Osteoporosis and gut microbiota, radiofrequency echographic multi-spectrometry and machine learning: one year in review 2023

M. Maffi, G. De Mattia, M. Mazzantini

Clinical and Experimental Rheumatology 2023

Competing interests: none declared.

Review

Osteoporosis is a prevalent bone disease with a relevant burden of mortality and comorbidity, especially due to fragility fractures occurring as a result of reduced bone mineral density. In this review we provide a critical digest of the most recent literature regarding the relationship between gut microbiota and osteoporosis and discuss the role of radiofrequency echographic multi-spectrometry (REMS) and machine learning in the diagnostic work-up and prevention of osteoporosis.

Introduction

Osteoporosis (OP) is a skeletal disease characterised by low bone mineral density (BMD) and deterioration of bone tissue microarchitecture. It is the most prevalent bone disorder in adults, especially among postmenopausal women, and is frequently associated with fragility fractures, resulting in increased morbidity and mortality, in a lower quality of life, and in a remarkable social and economic burden (1). The aim of this review was to provide an overview of recent advances in the field of OP. Given the large number of publications, we focused on three main topics: the role of gut microbiota in the pathogenesis of OP, and radiofrequency echographic multi-spectrometry (REMS) and machine learning (ML) technologies as recently developed tools for the screening and diagnosis of OP.

Medline databases (PubMed) were searched using the following keywords for studies published in 2022: “Gut microbiota AND osteoporosis”, “Radio frequency Echographic Multi Spectrometry”, “REMS”, “Machine Learning AND Osteoporosis”.

The relationship between gut microbiota and bone: do we have solid evidence?

The microorganisms that inhabit the human intestine are known as the gut microbiota, and their collective genome is called gut microbiome. The gut microbiota consists of over 10–100 trillion microbes (bacteria, archaea, viruses, fungi, protozoa and eukaryotes; their density and composition vary throughout the intestinal tract, the colon harbouring the most) that encode more than 3.3 million genes (2, 3). Firmicutes and Bacteroides phyla represent more than 90% of the intestinal microbiota (4). It has been well recognised that this enormous symbiotic population has multiple local effects, such as enhancing the extraction of energy from foods, increasing absorption of nutrients, helping in immune system development, and preventing the colonisation and invasion by pathogens. In addition to these local effects, it has been recently supposed that the gut microbiota may play a role in the systemic regulation of some physiological and pathological processes related to human health. For example, short-chain fatty acids (SCFAs) are metabolites of dietary fibres and are produced by microbiota in the large intestine; SCFAs seem to exert numerous effects such as improving calcium absorption, suppressing appetite, improving glucose tolerance, and affecting pro- and anti-inflammatory properties of immune cells. Microbiota has been reported also to influence the disease activity in some rheumatic conditions (5-7). Recent studies on animal models have reported that the gut microbiota may also contribute to the regulation of bone metabolism, which ignited enthusiasm
about the possibility of preventing bone loss by modulating the gut microbiota and even treating OP by mean of fecal microbiota transplantation (8). It is beyond the scope of this chapter to analyse the multiple mechanisms by which this may happen: the reader can find extensive reviews published in 2022 that examined the suggested connections between the gut microbiota and bone (9-16). Despite hundreds in vitro or animal studies illustrating an influence of the gut microbiota/microbiome on the skeleton, there are few studies involving humans. Therefore, the aim of this chapter is to provide a brief report on human studies published in 2022 which investigated the relationship between gut microbiota and bone, and to summarise their results in synthetic conclusion remarks.

Knowledge to date: human studies

The study that first examined the effects on bone of a probiotic was published by Nilsson in 2018 (17). In this proof of concept, double-blind, placebo-controlled study, women (mean age 76 years) with low BMD were randomised to receive orally 10^10 colony-forming units of Lactobacillus Reuteri 6475 daily or placebo for 12 months. Those receiving the probiotic supplement had a reduced loss of volumetric BMD at the tibia (assessed by high-resolution peripheral quantitative computed tomography) compared to those receiving placebo: mean relative change -0.83% (95% CI -1.47% to -0.19%) vs. -1.85% (95% CI -2.64% to -1.07%). Furthermore, the same authors in 2022 (18) investigated factors that could influence the response to the probiotic. They selected 20 women with the highest changes (good responders) and 20 with the lowest changes (poor responders) in the tibial BMD from the previous study and characterised the gut microbiome composition and function as well as serum metabolome in both groups. They found that gene richness of the gut microbiota was significantly higher (p<0.01) and inflammatory state was significantly improved (p<0.05) in the good responders at the end of the 12-months daily supplementation. In particular, the good responder group to Lactobacillus Reuteri showed decreased levels of serum inflammation marker ultrasensitive C-reactive protein, which has been reported to be associated to higher BMD (19, 20). Moreover, detrimental changes observed in the gut microbiota of poor responders, including the enrichment of E.Coli and its biofilm formation, were lessened in the good responders to L. Reuteri; and several SCFAs-producing bacterial species, including C. Acetobutylicum, A. Fermentans, A. Muciniphila, C. Catus and R. Bicirculus were more abundant in the good responder group than in the poor responder group at 12 months. In summary, L. Reuteri supplementation could prevent a deterioration of the gut microbiota and inflammatory status, which may explain the small but significant effect reported on bone density. SCFAs could influence bone formation. This has been suggested by a study (21) that investigated whether SCFAs are absorbed by osteoblasts and influence early osteoblastic differentiation using a pre-osteoblast murine cell line. In that model, acetate and propionate upregulated the osteoblast differentiation marker alkaline phosphatase; acetate also regulated alkaline phosphatase mRNA expression. If these results were confirmed in human studies, acetate and propionate could be useful for promoting bone formation.

The positive effects of Lactobacilli probiotics on bone were confirmed by other randomised, double-blind, placebo-controlled studies in humans (22-24). However, non-significant changes in bone density were observed in another RCT comparing probiotics to placebo (25).

In 2022, Orwell et al. (26) published the results of a study aimed at examining the associations of the fecal microbiome with measures of bone density, microarchitecture, and strength. They obtained stool samples from 831 participants in a longitudinal observational large study of community dwelling older men, aged 78 to 98 years; they performed 16S rRNA amplicon sequencing and tested for association between the abundance of microbial genera and bone measures obtained with dual-energy x-ray absorptiometry and high-resolution peripheral quantitative computed tomography (at distal radius, distal and diaphyseal tibia). The abundance of 4 bacterial genera were found to be weakly associated with the bone parameters selected (false discovery rate ≤0.05): Anaerofibrum (with lower radial and tibial density), Methanomassiliicoccus (greater distal cortical porosity), Ruminoclostridium 9 (with less distal tibial cortical porosity), and Tyzzerella (with greater tibial density measures). Also, Lactobacillus and Streptococcus were both associated with worse bone measures at radial and tibial sites, but only when a less strict criterion (false discovery rate ≤0.1) was adopted. As the Authors themselves recognise, these results must be interpreted with caution: the magnitudes of associations were not large, functional follow-up of the association is lacking, statistical power is low, findings relative to Lactobacillus are in contrast with previous findings. We find appropriate the Authors’ conclusion that “larger cohorts of men and women over wider age ranges and/or more causally incisive methods” are needed. An editorial about the study of Orwell brilliantly lists the difficulties in conducting down-to-earth studies in this field (27).

Among the many molecules derived from gut microbiota that may influence human health, trimethylamine N-oxide (TMAO) may adversely affect bone health by inducing oxidative stress (28). To test this hypothesis, Elam et al. (29) assessed TMAO plasma levels and hip fractures incidence during up to 26 years of follow-up in 5019 adults aged ≥65 years; total hip BMD was assessed by dual-energy x-ray absorptiometry in a subset of patients (n=1400). TMAO was not significantly associated with hip fracture: after controlling for covariates, hazard ratios per TMAO doubling were 1.00 (95% CI 0.92, 1.09) in women and 1.12 (0.95–1.33) in men. TMAO was also not associated with total hip BMD. These results are sharply in contrast with previous small longitudinal, cross-sectional, or retrospective studies (30-32).

The relationship between gut microbiota and bone is far from being clari-
fied. We eagerly await the results of the Prune study (33), a randomised controlled trial that assessed the effects on bone density, bone structure and strength of a 12-month daily prune consumption in postmenopausal women, which has the potential of manipulating the gut microbiota.

Take-home messages
• The gut microbiota may influence the bone (8-16).
• Probiotic supplementation may attenuate bone loss in postmenopausal women, but larger and long-term studies are needed (18-20).
• To date, large studies failed to find a sound relation between gut microbiome or its products and bone density or fractures. Currently there is no scientific basis to include gut microbiota modifications in our treatment strategy of OP (29-32).

Radiofrequency echographic multi-spectrometry (REMS): overcoming the limits of DXA scan
Dual-energy x-ray absorptiometry (DXA) scan at the femur and lumbar spine has traditionally been the gold standard method for the diagnosis of OP (34). Despite being widely available and providing high precision and reproducibility, DXA scan presents several limitations, most importantly the use of x-rays and the overestimation of BMD in patients with spinal osteoarthritis, previous vertebral fractures, or aortic calcifications (35), potentially resulting in misdiagnosis and undertreatment. A recent breakthrough in this regard is represented by REMS, which is a radiation-free tool analysing native ultrasound (US) waves acquired at lumbar vertebrae and/or femur. In contrast with conventional ultrasonography, REMS analyses native unfiltered US waves, thereby providing comprehensive information related to bone quality and quantity and obtaining T- and Z-scores through the comparison with age-, sex-, BMI- and site-matched reference curves created for healthy and pathologic bone tissue. REMS has been validated in mono- and multicentric studies as a tool for the diagnosis of OP (36-38), and it was demonstrated that it had a higher sensitivity than DXA in the detection of female subjects prone to fragility fractures in a follow-up period of 5 years (39). Furthermore, it was demonstrated that T-scores at the lumbar spine obtained through REMS in patients with osteoarthritis (OA) are significantly more reliable for the detection of OP in this category of patients, as compared to DXA scan (40).

An Italian study published in 2022 supported these findings by showing that in a cohort of 159 postmenopausal women (66.21±6.6 years) with previous vertebral fractures or lumbar spine OA, REMS classified as “osteoporotic” a larger percentage of patients with respect to DXA (35.1% vs. 9.3%, respectively, in the OA subjects; 58.7% vs. 23.3%, respectively, in the subjects with fractures) (41). Therefore, the capability to overcome artifacts such as OA and fractures, along with the absence of radiation and reduced costs, may lead to the use of REMS in general population screenings (37). Indeed, REMS can be employed even in subsets of patients for which DXA is not recommended, e.g. pregnant women, breast-feeding women and children (42). For instance, in an exploratory Italian prospective observational study using REMS on 78 pregnant women with uncomplicated pregnancy at 39.1±1.5 weeks have been found to have significantly lower femoral neck BMD values compared to a matched control group of non-pregnant women (0.769±0.094 g/cm² vs. 0.831±0.101 g/cm², p=0.0001), showing that REMS may be used to monitor pregnant women with OP or related risk factors (43).

Recently, REMS has been evaluated on patients with diseases increasing the risk for OP and fragility fractures. A study by Fassio et al. showed a promising agreement (Cohen’s κ correlation coefficient 0.663, p<0.01) between DXA- and REMS-derived BMD values for the worst site considered and in the consequent fracture risk assessment in a group of 41 patients affected by chronic kidney disease (CKD) receiving peritoneal dialysis (44). A recent study involving 90 female patients with type 2 diabetes mellitus (T2DM) showed that REMS was more reliable than DXA at diagnosing OP in this population (47% vs. 28%, respectively). Furthermore, lower values of BMD at the lumbar spine by REMS were significantly associated with a history of major fragility fractures in the T2DM population, whereas BMD by DXA were not (45). Another study found that in a cohort of 50 patients with anorexia nervosa, the subjects with previous vertebral fragility fractures presented lower values of BMD at total hip and lumbar spine by DXA and by REMS with respect to those without history of fractures; however, the difference was significant only for BMD at total hip as measured by REMS (46). However, it should be stated that there have been cases for which REMS was not as accurate as in the previously mentioned studies. For example, Lalli et al. assessed the accuracy of REMS in comparison with DXA in disuse-related OP in patients with spinal cord injury and found that REMS overestimated BMD at femoral neck and total hip in this category of patients, possibly due to the atrophy and myosteatosis of the iliopsoas muscle which may influence the ultrasound propagation between the probe and the femoral neck. However, the low number of patients warrant larger studies for confirmation (47).

Considering the high incidence of fragility fractures in the general population aged over 50 and the consequent significant economic burden for the healthcare systems (48), an accurate estimation of fracture risk is crucial and widely used tools in clinical practice for this purpose include FRAX, Garvan, QR fracture, Fra-HS and De-FRA (49). REMS has been employed for the development of another tool, called the Fragility Score (FS), which has been shown to have slightly better performance than DXA in discriminating patients with previous fractures (50). Interestingly, a prospective study involving 1989 patients recently demonstrated that FS is more capable of predicting fracture risk in both female and male subjects as compared to BMD T-score values obtained by either DXA or REMS (area under the curve (AUC) for lumbar spine adjusted for age and BMI: FS 0.715 vs. REMS BMD T-score...
Use of machine learning tools for the diagnosis and prevention of osteoporosis

Machine learning (ML) is a scientific discipline that focuses on how computers learn from data and become capable of building statistical models from massive datasets (52). ML technologies have recently demonstrated that they can also play a major role from a clinical perspective (53-55). Such technologies allow the creation of predictive models (e.g. logistic regression, artificial neural networks, eXtreme gradient boosting, support vector machine, stacking, random forest) which can be used to predict any range of outputs. Conventionally, ML is classified into supervised learning and unsupervised learning. The former approach is generally used for estimating risk and for creating automated models useful for interpreting instrumental examinations (e.g. evaluations of an EKG or a lung computed tomography), whereas the latter is used to search for the presence of specific patterns within a large amount of data (52). The advantages of ML include the ability to analyse different types of data and its potential for time- and cost-effectiveness (56-57).

Interestingly, ML techniques have been recently investigated for the estimation of fracture risk and for the diagnosis of OP from other imaging tools performed by the patient for other reasons. For example, in the work of Jang et al., the deep learning model “OsPor-screen” was trained in a supervised learning manner to recognize OP through retrospective analysis of chest radiographs, yielding promising results. A total of 13,026 chest x-rays and DXAs of individuals between 40 and 90 years of age were analysed, equally divided between “osteoporosis”, “osteopenia” and “normal findings”. The model was first trained to recognise osteoporosis patients belonging to the “Health Screening and Promotion Centre of Asan Medical Centre” cohort and was subsequently externally validated with the “Asan osteoporosis cohort dataset”. The ML model was shown to analyse an x-ray in less than 4 seconds with a sensitivity of 86.2% and specificity of 74.2% in recognising OP in the external validation. This study supports the use of the “OsPor-screen” model as a cost-effective method for opportunistic automated screening of patients with OP in clinical settings and without exposing the patient to additional radiation (67). Similar studies found that ML models were capable of screening for OP based on the analysis of lumbar spine and hip radiographs or orthopantomography (68-70). In fact, the presence of OP is often reported by the radiologist as a collateral finding on radiographs, thus being subject to interindividual variability. A similar problem occurs for the estimation of vertebral fractures with Genant semi-quantitative method; in this regard, ML technologies are also useful in the creation of automated screening tools for vertebral fracture assessment and identification of risk factors for fractures (71-72).

Indeed, the decision-tree-based model “LightGBM” designed by Microsoft Research Asia to clarify the relevant characteristics that determine fracture after surgically treated fragility fractures has been employed in a study involving 7000 patients, identifying rheumatoid arthritis (RA) and CKD as potential predictors of fracture. In fact, in the study of Shimizu et al., the “LightGBM” model was shown to be moderately accurate in predicting fractures, with an AUC of approximately 0.75. Furthermore, when compared to other ML models such as Artificial Neural Networks, “LightGBM” achieved a higher accuracy, suggesting the superiority of decision-tree-based models trained on table data, in this clinical context (72). ML models have also been tested in the context of opportunistic OP screening on images obtained from computed tomography (CT) scans of the chest and abdomen, demonstrating high specificity and sensitivity for OP assessment. Besides, it is also possible to integrate data from CT images with OP risk factors and routine laboratory tests data to build a hierarchical model to identify individuals with OP, as an alternative method to DXA. Liu et al. proposed a hierarchical model with three layers: a first layer consisting of only demographic characteristics, a second layer with only clinical data, and a third layer of CT images that partially or completely included the spine. Data of 2210 patients over age 40 were collected ret-
respectively, then six ML algorithms were used as classifiers to discriminate individuals between osteoporotic and non-osteoporotic; the results showed that the hierarchical model based on logistic regression had better performances, with an area under the receiver operating characteristic curve of 0.818, 0.838 and 0.962 for the three layers, respectively (73-74).

ML technologies can also be employed for the interpretation of complex and multifaceted information, such as that provided by high resolution peripheral quantitative computed tomography (HRpQCT), an imaging technique assessing trabecular and cortical microstructure of bone (75). In the work of Lu et al., an automatic high-performance diagnostic algorithm was proposed; it used various inputs (clinical data, bone mineral density measured at the femoral neck, and data from HRpQCT images of the tibia) to discriminate between patients with or without previous fragility fractures (76). The amount of information provided by HRpQCT, together with the computational capability of ML methods, allowed the identification of bone microarchitecture phenotypes. Whittier et al. employed “fuzzy c-means clustering”, an unsupervised ML method, which was able to identify three different clusters of bone microarchitecture: low-density, low-volume and healthy bone. Bone phenotypes were identified using cluster analysis and characteristics selected for clustering included height (value used as a surrogate for long bone length) and some HR-pQCT parameters measured at the radius and ulna. According to the authors, the low-density and low-volume phenotypes were those most associated with fragility fractures with a hazard ratio of 2.96 and 2.95, respectively. Therefore, this study suggests the utility of deep learning in assessing fracture risk associated with intrinsic phenotypic characteristics of bone, which DXA is unable to capture (77).

ML techniques are also useful in predicting the fracture risk of certain subpopulations of patients, such as in the case of elderly-onset RA (EORA) or breast cancer. For example, it has been shown that especially the random forest classifier model can accurately predict the occurrence of fractures in patients with RA and OP. In addition, Ji et al. constructed three models to predict OP, fragility fractures and survival in breast cancer patients, showing superiority over the FRAX and OSTa (Osteoporosis Self-assessment Tool for Asians) algorithms and thus identifying a new approach for screening this population at risk (78-79).

In conclusion, in the scientific literature, studies related to ML in the context of OP are increasing in number and depth, suggesting a growing interest in this field. The progressive refinement of ML technologies will hopefully allow their use in the daily clinical evaluation of the OP patient.

**Take-home messages**

- Machine Learning (ML) has emerged as a powerful tool in the clinical setting for creating predictive models that can be used to analyse different types of data. Relevant advantages include time- and cost-effectiveness (53-57).
- ML techniques are being used for the diagnosis and estimation of fracture risk in osteoporosis, and can use information from instrumental examinations already performed, providing a non-invasive method for screening and early detection of the disease (67-70).
- ML models are useful for automated screening tools for vertebral fracture assessment and identification of risk factors for refractures. Decision-tree-based models have been used for the prediction of refractures, as demonstrated by the “LightGBM” model. ML models can also be tested for opportunistic osteoporosis screening on images obtained from CT scans of the chest and abdomen (71-74).

**References**


16. DE SIRE A, DE SIRE R, CURCI C, CASTIGLIO-NE F, WAHLI W: Role of dietary supplements


42. STRÖM O, BORGSTRÖM F, KANIS JA et al: Osteoporosis: burden, health care provision and opportunities in the EU: a report pre-