed by the phlebotomist if the subject is also requiring blood tests for their clinical management (and passed to research staff promptly) or by the research team accredited in phlebotomy and on the delegation log.

The cannulation (if parenteral treatment is required) and withdrawal of blood will be performed by the Day Unit staff nurse or phlebotomist and passed to research staff promptly.

The bottles will be labelled temporarily with the patients' initials and study number and put in an opaque envelope to reduce UV exposure.

For serum; after 30minutes time has elapsed for coagulation the blood tube will be put in the centrifuge at 4 degrees Celsius, 3000rpm, for 20minutes.

For plasma; with as little delay as possible the blood tube will be put in the centrifuge at 4 degrees Celsius, 3000rpm for 15minutes.

The serum or plasma layer will be pipetted into 0.5ml aliquots labelled with initials, study number ending with sample aliquot number (i.e., P1-P8 or S1-S15) and date of sample.

The aliquots will be stored at -80 degrees Celsius and transferred into the SYNDMB biorepository racking system in the Medical School, Royal Hallamshire Hospital, Sheffield.

3 Appendix C: Metabolic Bone Centre Questionnaire (PD6421-PIL2472 August 2011, March 2014, September 2016)

Metabolic Bone Centre Questionnaire

Northern General Hospital, Tel: 0114 2715340

Name: _____

Please complete this form and bring it with you to your appointment. The information will be used to advise you and your doctor about your bone health. All answers will be treated in the strictest confidence. **If you need any help filling in this form please ask when you come for your appointment.**

When completing the form, please place a cross in the box to indicate your chosen answer like this: ☒

If you complete a box incorrectly, please black out the wrong one like this: ■ and put a cross in the correct box

Title	
Forename	
Surname	
Address	Postcode:
Telephone No	*Home* ☐☐☐☐☐☐☐☐☐☐☐ *Mobile* ☐☐☐☐☐☐☐☐☐☐☐
Main Occupation (before retirement)	
Date of Birth	☐☐ ☐☐ ☐☐☐☐ Age ☐☐ Gender ☐ F ☐ M
GP Name	
GP Address	
How tall are/were you as a young adult?	☐ Feet ☐☐ Inches Or ☐☐☐.☐☐ cm
If you think you have lost any height please tell us how much:	☐☐ Inches Or ☐☐ cm
Ethnic Background	☐ White ☐ Chinese/Oriental ☐ Black African ☐ Black Caribbean ☐ Asian ☐ Other (describe) _____ ☐ Arabic ☐ Mixed (describe) _____

Name: _____ DOB: D D M M Y Y Y Y

Q1 Bone Health

Have you ever broken (fractured) any bones?	☐ Yes ☐ No

If **Yes**, please fill in the table below

Which bone(s) have you broken (fractured)? *Shoulder, arm, elbow, wrist, hand, rib, spine, pelvis, hip, leg, ankle, foot* Other – please specify	**Age when broken** (approximate age if unsure)	**How severe was the incident that led to the broken bone?** ***(put a cross in one box)***		
		Minor e.g. tripped over, knocked into something or no injury at all	**Moderate** e.g. fall whilst running or from a low height such as a step	**Severe** e.g. road accident, fall from great height
	☐☐ years	☐	☐	☐
	☐☐ years	☐	☐	☐
	☐☐ years	☐	☐	☐
	☐☐ years	☐	☐	☐
	☐☐ years	☐	☐	☐

Please continue on the back page if you require more space

Q2 Back Health

Have you ever had severe back pain lasting for more than a few days? ☐ Yes ☐ No

If **Yes**, please describe **how** and **when** it started _____

_____

Have you had any x-rays or scans of your back? ☐ Yes ☐ No

If **Yes**, at which hospital and when? _____

Have you ever had an operation on your back? ☐ Yes ☐ No

If **Yes**, which year did you have the operation: ☐☐☐☐

Details of the operation: _____

_____

Name: _____ DOB: D D M M Y Y Y Y

Q3a About Your Health

Many medical conditions can affect bone health. Do you have or have you ever had any of the following conditions diagnosed by your GP or at the hospital? ***(Please put a cross in all that apply)***

- ○ Addison's disease
- ○ Ankylosing spondylitis
- ○ Anorexia or bulimia
- ○ Asthma
- ○ Avascular necrosis
- ○ Cancer (specify) _____ _____ _____
- ○ Coeliac disease
- ○ COPD, bronchitis or emphysema
- ○ Crohn's disease
- ○ Cystic fibrosis
- ○ Diabetes
- ○ Epilepsy
- ○ Indigestion, acid reflux, stomach ulcer
- ○ Kidney disease
- ○ Kidney stones
- ○ Liver disease
- ○ Lupus (SLE)
- ○ Myeloma
- ○ Multiple sclerosis
- ○ Osteogenesis imperfecta
- ○ Paget's disease
- ○ Parathyroid disease
- ○ Polymyalgia rheumatica
- ○ Rheumatoid arthritis
- ○ Stroke
- ○ Temporal arteritis
- ○ Thyroid disease:
  - ☐ Overactive
  - ☐ Underactive
- ○ Ulcerative colitis
- ○ Vasculitis
- ○ Vitamin D deficiency

Please list any other medical conditions you have:

_____

_____

Please continue on the back page if you require more space

Name: _____ DOB: D D M M Y Y Y Y

Q3b About Your Health

Have you had any of the following kinds of operations which may affect your scan results?	Year of Operation	Details
☐ Hip Replacement ☐ Right ☐ Left	Y Y Y Y / Y Y Y Y	
☐ Bowel or Stomach	Y Y Y Y	
☐ Endocrine (e.g. thyroid, parathyroid, adrenal, pituitary)	Y Y Y Y	
☐ Transplant (e.g. heart, lung, kidney, liver, bone marrow)	Y Y Y Y	
☐ Other: **Please tell us about any other operations you have had on the back page**		

Q4 Your Medication

Treatment for your bones - Please indicate if you have ever taken any of these treatments and approximate start and stop dates: **(*Please circle 'Take now' if you are currently taking medication*)**

○ Alendronate (alendronic acid, Fosamax)	Start : M M / Y Y Stop: M M / Y Y Take now
○ Denosumab (Prolia)	Start : M M / Y Y Stop: M M / Y Y Take now
○ Etidronate (Didronel PMO)	Start : M M / Y Y Stop: M M / Y Y Take now
○ Ibandronate (Bonviva, ibandronic acid)	Start : M M / Y Y Stop: M M / Y Y Take now
○ PTH (Preotact)	Start : M M / Y Y Stop: M M / Y Y Take now
○ Raloxifene (Evista)	Start : M M / Y Y Stop: M M / Y Y Take now
○ Risedronate (Actonel, risedronate sodium)	Start : M M / Y Y Stop: M M / Y Y Take now
○ Strontium ranelate (Protelos)	Start : M M / Y Y Stop: M M / Y Y Take now
○ Teriparatide (Forsteo)	Start : M M / Y Y Stop: M M / Y Y Take now
○ Vitamin D metabolite - (alfacalcidol, One-Alpha or calcitriol, Rocaltrol)	Start : M M / Y Y Stop: M M / Y Y Take now
○ Zoledronate (Aclasta, Zometa, zoledronic acid)	Start : M M / Y Y Stop: M M / Y Y Take now
○ Other Bone Treatments (specify): _____	Start : M M / Y Y Stop: M M / Y Y Take now

Name: _____ DOB: D D M M Y Y Y Y

Q5 Supplements For Your Bones

Do you take supplements of calcium with/without Vitamin D? ☐ Yes ☐ No

If **Yes**, what is the name of the supplement? _____

How often do you take your calcium supplement?	☐ Every day ☐ Most days ☐ Once or twice a week ☐ Occasionally
Do you take any of these supplements?	☐ Cod liver oil / fish oil ☐ Vitamin D ☐ Other(s): _____ _____ _____

Q6 Steroid Use

Have you ever had steroid treatment? ☐ Yes ☐ No

If **Yes,** what type(s) of treatment have you had? ***(Put a cross in all that apply)***

☐ Tablets ☐ Inhalers ☐ Creams ☐ Joint/muscle injections ☐ Injection into a vein

→ If you take **tablets** please indicate which type(s) and how often:

☐ Prednisolone ☐ Budesonide ☐ Hydrocortisone ☐ Dexamethasone

Daily dose ☐ Yes ☐ No

→ If **Yes**, please state the usual dose _____mg

"Booster" doses? ☐ Yes ☐ No

→ If **Yes**, how often do you have a booster?

☐ Less than once a year ☐ 2-3 times a year ☐ 4+ times a year

What is/was the reason for your steroid treatment? _____ _____

If you are taking steroids at the moment, how long have you taken them for? ☐☐ Years ☐☐ Months

If you are not taking treatment at the moment , how long ago was the last time you were treated with steroids? ☐☐ Years ☐☐ Months

Name: _____ DOB: D D M M Y Y Y Y

Q7 Treatment for Other Conditions

Are you being treated with any of the following medications which can affect bone health:

- ○ Anastrazole (Arimidex)
- ○ Exemestane (Aromasin)
- ○ Letrozole (Femara)
- ○ Goserelin (Zoladex injections)
- ○ Leuprorelin acetate (Prostap injections)
- ○ Treatment for stomach ulcer, indigestion or acid reflux – specify: _____
- ○ Phenytoin (Epanutin)
- ○ Phenobarbital
- ○ Carbamazepine (Tegretol)
- ○ Sodium valproate (Epilim)
- ○ Other treatment for epilepsy – specify: _____

Please write down all the medications you use. This includes tablets, medicines and injections. Include everything from your GP, from the hospital and those you buy yourself.
Instead of completing the table you can bring a copy of your prescription if you'd prefer but please also tell us about any medication that you buy yourself

Medication	Dose	What do you take this for?

Please continue on the back page if you require more space

Do you have any allergies or sensitivity to medication? ☐ Yes ☐ No

Is **Yes**, please give details _____

Name: _____ DOB: D D M M Y Y Y Y

Q8 About Your Lifestyle

How much milk do you have each day? (e.g. in drinks, cereal etc.)	☐ None	☐ Less than ½ pint
	☐ ½ to 1 pint	☐ More than 1 pint
How often do you eat dairy foods? (e.g. cheese, yoghurts)	☐ Never	☐ Less than once a week
	☐ Once/twice a week	☐ Most days/every day
In the summer how often do you spend 30 minutes out in the sunshine with your face and arms uncovered?	☐ Never	☐ Less than once a week
	☐ Once/twice a week	☐ Most days

Q9 Alcohol and Smoking

Do you ever drink alcohol? ☐ Yes ☐ No

If **Yes**, in a typical week on how many days do you have an alcoholic drink? *(put a cross in the number that applies)*

1 2 3 4 5 6 7

On each day that you drink, how many units*do you usually have? ☐☐ units

*one unit is ½ pint beer/lager, one small glass of wine, a small pub measure of spirits.

Do you smoke?

☐ Yes I smoke now

☐ No, I stopped within the past 5 years

☐ No, I stopped more than 5 years ago

☐ No, I have never smoked

If **Yes**, on average how many cigarettes do you smoke? ☐☐ per day

Q10 Falls

Have you had any falls in the last 6 months? ☐ Yes ☐ No

If **Yes**, how many? ☐ 1 ☐ 2-3 ☐ 4 or more

What caused your fall(s)? ☐ Felt dizzy/ lightheaded/ blacked out ☐ Other: _____

Name: _____ DOB: D D M M Y Y Y Y

Q11 Mobility

Do you use any assistive/ mobility aids?

☐ Yes ☐ No

If **Yes**, please indicate which:

	Please indicate how often		**Please indicate where**		
	Sometimes	Most or all of the time	Indoors Only	Outdoors Only	Indoors and Outdoors
☐ Wheelchair	☐	☐	☐	☐	☐
☐ Frame	☐	☐	☐	☐	☐
☐ Stick(s)	☐	☐	☐	☐	☐
☐ Crutches	☐	☐	☐	☐	☐
☐ Mobility Scooter / Motorised Wheelchair	☐	☐	☐	☐	☐

Q12 Exercise

Put a mark on the line below to indicate how active you are:

Immobile — Active every day, no formal exercise or sport — Daily strenuous exercise or sport

Q13 About Your Family

Do/did any of your close relatives *(parent, brother, sister, child)* have osteoporosis? ☐ Yes ☐ No ☐ Unsure

If any of your close relatives have had a broken hip please indicate who and when

Relationship to you (e.g. mother)	**Age at (first) hip fracture**
	☐☐ Years
	☐☐ Years
	☐☐ Years

Name: _____ DOB: D D M M Y Y Y Y

Q14 Women Only – Please complete this section

How old were you when your periods started? ☐☐ Years

Do you still have regular periods? (8 or more each year) ☐ Yes ☐ No

If **No**, how old were you when you had your last period? ☐☐ Years

Did your periods ever stop for more than 3 months? (except during pregnancy and at the menopause) ☐ Yes ☐ No

If **Yes**, please describe when and why: _____

_____

Have you ever used contraceptive injections (Depo Provera)? ☐ Yes ☐ No

If **Yes**, please indicate your approximate start and stop age: ☐☐ Start ☐☐ Stop ☐ Still using

Do you have menopausal symptoms now? ☐ Yes ☐ No

Have you ever taken Hormone Replacement Therapy (HRT)? ☐ Yes ☐ No

If **Yes,** please indicate your start and stop age: ☐☐ Start ☐☐ Stop ☐ Still using

Have you had a hysterectomy? ☐ Yes ☐ No

If **Yes**, how old were you? ☐☐ Years

Had your periods stopped before you had the hysterectomy? ☐ Yes ☐ No

Have you had either or both of your ovaries removed? ☐ Yes one ☐ Yes both ☐ No

If **Yes**, please indicate how old you were when they were removed: ☐☐ Age 1st ovary removed ☐☐ Age 2nd ovary removed

☐☐ Age if both were removed at the same time

Thank you for completing this questionnaire

Name: _____ DOB: D D M M Y Y Y Y

Continuation Sheet

Please use this page to give us any further information about any of the questions on the form.
Please include the question number with your answer.

Name: _____ DOB: D D M M Y Y Y Y

Please leave this box blank (office use only)

Appointment Date D D M M Y Y Y Y

Current Height ☐☐☐.☐ cm ☐ Feet ☐☐ Inches

Current Weight ☐☐☐.☐ kg ☐☐ Stone ☐☐ Pounds

LMP D D M M Y Y LMP Age ☐☐

Risk of pregnancy? ☐ Yes ☐ No

Comments: _____

_____

_____

NHS No ☐☐☐☐☐☐☐☐☐☐

STH No ☐☐☐☐☐☐

Date of previous measurements D D M M Y Y Y Y

Previous Height ☐☐☐.☐ cm ☐ Feet ☐☐ Inches

Previous Weight ☐☐☐.☐ kg ☐☐ Stone ☐☐ Pounds

Completed by: _____

4 Appendix D: Power Tables for Sample Size calculation

(Hess *et al.*, 2012)

4. Appendix D: Power Tables for Sample Size calculation

(Hess *et al.*, 2012)

Power tables for sensitivity

Prevalence	Performance Target	Expected Sensitivity	Precision	95% CI	Sample Size (Positive)	Total Sample Size
0.5	0.7	0.75	0.05	0.701 to 0.799	305	610
0.5	0.7	0.80	0.10	0.709 to 0.891	75	150
0.5	0.7	0.85	0.15	0.728 to 0.972	33	66
0.5	0.7	0.9	0.10	0.769 to 1.000	20	40
0.5	0.75	0.8	0.05	0.753 to 0.847	275	550
0.5	0.75	0.85	0.10	0.764 to 0.936	67	134
0.5	0.75	0.9	0.15	0.794 to 1.000	31	62

0.5	0.8	0.85	0.05	0.804 to 0.896	234	468
0.5	0.8	0.9	0.10	0.823 to 0.977	59	118

Power tables for specificity

Prevalence	**Performance Target**	**Expected Specificity**	**Precision**	**95% CI**	**Sample Size (Negative)**	**Total Sample Size**
0.5	0.5	0.55	0.05	0.5002 to 0.5998	384	768
0.5	0.5	0.6	0.10	0.502 to 0.698	96	192
0.5	0.5	0.65	0.15	0.506 to 0.794	42	84
0.5	0.5	0.7	0.20	0.517 to 0.883	24	48
0.5	0.5	0.75	0.25	0.531 to 0.969	15	30

0.5	0.5	0.8	0.30	0.564 to 1.000	11	22
0.5	0.6	0.65	0.05	0.601 to 0.699	362	724
0.5	0.6	0.7	0.10	0.605 to 0.795	89	178
0.5	0.6	0.75	0.15	0.614 to 0.886	39	78
0.5	0.6	0.8	0.20	0.633 to 0.967	22	44
0.5	0.7	0.75	0.05	0.701 to 0.799	305	610
0.5	0.7	0.80	0.10	0.709 to 0.891	75	150
0.5	0.7	0.85	0.15	0.728 to 0.972	33	66
0.5	0.7	0.9	0.20	0.769 to 1.031	20	40

0.5	0.75	0.8	0.05	0.753 to 0.847	275	550
0.5	0.75	0.85	0.10	0.764 to 0.936	67	134
0.5	0.75	0.9	0.15	0.794 to 1.000	31	62
0.5	0.8	0.85	0.05	0.804 to 0.896	234	468
0.5	0.8	0.9	0.10	0.823 to 0.977	59	118

5 Appendix E: Patient Information Sheet (PIS) V1.1 17NOV2015

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

The University Of Sheffield.

Academic Unit of Bone Metabolism

PARTICIPANT INFORMATION SHEET

South Yorkshire and North Derbyshire Musculoskeletal Biobank
Metabolic Bone Centre Collection
(non-fasting)

Protocol Reference: STH15691

Invitation

You are being invited to take part in a Musculoskeletal BioBank. Before you decide to take part it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Talk to others about the BioBank if you wish.

- **Part 1** tells you the purpose of this BioBank and what will happen to you if you take part.
- **Part 2** gives you more detailed information about the way the BioBank is being carried out.

Ask us if there is anything that is not clear, or if you would like more information. Take time to decide whether or not you wish to take part.

Part 1

What is the BioBank about?

As you may be aware a lot of research is currently being carried out worldwide into musculoskeletal diseases. Musculoskeletal diseases are those that affect the skeleton (bones) and the tissues around them, such as muscles and tendons. Examples of such diseases are osteoporosis, arthritis, Paget's disease and bone cancer. We would like to get a better understanding of these conditions so that we can develop new methods of diagnosis and treatment.

The Musculoskeletal BioBank is a resource for biological samples and clinical information for a number of research projects that study diseases of bone, joint and related soft tissues. Samples of blood and urine, and samples of diseased and normal tissues of patients are studied in order to gain knowledge of what causes musculoskeletal disease. Samples and clinical information, collected from patients looked after at the Sheffield Teaching Hospitals NHS Foundation Trust, are used for research carried out at the Academic Unit of Bone Metabolism and in other research centres in the UK and other countries. These studies are going on all of the time and require the collection and storage of tissue and blood samples over many years to provide a useful tissue resource.

Why have I been chosen to take part?

Patients who are being investigated or treated for a musculoskeletal disease at Sheffield Teaching Hospitals NHS Foundation Trust (which includes the Northern General Hospital, Royal Hallamshire

Hospital, and other hospitals), may be invited to take part. People who have previously participated in our research studies may also be asked to contribute to the BioBank.

Do I have to take part?

No. The choice to allow us to collect samples of your blood and urine for this project is entirely yours. If you decide to take part you can keep this information sheet and you will also be asked to sign a consent form releasing the sample for research use. You can still change your mind at any time and you don't have to explain why. Whatever you decide, it will not affect your treatment or any other part of the care you will receive in this hospital or any future care.

What will it involve if I decide to take part?

You will need to sign a consent form to allow us to take a sample of blood or urine.
The blood sample taken will be up to approximately 100mL (7 tablespoons). You may also be asked to provide a urine sample; the researcher will explain how this sample will be collected.

You will also be asked to complete some health questionnaires, and we will also collect data from your full hospital medical record to gather further information about your medical health. This information will be held on a restricted database within the BioBank and updated periodically after your sample donation by follow up reviews of your hospital medical records to gather up-to-date information about your medical health.

Will I receive any expenses or payments?

If you need to make an extra visit to take part in the BioBank research then you will be able to reclaim your expenses for travel to and from the hospital, or we can arrange taxi transport for you. However, we aim to arrange the research when you attend for your routine clinic visit.

What are the benefits and risks of taking part?

There are small risks from having blood samples taken. For most people, needle punctures for blood draws do not cause serious problems, however they may cause a bruise or a small amount of bleeding or pain at the needle site. Some people may feel faint. In very rare cases infection may occur.

There are no direct benefits to you, but the results of research using samples of blood and urine taken from you and others may help patients in the future. You are asked to donate your blood and urine freely for research and you will not receive a financial reward either now or in the future. Your samples will not be sold for profit to other researchers. Your samples may be used for research that may lead to the development of new drugs or therapies, which may eventually be marketed, and companies may sell these drugs for profit.

What if there is a problem?

Any complaint regarding the way you have been treated during the project will be addressed. If you have a complaint please firstly contact the research team on **0114 271 5240.** More detailed information is available in Part 2.

Will my taking part in the BioBank be kept confidential?

Yes, all the information about your participation in the BioBank will be kept confidential. Further details are included in Part 2.

This completes Part 1 of the information sheet.
If the information in Part 1 has interested you and you are considering taking part, please continue to read the additional information in Part 2 before making any decision.

Part 2

What will happen if I don't want to carry on with the BioBank?

You are free to withdraw your samples from the resource at any time, without saying why and we will destroy all of your samples and clinical records database. A log of your participation in the BioBank will be maintained for administration purposes. However if you change your mind a long time after donating the sample your blood or urine may already have been used for research. If you would like to withdraw consent for your samples to be used, please contact the research team on [telephone] or at the address given at the end of this leaflet. We will confirm in writing that any remaining samples stored in the BioBank have been destroyed.

What if there is a problem?

If you have a concern about any aspect of the BioBank, you should speak to the researchers who will do their best to answer your questions (contact the research team on 0114 271 5240).
Normal NHS complaints procedures will apply, in the unlikely event that something does go wrong and you are harmed during the research there are no special compensation arrangements. If you are harmed and this is due to someone's negligence then you may have grounds for legal action for compensation against Sheffield Teaching Hospitals NHS Foundation Trust but you may have to pay your legal costs. The normal National Health Service complaints procedures will still be available to you. In addition the University of Sheffield holds insurance against risk of claims against the University and its staff relating to clinical trials they design and undertake in their University employment.

Will my taking part in the BioBank be kept confidential?

If you consent to take part in the BioBank we would need to record some details of your medical history for the purpose of analysing the research results. We may also access your full medical records in the future for research purposes. This information will of course be strictly confidential and the people analysing your samples will not be able to identify you from the information they receive. You would also remain anonymous if we need to publish results, or share data or samples with other laboratories for research purposes.

What will happen to any samples that I give?

We will store the blood and urine samples indefinitely in the Sheffield Biorepository using a code, not your name, to tell us which sample is which. The information (data) about your general medical health will be kept in a similar database. Samples and data will be used by researchers for various projects, many of which are not yet known and will depend on the development of new research techniques in the future. However all research projects will be subject to approval by the BioBank Steering Committee and access to samples and data will be controlled. The Steering Committee consists of members of the public and independent researchers as well as researchers at the Academic Unit of Bone Metabolism.

Scientists may also perform genetic tests on your samples – the results of these tests may provide information on which genes cause arthritis, osteoporosis, Paget's disease and other musculoskeletal diseases. It is possible that future research carried out using your samples may not be restricted to the study of musculoskeletal diseases. Samples and information collected may be

transferred for research and analysis to investigators within and outside the European Economic Area, this may include commercial companies. Some countries outside Europe may not have laws which protect your privacy to the same extent as the Data Protection Act in the UK or European Law. However we will take all reasonable steps to protect your privacy; all samples will be anonymised and no personal information will be transferred outside of the Academic Unit of Bone Metabolism in Sheffield.

What will happen to the results of the research?

The results of investigations that we obtain will add to our overall understanding of musculoskeletal and other diseases. This information may help in the design of new ways to diagnose or treat in the future. The results we get will not change the treatment that you as an individual will get, and so you would not receive any of the individual research results that we find. However, the overall findings of our studies will be published in medical journals.

Very occasionally we may find that after testing your sample during our research, something unusual or abnormal is discovered which may affect your treatment. If this happens we would like to give this information back to the doctor in charge of your care. We ask your permission to provide this information to your General Practitioner or hospital doctor.

Who is organising and funding the BioBank?

The BioBank is being organised by the Academic Unit of Bone Metabolism which is jointly run and funded by the University of Sheffield and the Sheffield Teaching Hospitals NHS Foundation Trust. The lead researcher is Professor Mark Wilkinson who is Professor of Orthopaedics at the University of Sheffield and a Consultant Surgeon.

Who has reviewed the project?

This BioBank was reviewed by the Oxfordshire Research Ethics Committee. Research Ethics Committees are an independent group of people who review research projects in NHS patients, and act to protect the safety, dignity and well-being of research participants.

Contact for Further Information

If you would like further information or you have concerns about this project you can either:
Ask the person who has provided the leaflet to you or write to Professor M Wilkinson at the Academic Unit of Bone Metabolism, Sorby Wing, Northern General Hospital, Herries Road, Sheffield, S5 7AU or contact the research team on 0114 271 5240.

Thank you very much for taking the time to read this information sheet and for considering taking part in the BioBank

If you do decide to take part you will be given a copy of this information sheet and the signed consent form to take home with you.

6 Appendix F: Informed Consent Form (ICF) V1.0 31JAN2015

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

The University Of Sheffield.

Patient study ID: _____ Protocol Number: STH15691

CONSENT FORM

MBC patients

Title of Project: **South Yorkshire and North Derbyshire Musculoskeletal Biobank**
Chief Investigator: **Professor Mark Wilkinson**

Please place your initials in each of the boxes:

1. I confirm that I have read, understood and have had time to consider the information sheet dated …….. (version ……..) and have been given a copy to keep. I have had the opportunity to ask questions about this project and have had these answered satisfactorily. ☐
2. I agree to give samples of blood and urine for research as detailed in the patient information sheet. ☐
3. I agree that clinical data from my medical records and imaging will be stored by the South Yorkshire and North Derbyshire Musculoskeletal Biobank and this information may be updated periodically from the current hospital records. ☐
4. I understand how the sample(s) will be collected, that giving samples for this research is voluntary and that I am free to withdraw my approval for use of the sample(s) at any time without giving a reason and without my medical treatment or legal rights being affected. ☐
5. I understand and agree that parts of my medical information or samples may be passed to other organisations involved in the research on the understanding that my personal patient confidentiality will be maintained. This may include countries outside the European Economic Area. ☐
6. I understand that I will not benefit financially if this research leads to the development of a new treatment or medical test. ☐
7. I agree that the sample(s) I have given and the information gathered about me can be stored by the South Yorkshire and North Derbyshire Musculoskeletal Biobank for possible use in future projects, as described in the attached information sheet. I understand that some of these projects may be carried out by researchers other than the University of Sheffield, including researchers working for commercial companies. ☐
8. I agree that the sample(s) I give may be used for genetic research aimed at understanding which genes are related to bone health and disease. I understand that the results of these investigations are unlikely to change my future treatment in any way. ☐
9. I agree that findings of clinical significance may be fed back to the clinician in charge of my care. ☐

……	……	……
Name of patient	Date	Signature
……	……	……
Name of person taking consent (if different from Researcher)	Date	Signature
……	……	……
Researcher	Date	Signature

When completed: 1 copy to patient; 1 copy filed in medical notes; original for researcher.

Abbreviations

25OHD	25 hydroxyvitamin D
ALP	Alkaline Phosphatase
AUBM	Academic Unit of Bone Metabolism
AUC	Area Under the Curve
BAP	Bone Alkaline Phosphatase
BMD	Bone Mineral Density
BMI	Body Mass Index
BTM(s)	Bone Turnover Marker(s)
CI	Confidence Interval
CKD	Chronic Kidney Disease
CKD-MBD	CKD- Metabolic Bone Disease
CLSI	Clinical & Laboratory Standards Institute
CTX-MMP	CTX- Matrix Metalloproteinases
CVA	Coefficient of Variation interassay
CVI	Coefficient of Variation interindividual
DM	Diabetes Mellitus
DXA	Dual-energy x-ray absorptiometry
eGFR	Estimated Glomerular Filtration Rate
FRAS	Fracture Risk Assessment Service
FSH	Follicle Stimulating Hormone
GCIO	Glucocorticoid Induced Osteoporosis
GM	Geometric Mean
GnRH	Gonadotrophin Releasing Hormone
HRT	Hormone Replacement Therapy
ICF	Informed Consent Form
ICTP	C-terminal telopeptide of collagen I
IDS	Immunodiagnostic Systems
IFCC	International Federation of Clinical Chemistry
IOF	International Osteoporosis Federation

IQR	Interquartile Range
ISCD	International Society for Clinical Densitometry
IV	Intravenous
KDIGO	Kidney Disease Improving Global Outcomes
LSC	Least Significant Change
MBC	Metabolic Bone Centre
NMR	Nuclear Magnetic Resonance
NNT	Numbers Needed to Treat
NTX	N-terminal Crosslinked Telopeptides of Type 1 Collagen
OC	Osteocalcin
PICP	Procollagen I C-terminal Extension Peptide
PINP	Procollagen I N-terminal Extension Peptide
PIS	Patient Information Sheet
PTH	Parathyroid Hormone
QC	Quality Control
RANKL	RANK Ligand
RCT	Randomised Controlled Trial
REC	Research Ethics Committee
ROC	Receiver Operator Curve
SC	Subcutaneous
sCTX	serum C-terminal Crosslinked Telopeptides of Type 1 Collagen
SERM	Selective Estrogen Receptor Modulator
STH	Sheffield Teaching Hospitals
SYNDMB	South Yorkshire, North Derbyshire, Musculoskeletal Biobank
TNSALP	Tissue Non-Specific Alkaline Phosphatase
TRAP	Tartrate Resistant Acid Phosphatase
TSH	Thyroid Stimulating Hormone
uCTX	urine C-terminal Crosslinked Telopeptides of Type 1 Collagen
WHO	World Health Organisation

Bibliography

Andreasen, C. M., Delaisse, J. M., Cj van der Eerden, B., van Leeuwen, J. P., Ding, M. and Andersen, T. L. (2017) 'Understanding Age-Induced Cortical Porosity in Women: The Accumulation and Coalescence of Eroded Cavities Upon Existing Intracortical Canals Is the Main Contributor', *J Bone Miner Res*.

Bauer, D., Black, D., Bouxsein, M., Lui, L. Y., Cauley, J., de Papp, A., Grauer, A., Khosla, S., McCulloch, C., Eastell, R. and Foundation for the National Institutes of Health Bone Qua, T. (2018) 'Treatment-Related Changes in Bone Turnover and Fracture Risk Reduction in Clinical Trials of Anti-Resorptive Drugs: A Meta-Regression'.

Bauer, D. C., Garnero, P., Hochberg, M. C., Santora, A., Delmas, P., Ewing, S. K. and Black, D. M. (2006) 'Pretreatment Levels of Bone Turnover and the Antifracture Efficacy of Alendronate: The Fracture Intervention Trial', *Journal of Bone and Mineral Research,* 21(2), pp. 292-299.

Bergmann, P., Body, J. J., Boonen, S., Boutsen, Y., Devogelaer, J. P., Goemaere, S., Kaufman, J. M., Reginster, J. Y. and Gangji, V. (2009) 'Evidence-based guidelines for the use of biochemical markers of bone turnover in the selection and monitoring of bisphosphonate treatment in osteoporosis: a consensus document of the Belgian Bone Club', *International Journal of Clinical Practice,* 63(1), pp. 19-26.

Bhattoa, H. P., Cavalier, E., Eastell, R., Heijboer, A. C., Jørgensen, N. R., Makris, K., Ulmer, C. Z., Kanis, J. A., Cooper, C., Silverman, S. L. and Vasikaran, S. D. (2021) 'Analytical considerations and plans to standardize or harmonize assays for the reference bone turnover markers PINP and β-CTX in blood', *Clin Chim Acta,* 515, pp. 16-20.

Bianchi, M. L. (2015) 'Hypophosphatasia: an overview of the disease and its treatment', *Osteoporos Int,* 26(12), pp. 2743-57.

Bishop, N. (2016) 'Bone Material Properties in Osteogenesis Imperfecta', *Journal of Bone and Mineral Research,* 31(4), pp. 699-708.

Bjarnason, N. H., Sarkar, S., Duong, T., Mitlak, B., Delmas, P. D. and Christiansen, C. (2001) 'Six and Twelve Month Changes in Bone Turnover are Related to Reduction in Vertebral Fracture Risk During 3 Years of Raloxifene Treatment in Postmenopausal Osteoporosis', *With other Metabolic Bone Diseases,* 12(11), pp. 922-930.

Black, D. M., Cummings, S. R., Karpf, D. B., Cauley, J. A., Thompson, D. E., Nevitt, M. C., Bauer, D. C., Genant, H. K., Haskell, W. L., Marcus, R., Ott, S. M., Torner, J. C., Quandt, S. A., Reiss, T. F. and Ensrud, K. E. (1996) 'Randomised trial of effect of alendronate on risk of fracture in women with existing vertebral fractures', *The Lancet,* 348(9041), pp. 1535-1541.

Black, D. M., Delmas, P. D., Eastell, R., Reid, I. R., Boonen, S., Cauley, J. A., Cosman, F., Lakatos, P., Leung, P. C., Man, Z., Mautalen, C., Mesenbrink, P., Hu, H., Caminis, J., Tong, K., Rosario-Jansen, T., Krasnow, J., Hue, T. F., Sellmeyer, D., Eriksen, E. F. and Cummings, S. R. (2007) 'Once- Yearly Zoledronic Acid for Treatment of Postmenopausal Osteoporosis', *The New England Journal of Medicine,* 356(18), pp. 1809-1822.

Black, D. M., Reid, I. R., Cauley, J. A., Cosman, F., Leung, P. C., Lakatos, P., Lippuner, K., Cummings, S. R., Hue, T. F., Mukhopadhyay, A., Tan, M., Aftring, R. P. and Eastell, R. (2015) 'The Effect of 6 versus 9 Years of Zoledronic Acid Treatment in Osteoporosis: A Randomized Second Extension to the HORIZON-Pivotal Fracture Trial (PFT): EFFECT OF 6 VERSUS 9 YEARS OF ZOLEDRONIC ACID TREATMENT', *Journal of bone and mineral research,* 30(5), pp. 934-944.

Black, D. M., Thompson, D. E., Bauer, D. C., Ensrud, K., Musliner, T., Hochberg, M. C., Nevitt, M. C., Suryawanshi, S. and Cummings, S. R. (2000) 'Fracture Risk Reduction with Alendronate in Women with Osteoporosis: The Fracture Intervention Trial', *The Journal of Clinical Endocrinology & Metabolism,* 85(11), pp. 4118-4124.
Bone, H. G., Chapurlat, R., Brandi, M.-L., Brown, J. P., Czerwiński, E., Krieg, M.-A., Mellström, D., Radominski, S. C., Reginster, J.-Y., Resch, H., Ivorra, J. A. R., Roux, C., Vittinghoff, E., Daizadeh, N. S., Wang, A., Bradley, M. N., Franchimont, N., Geller, M. L., Wagman, R. B., Cummings, S. R. and Papapoulos, S. (2013) 'The Effect of Three or Six Years of Denosumab Exposure in Women With Postmenopausal Osteoporosis: Results From the FREEDOM Extension', *J Clin Endocrinol Metab,* 98(11), pp. 4483-4492.
Bonewald, L. F. (2011) 'The amazing osteocyte', *Journal of Bone and Mineral Research,* 26(2), pp. 229-238.
Booth, S. L., Centi, A., Smith, S. R. and Gundberg, C. (2013) 'The role of osteocalcin in human glucose metabolism: marker or mediator?', *Nat Rev Endocrinol,* 9(1), pp. 43-55.
Borgström, F., Karlsson, L., Ortsäter, G., Norton, N., Halbout, P., Cooper, C., Lorentzon, M., McCloskey, E. V., Harvey, N. C., Javaid, M. K., Kanis, J. A., Reginster, J.-Y. and Ferrari, S. (2020) 'Fragility fractures in Europe: burden, management and opportunities', *Arch Osteoporos,* 15(1), pp. 59-59.
Bouxsein, M. L., Eastell, R., Lui, L. Y., Wu, L. A., de Papp, A. E., Grauer, A., Marin, F., Cauley, J. A., Bauer, D. C. and Black, D. M. (2019) 'Change in Bone Density and Reduction in Fracture Risk: A Meta-Regression of Published Trials', *J Bone Miner Res,* 34(4), pp. 632-642.
Brady, J. J., Crowley, R. K., Murray, B. F., Kilbane, M. T., O'Keane, M. and McKenna, M. J. (2014) 'Limited utility of tartrate-resistant acid phosphatase isoform 5b in assessing response to therapy in osteoporosis', *Ir J Med Sci,* 183(1), pp. 47-52.
Buckley, L., Guyatt, G., Fink, H. A., Cannon, M., Grossman, J., Hansen, K. E., Humphrey, M. B., Lane, N. E., Magrey, M., Miller, M., Morrison, L., Rao, M., Robinson, A. B., Saha, S., Wolver, S., Bannuru, R. R., Vaysbrot, E., Osani, M., Turgunbaev, M., Miller, A. S. and McAlindon, T. (2017) '2017 American College of Rheumatology Guideline for the Prevention and Treatment of Glucocorticoid-Induced Osteoporosis', *Arthritis & Rheumatology,* 69(8), pp. 1521-1537.
Burch, J., Rice, S., Yang, H., Neilson, A., Stirk, L., Francis, R., Holloway, P., Selby, P. and Craig, D. (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 technology assessment (Winchester, England),* 18(11), pp. 1.
Burshell, A. L., Möricke, R., Correa-Rotter, R., Chen, P., Warner, M. R., Dalsky, G. P., Taylor, K. A. and Krege, J. H. (2010) 'Correlations between biochemical markers of bone turnover and bone density responses in patients with glucocorticoid-induced osteoporosis treated with teriparatide or alendronate', *Bone,* 46(4), pp. 935-939.
CALVO, M. S., Division of Endocrinology, M., and Internal Medicine (M.S.C., R.E.), Sections of Biostatistics (K.P.O., E.J.B.) and Clinical Chemistry (M.F.B.), Mayo Clinic, Mayo Foundation and Mayo Medical School Rochester, Minnesota 55905, EASTELL, R., Division of Endocrinology, M., and Internal Medicine (M.S.C., R.E.), Sections of Biostatistics (K.P.O., E.J.B.) and Clinical Chemistry (M.F.B.), Mayo Clinic, Mayo Foundation and Mayo Medical School Rochester, Minnesota 55905, OFFORD, K. P., Division of Endocrinology, M., and Internal Medicine (M.S.C., R.E.), Sections of Biostatistics (K.P.O., E.J.B.) and Clinical Chemistry (M.F.B.), Mayo Clinic, Mayo Foundation and Mayo Medical School Rochester,

Minnesota 55905, BERGSTRALH, E. J., Division of Endocrinology, M., and Internal Medicine (M.S.C., R.E.), Sections of Biostatistics (K.P.O., E.J.B.) and Clinical Chemistry (M.F.B.), Mayo Clinic, Mayo Foundation and Mayo Medical School Rochester, Minnesota 55905, BURRITT, M. F. and Division of Endocrinology, M., and Internal Medicine (M.S.C., R.E.), Sections of Biostatistics (K.P.O., E.J.B.) and Clinical Chemistry (M.F.B.), Mayo Clinic, Mayo Foundation and Mayo Medical School Rochester, Minnesota 55905 (1991) 'Circadian Variation in Ionized Calcium and Intact Parathyroid Hormone: Evidence for Sex Differences in Calcium Homeostasis', *The Journal of Clinical Endocrinology & Metabolism,* 72(1), pp. 69-76.
Calvo, M. S., Eyre, D. R. and Gundberg, C. M. (1996) 'Molecular basis and clinical application of biological markers of bone turnover', *Endocr Rev,* 17(4), pp. 333-368.
Carter, D. R., Stanford University, C., Beaupré, G. S. and System, V. P. A. H. C. (2000) *Skeletal Function and Form: Mechanobiology of Skeletal Development, Aging, and Regeneration.* Cambridge University Press.
Cavalier, R, E., N, R. J., K, M., S, T., S, V., JA, K., C, C., H, P. and HA, M. (2019) 'A multicenter study to evaluate harmonization of assays for N-terminal propeptide of type I procollagen (PINP): a report from the IFCC-IOF Joint Committee for Bone Metabolism', *Clinical chemistry and laboratory medicine,* 57(10).
Cavalier, E., Lukas, P., Bottani, M., Aarsand, A. K., Ceriotti, F., Coskun, A., Diaz-Garzon, J., Fernandez-Calle, P., Guerra, E., Locatelli, M., Sandberg, S., Carobene, A. and Metab, I. O. F. I. C. B. (2020) 'European Biological Variation Study (EuBIVAS): within- and between-subject biological variation estimates of β-isomerized C-terminal telopeptide of type I collagen (β-CTX), N-terminal propeptide of type I collagen (PINP), osteocalcin, intact fibroblast growth factor 23 and uncarboxylated-unphosphorylated matrix-Gla protein—a cooperation between the EFLM Working Group on Biological Variation and the International Osteoporosis Foundation-International Federation of Clinical Chemistry Committee on', *Osteoporos Int,* 31(8), pp. 1461-1470.
Chapurlat, R. D., Garnero, P., Brárt, G., Meunier, P. J. and Delmas, P. D. (2000) 'Serum type I collagen breakdown product (serum CTX) predicts hip fracture risk in elderly women:: The EPIDOS study', *Bone,* 27(2), pp. 283-286.
Chavassieux, P., Inserm UMR 1033 and Université de Lyon, L., France, Portero-Muzy, N., Inserm UMR 1033 and Université de Lyon, L., France, Roux, J.-P., Inserm UMR 1033 and Université de Lyon, L., France, Garnero, P., Inserm UMR 1033 and Université de Lyon, L., France, Chapurlat, R. and Inserm UMR 1033 and Université de Lyon, L., France (2017) 'Are Biochemical Markers of Bone Turnover Representative of Bone Histomorphometry in 370 Postmenopausal Women?', *The Journal of Clinical Endocrinology & Metabolism,* 100(12), pp. 4662-4668.
Chubb, S. A., Byrnes, E., Manning, L., Beilby, J. P., Ebeling, P. R., Vasikaran, S. D., Golledge, J., Flicker, L. and Yeap, B. B. (2015) 'Reference intervals for bone turnover markers and their association with incident hip fractures in older men: the Health in Men study', *J Clin Endocrinol Metab,* 100(1), pp. 90-9.
Chubb, S. A. P., Byrnes, E., Manning, L., Golledge, J., Ebeling, P. R., Flicker, L., Yeap, B. B. and Vasikaran, S. D. (2017) 'Bone turnover markers: Defining a therapeutic target', *Clinical Biochemistry,* 50(3), pp. 162-163.
Chubb, S. A. P., Mandelt, C. and Vasikaran, S. (2016) 'Comparison of clinical cut-points and treatment targets for urine NTX and plasma βCTX-I in osteoporosis', *Clinical Biochemistry,* 49(7), pp. 529-533.

Clowes, A. J., Peel, F. A. N. and Eastell, F. A. R. (2004) 'The Impact of Monitoring on Adherence and Persistence with Antiresorptive Treatment for Postmenopausal Osteoporosis: A Randomized Controlled Trial', *The Journal of Clinical Endocrinology & Metabolism,* 89(3), pp. 1117-1123.
CLSI (2004) 'EP05-A3: Evaluation of Precision of Quantitative Measurement Procedures; Approved Guidelineâ??Third Edition - ep05a3_sample.pdf', *CLSI Wayne PA*.
CLSI 2008. Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory, 3rd Edition EP28-A3c. Wayne, PA:Clinical and Laboratory Standards Institute.
Cooper, M. S., Blumsohn, A., Goddard, P. E., Bartlett, W. A., Shackleton, C. H., Eastell, R., Hewison, M. and Stewart, P. M. (2003) '11beta- hydroxysteroid dehydrogenase type 1 activity predicts the effects of glucocorticoids on bone', *The Journal of clinical endocrinology and metabolism,* 88(8), pp. 3874.
Cramer, J. A., Gold, D. T., Silverman, S. L. and Lewiecki, E. M. (2007) 'A systematic review of persistence and compliance with bisphosphonates for osteoporosis', *Osteoporos Int,* 18(8), pp. 1023-1031.
Crockett, J. C., Rogers, M. J., Coxon, F. P., Hocking, L. J. and Helfrich, M. H. (2011) 'Bone remodelling at a glance', *Journal of Cell Science,* 124(7), pp. 991-998.
Cui, L., Houston, D. A., Farquharson, C. and MacRae, V. E. (2016) 'Characterisation of matrix vesicles in skeletal and soft tissue mineralisation', *Bone,* 87, pp. 147-158.
Cummings, Sr., San Martin, J., McClung, M., Siris, E., Eastell, R., Reid, I. R., Delmas, P., Zoog, H., Austin, M., Wang, A., Kutilek, S., Adami, S., Zanchetta, J., Libanati, C., Siddhanti, S. and Christiansen, C. (2009) 'Denosumab for Prevention of Fractures in Postmenopausal Women with Osteoporosis', *N. Engl. J. Med.,* 361(8), pp. 756-765.
Cummings, S. R. (2017) 'Vertebral Fractures After Discontinuation of Denosumab: A Post Hoc Analysis of the Randomized Placebo-Controlled FREEDOM Trial and Its Extension', *Journal of bone and mineral research : JBMR.,* 33(2), pp. 190-198.
Cummings, S. R., Black, D. M., Thompson, D. E., Applegate, W. B., Barrett-Connor, E., Musliner, T. A., Palermo, L., Prineas, R., Rubin, S. M., Scott, J. C., Vogt, T., Wallace, R., Yates, A. J., Lacroix, A. Z. and for the Fracture Intervention Trial Research, G. (1998) 'Effect of Alendronate on Risk of Fracture in Women With Low Bone Density but Without Vertebral Fractures: Results From the Fracture Intervention Trial', *JAMA,* 280(24), pp. 2077-2082.
Davina, J. J., Priyadarssini, M., Rajappa, M., Parameswaran, S., Sahoo, J., Mohan Raj, P. S., Revathy, G., Palanivel, C. and Marella, M. G. (2017) 'Assessment of bone turnover markers to predict mineral and bone disorder in men with pre-dialysis non-diabetic chronic kidney disease', *Clin Chim Acta,* 469, pp. 195-200.
De castro machado, A. B., Hannon, R. and Eastell, R. (1999) 'Monitoring Alendronate Therapy for Osteoporosis', *J Bone Miner Res,* 14(4), pp. 602-608.
Delmas, P. D., Eastell, R., Garnero, P., Seibel, M. J. and Stepan, J. (2000) 'The Use of Biochemical Markers of Bone Turnover in Osteoporosis', *Osteoporosis International,* 11(6), pp. S2-S17.
Delmas, P. D., Munoz, F., Black, D. M., Cosman, F., Boonen, S., Watts, N. B., Kendler, D., Eriksen, E. F., Mesenbrink, P. G. and Eastell, R. (2009) 'Effects of Yearly Zoledronic Acid 5 mg on Bone Turnover Markers and Relation of PINP With Fracture Reduction in Postmenopausal Women With Osteoporosis', *J Bone Miner Res,* 24(9), pp. 1544-1551.
Delmas, P. D., Vrijens, B., Eastell, R., Roux, C., Pols, H. A. P., Ringe, J. D., Grauer, A., Cahall, D. and Watts, N. B. (2007) 'Effect of Monitoring Bone Turnover Markers on Persistence with

Risedronate Treatment of Postmenopausal Osteoporosis', *The Journal of Clinical Endocrinology & Metabolism,* 92(4), pp. 1296-1304.
Diez-Perez, A., Naylor, K. E., Abrahamsen, B., Agnusdei, D., Brandi, M. L., Cooper, C., Dennison, E., Eriksen, E. F., Gold, D. T., Guañabens, N., Hadji, P., Hiligsmann, M., Horne, R., Josse, R., Kanis, J., Obermayer-Pietsch, B., Prieto-Alhambra, D., Reginster, J.-Y., Rizzoli, R., Silverman, S., Zillikens, C. and Eastell, R. (2017) 'International Osteoporosis Foundation and European Calcified Tissue Society Working Group. Recommendations for the screening of adherence to oral bisphosphonates', *Osteoporos Int,* 28(3), pp. 767-774.
Dijk, F. S. v., Cobben, J. M., Kariminejad, A., Maugeri, A., Nikkels, P. G. J., Rijn, R. R. v. and Pals, G. (2011) 'Osteogenesis Imperfecta: A Review with Clinical Examples'.
Ducy, P. 2011. The role of osteocalcin in the endocrine cross- talk between bone remodelling and energy metabolism. *Clinical and Experimental Diabetes and Metabolism.* Berlin/Heidelberg.
Eastell, R. (2007) 'Aromatase inhibitors and bone', *Journal of Steroid Biochemistry and Molecular Biology,* 106(1), pp. 157-161.
Eastell, R., Barton, I., Hannon, R., Chines, A., Garnero, P. and Delmas, P. (2003) 'Relationship of Early Changes in Bone Resorption to the Reduction in Fracture Risk With Risedronate', *Journal of Bone and Mineral Research,* 18(6), pp. 1051-1056.
Eastell, R., Black, D. M., Lui, L., Chines, A., Marin, F., Khosla, S., Papp, A. E., Cauley, J. A., Mitlak, B., McCulloch, C. E., Vittinghoff, E. and Bauer, D. C. (2020) 'Treatment-related changes in bone turnover and fracture risk reduction in clinical trials of antiresorptive drugs: proportion of treatment effect explained'.
Eastell, R., Chen, P., Saag, K. G., Burshell, A. L., Wong, M., Warner, M. R. and Krege, J. H. (2010) 'Bone formation markers in patients with glucocorticoid-induced osteoporosis treated with teriparatide or alendronate', *Bone,* 46(4), pp. 929-934.
Eastell, R., Garnero, P., Audebert, C. and Cahall, D. L. (2012) 'Reference intervals of bone turnover markers in healthy premenopausal women: Results from a cross-sectional European study', *Bone,* 50(5), pp. 1141-1147.
Eastell, R., Mosekilde, L., Hodgson, S. F. and Riggs, B. L. (1990) 'Proportion of human vertebral body bone that is cancellous', *J Bone Miner Res,* 5(12), pp. 1237-41.
Eastell, R., Nickelsen, T., Marin, F., Barker, C., Hadji, P., Farrerons, J., Audran, M., Boonen, S., Brixen, K., Gomes, J. M., Obermayer-Pietsch, B., Avramidis, A., Sigurdsson, G. and Glüer, C. C. (2009) 'Sequential Treatment of Severe Postmenopausal Osteoporosis After Teriparatide: Final Results of the Randomized, Controlled European Study of Forsteo (EUROFORS', *Journal of Bone and Mineral Research,* 24(4), pp. 726-736.
Eastell, R., O'Neill, T. W., Hofbauer, L. C., Langdahl, B., Reid, I. R., Gold, D. T. and Cummings, S. R. (2016) 'Postmenopausal osteoporosis', *Nature Reviews Disease Primers,* 2, pp. 16069.
Eastell, R., Pigott, T., Gossiel, F., Naylor, K. E., Walsh, J. S. and Peel, N. F. (2017) 'Management of Endocrine Disease: Modern spectrum of bone turnover markers: are they clinically useful?'.
Eastell, R., Pigott, T., Gossiel, F., Naylor, K. E., Walsh, J. S. and Peel, N. F. A. (2018) 'DIAGNOSIS OF ENDOCRINE DISEASE: Bone turnover markers: are they clinically useful?', *Eur J Endocrinol,* 178(1), pp. R19-R31.
Eastell, R. and Szulc, P. (2017) 'Use of bone turnover markers in postmenopausal osteoporosis', *Lancet Diabetes Endocrinol,* 5(11), pp. 908-923.

Ebeling, P. R., Atley, L. M., Guthrie, J. R., Burger, H. G., Dennerstein, L., Hopper, J. L. and Wark, J. D. (1996) 'Bone turnover markers and bone density across the menopausal transition', *J Clin Endocrinol Metab,* 81(9), pp. 3366-3371.
Eekman, D. A., Bultink, I. E., Heijboer, A. C., Dijkmans, B. A. and Lems, W. F. (2011) 'Bone turnover is adequately suppressed in osteoporotic patients treated with bisphosphonates in daily practice', *BMC Musculoskeletal Disorders,* 12(1), pp. 167.
Ferrar, L., van der Hee, R. M., Berry, M., Watson, C., Miret, S., Wilkinson, J., Bradburn, M. and Eastell, R. (2011) 'Effects of calcium-fortified ice cream on markers of bone health', *Osteoporos Int,* 22(10), pp. 2721-2731.
Fink, E., Cormier, C., Steinmetz, P., Kindermans, C., Le Bouc, Y. and Souberbielle, J. C. (2000) 'Differences in the Capacity of Several Biochemical Bone Markers to Assess High Bone Turnover in Early Menopause and Response to Alendronate Therapy', *With other Metabolic Bone Diseases,* 11(4), pp. 295-303.
Garnero, P., Darte, C. and Delmas, P. D. (1999) 'A model to monitor the efficacy of alendronate treatment in women with osteoporosis using a biochemical marker of bone turnover', *Bone,* 24(6), pp. 603-609.
Glover, S. J., Eastell, R., McCloskey, E. V., Rogers, A., Garnero, P., Lowery, J., Belleli, R., Wright, T. M. and John, M. R. (2009) 'Rapid and robust response of biochemical markers of bone formation to teriparatide therapy', *Bone,* 45(6), pp. 1053-8.
Glover, S. J., Garnero, P., Naylor, K., Rogers, A. and Eastell, R. (2008) 'Establishing a reference range for bone turnover markers in young, healthy women', *Bone,* 42(4), pp. 623-630.
Gossiel, F., Altaher, H., Reid, D. M., Roux, C., Felsenberg, D., Glüer, C.-C. and Eastell, R. (2018) 'Bone turnover markers after the menopause: T-score approach', *Bone,* 111, pp. 44-48.
Gossiel, F., Hannon, R. A. and Eastell, R. (2009) 'Urinary excretion of C-telopeptide isomers in premenopausal and postmenopausal women', *Bone (New York, N.Y.),* 44, pp. S327-S327.
Greendale, G. A., Sowers, M., Han, W., Huang, M. H., Finkelstein, J. S., Crandall, C. J., Lee, J. S. and Karlamangla, A. S. (2012) 'Bone mineral density loss in relation to the final menstrual period in a multiethnic cohort: Results from the Study of Women's Health Across the Nation (SWAN)', *J Bone Miner Res,* 27(1), pp. 111-118.
Greenspan, L. S., Resnick, M. N. and Parker, A. R. (2005) 'Early Changes in Biochemical Markers of Bone Turnover Are Associated with Long-Term Changes in Bone Mineral Density in Elderly Women on Alendronate, Hormone Replacement Therapy, or Combination Therapy: A Three-Year, Double-Blind, Placebo-Controlled, Randomized Clinical Trial', *The Journal of Clinical Endocrinology & Metabolism,* 90(5), pp. 2762-2767.
Grey, A., Bolland, M., Wattie, D., Horne, A., Gamble, G. and Reid, I. R. (2010) 'Prolonged antiresorptive activity of zoledronate: A randomized, controlled trial', *J Bone Miner Res,* 25(10), pp. 2251-2255.
Hannemann, A., Friedrich, N., Spielhagen, C., Rettig, R., Ittermann, T., Nauck, M. and Wallaschofski, H. (2013) 'Reference intervals for serum osteocalcin concentrations in adult men and women from the study of health in Pomerania', *BMC Endocr Disord,* 13, pp. 11.
Hannon, R., Department of Human Metabolism and Clinical Biochemistry, U. o. S., Sheffield, United Kingdom, to:, A. r. r., Blumsohn, A., Department of Human Metabolism and Clinical Biochemistry, U. o. S., Sheffield, United Kingdom, Naylor, K., Department of Human Metabolism and Clinical Biochemistry, U. o. S., Sheffield, United Kingdom, Eastell, R. and Department of Human Metabolism and Clinical Biochemistry, U. o. S., Sheffield, United

Kingdom (1998) 'Response of Biochemical Markers of Bone Turnover to Hormone Replacement Therapy: Impact of Biological Variability', *Journal of Bone and Mineral Research,* 13(7), pp. 1124-1133.
Harris, S. T., Watts, N. B., Genant, H. K., McKeever, C. D., Hangartner, T., Keller, M., Chesnut Iii, C. H., Brown, J., Eriksen, E. F., Hoseyni, M. S., Axelrod, D. W., Miller, P. D. and for the Vertebral Efficacy With Risedronate Therapy Study, G. (1999) 'Effects of Risedronate Treatment on Vertebral and Nonvertebral Fractures in Women With Postmenopausal Osteoporosis: A Randomized Controlled Trial', *JAMA,* 282(14), pp. 1344-1352.
Henny, J. (2009) 'The IFCC recommendations for determining reference intervals: strengths and limitations', *J Lab Med,* 33((2)), pp. 45-51.
Hess, A. S., Shardell, M., Johnson, J. K., Thom, K. A., Strassle, P., Netzer, G. and Harris, A. D. (2012) 'Methods and recommendations for evaluating and reporting a new diagnostic test', *European Journal of Clinical Microbiology & Infectious Diseases,* 31(9), pp. 2111-2116.
Ingle, B. M., Hay, S. M., Bottjer, H. M. and Eastell, R. (1999) 'Changes in Bone Mass and Bone Turnover Following Distal Forearm Fracture', *Osteoporos Int,* 10(5), pp. 399-407.
Author (2017) *ISCD Bone Densitometry Precision Calculating Tool - ISCD*. Available at: https://iscd.org/learn/resources/calculators/precision-calculator/.
Ivaska, K. K., Gerdhem, P., Åkesson, K., Garnero, P. and Obrant, K. J. (2007) 'Effect of Fracture on Bone Turnover Markers: A Longitudinal Study Comparing Marker Levels Before and After Injury in 113 Elderly Women', *J Bone Miner Res,* 22(8), pp. 1155-1164.
Janghorbani, M., Van Dam, R. M., Willett, W. C. and Hu, F. B. (2007) 'Systematic review of type 1 and type 2 diabetes mellitus and risk of fracture', *Am J Epidemiol,* 166(5), pp. 495-505.
Jenkins, N., Black, M., Paul, E., Pasco, J. A., Kotowicz, M. A. and Schneider, H. G. (2013) 'Age-related reference intervals for bone turnover markers from an Australian reference population', *Bone,* 55(2), pp. 271-6.
Kanis, J. A., McCloskey, E., Branco, J., Brandi, M. L., Dennison, E., Devogelaer, J. P., Ferrari, S., Kaufman, J. M., Papapoulos, S., Reginster, J. Y. and Rizzoli, R. (2014) 'Goal- directed treatment of osteoporosis in Europe', *Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA,* 25(11), pp. 2533.
Ketteler, M., Block, G. A., Evenepoel, P., Fukagawa, M., Herzog, C. A., McCann, L., Moe, S. M., Shroff, R., Tonelli, M. A., Toussaint, N. D., Vervloet, M. G. and Leonard, M. B. (2017) 'Executive summary of the 2017 KDIGO Chronic Kidney Disease–Mineral and Bone Disorder (CKD-MBD) Guideline Update: what's changed and why it matters', *Kidney Int,* 92(1), pp. 26-36.
Khosla, S., Westendorf, J. J. and Mödder, U. I. (2010) 'Concise Review: Insights from Normal Bone Remodeling and Stem Cell-Based Therapies for Bone Repair', *Stem Cells,* 28(12), pp. 2124-2128.
Kim, S. Y., Ok, H. G., Birkenmaier, C. and Kim, K. H. (2017) 'Can denosumab be a substitute, competitor, or compliment to bisphosphonates?', *Korean Journal of Pain,* 30(2), pp. 86-92.
Levey, A. S., Stevens, L. A., Schmid, C. H., Zhang, Y., Castro, A. F., Feldman, H. I., Kusek, J. W., Eggers, P., Van Lente, F., Greene, T. and Coresh, J. (2009) 'A New Equation to Estimate Glomerular Filtration Rate', *Annals of internal medicine,* 150(9), pp. 604-612.
Long, F. (2011) 'Building strong bones: molecular regulation of the osteoblast lineage', *Nature Reviews Molecular Cell Biology,* 13(1), pp. 27.

Lorentzon, M., Sweden, S. A. G. M. M., Cummings, S. R. and USA, t. S. F. C. C. C. P. M. C. R. I. U. o. C. S. F. C. (2015) 'Osteoporosis: the evolution of a diagnosis', *Journal of Internal Medicine,* 277(6), pp. 650-661.
Makras, P., Polyzos, S. A., Papatheodorou, A., Kokkoris, P., Chatzifotiadis, D. and Anastasilakis, A. D. (2013) 'Parathyroid hormone changes following denosumab treatment in postmenopausal osteoporosis', *Clin Endocrinol (Oxf),* 79(4), pp. 499-503.
Marini, J. C., Forlino, A., Bächinger, H. P., Bishop, N. J., Byers, P. H., Paepe, A. D., Fassier, F., Fratzl-Zelman, N., Kozloff, K. M., Krakow, D., Montpetit, K. and Semler, O. (2017) 'Osteogenesis imperfecta'.
Michelsen, J., Wallaschofski, H., Friedrich, N., Spielhagen, C., Rettig, R., Ittermann, T., Nauck, M. and Hannemann, A. (2013) 'Reference intervals for serum concentrations of three bone turnover markers for men and women', *Bone,* 57(2), pp. 399-404.
Morovat, A., Catchpole, A., Meurisse, A., Carlisi, I., Bekaert, A. C., Rousselle, O., Paddon, M., James, T. and Cavalier, E. (2013) 'IDS iSYS automated intact procollagen- 1- Nterminus pro-peptide assay: method evaluation and reference intervals in adults and children'.
Morris, H. A., Eastell, R., Jorgensen, N. R., Cavalier, E., Vasikaran, S., Chubb, S. A. P., Kanis, J. A., Cooper, C. and Makris, K. (2017) 'Clinical usefulness of bone turnover marker concentrations in osteoporosis', *Clinica Chimica Acta,* 467(Supplement C), pp. 34-41.
Moss, D. W. (1997) 'Physicochemical and pathophysiological factors in the release of membrane- bound alkaline phosphatase from cells', *Clinica Chimica Acta,* 257(1), pp. 133-140.
Mundy, G. (1999) *Bone Remodelling and its Disorders.* London: Martin Dunitz.
Nair, S., Bhadricha, H., Hatkar, S., Kadam, S. S., Patil, A., Surve, S., Joshi, B., Khatkhatay, M. I. and Desai, M. (2020) 'Effect of Vitamin D Levels on Bone Remodeling in Healthy Women', *International journal of endocrinology and metabolism,* 18(2), pp. e100656-e100656.
Naylor, K. E., Bradburn, M., Paggiosi, M. A., Gossiel, F., Peel, N. F. A., McCloskey, E. V., Walsh, J. S. and Eastell, R. (2018) 'Effects of discontinuing oral bisphosphonate treatments for postmenopausal osteoporosis on bone turnover markers and bone density'.
Naylor, K. E., Clowes, J. A., Finigan, J., Paggiosi, M. A., Peel, N. F. A. and Eastell, R. (2010) 'The effect of cessation of raloxifene treatment on bone turnover in postmenopausal women', *Bone,* 46(3), pp. 592-597.
Naylor, K. E., Jacques, R. M., Paggiosi, M., Gossiel, F., Peel, N. F., McCloskey, E. V., Walsh, J. S. and Eastell, R. (2016) 'Response of bone turnover markers to three oral bisphosphonate therapies in postmenopausal osteoporosis: the TRIO study', *Osteoporos Int,* 27(1), pp. 21-31.
'NICE guideline CG146, Surveillance of osteoporosis: assessing the risk of fragility fracture', (2019).
'NICE, Bisphosphonates for treating osteoporosis | Guidance', (2019).
'NICE, Denosumab for the prevention of osteoporotic fractures in postmenopausal women | Guidance', (2010).
NIH, C. D. P. (2001) 'Osteoporosis Prevention, Diagnosis, and Therapy.[Report]', *JAMA*.
Novack, D. V. and Teitelbaum, S. L. (2008) 'The Osteoclast: Friend or Foe?', *Annual Review of Pathology: Mechanisms of Disease,* 3(1), pp. 457-484.
Oei, L., Rivadeneira Ramirez, F., Zillikens, C. and Oei, E. (2015) 'Diabetes, Diabetic Complications, and Fracture Risk', *Curr Osteoporos Rep,* 13(2), pp. 106-115.
Paggiosi, M. A., Peel, N., McCloskey, E., Walsh, J. S. and Eastell, R. (2014) 'Comparison of the effects of three oral bisphosphonate therapies on the peripheral skeleton in postmenopausal osteoporosis: the TRIO study', *Osteoporos Int,* 25(12), pp. 2729-2741.

Pazianas, M. and Miller, P. D. (2021) 'Osteoporosis and Chronic Kidney Disease-Mineral and Bone Disorder (CKD-MBD): Back to Basics', *Am J Kidney Dis*.
Peck, W. A., Burckhardt, P., Christiansen, C., Fleisch, H. A., Genant, H. K., Gennari, C., Martin, T. J., Martini, L., Morita, R., Ogata, E., Rapado, A., Shulman, L. E., Stern, P. H., Young, R. T. T., Barrett-Connor, E., Brandi, M. L., Chesnut, C. H., Delmas, P. D., Heaney, R. P. and et al. (1993) 'Consensus development conference: Diagnosis, prophylaxis, and treatment of osteoporosis', *Am J Med,* 94(6), pp. 646-650.
Ravn, P., Hosking, D., Thompson, D., Cizza, G., Wasnich, R. D., McClung, M., Yates, A. J., Bjarnason, N. H. and Christiansen, C. (1999) 'Monitoring of Alendronate Treatment and Prediction of Effect on Bone Mass by Biochemical Markers in the Early Postmenopausal Intervention Cohort Study', *The Journal of Clinical Endocrinology & Metabolism,* 84(7), pp. 2363-2368.
Russell, R. G. G., Croucher, P. I. and Rogers, M. J. (1999) 'Bisphosphonates: Pharmacology, Mechanisms of Action and Clinical Uses', *Osteoporos Int,* 9(S2), pp. S66-S80.
Saag, K., Cosman, F., De Villiers, T., Langdahl, B., Scott, B. B., Denker, A. E., Pong, A. and Santora, A. C. (2021) 'Early changes in bone turnover and bone mineral density after discontinuation of long-term oral bisphosphonates: a post hoc analysis', *Osteoporos Int*.
Salam, S., Gallagher, O., Gossiel, F., Paggiosi, M., Khwaja, A. and Eastell, R. (2018) 'Diagnostic accuracy of biomarkers and imaging for bone turnover in renal osteodystrophy', *J Am Soc Nephrol,* 29(5), pp. 1557-1565.
Scariano, J. K., Garry, P. J., Montoya, G. D., Wilson, J. M. and Baumgartner, R. N. (2001) 'Critical differences in the serial measurement of three biochemical markers of bone turnover in the sera of pre- and postmenopausal women', *Clin Biochem,* 34(8), pp. 639-644.
Schlemmer, A. and Hassager, C. (1999) 'Acute fasting diminishes the circadian rhythm of biochemical markers of bone resorption', *Eur J Endocrinol,* 140(4), pp. 332-337.
Sharir, A., Barak, M. M. and Shahar, R. (2008) 'Whole bone mechanics and mechanical testing', *The Veterinary Journal,* 177(1), pp. 8-17.
Shuhart, Yeap and al, e. (2019) 'Executive Summary of the 2019 ISCD Position Development Conference on Monitoring Treatment, DXA Cross-calibration and Least Significant Change, Spinal Cord Injury, Peri-prosthetic and Orthopedic Bone Health, Transgender Medicine, and Pediatrics | Elsevier Enhanced Reader'.
Starup-Linde, J. (2013) 'Diabetes, biochemical markers of bone turnover, diabetes control, and bone', *Front Endocrinol (Lausanne),* 4, pp. 21-21.
Starup-Linde, J., Lykkeboe, S., Gregersen, S., Hauge, E.-M., Langdahl, B. L., Handberg, A. and Vestergaard, P. (2016) 'Differences in biochemical bone markers by diabetes type and the impact of glucose', *Bone,* 83, pp. 149-155.
Stevens, P. E., Levin, A. and Members, K. D. I. G. O. C. K. D. G. D. W. G. (2013) 'Evaluation and management of chronic kidney disease: synopsis of the kidney disease: improving global outcomes 2012 clinical practice guideline', *Ann Intern Med,* 158(11), pp. 825-30.
Szulc, P. (2012) 'The role of bone turnover markers in monitoring treatment in postmenopausal osteoporosis', *Clinical Biochemistry,* 45(12), pp. 907-919.
Vasikaran, S., Eastell, R., Bruyere, O., Foldes, A. J., Garnero, P., Griesmacher, A., McClung, M., Morris, H. A., Silverman, S., Trenti, T., Wahl, D. A., Cooper, C. and Kanis, J. A. (2011) 'Markers of bone turnover for the prediction of fracture risk and monitoring of osteoporosis treatment: a need for international reference standards', *Osteoporos Int,* 22(2), pp. 391-420.

Vasikaran, S. and Paul Chubb, S. (2016) 'The use of biochemical markers of bone turnover in the clinical management of primary and secondary osteoporosis', *International Journal of Basic and Clinical Endocrinology,* 52(2), pp. 222-225.
Vasikaran, S. D. (2001) 'Bisphosphonates: an overview with special reference to alendronate', *Ann Clin Biochem,* 38(6), pp. 608-623.
Vasikaran, S. D., Bhattoa, H. P., Eastell, R., Heijboer, A. C., Jørgensen, N. R., Makris, K., Ulmer, C., Kanis, J. A., Cooper, C., Silverman, S. and Cavalier, E. (2020) 'Harmonization of commercial assays for PINP; the way forward', *Osteoporos Int,* 31(3), pp. 409-412.
Vescini, F. and Grimaldi, F. (2012) 'PTH 1-84: bone rebuilding as a target for the therapy of severe osteoporosis', *Clin Cases Miner Bone Metab,* 9(1), pp. 31-6.
Vestergaard, P. (2007) 'Discrepancies in bone mineral density and fracture risk in patients with type 1 and type 2 diabetes--a meta-analysis', *Osteoporos Int,* 18(4), pp. 427-44.
Vlot, M. C., Den Heijer, M., de Jongh, R. T., Vervloet, M. G., Lems, W. F., de Jonge, R., Obermayer-Pietsch, B. and Heijboer, A. C. (2018) 'Clinical utility of bone markers in various diseases', *Bone,* 114, pp. 215-225.
Walker, D. B. and Smith, S. (2017) 'Session Overview: Bone as an Endocrine Organ', *https://doi-org.sheffield.idm.oclc.org/10.1177/0192623317734812*.
Walsh, J. S., Evans, A. L., Bowles, S., Naylor, K. E., Jones, K. S., Schoenmakers, I., Jacques, R. M. and Eastell, R. (2016) 'Free 25-hydroxyvitamin D is low in obesity, but there are no adverse associations with bone health'.
Walsh, J. S., Gossiel, F., Scott, J. R., Paggiosi, M. A. and Eastell, R. (2017) 'Effect of age and gender on serum periostin: Relationship to cortical measures, bone turnover and hormones'.
Weitoft, T., Larsson, A., Saxne, T. and Rönnblom, L. (2005) 'Changes of cartilage and bone markers after intra-articular glucocorticoid treatment with and without postinjection rest in patients with rheumatoid arthritis', *Annals of the Rheumatic Diseases,* 64(12), pp. 1750.
Yamamoto, T., Taketsuna, M., Guo, X., Sato, M. and Sowa, H. (2014) 'The safety and effectiveness profile of daily teriparatide in a prospective observational study in Japanese patients with osteoporosis at high risk for fracture: interim report', *J Bone Miner Metab,* 32(6), pp. 699-708.
Yang, L., Peel, N., Clowes, J. A., McCloskey, E. V. and Eastell, R. (2009) 'Use of DXA-Based Structural Engineering Models of the Proximal Femur to Discriminate Hip Fracture', *J Bone Miner Res,* 24(1), pp. 33-42.
Yu, E. W., Thomas, B. J., Brown, J. K. and Finkelstein, J. S. (2012) 'Simulated increases in body fat and errors in bone mineral density measurements by DXA and QCT', *J Bone Miner Res,* 27(1), pp. 119-124.