E-ISSN: 2147-2653

Türk Osteoporoz Dergisi

TURKISH JOURNAL OF OSTEOPOROSIS

Cilt / Vol.: 31 Sayı / Issue: 3 Aralık / December 2025

www.turkosteoporozdergisi.org

TÜRKİYE OSTEOPOROZ DERNEĞİ
1998
www.osteoporoz.org.tr

Türk Osteoporoz Dergisi

TURKISH JOURNAL OF OSTEOPOROSIS

Türk Osteoporoz Dergisi

TURKISH JOURNAL OF OSTEOPOROSIS

Türkiye Osteoporoz Derneği "International Osteoporosis Foundation" Üyesidir.
"Web tabanlı danışma kurulu ve hakem listesi www.turkosteoporozdergisi.com adresinde yer almaktadır"
Derginin eski adı 'Osteoporoz Dünyasından Dergisi'dir.

Yayınevi İletişim/Publisher Contact
Adres/Address: Molla Gürani Mah. Kaçamak Sk. No: 21/1
34093 Fındıkzade-İstanbul-Türkiye
Telefon/Phone: +90 (530) 177 30 97
E-posta/E-mail: info@galenos.com.tr/yayin@galenos.com.tr
Web: www.galenos.com.tr
Yayıncı Sertifika No/Publisher Certificate Number: 14521

Online Yayınlanma Tarihi/Online Publishing Date: Aralık 2025/December 2025
E-ISSN: 2147-2653
Yılda üç kez yayımlanan süreli yayındır.
International periodical journal published three times in a year.

Türk Osteoporoz Dergisi

TURKISH JOURNAL OF OSTEOPOROSIS

Derginin "Yayın Etiği" ve "Yazarlara Bilgi" konularında bilgi almak için lütfen web sayfasına **(https://www.turkosteoporozdergisi.org/)** başvurun.

Derginin editöryal ve yayın süreçleri ile etik kuralları, ICMJE, COPE, WAME, CSE ve EASE gibi uluslararası kuruluşların kurallarına uygun olarak şekillendirilmektedir. Türk Osteoporoz Dergisi, **Emerging Sources Citation Index (ESCI), DOAJ, EBSCO Database, Gale/Cengage Learning, CINAHL, CABI, Embase, Scopus, ProQuest, J-Gate, IdealOnline, TÜBİTAK/ULAKBİM, Hinari, GOALI, ARDI, OARE, AGORA, Türk Medline** ve **Türkiye Atıf Dizini** tarafından indekslenmektedir.

Dergi, çevrimiçi olarak yayınlanmaktadır.

Sahibi: Türkiye Osteoporoz Derneği

Sorumlu Yönetici: Yeşim Kirazlı

Please refer to the journal's webpage **(https://www.turkosteoporozdergisi.org/)** for "Ethical Policy" and "Instructions to Authors".

The editorial and publication processes of the journal are shaped in accordance with the guidelines of the ICMJE, COPE, WAME, CSE and EASE. Turkish Journal Of Osteoporosis is indexed by the **Emerging Sources Citation Index (ESCI), DOAJ, EBSCO Database, Gale/Cengage Learning, CINAHL, CABI, Embase, Scopus, ProQuest, J-Gate, IdealOnline, TÜBİTAK/ULAKBİM, Hinari, GOALI, ARDI, OARE, AGORA, Turkish Medline** and **Turkish Citation Index**.

The journal is published online.

Owner: Turkish Osteoporosis Society

Responsible Manager: Yeşim Kirazlı

Türk Osteoporoz Dergisi
TURKISH JOURNAL OF OSTEOPOROSIS

İÇİNDEKİLER / CONTENTS

Derleme / *Review*

124 **Evaluation of the Effects of Menopause on the Musculoskeletal System with Experimental Rat Models**
Nesime Hamalı, Kezban Kartlaşmış, Ayşe Nur Çakır Güngör; Adana, Türkiye

Orijinal Araştırmalar / *Original Investigations*

130 **Can the Systemic Immune Inflammation Index (SII) Indicate Disease Activity in Patients with Rheumatoid Arthritis?**
Salim Mısırcı, Alev Alp, Büşra Başar Yılmaz; Bursa, Türkiye

138 **Osteoporotic Fracture Forecasting in Türkiye: Analysis of the Agreement and Correlation Between FRAX and Garvan Fracture Risk Calculator Tools**
Mustafa Hüseyin Temel, Seda Özcan İşler, Hatice Merve Beşışık Temel, Fatih Bağcıer, Emre Ata; İstanbul, Türkiye

145 **The Relationship Between Physical Performance with Dual-task Performance in Geriatric Individuals**
Merve Akış, Emine Atıcı, Berna Çağla Balkışlı; İstanbul, Türkiye

152 **Effects of Low-Dose Corticosteroids on Clinical Outcomes and Bone Mineral Density Parameters in Rheumatoid Arthritis Patients**
Sinem Kübra Beke, Hüseyin Kaplan, Taha Furkan Çakır, Gizem Cengiz, Senem Sas, İsa Cüce; Kayseri, Aksaray, Bursa, Türkiye

162 **A New Era in Digital Health: Evaluation of Artificial Intelligence Supported Chatbots (ChatGPT-4, BingAI, and Gemini) Responses on Osteoporosis**
Gamze Gül Güleç, Özge Özpolat Bulut, Fatih Bağıcer; Adana, Şanlıurfa, İstanbul, Türkiye

169 **Acute Impact of Spinal Manipulation on Pain and Muscle Mechanical Properties in Chronic Low Back Pain: An Assessor Blinded, Randomized Controlled Trial**
Ayça Aracı, Emine Eda Kurt, Serkan Taş; Antalya, Mersin, Türkiye

178 **Comparison of the Osteoporosis Knowledge Levels of Physiotherapy Students Pre-post Awareness Training: A Randomized Controlled Trial**
Bengisu Tüfekçi, Tuğba Birben Kurt, Ayşe Nur Şekeroğlu; Gaziantep, Rize, Osmaniye, Türkiye

189 **Performance Comparison of Gemini, DeepSeek, and ChatGPT-4o on American Board of Physical Medicine and Rehabilitation Board Exam Practice Questions**
Gonca Sağlam Akkaya, Hanife Baykal Şahin; Trabzon, Türkiye

Olgu Sunumları / *Case Reports*

195 **Combined Surgical and Concentrated Growth Factor Therapy in the Management of Stage II and III MRONJ Associated with Denosumab: A Report of Two Cases**
Mert Kırdemir, Muharrem Ergün Dudak, Burak Borlu; İzmir, Türkiye

201 **Osteopoikilosis as an Uncommon Cause of Extremity Pain: A Case Report**
Emre Fidan, Adil Can Karaoğlu, Ahmet Özdemir; Balıkesir, Türkiye

Editöre Mektup / *Letter to the Editor*

205 **Re-evaluating Cardiovascular Risk in Osteoporotic Patients: A Case for Individualized Romosozumab Therapy**
Ömer Faruk Bucak, Rümeysa Çalışkan, Mustafa Hüseyin Temel, Fatih Bağcıer, Evrim Coşkun; İstanbul, Türkiye

İndeks / *Index*

2025 Hakem Dizini - 2025 Referee Index
2025 Yazar Dizini - 2025 Author Index
2025 Konu Dizini - 2025 Subject Index

Türk Osteoporoz Dergisi
TURKISH JOURNAL OF OSTEOPOROSIS

Editörden / Editorial

Sevgili Meslektaşlarımız,

Değerli meslektaşlarımıza akademik çalışmalarını yayınlanmak üzere Emerging Sources Citation Index (ESCI) tarafından indekslenen dergimize düzenli olarak ilettikleri için çok teşekkür ederiz. Dergimiz kapsamında bu akademik çalışmaları yayınlamaktan ve çalışmalarımıza aynı hızla devam etmekten büyük mutluluk duyuyoruz.

26. Dünya Osteoporoz, Osteoartrit ve Kas İskelet Sistemi Hastalıkları Kongresi (WCO-IOF-ESCEO Congress) 16-19 Nisan 2026 tarihlerinde Prag-Çek Cumhuriyeti'nde yapılacaktır. Dünya çapındaki bu kongrede konularında önde gelen uluslararası bilim insanlarının konuşmacı olarak yer almasıyla, kas iskelet sistemi sağlığı ve hastalıklarındaki son araştırmalar, klinik bakış açıları ve terapötik ilerlemeler katılımcılara sunulma olanağı bulacaktır.

Siz değerli meslektaşlarımıza 2025 yılında mutluluğun ve başarının yaşamınızdan eksik olmamasını dileyerek sevgi ve saygılarımı sunarım.

Editör
Prof. Dr. Yeşim Kirazlı

Review / Derleme

DOI: 10.4274/tod.galenos.2025.74317
Turk J Osteoporos 2025;31(3):124-9

Evaluation of the Effects of Menopause on the Musculoskeletal System with Experimental Rat Models

Deneysel Sıçan Modelleri ile Menopozun Kas-iskelet Sistemi Üzerindeki Etkilerinin Değerlendirilmesi

Nesime Hamalı[1], Kezban Kartlaşmış[2], Ayşe Nur Çakır Güngör[3]

[1]Çukurova University Faculty of Medicine, Adana, Türkiye
[2]Çukurova University Institute of Health Sciences, Adana, Türkiye
[3]Çukurova University Faculty of Medicine, Department of Obstetrics and Gynaecology, Adana, Türkiye

Abstract

In pre-menopausal women, low bone mass and trauma fractures may be due to secondary causes such as estrogen deficiency, glucocorticoid exposure, malabsorption or hyperparathyroidism. However, during the menopausal transition and menopause, the major cause is a decrease in the hormone estrogen. The transition to menopause begins about four years before the last menstrual period. Over time, the production and secretion of estradiol, the potent estrogen, decreases in the ovaries. In the skeletal system, this deficiency causes loss of spongiosis and cortical bone; in the muscular system, it reduces muscle strength through muscle mass and contractile proteins. Micro- and macro-level losses in muscle and bone tissues, which are both endocrine and paracrine tissues, begin during perimenopause. This article discusses the evaluation of menopause-induced musculoskeletal changes in experimental rat models and different treatment approaches. The adverse effects of estrogen deficiency on muscle and bone have been investigated with various therapeutic strategies to address this deficiency. Several compounds such as myostatin, irisin, lycopene, nacre, superjami rice and annatto tocotrienol have the potential to improve muscle and bone health after menopause. Furthermore, it has been emphasized that estrogen plays important roles such as supporting muscle satellitic cells, improving mitochondrial function and inhibiting bone destruction. In each study, efforts to reduce the effects of menopause through different pathways and molecules were described. These findings offer significant potential for the protection of musculoskeletal health in the postmenopausal period and the development of treatment alternatives.

Keywords: Menopause, musculoskeletal system, rat models

Öz

Premenopozal dönemdeki kadınlarda düşük kemik kütlesi ve travma kaynaklı kırıklar, östrojen eksikliği, glukokortikoid maruziyeti, malabsorpsiyon veya hiperparatiroidizm gibi ikincil nedenlere bağlı olabilir. Ancak menopoz geçiş dönemi ve menopoz sırasında temel neden, östrojen hormonundaki azalmadır. Menopoza geçiş, son adet döneminden yaklaşık dört yıl önce başlar. Zamanla, güçlü bir östrojen olan estradiolün overlerdeki üretimi ve salınımı azalır. İskelet sisteminde bu eksiklik, süngerimsi ve kortikal kemik kaybına; kas sisteminde ise kas kütlesi ve kasılma proteinleri yoluyla kas gücünde azalmaya neden olur. Endokrin ve parakrin dokular olan kas ve kemik dokularındaki mikro ve makro düzeydeki kayıplar, perimenopoz döneminde başlar. Bu makalede, menopozun neden olduğu kas-iskelet değişikliklerinin deneysel sıçan modellerinde değerlendirilmesi ve farklı tedavi yaklaşımları ele alınmıştır. Östrojen eksikliğinin kas ve kemik üzerindeki olumsuz etkilerinin, bu eksikliği gidermek için çeşitli terapötik stratejilerle araştırıldığı bildirilmiştir. Myostatin, irisin, likopen, sedef, Superjami pirinci ve annatto tokotrienol gibi çeşitli bileşiklerin menopoz sonrası kas ve kemik sağlığını iyileştirme potansiyeline sahip olduğu ifade edilmiştir. Ayrıca, östrojenin kas satelit hücrelerini destekleme, mitokondriyal fonksiyonu iyileştirme ve kemik yıkımını önleme gibi önemli roller oynadığı vurgulanmıştır. Her bir çalışmada, menopozun etkilerini farklı yollar ve moleküller aracılığıyla azaltmaya yönelik çabaların ele alındığı belirtilmiştir. Bu bulguların, postmenopozal dönemde kas-iskelet sağlığının korunması ve tedavi alternatiflerinin geliştirilmesi açısından önemli bir potansiyele sahip olduğu ifade edilmiştir.

Anahtar kelimeler: Menopoz, kas-iskelet sistemi, sıçan modelleri

Corresponding Author/Sorumlu Yazar: Kezban Kartlaşmış MD, Çukurova University Institute of Health Sciences, Adana, Türkiye
E-mail: kzbn.krtlsms@gmail.com **ORCID ID:** orcid.org/0000-0001-5090-0013
Received/Geliş Tarihi: 27.01.2025 **Accepted/Kabul Tarihi:** 04.03.2025 **Epub:** 30.07.2025 **Publication Date/Yayınlanma Tarihi:** 05.12.2025

Cite this article as/Atıf: Hamalı N, Kartlaşmış K, Çakır Güngör AN. Evaluation of the effects of menopause on the musculoskeletal system with experimental rat models. Turk J Osteoporos. 2025;31(3):124-9

Introduction

Estrogens are a group of hormones that include E1, E2, E3 and E4 (1). They act through classical nuclear estrogen receptors (ERs) and the membrane-type G protein-coupled ER. ERs are "nuclear steroid receptors" in the NR3 class of the nuclear receptor superfamily. They act through the most familiar ERα (NR3A1) and ERβ (NR3A2) receptors, which have the same structural features but different distribution in tissues. They bind directly to DNA and regulate gene expression like ligands that activate transcription factors. The non-genomic effect of estrogen is to regulate gene expression by regulating protein kinase signaling cascades (2,3). Estrogen stimulates muscle stem cells (satellitic cells) proliferation and differentiation, thereby increasing muscle mass and strength to maintain muscle health. Collins et al. (4) conducted a study to show that estrogen deficiency negatively affects satellitic cells and impairs muscle fiber self-renewal and differentiation. They found that estradiol and ERα promote satellitic cells by downregulating genes associated with mitochondrial caspase-induced apoptosis.

Osteoblasts and osteoclasts are affected by systemic and sex hormones. ERα and ERβ receptors are found in both osteoblasts and osteoblast progenitors. Estrogen deficiency increases osteoblast apoptosis and decreases bone formation time and osteoclast apoptosis. Therefore, bone destruction continues for a longer period of time and thus bone remodelling imbalance occurs. Also during estrogen deficiency, an increase in osteocyte apoptosis is observed in humans and animals (1).

Bone tissue secretes osteokines (osteocalcin, fibroblast growth factor-23 and sclerostin) and muscle tissue secretes cytokines called myokines (interleukin-6, irisin and β-aminoisobutyric acid). These cytokines also produce autocrine and paracrine effects. This effect causes muscle and bone tissue to communicate and influence each other (1,2).

The subchondral bone regularly distributes the load and stress on the bone. Alterations here cause an uneven distribution, initiating or accelerating cartilage degeneration. Yang et al. (3) compared osteochondral changes by studying post-traumatic (PT) and oophorectomized (OVX) osteoarthritis (OA) models (5). Rapid cartilage degeneration and increased bone formation were observed in PT-OA, whereas only mild cartilage erosion, proteoglycan and significant bone loss were observed in OVX-OA. In addition, subchondral bone degradation in OVX-OA occurred 6 weeks before cartilage degeneration. Transforming growth factor (TGF)-β downregulation was observed in the osteochondral unit of OVX rats. Estrogen supplementation prevented subchondral bone loss, cartilage degradation and TGF-β downregulation. TGF-β has been shown to be a regulator of both subchondral bone and cartilage. Developing therapies targeting TGF-β in menopause-induced OA may be considered as a future treatment alternative by acting on both bone and cartilage tissue.

Many experimental studies have been conducted to reduce and stop the negative effects of menopause on the musculoskeletal system. In each of these studies, different active substances were used and effects were monitored through different pathways. Whether these active substances will be a treatment alternative in the future is a matter of debate. This article aims to comprehensively present the studies examining the effects of menopause on the musculoskeletal system in rats.

Active Ingredients and Potential Treatment Alternatives Aiming to Reduce the Effects of Menopause on the Musculoskeletal System

1.1. Exercise

Exercise-induced oxidative stress may differ between genders in humans and animals. In a 2017 study, these differences were evaluated in a male and female rat model that was jogged for 6 weeks (4). Exercise was shown to improve the oxidative capacity of mitochondrial function in both male and female rats, but this was more pronounced in males. In sedentary rats, oxidative stress resistance has been shown to be higher in female rats. It has been suggested that this may be due to the effects of estrogen. In another study, it was shown that estrogen deficiency in rats decreased mitochondrial respiratory complex I activity in muscle and this was corrected by estrogen treatment (6). In these studies, we see that estrogen hormone, which also affects mitochondria, which we know as the power plant; also affects the energy balance.

1.2. Myostatin

Myostatin (MSTN) is one of the key factors involved in communication between muscle and bone tissue. MSTN signaling, a member of the TGF-β superfamily, is thought to accompany aging in musculoskeletal tissues and lead to loss of muscle and bone mass. Tang et al. (7) established an osteoporosis (OP) model in OVX rats. Low-intensity pulsed ultrasound (LIPUS) application in OVX rats has been shown to have positive effects on MSTN expression inhibition and consequently prevention of bone loss and healing of bone damage. It has also been suggested that these effects may be related to MSTN/Wnt/β-catenin signaling pathways. According to the study, LIPUS application may be effective in the treatment of OP.

1.3. Irisin

The effect of exercise on the musculoskeletal system is closely related to Irisin. Irisin is a myokine that affects thermogenesis, energy expenditure and glucose homeostasis. It is released by exercise and targets bone tissue. Irisin's mechanism of action is mediated by peroxisome ɣ and its coactivator-1 α receptors that regulate thermogenesis via mitochondria. Irisin increases energy expenditure, promotes weight loss, and reduces diet-induced insulin resistance. Nyugen et al. (8) evaluated the effect of irisin injection on OP in OVX rats. It was suggested that intermittent treatment with irisin has a role in bone health and may be a valuable target for the treatment of postmenopausal OP.

1.4. Nacre

In the exoskeleton of mollusk's, the part consisting of calcium carbonate ($CaCO_3$) embedded in the organic matrix is called Nacre. The effect of Nacre on bone cells has been documented by *in vitro* studies Lin et al. (9) examined the effect of a 90-day long Nacre-supplemented diet on both bone mass and bone strength using 16-month-old C57BL/6 mice (10). Nacre supplementation increased bone strength by limiting gene expression related to osteoclast activity. It also reduced cortical bone loss by decreasing pore formation in cortical bone. It has also been shown that bone dynamics in trabecular bone, which decreases with aging, is preserved with Nacre diet (10).

1.5. Lycopene

Oxidative stress induced by reactive oxygen species is increased in aging or inflammatory conditions. This may adversely affect bone homeostasis. Carotenoids, which are antioxidant substances, have the potential to reduce these negative effects. Lycopene is one of the compounds in the carotenoid group. Oliveira et al. (11) evaluated the activity of osteoblasts cultured with lycopene from OVX rats for 8 weeks using *in vitro* analyses and also evaluated the femoral epiphyseal osteocytes and trabecular microarchitecture of 8-week-old ovectomised rats receiving lycopene by oral gavage using microtomography and stereological analyses. In addition to biochemical analyses, gene expression was investigated by molecular studies. The data obtained showed an increase in alkaline phosphatase in situ detection and mineralization in the group receiving lycopene. The previous report showed no difference in mineralization with the presence of lycopene after 14 days of culture. In conclusion, further studies are needed to clarify the issue. Quantitative expression of genes encoding important proteins and considered as bone markers was analysed using polymerase chain reaction. Lycopene significantly upregulated Runx2 and Bglap expression. Stereological analysis showed that lycopene treatment caused a significant increase in the number of osteoblasts, but no change in their volume, indicating that lycopene has an effect on osteoblastogenesis. There was also a decrease in the volume and number of osteoclasts. The number of osteocytes was significantly increased compared to the OVX group despite the decrease in their volume. No significant difference was found in microtomographic analyses (12).

1.6. Superjami

Superjami is a dark purple rice variety cultivated in Korea. It is characterized by its strong antioxidant activity as it is particularly rich in cyanidin-3-glucoside and anthocyanins. Chung et al. (13) evaluated the effects of ethanol extracts obtained from dark purple Superjami rice (*Oryza sativa* L. Cv. Superjami) bran on bone metabolism and antioxidant defense systems in OVX rats in an 8-week study. Bone turnover was significantly decreased in rats receiving Superjami rice bran extract supplementation as evidenced by the decrease in Alkaline Phosphatase, osteocalcin and CTx (C terminal telopeptide) amounts. The supplementation was shown to significantly improve bone metabolism and reduce bone loss in OVX rats. In addition, it was also found to significantly suppress oxidative stress and increase the activity of antioxidant enzymes. In another study conducted in 2019 and lasting 8 weeks, the effects of germinated and non-germinated Superjami diets were examined in OVX rats. The bioactive components contained between the two groups were found to be different. Both germinated and ungerminated Superjami rice flour were found to significantly reduce body weight gain, body fat, glucose and insulin levels and adipokine concentrations. In addition, it has been shown to significantly improve the antioxidant defense system and bone metabolism. In addition to all these, germinated Superjami has been shown to reduce body weight more than its ungerminated form and to have a greater effect on glucose homeostasis, antioxidative activities and bone metabolism. It has also been shown that germinated Superjami may be more effective and may be more useful in the treatment of menopausal hyperglycemia, oxidative stress and bone turnover imbalance (14).

1.7. Tocotrienol

Tocotrienol, which shows vitamin E activity, acts as an antioxidant. Since its bioavailability is low in oral administration, it is thought that solubilisation with self-emulsifying systems (SEDDS) facilitates its absorption by the lymphatic system and increases its bioavailability. Assuming that annatto (*Bixa orellana*) tocotrienol formulated with SEDDS has stronger effects on the skeleton than unformulated tocotrienol, the study was conducted in May 2021. The efficacy of SEDDS-formulated annatto-tocotrienol on bone parameters and oxidative stress markers was evaluated in OVX rats. In rats divided into 4 groups: OVX, OVX +unformulated AnTT, OVX +formulated AnTT-SEDDS and OVX +raloxifene (a selective ER modulator known to benefit the skeletal system). After 8 weeks of oral administration, blood levels of delta-tocotrienol and oxidative stress markers were analyzed, and microcomputed tomography, calcium content and biomechanical strength analyses were performed on the femur. Plasma delta-tocotrienol level was significantly higher in the AnTT-SEDDS group than in the AnTT group. AnTT-SEDDS group improved bone microarchitecture in rats by increasing trabecular thickness and trabecular number. However, these two parameters were not found to be dramatically different between OVX and AnTT groups. Both forms of annatto tocotrienol preserved femoral bone calcium content in rats. It is noteworthy that only the AnTT-SEDDS group significantly increased bone stiffness compared to the OVX group. The AnTT and AnTT-SEDDS groups did not attenuate the effects of OVX on bone mineral density (BMD). The group using raloxifene as a positive control showed similar skeletal effects as annatto tocotrienol. In this study, raloxifene treatment preserved bone microarchitecture, calcium content and strength. There was no noticeable change in BMD in this group. AnTT and AnTT-SEDDS groups increased superoxide dismutase and glutathione peroxidase activities in ovariectomised rats. However, it

decreased the level of malonaldehyde, a marker of oxidative stress, only in the AnTT-SEDDS group. In this study, the effects of annatto tocotrienol on the skeletal system were proven and no significant difference was observed between the SEDDS groups. This may be due to the fact that the experiments were not long-term experiments that would create a clear difference between the two groups or oral diet administration (15).

1.8. Nanoparticles

One of the recent studies involving modern technology is the experiment conducted by Guo et al. (16). Using a remote-controlled release system with Nanoparticles (NPs) targeting bone tissue, OP treatment was approached from a different direction. In recent years, with the use of drugs that act systematically on bone tissue (estrogen, calcitonin, bisphosphonates, raloxifene and RANK ligand inhibitors, etc.), new treatment methods have been needed due to serious side effects due to high doses and frequent administration. The aim of this system is to achieve a magnetically remotely controllable drug release under the guidance of an external magnetic field. An OVX rat model was used in a 3-month experiment to investigate the effect of NPs in anti-OP. As a result of the study, NPs showed good stability, good biocompatibility and high encapsulation ability for E2. In addition, the release of these NPs has a temperature-dependent effect. The system used improved OVX-induced bone loss, increased bone strength and enabled new bone formation in extra-skeletal tissues with fewer side effects. The absence of significant toxic side effects was also considered as an advantage (16,17).

1.9. Sirtuin

Many studies have been conducted on the potential role of the Sirtuin (SIRT) gene family, known as NAD^+-dependent class III deacetylase enzymes, in various diseases. However, little is currently known about the effect of mammalian SIRT1 on ageing. In a study, it was found that SIRT1 expression promotes healthy aging but does not improve longevity. In a mouse experiment, SIRT1 was observed to be partially protective against the development of pathologies typically associated with ageing, such as glucose intolerance, OP and wound healing (18).

1.10. Resveratrol

Resveratrol (R) is a phytoestrogen structurally similar to natural and synthetic estrogens. It can bind to ERs in bone and cartilage cells, providing chondroprotective and osteoprotective effects (19). R, a polyphenol compound found in peanuts, blueberries, grapes and other plants, is a natural activator of SIRT1 (20). The effect of R on bone mass in OVX rats and the role of SIRT1 in the maintenance of bone mass during perimenopause and early post-menopause were investigated (21). The rats were divided into 4 groups: 1- control group; 2-OVX; 3-OVX+low dose resveratrol; and 4-OVX + high dose resveratrol. One week after the surgery, rats in groups 3 and 4 were given R orally for 10 weeks. As a result, positive effects on serum osteoprotegrin, SIRT1 protein, Wnt/β-catenin signaling, BMD and bone microarchitecture in rat femurs and negative effects on RANKL were observed. In another study, control group, OVX and OVX groups given intravaginal R gel were formed and the experiment lasted for about 5 weeks. Knee joint tissues (articular cartilage, subchondral plate, subchondral bone) were evaluated by histomorphometry. In addition, mammalian target of rapamycin protein complex (mTOR), protein tyrosine phosphatase and tensin homologue, caspase 3 (cysteine proteinase involved in apoptosis) and [B-cell lymphoma/leukaemia-2 (BCL-2), antiapopitotic agent] expression in articular cartilage and subchondral bone were immunohistochemically evaluated. R treatment prevented weight gain by 17%. Trabecular bone degradation was attenuated due to upregulation of BCL-2 and downregulation of Casp-3. mTOR expression was downregulated. This effect prevented chondrocyte hypertrophy and maintained cartilage homeostasis. It was also found that intravaginal R treatment had systemic effects, decreased weight gain and increased oestradiol levels in OVX rats. Histological examination of joint tissues in OVX rats confirmed the protective effect of R treatment against degenerative changes in articular cartilage and trabecular bone resorption (19).

1.11. Edible Bird's Nest

In recent years, numerous studies have been conducted on the pharmacological effect and mechanism of Edible Bird's Nest (EBN). These studies are on antiviral effect, immune regulation, cognitive functions, neurodegenerative diseases and antioxidant issues. EBN has also been shown to control arthritis and may support the regeneration of chondrocytes (22). In a study conducted in 2019, administration of estrogen and EBN in OVX rats increased estrogen levels and ER expression in bone tissue, causing pro-osteoplastic hormone production and increased bone density (23). At this point, EBN can be considered as an alternative in hormone replacement therapies.

1.12. Opuntia ficus-indica

A study was conducted in 2020 to reveal the effect of Opuntia ficus-indica, which has been shown to have antioxidant and anti-inflammatory effects, on calcium bioavailability. In a 9-week period, the effect of O. ficus-indica in terms of physical, densitometric, biomechanical, microstructural and mineral content in bones in OVX rat model was evaluated. In rats receiving O. ficus-indica diet, calcium bioavailability and absorption, fracture resistance of bones and BMD were found to be high. No dose-based evaluation was performed in this study, but O. ficus-indica may be considered as a future therapeutic hope (24).

1.13. Parathormone (PTH) treatment

The effects of intermittent parathormone (PTH) treatment on vertebral body, tibia, BMD in distal femoral metaphysis, trabecular structures and femoral neck in OVX rats have been investigated in studies. In a 12-week study by Wang et al. (20), fracture resistance of the femoral neck was found to be highly correlated with two parameters, CSA (cross-sectional area) and bone strength index of the cortical bone in the femoral neck.

Using micro-CT, it has been demonstrated that intermittent PTH contributes to the cancellous and cortical bones of the femoral neck in OVX rats. Thus, it was shown that this method may be useful in the treatment of OP (22,24).

1.14. Abeliophyllum Distichum Nakai

Abeliophyllum distichum Nakai (AD), called Miseon, is one of the endemic species in Korea. In an 8-week experiment using OVX rat model, the effects and mechanism of orally administered AD-ethyl acetate fraction (EA) extract were investigated. Femur bone parameters were measured using micro-CT. AD-EA was found to inhibit osteoclast differentiation and bone resorption by inhibiting osteoclast-related gene expression via MAPK and c-fos/NFATc1 pathways. AD-EA also inhibited CTK and TRAcP activation, preventing bone loss caused by estrogen deficiency *in vivo* (25).

1.15. Marantodes pumilum (Blume) Kuntze

Wnt/β-catenin signaling is an important cellular pathway involved in osteoblast activation. This pathway is crucial for the regulation of bone formation and destruction. Marantodes pumilum (Blume) Kuntze (MP), also known as Kacip Fatimah, is a popular herb widely used in Southeast Asia. The protective effect of MP on bone has been attributed to phytoestrogenic compounds found in this plant. In the study, Marantodes pumilum leaf aqueous extract (MPLA) (50 mg/kg/day and 100 mg/kg/day) and estrogen were treated for 28 days. The results of the experiment showed that MPLA was able to ameliorate bone loss by increasing the level and distribution of osteoblastogenesis proteins in bone, and to maintain bone mass and collagen content close to normal in DM associated with estrogen deficiency. MPLA has also been found to prevent osteoblast apoptosis in the presence of sex steroid deficiency and DM. Treatment with 100 mg/kg/day MPLA had a greater effect than estrogen, whereas treatment with 50 mg/kg/day MPLA had less effect than estrogen (26).

1.16. Fructans

Fructans are indigestible carbohydrates found as storage polysaccharides in many higher plants. They also have numerous biological activities, including antioxidant, immunomodulatory, anti-inflammatory, anticancer, anti-hyperglycemic and prebiotic activities, providing many benefits to human health. Fructans also increase calcium absorption to prevent bone loss and OP (26). Estrogen deficiency during menopause as well as decreased dietary calcium intake cause serious problems on bone tissue. In the study by Topolska et al. (21) the effects of diet containing only fructan or enriched with "strawberry matrix" on bone tissue were investigated in calcium hypoalimentated OVX rats. As a result, inulin type fructans improved bone quality. Bone density was also increased in rats fed the given diet.

Conclusions and Future Perspective

The negative effects of menopause on the musculoskeletal system occur through estrogen deficiency and the resulting cellular and molecular mechanisms. Experimental studies focus on different treatment alternatives to reduce or stop these effects. In particular, various agents such as MSTN inhibitors, irisin, lycopene, nacre, superjami rice, and tocotrienol stand out as potential treatment options to improve postmenopausal musculoskeletal health. Each of these agents acts through different biochemical pathways targeting muscle and bone loss caused by estrogen deficiency. However, more research is needed before these substances can be put into clinical use. Although the experimental studies are promising, new studies are needed to fully understand and treat the effects of menopause on the musculoskeletal system.

Footnotes

Authorship Contributions

Concept: K.K., A.N.Ç.G., Design: K.K., A.N.Ç.G., Data Collection or Processing: N.H., Analysis or Interpretation: K.K., A.N.Ç.G., Literature Search: N.H., Writing: N.H., K.K.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declare that this study received no financial support.

References

1. Lingyun L, Tian L. Postmenopausal osteoporosis coexisting with sarcopenia: the role and mechanisms of estrogen. J Endocrinol. 2023;259:1-19.
2. Yoh K, Ikeda K, Horie K, Inoue S. Roles of estrogen, estrogen receptors, and estrogen-related receptors in skeletal muscle: regulation of mitochondrial function. Int J Mol Sci. 2023;24:1853.
3. Yang Z, Tan Q, Zhao Z, Niu G, Li S, Li W, et al. Distinct pathological changes of osteochondral units in early OVX-OA involving TGF-β signaling. Front Endocrinol. 2022;13:1074176.
4. Collins BC, Arpke RW, Larson AA, Baumann CW, Xie N, Cabelka CA, et al. Estrogen regulates the satellite cell compartment in females. Cell Rep. 2019;28:368-81.
5. Torres MJ, Ryan TE, Lin CT, Zeczycki TN, Neufer PD. Impact of 17β-estradiol on complex I kinetics and H_2O_2 production in liver and skeletal muscle mitochondria. J Biol Chem. 2018;293:16889-98.
6. Farhat F, Amérand A, Simon B, Guegueniat N, Moisan C. Gender-dependent differences of mitochondrial function and oxidative stress in rat skeletal muscle at rest and after exercise training. Redox Rep. 2017;22:508-14.
7. Tang L, Kang Y, Sun S, Zhao T, Cao W, Fan X, et al. Inhibition of MSTN signal pathway may participate in LIPUS preventing bone loss in ovariectomized rats. J Bone Miner Metab. 2020;38:14-26.
8. Nguyen DK, Vanden-Bossche A, Laroche N, Thomas M, Linossier MT, Peyroche S, et al. Dietary supplementation with nacre reduces cortical bone loss in aged female mice. Exp Gerontol. 2023;184:112337.
9. Lin JB, Wu H, Liu YL, Shaw PC, Li PB. N16 suppresses RANKL-mediated osteoclatogenesis by down-regulating RANK expression. Int J Biol Macromol. 2020;151:1154-62.

10. Semeghini MS, Scalize PH, Coelho MC, Fernandes, RR, Pitol DL, Tavares MS, et al. Lycopene prevents bone loss in ovariectomized rats and increases the number of osteocytes and osteoblasts. J Anat. 2022;241:729-40.
11. Oliveira GR, Vargas-Sanchez PK, Fernandes RR, Ricoldi MST, Semeghini MS, Pitol, DL, et al. Lycopene influences osteoblast functional activity and prevents femur bone loss in female rats submitted to an experimental model of osteoporosis. J Bone Miner Metab. 2019;37:658-67.
12. Chung SI, Ryu SN, Kang MY. Changes in bone metabolism and antioxidant defense systems in menopause-induced rats fed bran extract from dark purple rice (*Oryza sativa* L. Cv. Superjami). Nutrients. 2021;13:2926.
13. Chung SI, Ham TH, Kang MY. Effect of germinated pigmented rice "Superjami" on the glucose level, antioxidant defense system, and bone metabolism in menopausal rat model. Nutrients. 2019;11:2184.
14. Mohamad NV, Ima-Nirwana S, Chin KY. Therapeutic potential of annatto tocotrienol with self-emulsifying drug delivery system in a rat model of postmenopausal bone loss. Biomed Pharmacother. 2021;137:111368.
15. Wang SP, Chen YJ, Hsu CE, Chiu YC, Tsai MT, Hsu JT. Intermittent parathyroid hormone treatment affects the bone structural parameters and mechanical strength of the femoral neck after ovariectomy-induced osteoporosis in rats. BioMed Eng OnLine. 2022;21:6.
16. Guo Y, Liu Y, Shi C, Wu T, Cui Y, Wang S, et al. Remote-controllable bone-targeted delivery of estradiol for the treatment of ovariectomy-induced osteoporosis in rats. J Nanobiotechnol. 2021;19:1-18.
17. Herranz D, Muñoz-Martin M, Cañamero M, Mulero F, Martinez-Pastor B, Fernandez-Capetillo O, et al. Sirt1 improves healthy ageing and protects from metabolic syndrome-associated cancer. Nat Commun. 2010;12:3.
18. Zaychenko G, Belenichev I, Hnatiuk V, Doroshenko A, Sinitsyna O, Sulaieva O, et al. Protective effect of vaginal resveratrol administration on joint tissues in ovariectomized rats: targeting mTOR and caspase 3. Biomed Pharmacother. 2023;165:115176.
19. Koushki M, Amiri-Dashatan N, Ahmadi N, Abbaszadeh HA, Rezaei-Tavirani M. Resveratrol: a miraculous natural compound for diseases treatment. Food Sci Nutr. 2018;6:2473-90.
20. Wang X, Lu C, Chen Y, Wang Q, Bao X, Zhang Z, Huang X. Resveratrol promotes bone mass in ovariectomized rats and the SIRT1 rs7896005 SNP is associated with bone mass in women during perimenopause and early postmenopause. Climacteric. 2023;26:25-33.
21. Topolska K, Radzki RP, Filipiak-Florkiewicz A, Bienko M, Florkiewicz A, Cieslik E. Could fructan sources in strawberry matrix be more effective as a tool for improvement of bone structure than these compounds added to diet alone?-Study on osteopenic rat model. Ann Agric Environ Med. 2020;27:19-28.
22. Hou ZP, Tang SY, Ji HR, He PY, Li YH, Dong XL, et al. Edible bird's nest attenuates menopause-related bone degeneration in rats via increaing bone estrogen-receptor expression. Chin J Integr Med. 2021;27:280-5.
23. Nam DG, Yang HS, Bae UJ, Park E, Choi AJ, Choe JS. The cactus (*Opuntia ficus-indica*) cladodes and callus extracts: a study combined with lc-ms metabolic profiling, in-silico, and in-vitro analyses. Antioxidants (Basel). 2023;12:1329.
24. Kim EY, Kim JH, Kim M, Park JH, Sohn Y, Jung HS. Abeliophyllum distichum Nakai alleviates postmenopausal osteoporosis in ovariectomized rats and prevents RANKL-induced osteoclastogenesis in vitro. J Ethnopharmacol. 2020;257:112828.
25. Karim K, Giribabu N, Salleh N. Marantodes pumilum (blume) Kuntze (Kacip Fatimah) leaves aqueous extract prevents downregulation of Wnt/β-catenin pathway and upregulation of apoptosis in osteoblasts of estrogen-deficient, diabetes-induced rats. J Ethnopharmacol. 2021;280:114236.
26. Wang M, Cheong KL. Preparation, structural characterisation, and bioactivities of fructans: a review. Molecules. 2023;28:1613.

Original Investigation / Orijinal Araştırma

DOI: 10.4274/tod.galenos.2024.69862
Turk J Osteoporos 2025;31(3):130-7

Can the Systemic Immune Inflammation Index (SII) Indicate Disease Activity in Patients with Rheumatoid Arthritis?

Sistemik İmmün Enflamasyon İndeksi (Sİİ) Romatoid Artritli Hastalarda Hastalık Aktivitesini Gösterebilir mi?

Salim Mısırcı[1], Alev Alp[1], Büşra Başar Yılmaz[2]

[1]Bursa Uludağ University Faculty of Medicine, Department of Physical Medicine and Rehabilitation, Bursa, Türkiye
[2]Bursa Uludağ University Faculty of Medicine, Department of Biostatistics, Bursa, Türkiye

Abstract

Objective: This study aimed to investigate the clinical utility of the systemic immune-inflammation index (SII) as a potential marker for disease activity in patients with rheumatoid arthritis (RA).

Materials and Methods: A total of 104 RA patients aged between 18 and 65 years, along with 69 healthy controls, were included. RA patients were categorised into two subgroups based on the Disease Activity Score-28-Erythrocyte Sedimentation Rate (DAS-28-ESR): remission (DAS-28-ESR <2.6, n=51) and active disease (DAS-28-ESR >2.6, n=53). Associations between SII and conventional inflammatory markers as well as clinical disease activity indices were examined. Receiver operating characteristic analysis was conducted to assess the diagnostic performance of SII in distinguishing active RA.

Results: SII levels were significantly elevated in both the overall RA group and particularly in the active RA subgroup ($p<0.001$). SII showed strong positive correlations with C-reactive protein ($r_s=0.627$, $p<0.001$), ESR ($r_s=0.383$, $p<0.001$), DAS-28-ESR ($r_s=0.775$, $p<0.001$), the simplified disease activity index ($r_s=0.796$, $p<0.001$), and the clinical disease activity index ($r_s=0.798$, $p<0.001$). The most effective SII threshold for identifying active RA was 479.36×10^9/L, with an area under the curve of 0.968 (95% confidence interval: 0.914-0.993), yielding a sensitivity of 92.45% and specificity of 86.27%.

Conclusion: SII appears to be a valuable, accessible marker for assessing disease activity in patients with rheumatoid arthritis.

Keywords: Erythrocyte sedimentation rate, rheumatoid arthritis, systemic immune-inflammation index

Öz

Amaç: Bu çalışmada, sistemik immün enflamasyon indeksinin (SII), romatoid artrit (RA) hastalarında hastalık aktivitesini değerlendirmedeki yararlılığı araştırıldı.

Gereç ve Yöntem: Çalışmaya yaşları 18-65 arasında değişen 104 RA hastası ile 69 sağlıklı gönüllü dahil edildi. RA hastaları, hastalık aktivite skoru-28 (DAS-28-ESR skoru) <2,6 olan remisyon grubuna (n=51) ve >2,6 olan aktif RA grubuna (n=53) olmak üzere ikiye ayrıldı. SII ile enflamatuvar belirteçler ve hastalık aktivite indeksleri arasındaki ilişkiler analiz edildi. RA hastalık aktivitesinin değerlendirilmesinde SII'nın tanısal gücünü belirlemek amacıyla alıcı çalışma karakteristik eğrisi analizi uygulandı.

Bulgular: SII düzeyleri, hem genel RA grubunda hem de aktif hastalık grubunda anlamlı olarak yüksekti ($p<0,001$). SII; C-reaktif protein ($r_s=0,627$, $p<0,001$), ESR ($r_s=0,383$, $p<0,001$), DAS28-ESR ($r_s=0,775$, $p<0,001$), simplifiye hastalık aktivite indeksi ($r_s=0,796$, $p<0,001$) ve klinik hastalık aktivite indeksi ($r_s=0,798$, $p<0,001$) ile pozitif yönde anlamlı korelasyon gösterdi. RA hastalık aktivitesini belirlemede SII için en uygun eşik değerin $479,36\times10^9$/L olduğu belirlendi (eğri altında kalan alan: 0,968; %95 güven aralığı: 0,914-0,993; duyarlılık: %92,45; özgüllük: %86,27).

Sonuç: SII, romatoid artritli hastalarda hastalık aktivitesini değerlendirmek için değerli ve erişilebilir bir belirteç gibi görünmektedir.

Anahtar kelimeler: Eritrosit sedimentasyon hızı, romatoid artrit, sistemik immün-enflamasyon indeksi

Corresponding Author/Sorumlu Yazar: Salim Mısırcı MD, Bursa Uludağ University Faculty of Medicine, Department of Physical Medicine and Rehabilitation, Bursa, Türkiye
E-mail: dr.salim-misirci@hotmail.com **ORCID ID:** orcid.org/0000-0002-9362-1855

Received/Geliş Tarihi: 20.10.2023 **Accepted/Kabul Tarihi:** 17.02.2024 **Epub:** 17.07.2025 **Publication Date/Yayınlanma Tarihi:** 05.12.2025

Cite this article as/Atıf: Mısırcı S, Alp A, Başar Yılmaz B. Can the systemic immune inflammation index (SII) indicate disease activity in patients with rheumatoid arthritis? Turk J Osteoporos. 2025;31(3):130-7

This study was previously presented as an oral presentation at the Turkish Rheumatology Congress 2023.

Introduction

Patients with rheumatoid arthritis (RA) experience limits in their physical function and everyday activities as well as a loss of their ability to work due to the advancement of joint destruction (1). Extra-articular involvement such as rheumatoid nodules, vasculitis, cardiovascular disease, pulmonary disease, neurological disease, gastrointestinal disease, renal disease and hematological diseases can be seen in the course of RA. Although it can occur at any age, it is most common in women in the third to fifth decade of life (2,3). Its prevalence is expressed as approximately 5 in every 1000 adults worldwide (1).

In addition to the existing acute phase reactants routinely used to assess the extent of inflammation in RA, studies have also been conducted on markers calculated from complete blood count (CBC) results, such as platelet (PLT)-to-lymphocyte ratio (PLR) and neutrophil-to-lymphocyte ratio (NLR) (4-9). While there are studies demonstrating an association between disease activity in RA and NLR and PLR (5,9), there are also studies reporting insignificant results (6,8).

The indices for assessing disease activity in RA are recognised and routinely used (10-12). Systemic Immune Inflammation index (SII) is derived by multiplying the PLT and neutrophil counts and then dividing the result by the lymphocyte count (13). It was evaluated in subjects with malignancies, depression in diabetic patients, hypertension, aphthous stomatitis, psoriasis, hidradenitis suppurativa, interstitial lung disease, non-infectious uveitis and ulcerative colitis (14-24). Studies have investigated its potential as an innovative biomarker for the evaluate of disease activity in various rheumatologic diseases such as ankylosing spondylitis, Behçet's disease, Adult Still's disease (AOSD) and antineutrophil cytoplasmic antibody-associated vasculitis (25-29).

The relationship between RA activity and SII has not been sufficiently investigated in the literature (30-32). A CBC is ordered at every routine examination of patients with RA. The SII, which is calculated based on the parameters of the CBC, can reduce costs if it indicates disease activity without the need for an additional inflammatory marker. Based on this information, we wanted to evaluate the benefits of SII in RA.

Materials and Methods

The study population consisted of patients who had received treatment at the Outpatient Clinic for Physical Therapy and Rehabilitation at Bursa Uludağ University (ethical approval date: February 23, 2022, protocol code: 2022-4/24). In accordance with the classification criteria (33), 104 patients with ages between 18 and 65 years who had been diagnosed with RA were enrolled in the study. Patients with comorbidities were excluded. The control group consisted of 69 healthy volunteers. Those included in the study were interviewed face-to-face during their application to the outpatient clinic. Data was collected on the participants' age, gender, level of education, occupation, disease duration, current treatments and laboratory findings. NLR, PLR and SII values were calculated using the CBC results. The values of the clinical disease activity index (CDAI), the disease activity score 28-erythrocyte sedimentation rate (DAS28-ESR), and the simplified disease activity index (SDAI) were calculated to assess disease activity. Patients with a DAS28-ESR <2.6 were categorized in the remission group, whereas those with DAS28-ESR > of 2.6 were classified in the active disease group (10).

Statistical Analysis

All statistical analyses were performed using IBM SPSS Statistics for Windows, Version 26.0 (IBM Corp., Armonk, NY). Quantitative data were presented as mean ± standard deviation or as median (minimum–maximum), depending on distribution characteristics. The normality of continuous variables was assessed using both the Shapiro-Wilk and Kolmogorov-Smirnov tests. Categorical variables were reported as frequencies and percentages (n, %). Comparisons between groups were carried out using the chi-square test for categorical variables, and either the independent samples t-test or the Mann-Whitney U test for continuous variables, depending on the distribution. Correlation analysis was performed to assess the relationships between SII and C-reactive protein (CRP), ESR, NLR, PLR, DAS28-ESR, SDAI, and CDAI, and the Spearman correlation coefficient was calculated. Receiver operating characteristic (ROC) curve analysis was used to determine the optimal cut-off values for SII, ESR, CRP, NLR, and PLR. A p-value of less than 0.05 was considered statistically significant.

Sample size estimation was based on data from a prior study on the same subject (30). Power analysis was conducted at a 99% confidence level, with an effect size of 0.84 and a Type I error rate of 5%. This analysis indicated that a minimum of 17 participants per group was necessary.

Results

The RA group was not statistically different from the healthy volunteers in terms of demographics (p>0.05). SII, neutrophil count, NLR, and PLR values were significantly elevated in the RA group compared to healthy controls, whereas lymphocyte levels were reduced. No statistically significant difference was observed between the groups regarding PLT counts (p=0.067) (Table 1).

RA patients in remission were compared to those with active disease to evaluate group differences. No differences were found with regard to gender, age, disease duration, occupation, educational level, drug use and lymphocyte count (p>0.05). SII, neutrophil, PLT, CRP, ESR, NLR, PLR, DAS28, SDAI and CDAI values were higher in the active RA group when compared with the RA patients in remission (p<0.05) (Tables 2 and 3). Figure 1 shows the flow chart for both the healthy control group and the patients with RA.

A positive correlation was observed between SII and CRP levels. (Figure 2a), ESR (Figure 2b), DAS28-ESR (Figure 2c), SDAI (Figure 2d), and CDAI (Figure 2e) variables. Correlations of SII with CRP, ESR, DAS-28, SDAI, and CDAI in RA patients are shown in Figure 2.

Table 1. Clinical, laboratory and demographic parameters of healthy controls and RA

	RA (n=104)	Control (n=69)	p
Age (year)	53 (20-68)	54 (18-64)	0.808[m]
Gender, n (%)			
Female	79 (76)	51 (73.9)	0.760[χ2]
Male	25 (24)	18 (26.1)	
Level of education, n (%)			
Uneducated	2 (1.9)		
Primary education	66 (63.5)		
High school	23 (22.1)		
University	13 (12.5)		
Job, n (%)			
Housewife	56 (53.8)		
Retired	18 (17.3)		
Employee	17 (16.3)		
Officer	8 (7.7)		
Freelancer	5 (4.8)		
Medications, n (%)			
NSAIDs	4 (2.8)		
bDMARDs	22 (15.6)		
csDMARDs	86 (60.0)		
JAK inhibitors (tofacitinib, baricitinib)	9 (6.3)		
Glucocorticoids	22 (15.6)		
Neutrophils (×10⁹/L)	4.32 (1.93, 8.16)	3.73 (2.11, 6.67)	**<0.001[m]**
Lymphocytes (×10⁹/L)	2.20 (± SD: 0.58)	2.64 (± SD: 0.57)	**<0.00[t]**
PLT (×10⁹/L)	259.40 (9.00, 572.00)	237 (135.00, 339.00)	0.067[m]
NLR	1.97 (1.03, 5.22)	1.43 (0.79, 2.23)	**<0.001[m]**
PLR	119.46 (11.61, 303.80)	96.70 (39.82, 160.31)	**<0.001[m]**
SII (×10⁹/L)	500.32 (186.44, 1938.28)	355.16 (167.65, 532.23)	**<0.001[m]**

bDMARDs: Biological disease-modifying anti-rheumatic drugs, csDMARDs: Conventional synthetic disease-modifying anti-rheumatic drugs, JAK inhibitors: Janus kinase inhibitors, [m]: Mann-Whitney U test, NLR: Neutrophil-to-lymphocyte ratio, NSAIDs: Non-steroidal anti inflammatory drugs, PLR: Platelet-to-lymphocyte ratio, PLT: Platelet, RA: Rheumatoid arthritis, SD: Standard deviation, SII: Systemic Immune-Inflammation Index, [t]: Independent samples t-test, [χ2]: Chi-square test

Table 2. Comparisons of clinical data and demographics between the remission group and of patients with RA

	Active RA (n=53)	Remission RA (n=51)	p
Age (year)	57 (32-68)	52 (20-65)	0.251[m]
Gender, n (%)			
Female	39 (73.6)	40 (78.4)	0.563[χ2]
Male	14 (26.4)	11 (21.6)	
Disease duration (year)	10 (1-38)	11 (1-40)	0.386[m]
Level of education, n (%)			
Uneducated	2 (3.8)	0 (0.0)	0.528[f]
Primary education	34 (64.2)	32 (62.7)	
High school	12 (22.6)	11 (21.6)	
University	5 (9.4)	8 (15.7)	

Table 2. Continued

	Active RA (n=53)	Remission RA (n=51)	p
Job, n (%)			
Housewife	31 (58.5)	25 (49.0)	0.648[f]
Retired	8 (15.1)	10 (19.6)	
Employee	9 (17.0)	8 (15.7)	
Officer	4 (7.5)	4 (7.8)	
Freelancer	1 (1.9)	4 (7.8)	
Medications, n (%)			
NSAIDs	3 (3.8)	1 (1.6)	0.350[f]
bDMARDs	11 (13.8)	11 (17.5)	
csDMARDs	43 (53.8)	43 (68.3)	
JAK-inhibitors (tofacitinib, baricitinib)	7 (8.8)	2 (3.2)	
Glucocorticoids	16 (20)	6 (9.5)	

bDMARDs: Biological disease-modifying anti-rheumatic drugs, csDMARDs: Conventional synthetic disease-modifying anti-rheumatic drugs, [f]: Fisher's exact test, JAK inhibitors: Janus kinase inhibitors, [m]: Mann-Whitney U test, NSAIDs: Non-steroidal anti inflammatory drugs, RA: Rheumatoid arthritis, χ^2: Chi-square test

Table 3. Comparisons of laboratory parameters between the remission group and of active group patients with RA

	Active RA (n=53)	Remission RA (n=51)	p
Neutrophils (×10⁹/L)	5.39 (± SD: 1.37)	3.58 (± SD: 0.79)	**<0.001[t]**
Lymphocytes (×10⁹/L)	2.10 (± SD: 0.59)	2.29 (± SD: 0.56)	0.089[t]
PLT (×10⁹/L)	287.00 (178.30, 572.00)	236.00 (145.00, 361.00)	**<0.001[m]**
NLR	2.45 (1.31, 5.22)	1.50 (1.03, 2.60)	**<0.001[m]**
PLR	135.93 (82.42, 303.80)	106.07 (56.10, 201.42)	**<0.001[m]**
SII (×10⁹/L)	670.75 (436.60, 1938.28)	370.74 (186.44, 571.26)	**<0.001[m]**
CRP (mg/L)	11.40 (2.00, 162.20)	2.00 (2.00, 9.50)	**<0.001[m]**
ESR (mm/h)	25 (2, 73)	9 (2, 27)	**<0.001[m]**
DAS28-ESR	4.90 (2.70, 6.70)	2.00 (1.10, 2.50)	**<0.001[m]**
SDAI	25.20 (4.30, 79.00)	4.30 (2.10, 10.80)	**<0.001[m]**
CDAI	24.00 (4.00, 40.00)	4.00 (2.00, 10.00)	**<0.001[m]**

CDAI: Clinical disease activity index, CRP: C-reactive protein, csDMARDs: Conventional synthetic disease-modifying anti-rheumatic drugs, DAS28-ESR: Disease activity score 28-erythrocyte sedimentation rate, ESR: Erythrocyte sedimentation rate, [m]: Mann-Whitney U test, NLR: Neutrophil-to-lymphocyte ratio, PLR: Platelet-to-lymphocyte ratio, PLT: Platelet, RA: Rheumatoid arthritis, SDAI: Simplified disease activity index, SD: Standard deviation, SII: Systemic immune inflammation index, [t]: Independent samples t-test

Areas under the ROC curve (AUC) can be listed as; 0.968 for SII (Figure 3) [95% confidence interval (CI): 0.914-0.993], 0.898 for NLR (Figure 3) (95% CI: 0.823-0.948), 0.784 for PLR (Figure 3) (95% CI: 0.692-0.859), 0.864 for CRP (Figure 3) (95% CI: 0.783-0.923) and 0.765 for ESR (Figure 3) (95% CI: 0.672-0.843). It was determined that SII and NLR variables were better in determining disease activity than CRP and ESR. For the SII to evaluate disease activity, the ideal cut-off point was 479.36 (×10⁹/L) (sensitivity: 92.45%, specificity: 86.27%). ROC curve analyzes are shown in Figure 3.

Discussion

In the present study, RA patients exhibited elevated levels of SII, neutrophils, NLR, and PLR compared to healthy controls, while lymphocyte counts were notably reduced in the RA group. SII, neutrophil, PLT, CRP, ESR, NLR, PLR, DAS28, SDAI and CDAI values were elevated in the active disease group. A strong positive association was identified between SII and the variables CRP, ESR, DAS28-ESR, SDAI, and CDAI. It was determined that SII and NLR variables were better in determining disease activity than CRP and ESR with an effect size of 0.84.

ESR, CRP, DAS-28, SDAI and CDAI are valid laboratory and clinical variables used in determining RA disease activity. Recently, studies have been conducted on NLR, PLR and SII values calculated with CBC parameters, which is an easily accessible and inexpensive method routinely requested at every control of patients to determine RA disease activity (5-9,30-32).

Neutrophils, lymphocytes and PLTs have an important role in inflammation. Neutrophils play a role by activating antigen-presenting cells and producing pro-oxidant mediators and lytic enzymes (9). Although it is controversial in some sources whether PLTs play a pro-inflammatory or anti-inflammatory role in the pathogenesis of RA (9), they have recently been reported to be active in exacerbating and maintaining inflammation (7).

While elevated PLT levels are difficult to detect in the joints of patients with inactive RA, numerous PLT-specific proteins can be identified in the synovial fluid and serum of those with active disease. Furthermore, during active RA, increased T-cell infiltration into the synovium leads to a relative depletion of T-cells in the peripheral blood, which is reflected as a decreased lymphocyte count in CBC measurements.

Consequently, patients with RA in the active phase should expect an increase in neutrophil and PLT counts, while lymphocyte counts should decrease (7). In our study, neutrophil and PLT counts were significantly higher in the active RA group compared to the remission group. The lymphocyte counts decreased, although the difference observed did not reach statistical significance.

Our study showed that active RA patients had higher values than the remission group when the NLR and PLR variables were taken into account. The RA group also had higher values than the healthy controls. Since NLR and PLR are variables calculated by a formula including neutrophils, lymphocytes, and PLT, it can be predicted that they can reflect the active RA disease.

According to Jin et al. (5), the RA group had higher NLR and PLR values than the other groups with rheumatic diseases and the healthy control groups. In their ROC curve analysis (AUC: 0.831, cut-off point: 2.13, sensitivity: 76.7%, specificity: 75.9%) they pretended to demonstrate the effectiveness of NLR in evaluating the disease activation, although it is less valuable than CRP, but more valuable than ESR. Consistent with the present results,

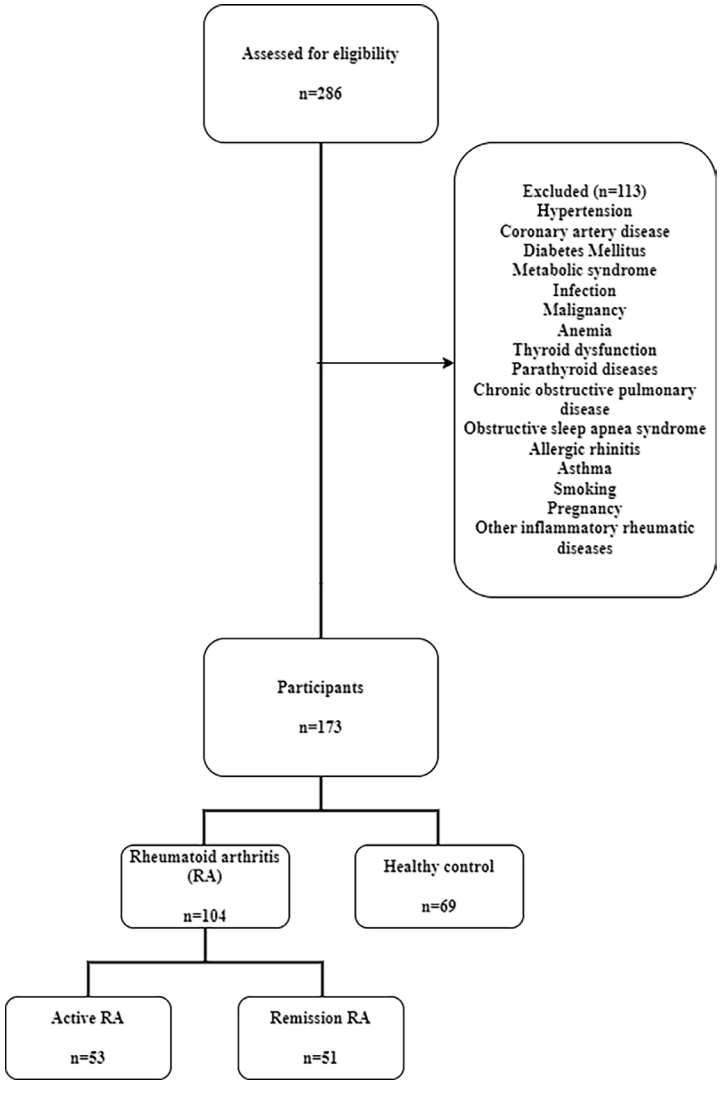

Figure 1. Flow chart
RA: Rheumatoid arthritis

both NLR and PLR values were high in our study. In the ROC curve analysis (AUC: 0.898, cut-off point: 2.08, sensitivity: 77.36%, specificity: 90.20%), the efficiency of NLR in determining RA disease activity was better than ESR and CRP. According to Chen et al. (7), patients with RA showed increased NLR and PLR values compared to controls. Analysis of the ROC curve showed that the cut-off value for PLR was 171.92 (AUC=0.676, sensitivity=61.28%, specificity=81.68%). A significant correlation was also found between PLR and DAS28. In conclusion, they noted that high PLR values are associated with an increased risk of RA. Erre et al. (9), stated that in a meta-analysis in which they included 13 of the studies published between 2015 and 2017, they confirmed that NLR and PLR values are associated with the presence of RA. Our results related to NLR and PLR values are compatible with the current literature.

SII variable can be more effective in demonstrating activity in RA patients than NLR and PLR, since neutrophil, lymphocyte and PLT values are calculated with a method that includes all three. SII was investigated in patients with malignant diseases (13-16). Apart from malignancies, studies have also been conducted in patients with aphthous stomatitis, psoriasis, hydradenitis suppurativa, interstitial lung disease, non-infectious uveitis and ulcerative colitis (19-24). The SII has been shown to be a useful index for monitoring disease activity in patients with ankylosing spondylitis, determining the activity of Behçet's disease, diagnosing AOSD and predicting the severity of psoriatic arthritis (25-27,34). It has also been reported to be an effective marker in studies on vasculitis (28,29).

There were three studies in the literature that investigated SII in RA patients (30-32). In one of these studies, some markers were calculated including SII in patients with active RA, AS and systemic lupus erythematosus. In contrast to healthy controls, RA patients had higher SII, which was statistically significant, but performed poorly in predicting RA disease activity (AUC: 0.622,

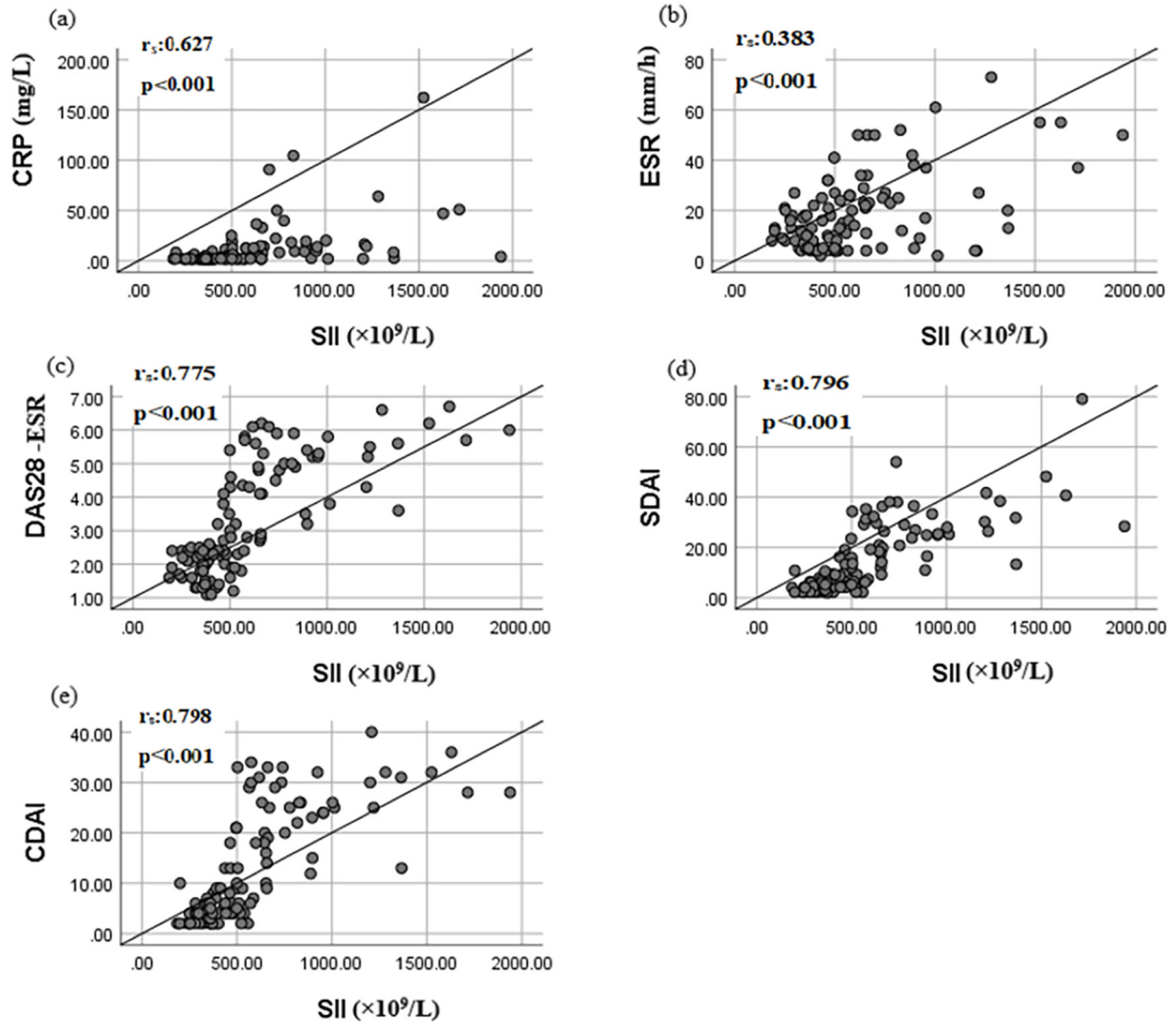

Figure 2. Correlations of SII with CRP (a), ESR (b), DAS28-ESR (c), SDAI (d) and CDAI (e)
CDAI: Clinical disease activity index, CRP: C-reactive protein, DAS28-ESR: Disease activity score 28-erythrocyte sedimentation rate, ESR: Erythrocyte sedimentation rate, RA: Rheumatoid arthritis, SDAI: Simplified disease activity index, SII: Systemic immune inflammation index, r_s: Spearmen correlation coefficient

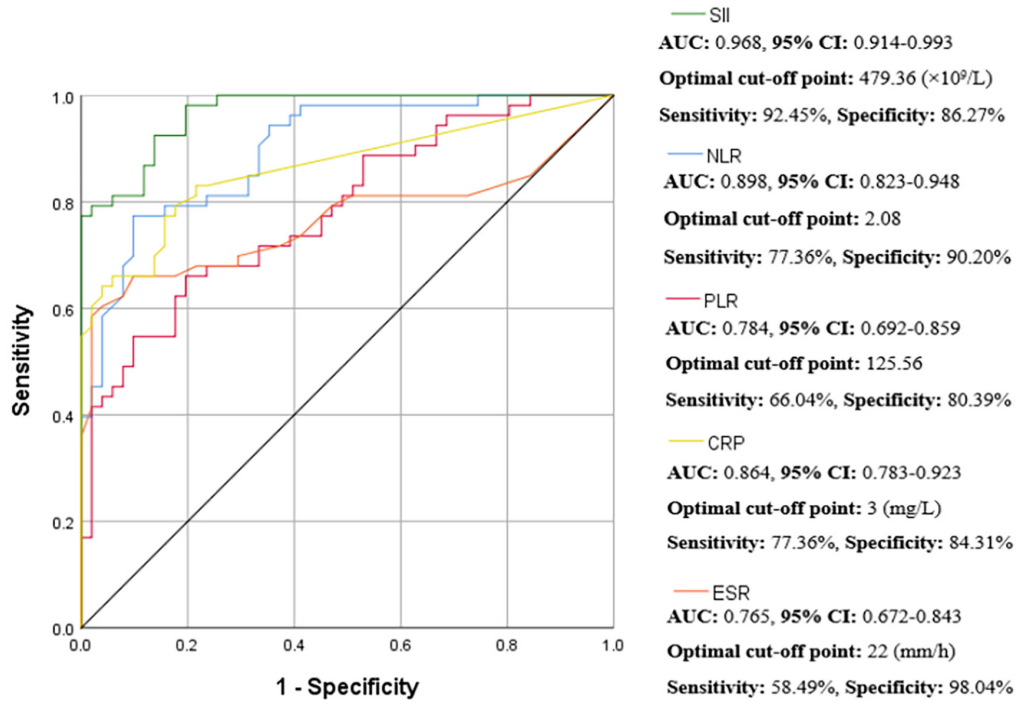

Figure 3. Receiver operating characteristic (ROC) curves
CRP: C-reactive protein, ESR: Erythrocyte sedimentation rate, NLR: Neutrophil-to-lymphocyte ratio, PLR: Platelet-to-lymphocyte ratio, SII: Systemic immune inflammation index, ROC: Receiver operating characteristic, AUC: Area under the curve

95% CI: 0.449-0.794, cut-off point: 691.55, sensitivity: 54.00%, specificity: 61.50%) (32). Compared to this study, SII was a good predictor of RA disease activity in our ROC curve analysis.

In another RA study on SII by Choe and Kim (31), evaluation with janus kinase (JAK) inhibitors was performed before and after 24 weeks of treatment. They determined that SII, NLR and PLR values in RA patients were higher than healthy controls in the initial evaluation. The present results are also consistent with our study. It was stated also that after treatment with JAK inhibitors for 24 weeks, there was a decrease in SII and NLR values.

A separate investigation carried out in Türkiye demonstrated elevated SII levels in RA patients relative to healthy controls, with even greater elevations observed in those with active disease compared to patients in remission. (p=0.002 and p=0.030, respectively) (30).

On the other hand, another study came to the conclusion that SII levels are not a reliable indicator of RA disease activity (32). Unlike previous studies, our research holds particular value due to the inclusion of a larger number of healthy controls and a well-balanced distribution between patients with active RA and those in remission and had SII values that were highly predictive of disease activity.

Study Limitations

The main limitation of our work is that it was a case control study conducted in a single centre. In addition, the SII scores (remission) of the patients in the active RA group could not be assessed after treatment. Future research with larger cohorts and prospective design is warranted to evaluate the efficacy of RA SII.

Conclusion

The SII could be an innovative indicator for evaluating disease activity in patients with RA.

Ethics

Ethics Committee Approval: The study population consisted of patients who had received treatment at the Outpatient Clinic for Physical Therapy and Rehabilitation at Bursa Uludağ University (ethical approval date: February 23, 2022, protocol code: 2022-4/24).

Informed Consent: Written informed consent was obtained from the patients.

Footnotes

Authorship Contributions

Surgical and Medical Practices: S.M., A.A., Concept: S.M., A.A., B.B.Y., Design: S.M., A.A., B.B.Y., Data Collection or Processing: S.M., A.A., Analysis or Interpretation: S.M., A.A., B.B.Y., Literature Search: S.M., A.A., Writing: S.M.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declare that this study received no financial support.

References

1. Aletaha D, Smolen JS. Diagnosis and management of rheumatoid arthritis: a review. JAMA. 2018;320:1360-72.
2. Figus FA, Piga M, Azzolin I, McConnell R, Iagnocco A. Rheumatoid arthritis: extra-articular manifestations and comorbidities. Autoimmun Rev. 2021;20:102776.
3. Ma Y, Hong FF, Yang SL. Role of prostaglandins in rheumatoid arthritis. Clin Exp Rheumatol. 2021;39:162-72.
4. Chandrashekara S, Sachin S. Measures in rheumatoid arthritis: are we measuring too many parameters. Int J Rheum Dis. 2012;15:239-48.
5. Jin Z, Cai G, Zhang P, Li X, Yao S, Zhuang L, et al. The value of the neutrophil-to-lymphocyte ratio and platelet-to-lymphocyte ratio as complementary diagnostic tools in the diagnosis of rheumatoid arthritis: a multicenter retrospective study. J Clin Lab Anal. 2021;35:e23569.
6. Lijuan W, Yuting Z, Chaoyang L, Ju Y. Neutrophil-lymphocyte, platelet-lymphocyte and lymphocyte-monocyte ratios may not be useful markers to assess disease activity in rheumatoid arthritis: a STROBE-compliant article. Medicine (Baltimore). 2021;100:e27631.
7. Chen Q, Chen DY, Xu XZ, Liu YY, Yin TT, Li D. Platelet/lymphocyte, lymphocyte/monocyte, and neutrophil/lymphocyte ratios as biomarkers in patients with rheumatoid arthritis and rheumatoid arthritis-associated interstitial lung disease. Med Sci Monit. 2019;25:6474-81.
8. Erre GL, Buscetta G, Mangoni AA, Castagna F, Paliogiannis P, Oggiano M, et al. Diagnostic accuracy of different blood cells-derived indexes in rheumatoid arthritis: a cross-sectional study. Medicine (Baltimore). 2020;99:e22557.
9. Erre GL, Paliogiannis P, Castagna F, Mangoni AA, Carru C, Passiu G, et al. Meta-analysis of neutrophil-to-lymphocyte and platelet-to-lymphocyte ratio in rheumatoid arthritis. Eur J Clin Invest. 2019;49:e13037.
10. van Riel PL. The development of the disease activity score (DAS) and the disease activity score using 28 joint counts (DAS28). Clin Exp Rheumatol. 2014;32(5 Suppl 85):S-65-74.
11. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G, et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. Rheumatology (Oxford). 2003;42:244-57.
12. Smolen JS, Aletaha D. Scores for all seasons: SDAI and CDAI. Clin Exp Rheumatol. 2014;32(5 Suppl 85):S-75-9.
13. Hu B, Yang XR, Xu Y, Sun YF, Sun C, Guo W, et al. Systemic immune-inflammation index predicts prognosis of patients after curative resection for hepatocellular carcinoma. Clin Cancer Res. 2014;20:6212-22.
14. Liu J, Li S, Zhang S, Liu Y, Ma L, Zhu J, et al. Systemic immune-inflammation index, neutrophil-to-lymphocyte ratio, platelet-to-lymphocyte ratio can predict clinical outcomes in patients with metastatic non-small-cell lung cancer treated with nivolumab. J Clin Lab Anal. 2019;33:e22964.
15. Chen JH, Zhai ET, Yuan YJ, Wu KM, Xu JB, Peng JJ, et al. Systemic immune-inflammation index for predicting prognosis of colorectal cancer. World J Gastroenterol. 2017;23:6261-72.
16. Jomrich G, Gruber ES, Winkler D, Hollenstein M, Gnant M, Sahora K, et al. Systemic immune-inflammation index (SII) predicts poor survival in pancreatic cancer patients undergoing resection. J Gastrointest Surg. 2020;24:610-8.
17. Wang J, Zhou D, Dai Z, Li X. Association between systemic immune-inflammation index and diabetic depression. Clin Interv Aging. 2021;16:97-105.
18. Çırakoğlu ÖF, Yılmaz AS. Systemic immune-inflammation index is associated with increased carotid intima-media thickness in hypertensive patients. Clin Exp Hypertens. 2021;43:565-71.
19. Tanacan E, Dinçer Rota D, Oktem R, Erdogan FG. The correlation of systemic immune-inflammation index, neutrophil-to-lymphocyte ratio, derived neutrophil-to-lymphocyte ratio, and platelet-to-lymphocyte ratio with disease severity in recurrent aphthous stomatitis. J Cosmet Dermatol. 2022;21:4858-63.
20. Dincer Rota D, Tanacan E. The utility of systemic-immune inflammation index for predicting the disease activation in patients with psoriasis. Int J Clin Pract. 2021;75:e14101.
21. Gambichler T, Hessam S, Cramer P, Abu Rached N, Bechara FG. Complete blood collection-based systemic inflammation biomarkers for patients with hidradenitis suppurativa. J Eur Acad Dermatol Venereol. 2022;36:1593-6.
22. Ruta VM, Man AM, Alexescu TG, Motoc NS, Tarmure S, Ungur RA, et al. Neutrophil-to-lymphocyte ratio and systemic immune-inflammation index-biomarkers in interstitial lung disease. Medicina (Kaunas). 2020;56:381.
23. Kurtul BE, Cakmak AI, Elbeyli A, Ozcan SC, Ozarslan Ozcan D, Kimyon G. Evaluation of systemic immune-inflammation index level as a novel marker for severity of noninfectious uveitis. Int Ophthalmol. 2021;41:3615-22.
24. Pakoz ZB, Ustaoglu M, Vatansever S, Yuksel ES, Topal F. serum immune-inflammation index assessment in the patients with ulcerative colitis. Gastroenterol Res Pract. 2022;2022:9987214.
25. Wu J, Yan L, Chai K. Systemic immune-inflammation index is associated with disease activity in patients with ankylosing spondylitis. J Clin Lab Anal. 2021;35:e23964.
26. Tanacan E, Dincer D, Erdogan FG, Gurler A. A cutoff value for the systemic immune-inflammation index in determining activity of Behçet disease. Clin Exp Dermatol. 2021;46:286-91.
27. Kim JW, Jung JY, Suh CH, Kim HA. Systemic immune-inflammation index combined with ferritin can serve as a reliable assessment score for adult-onset Still's disease. Clin Rheumatol. 2021;40:661-8.
28. Kim Y, Choi H, Jung SM, Song JJ, Park YB, Lee SW. Systemic immune-inflammation index could estimate the cross-sectional high activity and the poor outcomes in immunosuppressive drug-naïve patients with antineutrophil cytoplasmic antibody-associated vasculitis. Nephrology (Carlton). 2019;24:711-7.
29. Chen JB, Tang R, Zhong Y, Zhou YO, Zuo X, Luo H, et al. Systemic immune-inflammation index predicts a reduced risk of end-stage renal disease in Chinese patients with myeloperoxidase-anti-neutrophil cytoplasmic antibody-associated vasculitis: a retrospective observational study. Exp Ther Med. 2021;22:989.
30. Satis S. New inflammatory marker associated with disease activity in rheumatoid arthritis: the systemic immune-inflammation index. Curr Health Sci J. 2021;47:553-7.
31. Choe JY, Kim SK. Association between hematological indicesand disease activity in patients with rheumatoid arthritis treated with janus kinase inhibitors for 24 weeks. Medicina (Kaunas). 2022;58:426.
32. Taha SI, Samaan SF, Ibrahim RA, Moustafa NM, El-Sehsah EM, Youssef MK. Can complete blood count picture tell us more about the activity of rheumatological diseases? Clin Med Insights Arthritis Musculoskelet Disord. 2022;15:11795441221089182.
33. Kay J, Upchurch KS. ACR/EULAR 2010 rheumatoid arthritis classification criteria. Rheumatology (Oxford). 2012;51(Suppl 6:vi5-9).
34. Yorulmaz A, Hayran Y, Akpinar U, Yalcin B. Systemic immune-inflammation index (SII) predicts increased severity in psoriasis and psoriatic arthritis. Curr Health Sci J. 2020;46:352-7.

Original Investigation / Orijinal Araştırma

DOI: 10.4274/tod.galenos.2025.09475
Turk J Osteoporos 2025;31(3):138-44

Osteoporotic Fracture Forecasting in Türkiye: Analysis of the Agreement and Correlation Between FRAX and Garvan Fracture Risk Calculator Tools

Türk Postmenopozal Kadınlarda Osteoporotik Kırık Riski Değerlendirme Yaklaşımları: FRAX ve Garvan Kırık Risk Hesaplama Araçları Arasındaki Uyum ve Korelasyonun Analizi

Mustafa Hüseyin Temel[1], Seda Özcan İşler[2], Hatice Merve Beşışık Temel[3], Fatih Bağcıer[4], Emre Ata[1]

[1]University of Health Sciences, Sultan 2. Abdulhamid Han Training and Research Hospital, Department of Physical Medicine and Rehabilitation, İstanbul, Türkiye
[2]Üsküdar State Hospital, Clinic of Physical Medicine and Rehabilitation, İstanbul, Türkiye
[3]İstanbul Medeniyet University, Göztepe Prof. Dr. Süleyman Yalçın City Hospital, Department of Internal Medicine, İstanbul, Türkiye
[4]Başakşehir Çam ve Sakura City Hospital, Department of Physical Medicine and Rehabilitation, İstanbul, Türkiye

Abstract

Objective: To evaluate the correlation and agreement between fracture risk assessment tool (FRAX) and the Garvan fracture risk calculator (GFRC) in estimating the 10-year hip fracture risk in postmenopausal Turkish women with osteoporosis.
Materials and Methods: A retrospective cross-sectional study was conducted in the Üsküdar State Hospital. Medical records of 347 postmenopausal women aged between 50 and 90 were analyzed. Data on clinical risk factors were collected, and fracture probabilities were calculated using FRAX and GFRC tools. Spearman's rank correlation test was used to assess correlation, Wilcoxon signed-rank test was used to compare means, while the Interclass Correlation Coefficient (ICC) was calculated using two-way mixed effects model (ICC3) to evaluate agreement.
Results: FRAX 10-year hip fracture risk scores were significantly higher than GFRC 10-year hip fracture scores ($p<0.001$). The Spearman's correlation between FRAX and GFRC 10-year hip fracture risk scores was found to be strong ($r=0.821$, $p<0.001$). The ICC3 value was 0.054 [95% confidence interval (0.02, 0.11)].
Conclusion: In its current form, GFRC should be used complementarily rather than interchangeably for fracture risk prediction in postmenopausal Turkish women and should be reserved for specific patient populations. The results underline the need for population-specific calibrations for GFRC to improve predictive accuracy and clinical utility.
Keywords: Osteoporosis, fracture risk assessment tool (FRAX), Garvan fracture risk calculator (GFRC), postmenopausal women, hip fracture prediction

Öz

Amaç: Bu çalışmanın amacı postmenapozal osteoporoz tanılı Türk kadınlarında kırık risk değerlendirme aracı (FRAX) ve Garvan kırık risk hesaplayıcı (GFRC) arasında 10 yıllık kalça kırığı riski tahmininde korelasyon ve uyumu değerlendirmektir.
Gereç ve Yöntem: Bu çalışma Üsküdar Devlet Hastanesi'nde retrospektif kesitsel bir çalışma olarak gerçekleştirildi. Yaşları 50 ile 90 arasında değişen 347 postmenopozal osteoporoz tanılı kadın hastanın tıbbi kayıtları analiz edildi. Klinik risk faktörlerine ilişkin veriler toplandı ve FRAX ile GFRC araçları kullanılarak 10 yıllık kalça kırığı olasılıkları hesaplandı. Korelasyonu değerlendirmek için Spearman korelasyon testi, ortalamaları karşılaştırmak için Wilcoxon testi, uyumu değerlendirmek için ise iki yönlü karışık modelin Interclass Correlation Coefficient (ICC3) hesaplandı.
Bulgular: FRAX ile hesaplanan 10 yıllık kalça kırığı risk skorları, GFRC ile hesaplanan skorlarla karşılaştırıldığında istatistiksel olarak anlamlı şekilde daha yüksek bulundu ($p<0,001$). FRAX ve GFRC 10 yıllık kalça kırığı risk skorları arasında istatistiksel olarak anlamlı bir şekilde güçlü korelasyon bulundu ($r=0,821$, $p<0,001$). ICC3 değeri 0,054 [95% güven aralığı (0,02, 0,11)] olarak hesaplandı.

Corresponding Author/Sorumlu Yazar: Mustafa Hüseyin Temel MD, University of Health Sciences, Sultan 2. Abdulhamid Han Training and Research Hospital, Department of Physical Medicine and Rehabilitation, İstanbul, Türkiye
E-mail: mhuseyintemel@gmail.com **ORCID ID:** orcid.org/0000-0003-0256-5833
Received/Geliş Tarihi: 19.01.2025 **Accepted/Kabul Tarihi:** 04.02.2025 **Epub:** 06.10.2025 **Publication Date/Yayınlanma Tarihi:** 05.12.2025

Cite this article as/Atıf: Temel MH, Özcan İşler S, Beşışık Temel HM, Bağcıer F, Ata E. Osteoporotic fracture forecasting in Türkiye: analysis of the agreement and correlation between FRAX and Garvan fracture risk calculator tools. Turk J Osteoporos. 2025;31(3):138-44

Öz

Sonuç: Mevcut haliyle GFRC kırık riski tahmininde tek başına FRAX yerine kullanılmamalı, klinik süreçlerde mevcut algoritmalar açısından tamamlayıcı ve belirli hasta popülasyonlarına özel bir araç olarak değerlendirilmelidir. Çalışmanın sonucunda GFRC'nin daha doğru ve klinik olarak kullanışlı sonuçlar verebilmesi adına popülasyona özgü kalibrasyonlara ihtiyacı olduğu öngörülebilir.
Anahtar kelimeler: Osteoporoz, kırık risk değerlendirme aracı (FRAX), Garvan kırık risk hesaplayıcı (GFRC), postmenopozal kadın, kalça kırığı tahmini

Introduction

Osteoporosis is a bone disorder that develops due to the loss of bone mass and the alteration of bone structure, resulting in more fragile and fracture-prone bones (1). The disease is a matter of concern for more than 200 million people with this condition, as well as health professionals and insurers, because of its clinical consequences and economic costs (2). Moreover, the burden of the disease extends to increased healthcare expenditures, decline in productivity, and reduced quality of life for the affected population, primarily the elderly (3). The issue that needs to be addressed by policymakers is this increasing challenge that is expected to worsen because of the demographic transitions that occur with the population getting older (4).

Osteoporotic fractures, which are referred to as the most severe manifestations of osteoporosis, are hip, spine, and wrist fractures, with each of them causing considerable morbidity, mortality, and healthcare costs (5). Among them, hip fractures, which are the most important, usually result in long-term disability and higher mortality rates. The assessment of fracture risk is an essential tool in the prevention of healthcare strategies, and predictive algorithms are of paramount importance in the early detection of the individuals who are at risk (6).

The fracture risk assessment tool (FRAX) is one widely used algorithm designed to estimate the 10-year probability of fractures based on clinical risk factors and bone mineral density (BMD) (7). FRAX's utility lies in its integration of patient-specific risk factors with global epidemiological data, enabling clinicians to make informed treatment decisions. However, its reliability varies by population, emphasizing the need for validation studies across diverse demographics (8). Although FRAXplus holds great promise for the future, it is currently in beta testing and there is a risk that its widespread use may be limited by its cost (9).

The Garvan fracture risk calculator (GFRC), another predictive tool, includes additional variables such as fall history to enhance fracture risk assessment. While it offers a broader risk profile, the literature suggests that its clinical application requires further validation to confirm its generalizability and accuracy across different cultural settings (10).

Evaluating the correlation and agreement between different fracture prediction methods, like FRAX and GFRC within cultural contexts is essential for improving the precision of osteoporosis management globally. Such studies have the possibility to refine risk stratification and ensure the appropriateness of interventions tailored to specific populations (11). In the light of all this information, the aim of this study was to investigate the agreement and correlation levels of FRAX and GFRC in the Turkish population. To the best of our knowledge, this is the first study to compare FRAX and GFRC methods among the specific population of postmenopausal Turkish women with osteoporosis.

Materials and Methods

This retrospective cross-sectional study was conducted between Üsküdar State Hospital. The study was conducted in accordance with the Declaration of Helsinki and was approved by the University of Health Sciences Türkiye, Zeynep Kamil Women and Children Diseases Training and Research Hospital Ethics Committee (approval no: 11 date: 11.01.2023).

Patients under the age of 50 were not included due to the inability to calculate the GFRC and above the age of 90 due to inability to calculate the FRAX score. Due to potential differences in technicians and equipment, T-score and BMD were not used in the risk calculations. The medical records of postmenopausal female patients diagnosed with osteoporosis and followed up at the Agreement of FRAX and Garvan in Fracture Risk between September 1, 2023 and September 1, 2024, were reviewed. The data collected from the patient files included age, height, weight, current smoking status, high alcohol consumption (≥3 units/day), history of previous fragility fractures, number of falls in the past year, parental history of hip fracture, current or past use of glucocorticoids, presence of rheumatoid arthritis, type 1 diabetes, untreated long-standing hyperthyroidism, hypogonadism or premature menopause, chronic malnutrition or malabsorption, and chronic liver disease.

Before the calculations, the browser was operated in incognito mode to mitigate potential interference and bias. FRAX score computations were performed using the online tool at https://frax.shef.ac.uk/FRAX/tool.aspx?lang=tu, and the 10-year probability of hip scores were recorded. Similarly, GFRC calculations were carried out via the website https://fracturerisk calculator.com.au/calculator/ and 10-year risk of hip fracture scores have been documented. Due to the risk of time and the coverage of fracture locations differences adversely impacting the analysis, the 10-year probability of major osteoporotic fracture from FRAX and the 5-year risk scores and the risk of any fracture from GFRC were excluded from the study. The browser was restarted for each calculation to ensure consistency and avoid bias.

Statistical Analysis

The behavior of quantitative variables was assessed using centralization and variance measurements: Mean ± standard deviation. Wilcoxon signed-rank test was used to compare the distribution of continuous variables between two dependent groups that have non-normal distributed data. The Shapiro-Wilk test was used to determine the normality of the distribution. Spearman's rank correlation test was used to investigate a possible correlation. A correlation between 0.10 and 0.39 is considered weak, between 0.40 and 0.69 is considered moderate, between 0.70 and 0.89 is considered strong, and between 0.90 and 1 is considered very strong (12). For interclass correlation coefficient (ICC) values; values less than 0.5 are considered indicative of poor reliability, values between 0.5 and 0.75 are indicative of moderate reliability, values between 0.75 and 0.9 are indicative of good reliability, and values greater than 0.90 are indicative of excellent reliability (13). A significance level of p=0.05 was set for all analyses. Statistical analyses were conducted using the IBM SPSS (Statistical Package for the Social Sciences, Version 27.0, Armonk, NY, IBM Corp.) software package and Python version 3.11.10 (Python Software Foundation), utilizing "pandas", "pyreadstat" and "pingouin" librarioes for statistical analysis. "Plotly" library was used for visualizations.

Results

A total number of 347 patients were included in this study. Table 1 and Table 2 summarize the demographic and disease specific data of the participants.

The Saphiro-Willk test results demonstrated that FRAX 10-year probability of hip fracture and GFRC 10-year risk of hip osteoporotic fracture scores deviated from a normal distribution (p<0.001, p<0.001, respectively).

Figure 1 visually compares the distribution of FRAX 10-year probability of hip fracture and GFRC 10-year risk of hip osteoporotic fracture scores, with corresponding kernel density estimates providing a smoothed representation of the data. The results of the Wilcoxon signed-rank test indicated a statistically significant difference between the two measures, with FRAX scores being significantly higher (p<0.001).

Figure 2 presents a scatterplot illustrating the relationship between FRAX 10-year probability of hip fracture and GFRC 10-year risk of hip osteoporotic fracture scores, with a locally weighted scatterplot smoothing (LOWESS) curve applied to enhance visualization. The Spearman rank correlation analysis demonstrated a strong positive correlation between the two risk scores (r=0.821, p<0.001). The LOWESS curve revealed a non-linear relationship and variation across different score ranges.

An ICC calculation using two-way mixed effects model (ICC3) was performed to assess the agreement between the FRAX 10-year probability of hip fracture score and GFRC 10-year risk of hip osteoporotic fracture scores. The ICC3 value was determined to be 0.054 [95% confidence interval (0.02, 0.11)] (p<0.001).

Discussion

The data collated from the current research revealed a strong rank-order correlation between FRAX and GFRC instruments in postmenopausal Turkish women with osteoporosis, while the same instruments estimated fracture risks with moderate to low agreement. Also, it was observed that GFRC demonstrated lower scores than FRAX. Based on this information, it is evident that the Garvan fracture risk score must be used with the utmost care when determining fracture risk in this group of women and only for those who show specific characteristics, like increased risk of falls, and it should be taken as a supplementary measure.

The FRAX tool is an important instrument for assessing fracture risk in the Turkish population, as it predicts the 10-year probability of major osteoporotic and hip fractures using clinical risk factors, with or without BMD inputs. It is beneficial for risk stratification and cost-efficient population screening, especially in settings where diagnostic tools like DEXA scans are not widely available (14). A study on Turkish postmenopausal women with osteopenia demonstrated moderate agreement between FRAX predictions made with and without BMD, showcasing its utility even in resource-constrained environments (15). Additionally, FRAX models are calibrated to country-specific fracture and mortality rates, ensuring their relevance to local populations like Türkiye (16). Globally, FRAX is a validated and widely recognized reference tool, often incorporated into clinical guidelines to guide treatment strategies (17). It provides a robust framework for assessing fracture risk and planning prevention strategies for the Turkish population.

On the other hand, the GFRC is another vital instrument in detecting individual fracture risks since it focuses on clinical risk

Table 1. Mean, standard deviation, minimum and maximum values of age, height, weight, body mass index, FRAX 10-year probability of hip fracture score, and GFRC 10-year risk of hip osteoporotic fracture score of the participants

Variable	Mean	SD	Minimum	Maximum
Age	67.14	8.05	51.0	83.0
Height	154.08	6.60	135.0	170.0
Weight	62.46	9.21	44.0	93.0
Body mass index	26.42	4.44	18.40	40.05
FRAX 10-year probability of hip fracture	4.74	7.04	0.4	38.0
GFRC 10-year risk of hip osteoporotic fracture	1.44	1.61	0.1	9.0

FRAX: Fracture risk assessment tool (FRAX), GFRC: Garvan fracture risk calculator, SD: Standard deviation

factors such as fall history, and the number of prior fractures compared to FRAX. Because of this, the GFRC is especially helpful for evaluating patients who are vulnerable to falls, like the elderly, in whom FRAX may be less accurate (18). Also, the literature suggests that GFRC may be more advantageous in some cases than FRAX because of its flexibility, as it offers a single 10-year risk assessment and a 5-year risk assessment, which can be a plus for individuals (19). GFRC has also been observed in literature to be more sensitive in predicting future fractures in people with a fracture history than FRAX (20). In such cases, GFRC might be a better option for the "any fracture" prediction, even if it is usually better for predicting hip fractures (21). Addressing the specific limitations of FRAX, for instance, the omission of fall risk and offering personalized fracture predictions, GFRC stands as a valuable alternative for population subgroups in fracture risk evaluations.

The variance in fracture risk estimates predicted by the FRAX and GFRC tools among Turkish postmenopausal females remains in line with earlier research results. However, contrary to previous literature, FRAX scores were significantly higher than GFRC. This difference may be ascribed to several variables, such as methodological or demographic-specific factors. For example, FRAX omits some risk factors, like a history of falls, while GFRC adds them, which makes it more sensitive to fracture risk prediction in populations where falls are prevalent (22). This has been shown in a study conducted with an Australian cohort where GFRC has been more precise in forecasting fracture risks in patients having a history of falls (21). Conversely, FRAX, being aligned with national fracture epidemiology, usually issues conservative risk estimates, while GFRC might overestimate fracture risks as it is based on generalizations that are not accurate in certain cases, which has been observed in New Zealand cohort (23).

Another reason for the low agreement between the two models could be the unmeasured variables that are of importance for the prediction of fractures. The FRAX and GFRC both overlook the cortical bone properties that have been shown to independently predict fractures (24). In addition to that, comorbidities, medications, and physical activity levels may also be factors impacting the fracture risk besides being inconsistently considered in both models (23,25). The markers of bone turnover may also have the potential to make predictions more accurate; however, they have not yet been included in the routine assessments (26).

The mismatch between FRAX and GFRC predictions can be worked on through some possible improvements, such as making the models more inclusive and more adaptable. For instance, the addition of extra risk factors such as falls history, cortical bone properties, and bone turnover markers into FRAX and GFRC could lead to these models' better predictive accuracy

Table 2. Marital status, occupation, alcohol consumption, smoking status, fracture history, family fracture history, corticosteroid use, and total number of fragility fractures after age 50 of the participants

Variable	Category	Frequency	Percentage
Marital status	Married	220	63.40%
	Single	38	10.95%
	Widowed	89	25.65%
Occupation	Unemployed	328	94.52%
	Worker	4	1.15%
	Civil servant	8	2.31%
	Freelance	7	2.02%
Alcohol consumption (≥3 units/day)	Yes	4	1.15%
	No	343	98.85%
Current smoking	Yes	76	21.90%
	No	271	78.10%
Fracture history	Yes	90	25.94%
	No	257	74.06%
Family fracture history	Yes	80	23.05%
	No	267	76.95%
Corticosteroid use	Yes	20	5.76%
	No	327	94.24%
Total number of fragility fractures after age of 50	0	257	74.06%
	1	46	13.26%
	2	40	11.53%
	3	4	1.15%

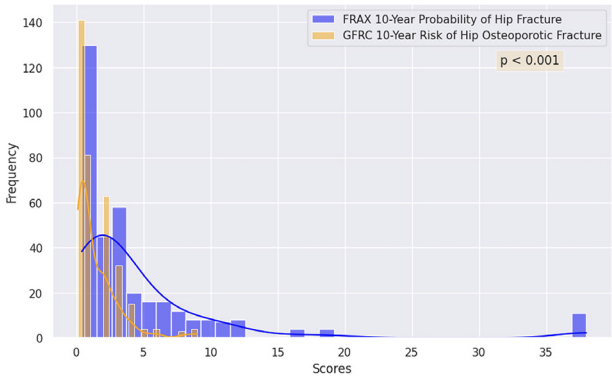

Figure 1. Comparison of FRAX 10-year probability of hip fracture and GFRC 10-year probability of hip fracture scores of the participants
FRAX: Fracture risk assessment tool (FRAX), GFRC: Garvan fracture risk calculator

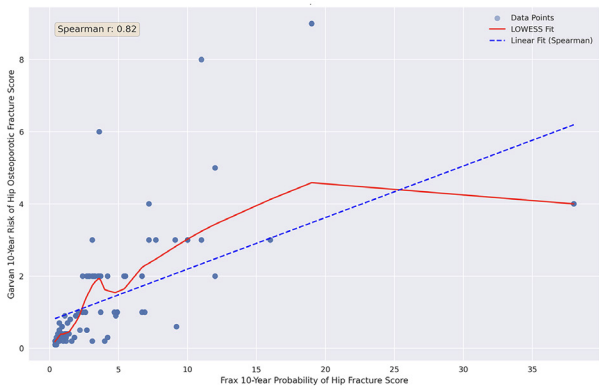

Figure 2. Scatter plot showing the relationship between the FRAX 10-year probability of hip fracture and GFRC 10-year probability of hip fracture scores of the participants
LOWESS: Locally weighted scatterplot smoothing, FRAX: Fracture risk assessment tool (FRAX), GFRC: Garvan fracture risk calculator

[Billington et al. (18), (24,26)]. Also, calibration of the GFRC model to the national population's epidemiological data, similar to the design of FRAX, could potentially minimize differences in the risk estimations for certain communities (25) Moreover, incorporating other imaging technologies, such as trabecular bone score, would not only enhance the information but also present a small-scale assessment of skeletal health alongside using the existing models (22,27). Regular validation of these tools via recent data, as well as the implementation of machine learning for the dynamic improvement of algorithms, can also be the way to cope with the discrepancies (21,28).

Implementing country-specific calibrations into osteoporotic risk calculation tools is highly necessary since there are different fracture risks and mortality rates among the various populations due to different factors such as genetics, lifestyle, and healthcare infrastructure. For example, different rates of fracture are observed in different regions of the world, and these differences are attributed to the levels of calcium and vitamin D intake, physical activity, and genetic predisposition, which can lead to problems of under/overestimation if tools are not calibrated locally (29). In addition, mortality rates also impact the predictive value of fracture risk calculations. For instance, the FRAX tool incorporates competing death risks into its model, which varies across regions based on the healthcare available and economic factors (20). Besides that, cultural and environmental factors, such as fall risks, smoking, alcohol consumption, and sunlight exposure, affect fracture incidence, too; thus, localized correction becomes necessary (23). Without country-specific data, fracture risk tools may misclassify patients, leading to inappropriate treatment decisions or resource allocation. Therefore, calibrations tailored to specific populations ensure more accurate and clinically relevant risk predictions and interventions.

Study Limitations

This study has several limitations. The calculations might have been error-prone due to the lack of BMD values. The research was centered only on postmenopausal osteoporosis individuals who were more than 50 years old, which might affect its applicability to broader demographic groups. Moreover, the retrospective character of this investigation might bring about biases due to the uncertain data, reliance on records, and the difficulty of proving causation between the exposure and the outcome. Furthermore, the lack of utilization of FRAXplus in this study represents an additional limitation, as its inclusion could have contributed to more boarder results. Future studies should try to implement prospective approaches and insert measurements of BMD to raise the reliability of their findings. The conduction of research utilizing larger and more assorted populations is fundamentally necessary for the generalization and relevance of the discoveries. Additionally, incorporating other risk calculation tools like FRAXplus in future investigations could play a significant role for border and more generalizable results.

Conclusion

In conclusion, the strong correlation yet low concordance between the FRAX and GFRC may indicate their complementary nature rather than interchangeability. Under the current circumstances, the GFRC system must be run as an additional tool and applied only in a limited way to specialized patient groups within the postmenopausal Turkish patients. The results may also hint the potential need for local calibrations to be established to ensure the accuracy and relevance of predictive algorithms within specific populations, such as Türkiye. Overcoming these limitations via specific modifications and broader studies has the potential to improve the Garvan tool in its application to the Turkish population, hence the development of more accurate and culturally relevant medical interventions for osteoporosis management strategies.

Ethics

Ethics Committee Approval: The study was conducted in accordance with the Declaration of Helsinki and was approved by the University of Health Sciences Türkiye, Zeynep Kamil

Women and Children Diseases Training and Research Hospital Ethics Committee (approval no: 11 date: 11.01.2023).

Informed Consent: This retrospective cross-sectional study was conducted between Üsküdar State Hospital.

Acknowledgments

The manuscript preparation process involved using Jenni AI, a generative AI tool, to enhance the readability and the language of the text. This technology was applied solely to improve grammar, syntax, and clarity Without influencing the scientific content, data interpretation, or conclusions. The authors thoroughly reviewed, verified, and edited all AI-generated outputs to ensure accuracy, completeness, and neutrality, as AI-generated content may contain inaccuracies, omissions, or biases. The tool was used under strict human oversight, with full accountability for the final manuscript resting with the authors.

Footnotes

Authorship Contributions

Surgical and Medical Practices: S.Ö.İ., Concept: M.H.T., S.Ö.İ., Design: M.H.T., H.M.B.T., Data Collection or Processing: M.H.T., H.M.B.T., Analysis or Interpretation: M.H.T., F.B., Literature Search: M.H.T., F.B., E.A., Writing: M.H.T., F.B.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Salari N, Ghasemi H, Mohammadi L, Behzadi MH, Rabieenia E, Shohaimi S, et al. The global prevalence of osteoporosis in the world: a comprehensive systematic review and meta-analysis. J Orthop Surg Res. 2021;16:1-11.
2. Reginster J-Y, Burlet N. Osteoporosis: a still increasing prevalence. Bone. 2005;38(Suppl 1):S4-9.
3. Harvey NC, Dennison EM, Cooper C. Osteoporosis: impact on health and economics. Nat Rev Rheumatol. 2010;6:99-105.
4. Ström O, Borgström F, Kanis JA, Compston J, Cooper C, McCloskey EV, et al. Osteoporosis: burden, health care provision and opportunities in the EU: a report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). Arch Osteoporos. 2011;6:59-155.
5. Pisani P, Renna MD, Conversano F, Casciaro E, Di Paola M, Quarta E, et al. Major osteoporotic fragility fractures: risk factor updates and societal impact. World J Orthop. 2016;7:171-81.
6. Mauck KF, Clarke BL. Diagnosis, screening, prevention, and treatment of osteoporosis. Mayo Clin Proc. 2006;81:662-72.
7. Kanis JA, Johnell O, Odén A, Johansson H, McCloskey EV. FRAX™ and the assessment of fracture probability in men and women from the UK. Osteoporos Int. 2008;19:385-97.
8. Sànchez-Riera L, Wilson NM, Kamalaraj N, Nolla JM, Kok C, Li Y, et al. Osteoporosis and fragility fractures. Best Pract Res Clin Rheumatol. 2010;24:793-810.
9. Zerikly R, Demetriou EW. Use of Fracture risk assessment tool in clinical practice and fracture risk assessment tool future directions. Womens Health (Lond). 2024;20:17455057241231387.
10. Nguyen ND, Frost SA, Center JR, Eisman JA, Nguyen TV. Development of prognostic nomograms for individualizing 5-year and 10-year fracture risks. Osteoporos Int. 2008;19:1431-44.
11. Wright NC, Looker AC, Saag KG, Curtis JR, Delzell ES, Randall S, et al. The recent prevalence of osteoporosis and low bone mass in the United States based on bone mineral density at the femoral neck or lumbar spine. J Bone Miner Res. 2014;29:2520-6.
12. Schober P, Vetter TR. Correlation analysis in medical research. Anesth Analg. 2020;130:332-3.
13. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15:155-63.
14. Kavroudakis E, Karampinas P, Eyaggelopoulos D, Vlamis J, Papaioanoy N, Liritis G. FRAX® (fracture risk assessment tool): a prognostic tool of osteoporotic fractures. Acta Orthop Traumatol Hellenica. 2012;63:36-40.
15. Imerci A, Kılınç CY, Aydogan NH, Karalezli MN, Savran A. Fracture risk assessment tool (FRAX®) results calculated with and without bone mineral density values for evaluating fracture risk in postmenopausal women with osteopenia. J Clin Densitom. 2018;21:464-71.
16. McCloskey EV, Johansson H, Odén A, Kanis JA. From relative risk to absolute fracture risk calculation: the FRAX algorithm. Curr Osteoporos Rep. 2009;7:77-83.
17. McCloskey EV, Harvey NC, Johansson H, Lorentzon M, Liu E, Vandenput L, et al. Fracture risk assessment by the FRAX model. Climacteric. 2021;25(Suppl 1):22-8.
18. Billington EO, Gamble GD, Reid IR. Reasons for discrepancies in hip fracture risk estimates using FRAX and Garvan calculators. Maturitas. 2016;85:11-8.
19. van Geel TA, Eisman JA, Geusens PP, van den Bergh JP, Center JR, Dinant GJ. The utility of absolute risk prediction using FRAX® and Garvan fracture risk calculator in daily practice. Maturitas. 2014;77:174-9.
20. Baleanu F, Iconaru L, Charles A, Kinnard V, Fils JF, Moreau M, et al. Independent external validation of FRAX and Garvan fracture risk calculators: a sub-study of the FRISBEE cohort. JBMR Plus. 2021;5:e10471.
21. Sandhu SK, Nguyen ND, Center JR, Pocock NA, Eisman JA, Nguyen TV. Prognosis of fracture: evaluation of predictive accuracy of the FRAX™ algorithm and Garvan nomogram. Osteoporos Int. 2010;21:863-71.
22. Edwards MH, Jameson KA, Denison HJ, Harvey NC, Sayer AA, Dennison EM, et al. Clinical risk factors, bone density, and fall history in predicting incident fracture among men and women. Bone. 2013;52:541-7.
23. Bolland MJ, Siu AT, Mason BH, Horne AM, Ames RW, Grey A, et al. Evaluation of the FRAX and Garvan fracture risk calculators in older women. J Bone Miner Res. 2011;26:2770-7.
24. Kral R, Osima M, Borgen TT, Vestgaard R, Richardsen E, Bjørnerem Å. Increased cortical porosity and reduced cortical thickness of the proximal femur are associated with nonvertebral fracture independent of FRAX and Garvan estimates in postmenopausal women. PLoS One. 2017;12:e0175919.
25. Crandall CJ, Larson JC, LaCroix AZ, Cauley JA, LeBoff MS, Li W, et al. Predicting fracture risk in younger postmenopausal women: comparison of the Garvan and FRAX risk calculators in the women's health initiative study. J Gen Intern Med. 2018;34:235-42.
26. Uludağ M, Akarırmak Ü, Özbayrak S, Bolayırlı M, Tüzün Ş. The relationship between the FRAX tool and bone turnover markers in postmenopausal osteoporosis. Turk J Osteoporos. 2013;19:38-41.

27. Felicia B, Michel M, Alexia C, Laura I, Rafik K, Murielle S, et al. Fragility fractures in postmenopausal women: development of 5-year prediction models using the FRISBEE study. Endocr Abstr. 2021;74:EA1396.
28. Sambrook PN, Flahive J, Hooven FH, Boonen S, Chapurlat R, Lindsay R, et al. Predicting fractures in an international cohort using risk factor algorithms without BMD. J Bone Miner Res. 2011;26:2770-7.
29. Kanis JA, Harvey NC, Johansson H, Odén A, McCloskey EV, Leslie WD. Overview of fracture prediction tools. J Clin Densitom. 2017;20:444-50.

DOI: 10.4274/tod.galenos.2025.90018
Turk J Osteoporos 2025;31(3):145-51

The Relationship Between Physical Performance with Dual-task Performance in Geriatric Individuals

Geriatrik Bireylerde Fiziksel Performans ile İkili Görev Performansı Arasındaki İlişki

Merve Akış[1], Emine Atıcı[2], Berna Çağla Balkışlı[2]

[1]İstanbul Okan University, Institute of Postgraduate Education, Department of Physiotherapy and Rehabilitation, İstanbul, Türkiye
[2]İstanbul Okan University Faculty of Health Sciences, Department of Physiotherapy and Rehabilitation, İstanbul, Türkiye

Abstract

Objective: Maintaining functional independence in older adults requires an intricate balance between physical and cognitive abilities. While physical performance is known to support overall health, its relationship with dual-task performance -an essential component of daily life activities- remains a critical area of investigation. This study aimed to investigate the association between physical performance and dual-task performance and to compare the effects of different levels of physical performance on dual-task execution in geriatric individuals.

Materials and Methods: A total of 79 geriatric individuals (mean age=68.68±4.42) were included in the study. Their physical performances were evaluated through the Alusti test, Timed Up and Go test (TUGT), 30-Second Sit-to-Stand test (30s STST) and 10-meter Walk test. Dual-task performance measurements were assessed using the TUGT, 30s STST and 10-meter Walk test (using motor and cognitive task) and the dual-task questionnaire.

Results: There was a low correlation between the Alusti test and Dual-task scale (r=-0.222; p=0.048), while a moderate-level correlation between TUGT (r=0.339; p=0.001), 30s STST (r=-0.336; p=0.002), and the 10-meter Walk test (r=0.365; p=0.001). When individuals were divided into two groups based on Alusti test scores (good mobility and very good mobility), the mean of 30s STST motor and -cognitive tasks were statistically significantly higher in the very good mobility group compared to good mobility (p=0.026, p=0.005; respectively).

Conclusion: The findings suggest that dual-task performance is closely linked to physical function in older adults, with higher physical performance associated with improved dual-task execution. Given the increasing importance of maintaining cognitive-motor abilities for aging populations, targeted physical activity interventions may help mitigate declines in dual-task performance, ultimately promoting safer mobility and greater independence in daily life. Future research should further explore the mechanisms underlying these interactions to develop effective strategies for cognitive-motor preservation in geriatric care.

Keywords: Physical performance, geriatrics, task performance, cognition, motor skills, balance

Öz

Amaç: Geriatrik bireylerde fonksiyonel bağımsızlığın korunması, fiziksel ve bilişsel yetilerin dengeli bir şekilde sürdürülmesini gerektirir. Fiziksel performansın genel sağlık üzerindeki olumlu etkileri bilinse de, ikili görev performansı ile olan ilişkisi -günlük yaşam aktivitelerinin önemli bir bileşeni- hala araştırılması gereken kritik bir konudur. Bu çalışma, fiziksel performans ile ikili görev performansı arasındaki ilişkiyi incelemeyi ve farklı fiziksel performans seviyelerinin ikili görev yürütme üzerindeki etkilerini karşılaştırmayı amaçlamaktadır.

Gereç ve Yöntem: Çalışmaya toplam 79 geriatrik birey (ortalama yaş=68,68±4,42) dahil edilmiştir. Fiziksel performansları Alusti testi, Zamanlı Kalk ve Yürü testi (ZKYT), 30 Saniye Otur-Kalk testi (30s OKT) ve 10 Metre Yürüme testi ile değerlendirilmiştir. İkili görev performansı ölçümleri ZKYT, 30s OKT ve 10 metre Yürüme testi (motor ve bilişsel görev kullanılarak) ile ikili görev ölçeği kullanılarak değerlendirilmiştir.

Bulgular: Alusti testi ile İkili Görev ölçeği arasında düşük düzeyde bir korelasyon (r=-0,222; p=0,048) bulunurken, ZKYT (r=0,339; p=0,001), 30s OKT (r=-0,336; p=0,002) ve 10 metre Yürüme testi (r=0,365; p=0,001) ile orta düzeyde korelasyon görülmüştür. Alusti test skorlarına göre bireyler iki gruba ayrıldığında (iyi mobilite ve çok iyi mobilite), çok iyi mobilite grubunda 30s OKT motor ve bilişsel görev ortalamalarının iyi mobilite grubuna kıyasla istatistiksel olarak anlamlı derecede yüksek olduğu bulunmuştur (sırasıyla p=0,026, p=0,005).

Corresponding Author/Sorumlu Yazar: Berna Çağla Balkışlı Asst. Prof., İstanbul Okan University Faculty of Health Sciences, Department of Physiotherapy and Rehabilitation, İstanbul, Türkiye
E-mail: ptcaglacaglayan@gmail.com **ORCID ID:** orcid.org/0000-0002-2559-9756
Received/Geliş Tarihi: 19.02.2025 **Accepted/Kabul Tarihi:** 08.04.2025 **Epub:** 14.07.2025 **Publication Date/Yayınlanma Tarihi:** 05.12.2025

Cite this article as/Atıf: Akış M, Atıcı E, Balkışlı BÇ. The relationship between physical performance with dual-task performance in geriatric individuals. Turk J Osteoporos. 2025;31(3):145-51

Öz

Sonuç: Bulgular, ikili görev performansının yaşlı bireylerde fiziksel fonksiyon ile yakından ilişkili olduğunu ve daha yüksek fiziksel performansın daha iyi ikili görev yürütme ile bağlantılı olduğunu göstermektedir. Yaşlanan nüfus için bilişsel-motor yetilerin korunmasının artan önemi göz önüne alındığında, hedefe yönelik fiziksel aktivite müdahaleleri, ikili görev performansındaki düşüşleri azaltmaya yardımcı olabilir ve böylece daha güvenli hareketlilik ile günlük yaşamda daha büyük bir bağımsızlık sağlayabilir. Gelecekteki araştırmalar, bu etkileşimlerin altında yatan mekanizmaları daha ayrıntılı inceleyerek geriatrik bakımda bilişsel-motor korunmaya yönelik etkili stratejiler geliştirmelidir.
Anahtar kelimeler: Fiziksel performans, geriatri, görev performansı, biliş, motor beceriler, denge

Introduction

With ageing, structural and functional changes occur in many organ systems, leading to reductions in reserves. In the case of physical inactivity, the physiological changes associated with the ageing process accelerate, the frequency of chronic diseases increases, and life expectancy shortens (1). Therefore, the decrease in physical performance with advancing age can impact health and well-being, predicting disability and mortality in older adults (2,3). The importance of physical performance is emphasized in the relationship between physical disability and cognitive impairment (4). In addition to physical performance, cognitive impairment in the older people also diminish the quality of life and daily life activities (5).

The identification of age-related important conditions involves recognizing the loss of muscle strength and deterioration in physical performance as crucial elements. In the literature, the evaluation of these elements often includes grip strength, speed of walking, Timed Up and Go test (TUG), and batteries (6). Besides, the Alusti test, created by Calvo Aguirrey et al. (8) and rooted in existing assessments, offers a quick evaluation of physical performances in older adults with diverse functional and cognitive level capacities without causing exhaustion to the patient (7,8).

Dual-tasking refers to the capability of executing two or more tasks at the same time. It is known that different age groups exhibit varying performance outcomes for tasks that need to be accomplished simultaneously. It has been observed that dual-task conditions enhance both cognitive and motor performance, improving standing and walking performance (9). Previous studies employed various primary mobility tasks, such as balance or walking, alongside a secondary task, which could be cognitive (e.g., an arithmetic task) or motor-related (e.g., carrying an object) (10). These results highlight a possible cognitive-motor interaction that leads to a decrease in mobility performance and/or dual-task performance in older adults compared to younger individuals. Additionally, reduced performance in secondary tasks (both cognitive and motor) was observed in geriatric adults (11-13).

The literature consistently demonstrates that higher levels of physical activity correlate with better functional status and lower risks of cognitive decline and dementia (14). Assessing only a single performance parameter, typically mobility tasks, may result in misleading conclusions about the impact of dual-task performance, particularly in the geriatric population, where both cognitive and motor components may be affected. For this reason, this study aims to contribute to the growing body of evidence supporting physical activity as a cornerstone of aging, emphasizing the necessity for tailored assessments that address both physical and cognitive performance in the geriatric population. Specifically, the present study investigates the relationship between physical performance and dual-task performance, while also comparing the varying levels of physical performance among older adults in relation to their ability to manage dual tasks.

Materials and Methods

Study Design

The research was designed as a cross-sectional research.

Participants

The study's sample size was determined using the G*Power software. Using the dual task performance values obtained from a similar study, it was determined that the study should include at least 55 participants, ensuring a 95% confidence interval and an 80.3% statistical power level.

The research was carried out at a University Training and Research Hospital, involving a total of 79 geriatric individuals.

Inclusion criteria were (1) being between the ages of 65-85, (2) voluntary participation in the study, (3) ability to read and write, (4) Scoring 21 points or above on the Montreal Cognitive Assessment scale.

Exclusion criteria were (1) the presence of cardiovascular, pulmonary, orthopedic, or neurological disorders that significantly affect their physical condition.

The ethical approval for the study was obtained from the Local Ethical Committee for Research in Natural, Social, and Non-Invasive Health Sciences (decision no: 144, date: 10.11.2021). All patients were informed about the research before they started the study and obtained signed written and verbal consent for participation. All protocols followed were in compliance with the ethical standards set by the committee responsible for human experimentation and the principles of the Helsinki Declaration.

Assessment

All evaluations were completed by the same researcher. Demographic characteristics of geriatric individuals were recorded. Their physical performances were assessed using the Alusti test, timed up and go, the 30-second sit-to-stand tests and

10-meter walk test; and dual-task performance measurements were assessed by adding both motor and cognitive tasks to the tests used to evaluate physical performance, and using the dual Task scale.

Physical Performance Assessment

Alusti Test: This test is commonly used to assess physical performance in geriatric individuals. Comprising 10 items, the total score of the test was evaluated on a scale of 0 to 100. The assessment began with the individual lying supine in bed, evaluating the range of joint motion in four extremities (item 1), muscle strength of extremities (item 2), ability to transition from lying down to sitting position (item 3), and sitting balance in the sitting position (item 4). Additionally, the ability to stand up from a sitting position (item 5), standing (item 6), walking (item 7), walking distance (item 8), eyes closed tandem stance (item 9), and eyes closed single-leg stance (item 10) were assessed. The total score provides information about the individual's level of mobility. Based on the total score, the evaluation is as follows: 0-30: Completely dependent mobility, 31-40: Severely dependent mobility, 41-50: Moderately dependent mobility, 51-60: Mildly dependent mobility, 61-75: Good level of mobility, 76-90: Very good level of mobility, 91-100: Excellent level of mobility (8).

The Timed Up and Go Test (TUGT): It is used to assess an individual's dynamic balance, walking speed, and mobility. In the test, the individual is instructed to stand up from a standard chair with back support upon the starting command, walk a distance of 3 meters at their normal pace, return, and then sit back down in the chair. The time is measured with a stopwatch and recorded in seconds. During the test, the use of walking aids is allowed if the individual normally uses them. Completing the TUGT in 14 seconds or more is identified as a risk of falling (15).

The 30-Second Chair Stand Test: This test is utilized to assess individuals' physical performance levels. This test evaluates lower extremity muscle function, muscle endurance, sit-to-stand activity, lower extremity strength, and dynamic balance. During the test, the person's ability to stand up from and sit back down on a chair within 30 seconds is recorded. If the individual records a score of 10 or fewer sit-to-stand repetitions within the 30-second timeframe, it indicates lower extremity weakness (16).

The 10-Meter Walk Test: It evaluates an individual's normal walking speed. In this test, the person is asked to walk in a predefined 10-meter area at their usual walking pace in daily life. If the individual use a walking aid, it is allowed. The timing is initiated when the person's foot crossed the starting line and stopped when they crossed the finish line, with the total time recorded. Two measurements are taken, and the best result is recorded in seconds (17).

Dual-Task Assessment

Dual Task Questionnaire: It assesses the difficulties individuals face when performing dual-task activities encountered in their daily lives. It consists of 10 questions, each scored on a scale of very often (4), often (3), occasionally (2), rarely (1), and never (0). Responses to all questions are summed up to calculate a total score, which is then divided by 10. An increase score indicates greater difficulty in behaviors that require dual-task performance (18).

Motor-Motor Tasks: To evaluate the dual-task motor-motor performance of the participants, during the TUGT, 30-second sit-to-stand test, and 10-meter walk test, they were given the task of carrying a 0.5-liter bottle full of water in their hands. For motor-motor assessments something was carried in previous studies (19).

Motor-Cognitive Tasks: To assess cognitive performance, tasks involving the sequential recitation of months (from January to December) were assigned during the TUGT, 30-second sit-to-stand test, and 10-meter walking test. Studies showed that tasks requiring verbal processing, such as counting or reciting sequences, can impose varying levels of cognitive load that affect gait parameters, including speed and stability (20).

Mean time was recorded in seconds in TUGT and 10-meter walk test and number of repetitions in The 30-Second Chair Stand test. In order to minimize the effects of fatigue, rest intervals were given. Patients' failed attempts were repeated (inaccuracy in the order of months).

The dual-task performances were calculated separately for motor and cognitive-added tasks.

The dual-task effect (DTE) was calculated according to the following formula (20):

DTE = [(dual-task performance- single task performance)/single task performance]x100.

A higher DTE indicates a greater decline in performance under dual-task conditions.

Statistical Analysis

The data analysis was conducted using IBM SPSS Statistics 25 software (IBM Corp., Armonk, NY, USA). Descriptive statistics were calculated for both categorical and continuous variables in the study. The homogeneity of variances, required for parametric tests, was assessed using the Levene test. The normality assumption was evaluated with the Shapiro-Wilk test. For comparing two groups, the Student's t-test was used when parametric test assumptions were met; otherwise, the Mann-Whitney U test was applied. The relationship between continuous variables was examined using the Pearson Correlation Coefficient when parametric assumptions were satisfied, and the Spearman correlation coefficient when they were not. A p-value of <0.05 was considered statistically significant.

Results

At the beginning, a total of 84 geriatric individuals who voluntarily participated in the study for treatment at the clinic were assessed. Five individuals were excluded from the study as they did not meet the inclusion criteria. The study was completed with a total of 79 individuals (40 male, 39 female), with an average age of 68.68±4.42. Demographic data and the

mean scores of individuals on the scales and tests are presented in Table 1.

When comparing dual-task questionnaire and physical performance, The dual-task questionnaire showed a low correlation with the Alusti test (r=-0.290, p=0.009) and 30-second sit-to-stand test (r=-0.267, p=0.017), while it showed a moderate correlation with the 10-meter walk test (r=0.359, p=0.001) and TUGT (r=0.454, p<0.001) (Table 2).

When comparing physical performance and dual-task performances (both motor and cognitive tasks for each test), a moderate correlation was found between the Alusti test and the 30-second sit-to-stand test (r=-0.381, p=0.001; r=-0.407, p<0.001, respectively). Additionally, TUGT-motor/cognitive tasks had a low to moderate correlations with both 10-meter walk test (motor/cognitive) (r=-0.340, p=0.002; r=-0.232, p=0.039, respectively) and 30-second sit-to-stand test (motor /cognitive) (r=0.411, p<0.001; r=0.232, p=0.040, respectively) (Table 3).

When individuals were grouped into Alusti test categories as good mobility and very good mobility, the mean of 30-second the sit-to-stand test- motor and -cognitive tasks were statistically significantly higher in the very good mobility group compared to good mobility (p=0.026, p=0.005; respectively). No differences were observed among the groups for other dual-task performance parameters (p>0.05) (Table 4).

Table 1. Demographic data and the mean scores of measurements of individuals (n=79)

Variables	**Mean ± SD**
Age (years)	68.68±4.42
MOCA	23.91±2.27
Alusti test	79.39±6.09
TUG	11.67±1.86
30s STST	11.29±2.72
10-meter Walk test	20.77±2.88
	n (%)
Gender	
Male	40 (51)
Female	39 (49)
Dominant side	
Right	64 (81)
Left	15 (19)
History of fall	
None	23 (29)
1-2 times	41 (52)
More than 2	15 (19)
Smoking	
Yes	16 (20)
No	63 (80)
Assistive device usage	
Yes	0
No	70 (100)

TUG: Timed up and go test, 30s STST: 30-second sit-to-stand test, SD: Standard deviation, MOCA: Montreal cognitive assessment

Discussion

This cross-sectional study showed that the physical performances in geriatric individuals are associated with their dual-task performance. Additionally, individuals with very good mobility observed to exhibit better performance in sit-to-stand tasks with added motor and cognitive tasks.

With ageing, loss of muscle strength and decreased balance cause older adults to be less able to automatically adapt their balance and movement abilities to environmental stimuli. They need to be more cautious to maintain their balance. Consequently, walking speed decreases, and the risk of falls increases (21). The decline in physical performance adversely affects individuals' daily life activities. Decreased balance, inactivity, and increased fear of falling are associated with this decline. Diminished physical performance in older individuals has led to serious issues, particularly deterioration in physical and cognitive functions, progressing with age and contributing to falls (5). Therefore, when assessing physical performance in older individuals, their cognitive capacities should also be evaluated. The physical performance of the participants in our study was evaluated with the Alusti test, which has been previously used in this population. When the participants were divided into two groups based on their scores on this test, it was observed that the participants had good or very good levels of mobility. This may explain the low-moderate correlations we obtained.

In a study examining the dual-task performance in older adults, it was noted that the performance of the TUG test, slowing down particularly during the returning to the chair and sitting, was affected when performed concurrently with an additional cognitive task. The study was recorded that the performance of the TUG test with an added cognitive task was worse compared to performing the TUG test alone (22). Sertel et al. (23), in their study investigating the impact of additional cognitive and motor tasks on balance in older adults, reported increased performance times in the one-leg standing test, TUG, and sit-to-stand test when performed with additional tasks. Therefore, they concluded that task addition impairs balance performance in older adults. Similarly, our study revealed a

Table 2. Correlation between physical performance tests and dual-task questionnaire

	Dual-task questionnaire	
	r	**p***
Alusti test	-0.290	**0.009**
TUG	0.454	**<0.001**
30s STST	-0.267	**0.017**
10-meter Walk test	0.359	**0.001**

TUG: Timed up and go test, 30s STST: 30-second sit-to-stand test
*: Pearson correlation

Table 3. Correlation between physical performance and dual task performances in motor and cognitive tasks

		Alusti test		TUG		10-meter Walk test		30s STST	
		r	p	r	p	r	p	r	p
TUG	Cognitive	-0.033	0.770	**0.228**	**0.043**	**-0.340**	**0.002**	**0.411**	**<0.001**
	Motor	-0.033	0.770	**0.333**	**0.003**	**-0.232**	**0.039**	**0.232**	**0.040**
10-meter Walk test	Cognitive	-0.037	0.747	-0.108	0.343	0.189	0.096	0.077	0.500
	Motor	-0.125	0.272	-0.133	0.243	0.119	0.295	-0.051	0.658
30s STST	Cognitive	**-0.381**	**0.001**	-0.125	0.271	-0.099	0.384	0.217	0.055
	Motor	**-0.407**	**<0.001**	-0.024	0.833	-0.072	0.526	0.105	0.358

TUG: Timed up and go test, 30s STST: 30-second sit-to-stand test
*: Pearson correlation

Table 4. Comparison of cognitive and motor measurements of physical performance tests according to Alusti test categories

	Alusti test		P*
	Mean ± SD Median (min-max)		
	Good mobility	Very good mobility	
TUG Cognitive	-10.07±24.11 -7.1 (-100-22.2)	-7.99±17.21 -7.1 (-87.5-10)	0.659
TUG Motor	-7.42±20.88 0 (-87.5-7.7)	-5.31±16.89 0 (-87.5-18.2)	0.670
10-meter Walk test Cognitive	-2.14±6.77 0 (-14.3-12.5)	-3.04±5.15 -4.2 (-25-5.6)	0.476
10-meter Walk test Motor	-0.28±4.75 0 (-9.5-10)	-1.52±3.7 0 (-9.5-6.3)	0.243
30s STST Cognitive	0.40±6.99 0 (-22.2-10.5)	-2.88±8.01 0 (-27.3-18.8)	**0.026 ***
30s STST Motor	5.51±12.66 10 (-33.3-30)	-0.92±9.64 0 (-27.3-18.8)	**0.005 ***

TUG: Timed up and go test, 30s STST: 30-second sit-to-stand test, SD: Standard deviation, min-max: Minimum-maximum
*: Mann-Whitney U test

correlation between dual-task performances assessed through both questionnaire and performance tests with the TUG test. As dual-task performance deteriorates, physical performance may be affected or performing the second task to prevent performance impairment may become more challenging.

The 30-second sit-to-stand test explained as a suitable method to assess physical function and screen the risk of falls in older individuals due to the recurrent falls being a risk factor for lower extremity weakness (24). Our study also demonstrated a correlation between the dual-task questionnaire and the 30-second sit-to-stand test. The decrease in muscle strength may be associated with dual-task performance deficits. Additionally, in this test where motor and cognitive dual tasks are added, individuals with good mobility performed better.

In the study conducted by Soumaré et al. (25), it was stated that there is a relationship between cognitive function and walking speed in geriatric individuals. The assessed cognitive function was specifically associated with psychomotor speed and verbal fluency, correlating with a decrease in walking speed over time. Another study found that individuals with faster walking speeds exhibited better cognitive performance across various domains, including executive function, memory, and processing speed (26). A systematic review and meta-analysis revealed that adding a dual task significantly reduces walking speed in community-dwelling older adults. Furthermore, it highlighted that the assessment of walking with a dual task is essential as part of the standard clinical evaluation for older individuals (27). In our study, a positive correlation was observed between single-task 10-meter walking test and dual-task questionnaire scores; however, no relationship was observed in dual-task performances. This may be attributed to the participants included in the study having good physical performance.

There are also studies in the literature where dual-task experiments were conducted using laboratory and clinical measurements. Coelho et al. (28) investigated the standing balance of older adults living in the community under both single and dual-task conditions through sway analysis [center of pressure (COP) sway parameters]. Generally, it was found

that older adults have poorer standing balance compared to younger individuals. Performing a secondary task had increased swaying and had an impact on the standing balance of older adults. Older adults required more attention to maintain postural stability compared to young individuals. Similarly, in another study, it was reported that dual-tasking may increase the amount of sway activity assessed by the COP and impede balance control (29).

Various cognitive tasks, such as serial subtraction, word production, and responding to stimuli, have been used in dual-task gait research (20). Beauchet et al. (30) found that older adults exhibited a significant increase in gait variability during dual-task walking with an arithmetic task, but not with a verbal fluency task. Different attentional resources may have varying levels of sensitivity to brain damage or aging, which can lead to task-specific dual-task interference in different populations (31). The motor and cognitive tasks included in our study were limited. Future research should explore the effects of a broader range of tasks, including attention, inhibitory control, working memory, and cognitive flexibility.

The strength of this research lies in its integration of both mobility and cognitive aspects, providing a more comprehensive evaluation of dual-task performance, unlike many studies that focus solely on mobility or cognitive function. By including both motor-motor and motor-cognitive dual-task conditions, the study offers a nuanced understanding of how different types of secondary tasks impact mobility in geriatric individuals. Furthermore, the use of the Alusti test, a novel and efficient assessment tool designed for older adults, enhances the practicality of the study in clinical settings. These findings contribute to the growing body of evidence supporting the inclusion of dual-task assessments in standard geriatric evaluations, encouraging a more holistic approach to fall prevention and the promotion of functional independence.

Study Limitations

There are several limitations to this study. One limitation is the inability to assess physical performance measures, such as balance and muscle strength, using computerized systems, which could have provided more precise data. Additionally, since the study included only one group, cutoff values could not be established. Another limitation is that the participants generally had good physical performance, which may have resulted in weaker correlations in some tests. Furthermore, the variety of cognitive and motor tasks used in dual-task performance assessments was limited, potentially restricting the scope of the findings. Future studies should utilize computerized systems for more precise physical performance assessments and include a more diverse sample to establish cutoff values. Additionally, expanding dual-task conditions, conducting longitudinal research, and considering gender differences, real-world environments, and multiple assessors would enhance the reliability and applicability of findings.

Conclusion

The dual-task capabilities of older people individuals correlate with their physical performance. Enhancements in the physical performance of geriatric individuals could potentially result in improved dual-task performances, attributed to variations in cognitive skills observed as part of the natural ageing process.

Ethics

Ethics Committee Approval: The ethical approval for the study was obtained from the Local Ethical Committee for Research in Natural, Social, and Non-Invasive Health Sciences (decision no: 144, date: 10.11.2021).

Informed Consent: All patients were informed about the research before they started the study and obtained signed written and verbal consent for participation.

Footnotes

Authorship Contributions

Concept: E.A., Design: E.A., Data Collection or Processing: M.A., Analysis or Interpretation: M.A., Literature Search: B.Ç.B., Writing: B.Ç.B.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declare that this study received no financial support.

References

1. Netz Y, Wu MJ, Becker BJ, Tenenbaum G. Physical activity and psychological well-being in advanced age: a meta-analysis of intervention studies. Psychol Aging. 2005;20:272-84.
2. Denison HJ, Jameson KA, Sayer AA, Patel HP, Edwards MH, Arora T, et al. Poor sleep quality and physical performance in older adults. Sleep Health. 2021;7:205-11.
3. Spartano NL, Lyass A, Larson MG, Tran T, Andersson C, Blease SJ, et al. Objective physical activity and physical performance in middle-aged and older adults. Exp Gerontol. 2019;119:203-11.
4. Facal D, Maseda A, Pereiro AX, Gandoy-Crego M, Lorenzo-López L, Yanguas J, et al. Cognitive frailty: a conceptual systematic review and an operational proposal for future research. Maturitas. 2019;121:48-56.
5. Ikegami S, Takahashi J, Uehara M, Tokida R, Nishimura H, Sakai A, et al. Physical performance reflects cognitive function, fall risk, and quality of life in community-dwelling older people. Sci Rep. 2019;9:12242.
6. Patrizio E, Calvani R, Marzetti E, Cesari M. Physical functional assessment in older adults. J Frailty Aging. 2021;10:141-9.
7. Kesikbaş B, Atıcı E, Navarro JA. Turkish version of functional evaluation of physical performance for the geriatric population: a reliability and validity study. Turk J Osteoporos. 2023;29:96-102.
8. Calvo Aguirrey JJ, Alustiza Navarro J, Uranga Zaldúa J, Sarasqueta Eizaguirre C, Bueno Yáñez O. Test Alusti: nueva escala de valoración del rendimiento físico para la población geriátrica [Alusti test: new scale for assessment of physical performance in the geriatric population]. Rev Esp Geriatr Gerontol. 2018;53:255-61. Spanish.
9. Ruffieux J, Keller M, Lauber B, Taube W. Changes in standing and walking performance under dual-task conditions across the lifespan. Sports Med. 2015;45:1739-58.

10. Brustio PR, Rabaglietti E, Formica S, Liubicich ME. Dual-task training in older adults: the effect of additional motor tasks on mobility performance. Arch Gerontol Geriatr. 2018;75:119-24.
11. Chu YH, Tang PF, Peng YC, Chen HY. Meta-analysis of type and complexity of a secondary task during walking on the prediction of elderly falls. Geriatr Gerontol Int. 2013;13:289-97.
12. Brustio PR, Magistro D, Zecca M, Rabaglietti E, Liubicich ME. Age-related decrements in dual-task performance: Comparison of different mobility and cognitive tasks. A cross sectional study. PLoS One. 2017;12:e0181698.
13. Nordin E, Moe-Nilssen R, Ramnemark A, Lundin-Olsson L. Changes in step-width during dual-task walking predicts falls. Gait Posture. 2010;32:92-7.
14. Guo J, Huang X, Dou L, Yan M, Shen T, Tang W, et al. Aging and aging-related diseases: from molecular mechanisms to interventions and treatments. Signal Transduct Target Ther. 2022;7:391.
15. Bohannon RW. Reference values for the timed up and go test: a descriptive meta-analysis. J Geriatr Phys Ther. 2006;29:64-8.
16. Purath J, Buchholz SW, Kark DL. Physical fitness assessment of older adults in the primary care setting. J Am Acad Nurse Pract. 2009;21:101-7.
17. Peters DM, Fritz SL, Krotish DE. Assessing the reliability and validity of a shorter walk test compared with the 10-Meter Walk Test for measurements of gait speed in healthy, older adults. J Geriatr Phys Ther. 2013;36:24-30.
18. Abasıyanık Z, Veldkamp R, Fostier A, Van Goubergen C, Kalron A, Feys P. Patient-reported outcome measures for assessing dual-task performance in daily life: a review of current ınstruments, use, and measurement properties. Int J Environ Res Public Health. 2022;19:15029.
19. Kayabinar B, Alemdaroğlu-Gürbüz İ, Yilmaz Ö. The effects of virtual reality augmented robot-assisted gait training on dual-task performance and functional measures in chronic stroke: a randomized controlled single-blind trial. Eur J Phys Rehabil Med. 2021;57:227-37.
20. Goh HT, Pearce M, Vas A. Task matters: an investigation on the effect of different secondary tasks on dual-task gait in older adults. BMC Geriatr. 2021;21:510.
21. Shumway-Cook A, Woollacott M, Kerns KA, Baldwin M. The effects of two types of cognitive tasks on postural stability in older adults with and without a history of falls. J Gerontol A Biol Sci Med Sci. 1997;52:M232-40.
22. Smith E, Walsh L, Doyle J, Greene B, Blake C. Effect of a dual task on quantitative timed up and go performance in community-dwelling older adults: a preliminary study. Geriatr Gerontol Int. 2017;17:1176-82.
23. Sertel M, Sakızlı E, Bezgin S, Savcun Demirci C, Yıldırım Şahan T, Kurtoğlu F. The effect of single-tasks and dual-tasks on balance in older adults. Cogent Soc Sci. 2017;3.
24. Applebaum EV, Breton D, Feng ZW, Ta AT, Walsh K, Chassé K, et al. Modified 30-second Sit to Stand test predicts falls in a cohort of institutionalized older veterans. PLoS One. 2017;12:e0176946.
25. Soumaré A, Tavernier B, Alpérovitch A, Tzourio C, Elbaz A. A cross-sectional and longitudinal study of the relationship between walking speed and cognitive function in community-dwelling elderly people. J Gerontol A Biol Sci Med Sci. 2009;64:1058-65.
26. Gale CR, Allerhand M, Sayer AA, Cooper C, Deary IJ. The dynamic relationship between cognitive function and walking speed: the English longitudinal study of ageing. Age (Dordr). 2014;36:9682.
27. Smith E, Cusack T, Blake C. The effect of a dual task on gait speed in community dwelling older adults: a systematic review and meta-analysis. Gait Posture. 2016;44:250-8.
28. Coelho T, Fernandes Â, Santos R, Paúl C, Fernandes L. Quality of standing balance in community-dwelling elderly: Age-related differences in single and dual task conditions. Arch Gerontol Geriatr. 2016;67:34-9.
29. Drozdova-Statkevičienė M, Česnaitienė VJ, Pukėnas K, Levin O, Masiulis N. Sway regularity and sway activity in older adults' upright stance are differentially affected by dual task. Neurosci Lett. 2018;666:153-7.
30. Beauchet O, Dubost V, Aminian K, Gonthier R, Kressig RW. Dual-task-related gait changes in the elderly: does the type of cognitive task matter? J Mot Behav. 2005;37:259-64.
31. Howell DR, Osternig LR, Koester MC, Chou LS. The effect of cognitive task complexity on gait stability in adolescents following concussion. Exp Brain Res. 2014;232:1773-82.

Original Investigation / Orijinal Araştırma

DOI: 10.4274/tod.galenos.2025.94809
Turk J Osteoporos 2025;31(3):152-61

Effects of Low-Dose Corticosteroids on Clinical Outcomes and Bone Mineral Density Parameters in Rheumatoid Arthritis Patients

Romatoid Artritli Hastalarda Düşük Doz Kortikosteroidin Klinik Bulgular ve Kemik Mineral Yoğunluğu Üzerine Etkileri

Sinem Kübra Beke[1], Hüseyin Kaplan[2], Taha Furkan Çakır[3], Gizem Cengiz[1], Senem Sas[4], İsa Cüce[3]

[1]Erciyes University Faculty of Medicine, Department of Physical Medicine and Rehabilitation, Division of Rheumatology, Kayseri, Türkiye

[2]Aksaray Training and Research Hospital, Clinic of Rheumatology, Aksaray, Türkiye

[3]Erciyes University Faculty of Medicine, Department of Physical Medicine and Rehabilitation, Kayseri, Türkiye

[4]University of Health Sciences Türkiye, Bursa Yüksek İhtisas Training and Research Hospital, Department of Physical Medicine and Rehabilitation, Division of Rheumatology, Bursa, Türkiye

Abstract

Objective: Rheumatoid arthritis (RA) is a chronic autoimmune disease that primarily affects the joints and frequently leads to bone loss. Glucocorticoids (GCs), widely prescribed for their anti-inflammatory properties in RA, have a controversial impact on bone health when used in low doses. This study evaluates the effect of low-dose GC use on bone mineral density (BMD) and the risk of osteoporosis (OP) in RA patients.
Materials and Methods: This study was conducted at the Rheumatology Outpatient Clinic of Erciyes University Faculty of Medicine. A total of 74 RA patients and 39 healthy controls (≥18 years) were included. Demographic data [age, gender, body mass index (BMI)] and clinical assessments (disease activity score-28), RA quality of life questionnaire was recorded using a standardized form.
Results: RA patients had significantly lower BMD and T-scores at the lumbar spine and femur compared to healthy controls ($p<0.001$). Among RA patients, those taking corticosteroids had lower BMI ($p=0.012$) and femur T-scores ($p=0.011$); however, no significant differences were observed in other clinical or laboratory parameters. Logistic regression analysis determined that age is an independent risk factor for OP [odds ratio (OR)=1.13, 95% confidence interval (CI): 1.04-1.24, $p=0.005$] and the use of low-dose GC was negatively associated with the risk of OP (OR=0.36, 95% CI: 0.11-0.99, $p=0.048$).
Conclusion: This study hypothesizes that low-dose GC therapy does not increase OP risk in RA patients and may even reduce it in some cases, with age being a key predictor, highlighting the importance of personalized treatment strategies for bone health.
Keywords: Rheumatoid arthritis, glucocorticoids, low-dose corticosteroids, bone mineral density, osteoporosis

Öz

Amaç: Romatoid artrit (RA), öncelikli olarak eklemleri etkileyen kronik bir otoimmün hastalık olup sıklıkla kemik kaybına yol açar. RA tedavisinde anti-enflamatuvar özellikleri nedeniyle yaygın olarak reçete edilen glukokortikoidlerin (GK), düşük dozda kullanıldıklarında kemik sağlığı üzerindeki etkileri tartışmalıdır. Bu çalışma, RA hastalarında düşük doz GK kullanımının kemik mineral yoğunluğu (KMY) ve osteoporoz (OP) riski üzerindeki etkilerini değerlendirmeyi amaçlamaktadır.
Gereç ve Yöntem: Bu çalışma, Erciyes Üniversitesi Tıp Fakültesi Romatoloji Polikliniği'nde yürütüldü. Çalışmaya 74 RA hastası ve 39 sağlıklı kontrol bireyi (≥18 yaş) dahil edildi. Katılımcıların demografik verileri [yaş, cinsiyet, vücut kitle indeksi (VKİ)] ve klinik değerlendirmeleri (hastalık aktivite skoru, RA yaşam kalitesi ölçeği) standart bir form aracılığıyla kaydedildi.
Bulgular: RA hastalarının, sağlıklı kontrollere kıyasla lomber omurga ve femur bölgelerinde anlamlı derecede daha düşük KMY ve T-skorlarına sahip olduğu saptandı ($p<0,001$). RA hastaları arasında kortikosteroid kullananlarda VKİ ($p=0,012$) ve femur T-skorları ($p=0,011$) daha düşük bulunurken, diğer klinik veya laboratuvar parametrelerde anlamlı bir fark gözlenmedi. Lojistik regresyon analizine göre yaş, OP için bağımsız bir risk faktörü olarak belirlendi [risk oranı (RO)=1,13, %95 güven aralığı (GA): 1,04-1,24, $p=0,005$] ve düşük doz GK kullanımı OP riskiyle negatif yönde ilişkili bulundu (RO=0,36, %95 GA: 0,11-0,99, $p=0,048$).
Sonuç: Bu çalışma, düşük doz GK tedavisinin RA hastalarında OP riskini artırmadığı ve bazı durumlarda bu riski azaltabileceğini göstermektedir. Yaş, OP açısından temel belirleyicilerden biri olarak öne çıkmakta ve kemik sağlığını korumada kişiselleştirilmiş tedavi stratejilerinin önemini vurgulamaktadır.
Anahtar Kelimeler: Romatoid artrit, glukokortikoidler, düşük doz kortikosteroidler, kemik mineral yoğunluğu, osteoporoz

Sorumlu Yazar/Corresponding Author: Sinem Kübra Beke MD, Erciyes University Faculty of Medicine, Department of Physical Medicine and Rehabilitation, Division of Rheumatology, Kayseri, Türkiye
E-posta: drsinemkubrabeke@gmail.com **ORCID ID:** orcid.org/0000-0002-4445-0551
Geliş Tarihi/Received: 24.03.2025 **Kabul Tarihi/Accepted:** 29.05.2025 **Epub:** 14.07.2025 **Yayınlanma Tarihi/Publication Date:** 05.12.2025

Atıf/Cite this article as: Beke SK, Kaplan H, Çakır TF, Cengiz G, Sas S, Cüce İ. Effects of low-dose corticosteroids on clinical outcomes and bone mineral density parameters in rheumatoid arthritis patients. Turk J Osteoporos. 2025;31(3):152-61

Introduction

Rheumatoid arthritis (RA) is a chronic autoimmune disease that predominantly targets the joints. It is characterized by a progressive symmetric inflammation of affected joints causing cartilage destruction, bone erosion, and disability (1). The worldwide prevalence of RA is estimated to be between 0.4% and 1.3%, depending on demographic and regional factors such as sex and age. The occurrence of RA is significantly higher in women, approximately two to three times that in men, and peaks during the sixth decade of life (2).

Osteoporosis (OP) is considered a significant extra-articular complication of RA, contributing to increased morbidity and fracture risk. OP is characterized as a progressive metabolic bone disease, leading to diminished bone mass and compromised microarchitecture of bone tissue. This results in elevated bone fragility and an increased fracture propensity (3). OP is more prevalent in RA patients than in the general population, contributing to increased morbidity and fracture risk (4,5). Bone loss due to RA is associated with some factors. Firstly, it has been demonstrated that this is associated with OP mechanisms. Secondly, it has been shown that the disease is associated with additional disease-specific factors. These additional factors include systemic inflammation, reduced physical activity, and inadequacies in disease management (6). Furthermore, the risk of fracture is higher in patients suffering from RA than in the general population, even in cases where bone density is preserved. The following general risk factors for OP have been identified: advanced age, low body mass index (BMI), female gender, previous history of fracture, menopause, hypogonadism or low vitamin D level, smoking, and alcohol use (7). However, in RA patients, factors such as chronic inflammation due to disease activity, long disease duration, presence of erosive disease, anti-citrullinated protein antibody (ACPA) positivity and physical activity limitation stemming from joint damage further accelerate bone loss (8).

Disease-modifying anti-rheumatic drugs (DMARDs) and glucocorticoids (GCs) utilized in the management of the disease also exert direct or indirect effects on bone metabolism. GCs are widely used in the management of RA and provide rapid symptomatic relief through their anti-inflammatory effects. GCs are utilized as bridge therapy in RA treatment or until the effect of DMARDs occurs, and they are effective in slowing disease progression and increasing the number of patients reaching remission. They are included in guidelines for the management of RA. Long-term or high-dose use of GCs has been associated with osteoporotic bone loss (9,10). However, the use of low-dose GCs may suppress osteoclast activity by controlling inflammation and thus partially offset bone loss (11). The stabilizing effect of GCs has been shown to depend on factors such as disease activity, dose of GCs administered, and duration of treatment. In addition, increasing the cumulative dose of GCs may double the risk of OP. Continuous oral GC use has been linked to dose-dependent bone loss and a higher risk of fractures, particularly within the first 3-6 months of treatment (12). Consequently, it is recommended that GCs be tapered as soon as clinically feasible (13).

The findings on the effects of low-dose GC use on bone health in RA patients have been inconclusive. Some studies have found a direct association between low-dose GC use and low bone mass (14-16). While others have not found this link (17,18). The observed variations may be attributed to factors such as the study population, the duration of GC use, dosage discrepancies, and additional variables impacting bone health. The present study was undertaken to conduct a comprehensive evaluation of bone health among individuals with RA receiving low-dose GC treatment. In this regard, the study's primary aim was to enhance the existing body of knowledge concerning the impact of these pharmaceutical agents on bone health. This objective was pursued through a meticulous analysis of bone density, the risk of OP, and the interplay of associated clinical factors.

Materials and Methods

Study Participants and Design

This prospective observational study was conducted in the Rheumatology Outpatient Clinic of Erciyes University Faculty of Medicine, Department of Physical Medicine and Rehabilitation.

Participants aged 18 years or older who met the RA classification criteria proposed by American College of Rheumatology/ European League Against Rheumatism (EULAR) in 2010 (19), and who agreed to participate in the study were included in the study after written informed consent was obtained. The clinical and demographic data of the patients were analyzed prospectively, while the results of the blood and imaging studies were evaluated retrospectively. In this study, a certain threshold was set for the evaluation of patients using GCs. Patients using 5 mg prednisolone or an equivalent dose of GCs for at least three months were classified as the low-dose corticosteroid group (20). Also demographic characteristics such as gender, age, and BMI, clinical data such as disease duration, comorbidities (diabetes, hypertension, cardiovascular diseases), smoking, drugs used in RA treatment, GC use and fracture history were evaluated with a standardized form.

Assessment of disease activity was performed using the disease activity score-28 (DAS-28), which is a tool designed to assess disease activity in RA across 28 joints. The DAS-28 score is calculated based on the count of tender and swollen joints, levels of erythrocyte sedimentation rate or C-reactive protein value, and the patient's global visual analog scale (VAS) assessment (21). The assessment of fatigue levels was conducted utilizing a VAS, wherein participants were invited to indicate their level of fatigue on a scale ranging from 0-10, with 0 representing no fatigue and 10 representing unbearable fatigue. The evaluation of quality of life was facilitated by the rheumatoid arthritis quality of life scale (RAQoL). The RAQoL has been meticulously developed to assess how RA influences daily life activities, individual skills, coping mechanisms and social relationships. The total score obtained from the scale ranges from 0-30, with higher scores indicating deterioration in quality of life (22). There is a Turkish validity and reliability study of the RAQoL scale (23).

Assessment of Bone Mineral Density

An investigation into body composition and bone mineral density (BMD) assessment was conducted, with BMD measurements being evaluated in the lumbar spine (L1-L4) and femur using dual-energy X-ray absorptiometry (DEXA). The DEXA measurements were performed at Erciyes University Faculty of Medicine, Department of Nuclear Medicine, with the data obtained using the Hologic QDR series X-ray bone densitometer (USA) being analyzed from the system. The diagnosis of OP was made according to the World Health Organization criteria, with a T-score ≤-2.5 being indicative of the condition (24).

The Erciyes University Faculty of Medicine Clinical Research Ethics Committee reviewed and approved the study protocol (approval no: 2021/613, date: 22.09.2021). The study was conducted following the tenets of the Declaration of Helsinki.

Statistical Analyses

IBM SPSS 23.0 software was used for data analysis. The Shapiro-Wilk test was employed to assess the normal distribution of data. To assess differences in categorical variables across groups, chi-square and Fisher's exact tests were applied. Normally distributed variables were compared using the independent samples t-test, whereas the Mann-Whitney U test was applied for data that did not follow a normal distribution. Logistic regression analysis (univariate and multiple models) was performed to identify independent predictors of OP in RA patients. Of the 74 patients included in the study, 39 who were diagnosed with OP were included in the regression analysis. Potential and statistically significant predictor variables in simple comparison or univariate analysis were included in the model. Adjusted odds ratios (ORs) with 95% confidence intervals (CIs) were calculated for each. Statistical significance was set at $p<0.05$ unless otherwise stated.

Results

The present study comprised a total of 74 patients and 39 healthy control subjects. The mean age of the patients was 56.81±7 years and 54.07±7.44 years in the control group. There was no statistically significant difference in age between the two groups (p=0.070). The study found that 97.3% (n=72) of the patients and 94.9% (n=37) of the control group were female, and no significant difference was found between the groups in terms of gender distribution (p=0.607). Similarly, no significant difference was found between the two groups in terms of BMI (p=0.146). The mean disease duration in the patient group was 14.02±6.20 years, with rheumatoid factor (RF) positivity detected in 74.3% and ACPA positivity in 81.7% of the patients. The prevalence of comorbidities among the patient population was found to be 66.7%, with the most prevalent being hypertension (34.6%) and neuropathic pain (21.0%). Regarding treatment modalities, 44.4% of patients were administered GCs, while 43.2% were undergoing biological DMARD therapy. Furthermore, 48.8% of the patient group were undergoing treatment for OP. The mean vitamin D levels were 22.95 ng/mL (17.55-32) in the patient group and 24.2 ng/mL (15.50-31.3) in the control group, and no significant difference was observed between the two groups (p=0.914).

Concerning BMD, the patient group exhibited statistically significantly lower T-scores. Notably, the T-scores in the L1-L4 vertebrae and femur regions demonstrated a marked decrease compared to the control group ($p<0.001$). Comprehensive data are delineated in Table 1.

Table 1. Demographic, clinical, and bone mineral density characteristics of patients with rheumatoid arthritis and healthy controls

	Patients (n=74)	Controls (n=39)	p
Age, year	56.81±7	54.07±7.44	0.070[β]
Female gender, n (%)	72 (97.3)	37 (94.9)	0.607
BMI (kg/m²)	31.8±5.6	33.4±5.3	0.146*
Disease duration, years	14.02±6.20		
RF positivity, n (%)	55 (74.3)		
RF titer	31.75 (12.88-77.40)		
ACPA positivity, n (%)	60 (81.7)		
ACPA titer	241.5 (33.12-500)		
Comorbidities			
Yes, n %	54 (66.7)		
Cardiac disease, n (%)	13 (16)		
Pulmonary disease, n (%)	16 (19.8)		
Hypertension (HT), n (%)	28 (34.6)		
Neuropathy, n (%)	17 (21.0)		
Diabetes mellitus (DM), n (%)	9 (11.1)		
Hypothyroidism, n (%)	9 (11.1)		
Current steroid dose (mg)	0 (0-5)		

Table 1. Continued

	Patients (n=74)	Controls (n=39)	p
Total steroid duration (months)	12 (3.75-30)		
Tender joint count	4.10±4.32		
Swollen joint count	0.45±1.02		
Patient global score	5.04±1.90		
Physician global score	4.36±1.86		
CRP (mg/L)	3.67 (1.5-7.71)		
ESR (mm/h)	23 (15-32.5)		
DAS28-ESR	3.43±1.04		
DAS28-CRP	3.02 (2-3.61)		
VAS Fatigue	5 (4-7)		
RAQoL	15.36±7		
Treatments			
Steroid use, n (%)	36 (44.4)		
MTX use, n (%)	61 (75.3)		
HCQ use, n (%)	57 (77)		
SSZ use, n (%)	34 (45.9)		
LEF use, n (%)	31 (41.9)		
bDMARD use, n (%)	32 (43.2)		
Anti-TNF use, n (%)	14 (18.9)		
JAK inhibitors use, n (%)	5 (6.8)		
Abatacept use, n (%)	3 (4.1)		
IL-6 inhibitor use, n (%)	4 (5.4)		
OP treatment, n (%)	39 (48.8)		
OP diagnosis duration (months)	18.5 (0-89)		
OP medications used			
Bisphosphonates use, n (%)	24 (29.6)		
Denosumab use, n (%)	16 (19.8)		
Vitamin D (ng/mL)	22.95 (17.55-32)	24.2 (15.50-31.3)	0.914
L1 BMD (g/cm²)	0.860±0.15	0.96±0.12	0.000*
L2 BMD (g/cm²)	0.90±0.16	1.01±0.12	0.000$^{\beta}$
L3 BMD (g/cm²)	1.22±2.33	1.02±0.21	0.005$^{\beta}$
L4 BMD (g/cm²)	0.92±0.14	1.0±0.21	0.001$^{\beta}$
L1-L4 BMD (g/cm²)	0.89±0.15	1.01±0.12	0.000$^{\beta}$
L1 T-score	-1.34±1.31	-0.27±1.02	0.000*
L2 T-score	-1.18±1.43	-0.21±1.11	0.000$^{\beta}$
L3 T-score	-1.3±1.19	-0.35±1.28	0.000
L4 T-score	-1.4±-1.28	-0.20±1.12	0.000$^{\beta}$
L1-L4 T-score	-1.3±1.20	-0.31±1.11	0.000$^{\beta}$
Femoral neck BMD (g/cm²)	0.73±0.11	0.80±0.27	0.000$^{\beta}$
Total femur BMD (g/cm²)	0.84±0.168	0.95±0.17	0.000$^{\beta}$
Femoral neck T-score	-0.99±1.06	-0.22±1.05	0.000
Total femur T-score	-0.79±0.98	0.26±0.96	0.000*

n: Number of patients, RF: Rheumatoid factor, ACPA: Anti-citrullinated protein antibody, BMI: Body mass index, CRP: C-reactive protein, ESR: Erythrocyte sedimentation rate, DAS28: Disease activity score in 28 joints, VAS: Visual analog scale, RAQoL: Rheumatoid arthritis quality of life, MTX: Methotrexate, HCQ: Hydroxychloroquine, SSZ: Sulfasalazine, LEF: Leflunomide, bDMARD: Biologic disease-modifying antirheumatic drug, TNF: Tumor necrosis factor, JAK: Janus kinase, IL-6: Interleukin-6, BMD: Bone mineral density, T-score: Standard deviation score used to assess bone density, OP: Osteoporosis, *: Independent Samples t-test, $^{\beta}$: Mann-Whitney U test

A comparison was made between individuals who used corticosteroids (n=36) and those who did not use corticosteroids (n=38). The results revealed that BMI (p=0.012) and total femur T-score (p=0.011) were significantly lower in the corticosteroid group. However, no significant differences were observed between the two groups for other demographic, clinical and laboratory parameters. A comprehensive overview of the findings is provided in Table 2.

In the present study, a comparison was made between RA patients diagnosed with OP who used corticosteroids (n=24) and those who did not use corticosteroids (n=15). No significant difference was found in terms of demographic, clinical and laboratory characteristics, and DEXA results (Table 3).

Logistic regression analysis was performed to evaluate the factors affecting the presence of OP in both univariate and multivariate models. Univariate analysis revealed a significant

Table 2. Comparison of bone mineral density and t-scores in rheumatoid arthritis patients with and without corticosteroid use

	Patients without corticosteroid (n=38)	Patients with low-dose corticosteroid (n=36)	p
Age (years)	55.7±7.4	57.9±6.4	0.197$^{\beta}$
Female gender, n (%)	37 (97.4)	35 (97.2)	0.969
BMI (kg/m²)	33.35±5.20	30.13±5.52	**0.012***
Disease duration (years)	14.23±6	13.4±6.60	0.564*
CRP (mg/L)	5.16±7.17	7.48±8.08	0.241$^{\beta}$
ESR (mm/h)	26.27±12.9	23.9±15	0.390$^{\beta}$
DAS28-ESR	3.39±1.02	3.5±1.08	0.567*
DAS28-CRP	2.89±1.22	2.98±1	0.364$^{\beta}$
VAS fatigue	5.35±2	5.5±1.72	0.542*
RAQOL	14.21±6.46	16.57±7.40	0.132*
Vitamin D level (ng/mL)	24.83±9.38	24.7±11.05	0.491*
Fracture presence, n (%)	3 (8.3)	4 (12.1)	0.603
Hip fracture presence, n (%)	0 (0)	1 (3)	0.293
Smoking status			
Never	28 (77.8)	24 (72.7)	0.874
Former smoker	5 (13.9)	6 (18.2)	
Current smoker	3 (8.3)	3 (9.1)	
Exercise frequency			
None	14 (38.9)	20 (60.06)	0.107
1-5 days per week	20 (55.6)	10 (30.3)	
≥5 days per week	2 (5.6)	3 (9.1)	
L1 BMD (g/cm²)	0.86±0.14	0.85±0.16	0.777*
L2 BMD (g/cm²)	0.91±0.16	0.90±0.16	0.461*
L3 BMD (g/cm²)	0.95±0.13	0.98±0.23	0.737$^{\beta}$
L4 BMD (g/cm²)	0.93±0.15	0.90±0.14	0.363*
L1-L4 BMD (g/cm²)	0.90±0.16	0.90±0.14	0.830*
L1 T-score	-1.25±1.25	-1.43±1.4	0.478*
L2 T-score	-1.05±1.5	-1.31±1.41	0.392*
L3 T-score	-1.23±1.19	-1.33±1.21	0.651*
L4 T-score	-1.20±1.31	-1.4±1.25	0.388*
L1-L4 T-score	-1.21±1.21	-1.40±1.22	0.429*
Femoral neck BMD (g/cm²)	0.75±0.115	0.70±0.09	0.33*
Total femur BMD (g/cm²)	0.87±0.12	0.79±0.19	0.069$^{\beta}$
Femoral neck T-score	0.88±0.13	0.81±0.12	0.13*
Total femur T-score	-0.5±0.94	-1.09±0.94	**0.011***

n: Number of patients, BMI: Body mass index, CRP: C-reactive protein, ESR: Erythrocyte sedimentation rate, DAS28: Disease activity score in 28 joints, VAS: Visual analog scale, RAQoL: Rheumatoid arthritis quality of life, BMD: Bone mineral density, *: Independent samples t-test, $^{\beta}$: Mann-Whitney U test

association between the risk of OP and age and corticosteroid treatment. The risk of OP increased significantly with increasing age (OR=1.15, 95% CI: 1.06-1.25, p=0.001). Conversely, patients receiving corticosteroid therapy exhibited a significantly lower risk of OP (OR=0.34, 95% CI: 0.13-0.88, p=0.027). The analysis revealed that factors such as gender, BMI, disease duration, vitamin D levels, and fracture history did not attain statistical significance in their association with OP (p>0.05). However, no significant association between ACPA and RF and BMD was found in our study (data not shown).

Multivariate analysis indicated that age and corticosteroid treatment were independent risk factors for OP. The risk of OP increased with age (OR=1.13, 95% CI: 1.04-1.24, p=0.005), while the risk remained significantly lower in patients receiving corticosteroid therapy (OR=0.36, 95% CI: 0.11-0.99, p=0.048). The remaining variables exhibited no statistically significant impact on OP (Table 4).

Discussion

The present study evaluated the effects of low-dose corticosteroid use on BMD, OP risk, and clinical parameters in patients with RA. A comparison was made between the BMD values measured by DEXA in RA patients and those in a healthy control group. It was found that BMD values were significantly lower and T-scores decreased in the L1-L4 vertebrae and femur regions in RA patients. Furthermore, in the RA patient group, BMI and total femur T-score were lower in patients using corticosteroids. However, no significant differences were observed between the two groups concerning other demographic, clinical, and laboratory variables. Logistic regression analysis revealed that age emerged as an independent risk factor for OP, while corticosteroid use was not found to increase the risk of OP, but was inversely associated with the likelihood of its development.

The risk of OP is elevated in patients suffering from RA in comparison to the general population, and this increase can

Table 3. Comparison of clinical and bone mineral density parameters between rheumatoid arthritis patients with osteoporosis using and not using corticosteroids

	GC users (n=24)	Non-GC users (n=15)	p
Age (years)	59±6.46	60±3.83	0.618$^{\beta}$
Female gender, n (%)	23 (95.8)	15 (100)	0.423
BMI (kg/m^2)	30.6±5.5	33.6±5.6	0.064*
Disease duration (years)	14.14±6.48	15.06±5.20	0.857*
RF positivity, n (%)	16 (66.7)	10 (66.7)	1
ACPA positivity, n (%)	20 (83.3)	10 (66.7)	0.229
Vitamin D level (ng/mL)	27.39±11.15	24.86±9.80	0.467$^{\beta}$
Fracture presence, n (%)	3 (13.6)	1 (6.7)	0.503
Hip fracture presence, n (%)	1 (4.5)	0 (0)	0.403
L1 BMD (g/cm^2)	0.76±0.10	0.73±0.09	0.282*
L2 BMD (g/cm^2)	0.805±0.105	0.781±0.104	0.377*
L3 BMD (g/cm^2)	1.814±4.32	0.828±0.089	0.172$^{\beta}$
L4 BMD (g/cm^2)	0.856±0.109	0.819±0.109	0.336*
L1-L4 BMD (g/cm^2)	0.822±0.81	0.794±0.085	0.248*
L1 T-score	-2. 17±0.91	-2.3±0.88	0.540*
L2 T-score	-2. 05±1.01	-2.25±0.97	0.424*
L3 T-score	-2.02±0.78	-2.32±0.80	0.206*
L4 T-score	-1.86±1	-2.2±0.99	0.330*
L1-L4 T-score	-2.05±0.76	-2.3±0.77	0.280*
Femoral neck BMD (g/cm^2)	0.67±0.07	0.71±0.11	0.279*
Total femur BMD (g/cm^2)	0.77±0.11	0.82±0.09	0.156*
Femoral neck T-score	-1.38±1.13	-0.95±1.17	0.133$^{\beta}$
Total femur T-score	-1.4±0.92	-0.1±0.75	0.165*

n: Number of patients, BMI: Body mass index, RF: Rheumatoid factor, ACPA: Anti-citrullinated protein antibody, BMD: Bone mineral density, GC: Glucocorticoid, *: Independent samples t-test, β: Mann-Whitney U test

Table 4. Univariate and multivariate logistic regression analysis of factors associated with osteoporosis

	Osteoporosis			
	Univariate		Multivariate	
	OR (95% CI)	p	OR (95% CI)	p
Age (years)	1.15 (1.06-1.25)	0.001	1.13 (1.04-1.24)	0.005
Gender (female)	1.15 (0.07-19)	0.922		
BMI (kg/m²)	0.97 (0.89-1.06)	0.574		
Disease duration (years)	1.03 (0.97-1.10)	0.252	1.04 (0.95-1.13)	0.393
Vitamin D level (ng/mL)	1.03 (0.98-1.08)	0.208	1.03 (0.97-1.09)	0.294
Fracture presence, n (%)	0.85 (0.17-4.13)	0.844		
Steroid use, n (%)	0.34 (0.13-0.88)	0.027	0.36 (0.11-0.99)	0.048

OR: Odds ratio, CI: Confidence interval, BMI: Body mass index

be attributed to several interconnected mechanisms. Immune-mediated processes and inflammatory cytokines contribute to the development of OP by directly affecting bone metabolism. Chronic inflammation in RA causes bone loss by increasing osteoclast activity driven by synovial fibroblasts, macrophages, and Th17 cells. This process promotes osteoclastogenesis through increased expression of RANKL and suppression of OPG. Furthermore, the Wnt pathway, which is critical for osteoblast activity, is inhibited by increased DKK-1 and sclerostin levels (25). This results in a reduction in the activity of osteoblasts, leading to insufficient bone formation.

Pro-inflammatory cytokines (e.g., tumor necrosis factor-alpha, interleukin-6, and interleukin-1) play a key role in RA-related bone loss. These cytokines enhance osteoclast activity through increased RANKL expression while concurrently suppressing osteoblast function by inhibiting the Wnt signaling pathway. This imbalance causes reduced bone formation and increased bone resorption, contributing to OP in RA patients (26). A Mendelian randomization study has indicated that the association between RA and OP/BMD may be predominantly attributable to secondary factors. These factors encompass the utilization of antirheumatic therapy and diminished physical activity. It is conceivable that prior observational studies may have been influenced by these variables (27).

GCs cause bone loss by promoting osteoblast and osteocyte apoptosis, inhibiting osteoblast formation, and increasing osteoclast activity. GCs impair osteoblast function and negatively affect bone turnover by suppressing the WNT and RANK-RANKL-osteoprotegerin pathways. Increased levels of RANKL lead to bone destruction by increasing osteoclast activity, while inhibition of the WNT pathway reduces osteoblast activity (26,28). GCs inhibit osteoblast function by suppressing IGF-1 transcription. They also affect bone mineralization by reducing intestinal calcium absorption and causing loss of muscle mass. These mechanisms play a central role in the development of GC-induced OP (29,30). The impact of low-dose corticosteroids on bone health remains controversial, with conflicting evidence in the literature (31). While some studies suggest that corticosteroids may increase bone loss, other studies show that this effect is not evident with low-dose and short-term use (32). In addition, corticosteroid use is generally associated with a more severe disease course and poorer functional status, suggesting an association between the degree of systemic inflammation and fracture risk. Some recent studies have suggested that low-dose, short-term systemic GC use may stabilize bone metabolism by controlling inflammation rather than increasing the risk of bone loss (5).

Previous studies have reported that the use of corticosteroids causes a loss of BMD, particularly in the hip. A meta-analysis showed that RA patients treated with GC had decreased BMD at the femoral neck and spine and higher fracture rates compared with non-GC or healthy controls (33). Hall et al. (34) observed a decrease in BMD at the spine and hip in RA patients using steroids, but this decrease was more pronounced in the hip region. Martin et al. (14) demonstrated a decrease in appendicular (hip region) bone mass in RA patients receiving low-dose corticosteroids. In addition, another study suggested that the most adverse effect of corticosteroids was seen in the hip rather than the spine, and that the reliability of spinal BMD measurements may be reduced by comorbidities such as osteoarthritis (35). Our study also supports these findings because the total femur T-score was significantly lower in patients using GCs compared to the RA group. This suggests that GCs may increase bone loss, particularly in the hip, and bone health in the femur may be more at risk. However, it should be noted that in the multivariate regression analysis of our study, the use of GCs was found to be a protective factor against OP. This finding suggests that the anti-inflammatory effects of GCs may have a potential protective role against bone loss. Chronic inflammation itself is an important factor that can cause bone loss in RA, and GCs may indirectly limit bone loss by suppressing this inflammation. In addition, patients on corticosteroids are more frequently screened for OP and may be more likely to receive OP treatment, which may also explain the protective effect.

One study showed that RA patients had lower BMD compared with an age-matched non-RA control group. BMD values in the RA control group were found to be higher than those in the RA group on steroids, despite longer disease duration. In

particular, it has been reported that the synergistic effect of methotrexate and low-dose GCs may contribute to bone loss, but GCs combined with DMARDs may have beneficial effects in terms of preserving bone mass (36). Significant BMD loss at the femoral neck and lumbar spine was found in RA patients with and without low-dose prednisolone in a study of 84 RA patients (33). Another study found no significant difference in BMD loss between the two groups during a mean follow-up of 89.6 months and concluded that low-dose oral GCs do not increase the overall risk of OP in RA patients (17). Similarly, another study found that low-dose GCs were not directly associated with BMD (37) but cumulative prednisone levels were higher in osteopenic women (38).

However, a study using the Clinical Practice Research Datalink found no difference in the risk of osteoporotic fractures between current low-dose oral GC use (≤7.5 mg PED/day) and past GC use in RA patients. However, low-dose GC use was associated with an increased risk of clinical vertebral fracture (39). In another study, no increased risk of OP fracture was found in RA patients with an average daily dose ≤5.0 mg. However, an increased risk of OP fracture was reported for doses >5.0 mg (40). In addition, a review conducted by the EULAR showed that reviews of the harms of long-term GC therapy (including OP and OP fractures) in RA patients could not reach a definitive conclusion for the dose range of 5-10 mg PED daily (41). These conflicting results suggest that the effects of GCs on bone health should be assessed not only by BMD measurements, but also by fracture risk and bone turnover markers. Findings suggest that osteoporotic fractures may occur early in women with RA, independent of GC use (42). A previous study reported that the risk of osteoporotic fracture was higher in patients using GCs, even in the presence of osteopenia, compared with the control group. It was emphasized that this increased risk was also seen in patients using low doses of GCs and that the adverse effects of GCs on bone were not only due to high-dose use (20). There are studies suggesting that bone quality, which cannot be assessed with current techniques, may be more affected by GCs than BMD. Van Everdingen et al. (43) reported an increased rate of vertebral fractures despite no significant difference in BMD. Various methods such as magnetic resonance imaging, quantitative computed tomography, and nuclear magnetic resonance have been proposed to measure bone quality (44). However, because BMD measurement is inexpensive and widely available, it is the most commonly used method in clinical practice. Therefore, almost all diagnostic criteria for OP and osteopenia are based on BMD (24). This suggests that although BMD is an important indirect measure of fracture risk, the risk of fracture from corticosteroid use may be high, especially in postmenopausal patients. Therefore, a more comprehensive assessment of the effects of GCs on bone must take into account the incidence of fractures and structural changes in bone microstructure. It has been suggested that patients on GCs have a higher risk of osteoporotic fractures than controls, despite having higher BMD (45). A meta-analysis of observational studies showed that the incidence of hip and vertebral fractures after GC use was higher than estimates based on BMD reduction alone (46). The fact that GCs induces micro-architectural changes in certain active bone zones may explain this (47). However, these changes are not directly reflected by a decrease in BMD. In our study, although we evaluated the incidence of fractures, we did not find a significant difference between the groups in terms of hip fractures. However, the limited number of participants makes it difficult to make a definitive judgment on this issue. Larger studies with longer follow-ups are needed to better understand the changes in bone microstructure, especially in patients using GCs. In the future, analyses based not only on BMD measurements but also on advanced techniques to assess bone microarchitecture may help to predict fracture risk more accurately.

Body composition and age are closely related to BMD. Previous studies have shown that lean body mass and fat mass have beneficial effects on bone mass. In particular, it has been reported that lean body mass may support bone mass through mechanical loading forces on bone, whereas fat mass may influence osteoblast/osteoclast function by increasing BMD through hormonal metabolism of adipocytes (37). In the present study, there was no significant difference in BMI between the healthy control group and the RA patients, while it was observed that the BMI of RA patients using GCs was lower. Previous studies have shown that fat and lean body mass are significantly lower in osteoporotic women (37). The metabolic effects of GC use and loss of lean mass through increased muscle protein breakdown may be one of the mechanisms that explain the lower BMI. In addition, the low levels of physical activity commonly observed in RA patients may contribute to both bone loss and loss of lean mass.

The relationship between age and the loss of bone mass in people with RA has been shown in previous studies (31). This association is expected because older people have a longer exposure to risk factors that can lead to bone loss. These factors include GC use, low estrogen levels, prolonged immobility, and increased inflammation. Similarly, in our study, we used logistic regression analysis to show that older age is an independent risk factor for OP. This finding highlights the determinant effect of age on bone health and supports the importance of early OP screening in RA patients.

Previous studies have suggested that RF may exert a titration-dependent effect on BMD, and that the effect may be more pronounced in combination with ACPA by creating a subclinical inflammatory environment due to increased immune complex activity. Furthermore, low BMD, particularly in the femoral neck, has been reported to be associated with the presence of ACPA and RF. Therefore, careful monitoring of OP in seropositive RA patients is recommended (48). However, no significant association between ACPA and RF and BMD was found in our study.

A multifaceted approach, including lifestyle changes, optimal nutrition, and structured exercise programs, is essential to

reduce the risk of OP and fractures in people with RA. Studies have shown that regular weight-bearing and resistance exercise improves bone and muscle health and reduces the risk of fractures. These non-pharmacological strategies, when combined with appropriate medical treatment, can help people with RA maintain bone strength and overall musculoskeletal health (49). However, our study observed that approximately half of the RA patients did not engage in regular exercise. This finding highlights that physical inactivity remains a significant issue among RA patients and represents a contributing factor to the increased risk of OP. Lack of exercise can lead to a reduction in lean body mass, decreased muscle strength, and accelerated bone loss, thereby increasing the risk of falls and fractures through a critical pathophysiological mechanism. Therefore, strategies for OP management in RA patients should not be limited to pharmacological treatment alone but should also incorporate structured exercise programs and lifestyle modifications.

The study used a prospective design and analyzed the patients' clinical and laboratory data in detail. Factors that may influence the risk of OP, such as vitamin D levels, smoking, and exercise frequency, were recorded, and the groups were directly compared with regard to GC use. In addition, OP risk factors were evaluated using both univariate and multivariate models, and new data were presented that may contribute to studies in this area, which have so far produced conflicting results.

Study Limitations

Our study has some limitations. The relatively small number of patients may limit the generalizability of the results, so studies with larger cohorts are needed. In addition, as this is an observational study, it is not possible to establish a causal relationship with certainty. The long-term effects of the duration and dose of corticosteroid use could not be fully assessed in this study, so this effect should be investigated in more detail in future long-term follow-up studies. In addition, dietary intake and dietary characteristics were not assessed, so these factors should be investigated in future studies for their possible effects on OP risk.

Conclusion

Our findings suggest that low-dose GC use does not elevate OP risk in RA patients and may, in some cases, be linked to a lower risk. However, further large-scale longitudinal studies are needed to confirm this association. Age stands out as one of the strongest predictors of OP. Adopting individualized treatment approaches to effectively manage OP in RA patients is considered an important strategy for maintaining bone health.

Acknowledgment

We would like to thank the Proofreading & Editing Office of the Dean for Research at Erciyes University for the copyediting and proofreading service for this manuscript.

Ethics

Ethics Committee Approval: The Erciyes University Faculty of Medicine Clinical Research Ethics Committee reviewed and approved the study protocol (approval no: 2021/613, date: 22.09.2021).

Informed Consent: Who agreed to participate in the study were included in the study after written informed consent was obtained.

Footnotes

Authorship Contributions

Concept: S.K.B., H.K., T.F.Ç., G.C., S.S., İ.C., Design: S.K.B., H.K., T.F.Ç., G.C., S.S., İ.C., Data Collection or Processing: S.K.B., H.K., T.F.Ç., G.C., S.S., İ.C., Analysis or Interpretation: S.K.B., H.K., T.F.Ç., G.C., S.S., İ.C., Literature Search: S.K.B., H.K., T.F.Ç., G.C., S.S., İ.C., Writing: S.K.B., H.K., T.F.Ç., G.C., S.S., İ.C.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declare that this study received no financial support.

References

1. Aletaha D, Smolen JS. Diagnosis and management of rheumatoid arthritis: a review. JAMA. 2018;320:1360-72.
2. Lin YJ, Anzaghe M, Schülke S. Update on the pathomechanism, diagnosis, and treatment options for rheumatoid arthritis. Cells. 2020;9:880.
3. Gómez O, Talero AP, Zanchetta MB, Madeira M, Moreira CA, Campusano C, et al. Diagnostic, treatment, and follow-up of osteoporosis-position statement of the Latin American Federation of Endocrinology. Arch Osteoporos. 2021;16:114.
4. Hauser B, Riches PL, Wilson JF, Horne AE, Ralston SH. Prevalence and clinical prediction of osteoporosis in a contemporary cohort of patients with rheumatoid arthritis. Rheumatology. 2014;53:1759-66.
5. Fardellone P, Salawati E, Le Monnier L, Goëb V. Bone loss, osteoporosis, and fractures in patients with rheumatoid arthritis: a review. J Clin Med. 2020;9:3361.
6. Coulson KA, Reed G, Gilliam BE, Kremer JM, Pepmueller PH. Factors influencing fracture risk, T score, and management of osteoporosis in patients with rheumatoid arthritis in the Consortium of Rheumatology Researchers of North America (CORRONA) registry. J Clin Rheumatol. 2009;15:155-60.
7. Wysham KD, Baker JF, Shoback DM. Osteoporosis and fractures in rheumatoid arthritis. Curr Opin Rheumatol. 2021;33:270-6.
8. Baker R, Narla R, Baker JF, Wysham KD. Risk factors for osteoporosis and fractures in rheumatoid arthritis. Best Pract Res Clin Rheumatol. 2022;36:101773.
9. Güler-Yüksel M, Hoes JN, Bultink IE, Lems WF. Glucocorticoids, inflammation and bone. Calcif Tissue Int. 2018;102:592-606.
10. Fenton C, Webster J, Martin C, Fareed S, Wehmeyer C, Mackie H, et al. Therapeutic glucocorticoids prevent bone loss but drive muscle wasting when administered in chronic polyarthritis. Arthritis Res Ther. 2019;21:1-12.
11. Gøtzsche PC, Johansen HK. Short-term low-dose corticosteroids vs placebo and nonsteroidal antiinflammatory drugs in rheumatoid arthritis. Cochrane Database Syst Rev. 2004;2005:CD000189
12. Da Silva JA, Jacobs JW, Kirwan JR, Boers M, Saag KG, Inês LB, et al. Safety of low-dose glucocorticoid treatment in rheumatoid arthritis: published evidence and prospective trial data. Ann Rheum Dis. 2006;65:285-93.

13. Humphrey MB, Russell L, Danila MI, Fink HA, Guyatt G, Cannon M, et al. 2022 American College of Rheumatology guideline for the prevention and treatment of glucocorticoid-induced osteoporosis. Arthritis Rheumatol. 2023;75:2088-102.
14. Martin J, Munro R, Campbell M, Reid D. Effects of disease and corticosteroids on appendicular bone mass in postmenopausal women with rheumatoid arthritis: comparison with axial measurements. Br J Rheumatol. 1997;36:43-9.
15. Saario R, Sonninen P, Möttönen T, Viikari J, Toivanen A. Bone mineral density of the lumbar spine in patients with advanced rheumatoid arthritis: influence of functional capacity and corticosteroid use. Scand J Rheumatol. 1999;28:363-7.
16. Haugeberg G, Ørstavik RE, Uhlig T, Falch JA, Halse JI, Kvien TK. Bone loss in patients with rheumatoid arthritis: results from a population-based cohort of 366 patients followed up for two years. Arthritis Rheum. 2002;46:1720-8.
17. Sambrook P, Eisman J, Yeates M, Pocock N, Eberl S, Champion G. Osteoporosis in rheumatoid arthritis: safety of low-dose corticosteroids. Ann Rheum Dis. 1986;45:950-3.
18. Sambrook P, Raj A, Hunter D, Naganathan V, Mason R, Robinson B. Osteoporosis with low-dose corticosteroids: contribution of underlying disease effects and discriminatory ability of ultrasound versus bone densitometry. J Rheumatol. 2001;28:1063-7.
19. Initiative C. 2010 rheumatoid arthritis classification criteria. Arthritis Rheum. 2010;62:2569-81.
20. Park SY, Ahn SH, Bae GH, Jang S, Kwak MK, Kim HY, et al. Low-dose glucocorticoid increases the risk of fracture in postmenopausal women with low bone mass: a retrospective cohort study. Osteoporos Int. 2024;35:1779-87.
21. Prevoo M, Van'T Hof MA, Kuper H, Van Leeuwen M, Van De Putte L, Van Riel P. Modified disease activity scores that include twenty-eight-joint counts: development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum. 1995;38:44-8.
22. De Jong Z, Van der Heijde D, McKenna S, Whalley D. The reliability and construct validity of the RAQoL: a rheumatoid arthritis-specific quality of life instrument. Br J Rheumatol. 1997;36:878-83.
23. Kutlay S, Küçükdeveci AA, Gönül D, Tennant A. Adaptation and validation of the Turkish version of the Rheumatoid Arthritis Quality of Life Scale. Rheumatol Int. 2003;23:21-6.
24. World Health Organization. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis: report of a WHO Study Group [meeting held in Rome, 22-25 June 1992]. Geneva: World Health Organization; 1994.
25. Redlich K, Smolen JS. Inflammatory bone loss: pathogenesis and therapeutic intervention. Nat Rev Drug Discov. 2012;11:234-50.
26. Choi YJ, Chung YS, Suh CH, Jung JY, Kim HA. Trabecular bone score as a supplementary tool for the discrimination of osteoporotic fractures in postmenopausal women with rheumatoid arthritis. Medicine (Baltimore). 2017;96:e8661.
27. Liu YQ, Liu Y, Chen ZY, Li H, Xiao T. Rheumatoid arthritis and osteoporosis: a bi-directional Mendelian randomization study. Aging (Albany NY). 2021;13:14109.
28. Piemontese M, Xiong J, Fujiwara Y, Thostenson JD, O'Brien CA. Cortical bone loss caused by glucocorticoid excess requires RANKL production by osteocytes and is associated with reduced OPG expression in mice. Am J Physiol Endocrinol Metab. 2016;311:E587-93.
29. Delany AM, Durant D, Canalis E. Glucocorticoid suppression of IGF I transcription in osteoblasts. Mol Endocrinol. 2001;15:1781-9.
30. Sato AY, Richardson D, Cregor M, Davis HM, Au ED, McAndrews K, et al. Glucocorticoids induce bone and muscle atrophy by tissue-specific mechanisms upstream of E3 ubiquitin ligases. Endocrinology. 2017;158:664-77.
31. Lanyon L, Skerry T. Perspective: postmenopausal osteoporosis as a failure of bone's adaptation to functional loading: a hypothesis. J Bone Miner Res. 2001;16:1937-47.
32. Engvall I, Brismar K, Hafström I, Tengstrand B. Treatment with low-dose prednisolone is associated with altered body composition but no difference in bone mineral density in rheumatoid arthritis patients: a controlled cross-sectional study. Scand J Rheumatol. 2011;40:161-8.
33. Wang Y, Zhao R, Gu Z, Dong C, Guo G, Li L. Effects of glucocorticoids on osteoporosis in rheumatoid arthritis: a systematic review and meta-analysis. Osteoporos Int. 2020;31:1401-9.
34. Hall GM, Spector TD, Jane Griffin A, Jawad ASM, Hall ML, Doyle DV. The effect of rheumatoid arthritis and steroid therapy on bone density in postmenopausal women. Arthritis Rheum. 1993;36:1510-6.
35. Chung CP, Russell AS, Segami MI, Ugarte CA. The effect of low-dose prednisone on bone mineral density in Peruvian rheumatoid arthritis patients. Rheumatol Int. 2005;25:114-7.
36. Heidari B, Heidari P, Hajian-Tilaki K, Bayani MA, Babaei M. Effect of long-term low-dose prednisolone administration on bone mineral density: relating to non-compliant women with rheumatoid arthritis. Caspian J Intern Med. 2018;9:171.
37. Sarkis KS, Salvador MB, Pinheiro MM, Silva RG, Zerbini CA, Martini LA. Association between osteoporosis and rheumatoid arthritis in women: a cross-sectional study. Sao Paulo Med J. 2009;127:216-22.
38. Dykman TR, Gluck OS, Murphy WA, Hahn TJ, Hahn BH. Evaluation of factors associated with glucocorticoid-induced osteopenia in patients with rheumatic diseases. Arthritis Rheum. 1985;28:361-8.
39. Abtahi S, Driessen JH, Burden AM, Souverein PC, van den Bergh JP, van Staa TP, et al. Low-dose oral glucocorticoid therapy and risk of osteoporotic fractures in patients with rheumatoid arthritis: a cohort study using the Clinical Practice Research Datalink. Rheumatology. 2022;61:1448-58.
40. Robinson DE, Van Staa TP, Dennison EM, Cooper C, Dixon WG. The limitations of using simple definitions of glucocorticoid exposure to predict fracture risk: a cohort study. Bone. 2018;117:83-90.
41. Strehl C, Bijlsma JW, De Wit M, Boers M, Caeyers N, Cutolo M, et al. Defining conditions where long-term glucocorticoid treatment has an acceptably low level of harm to facilitate implementation of existing recommendations: viewpoints from an EULAR task force. Ann Rheum Dis. 2016;75:952-7.
42. Ketabforoush AHME, Aleahmad M, Qorbani M, Mehrpoor G, Afrashteh S, Mardi S, et al. Bone mineral density status in patients with recent-onset rheumatoid arthritis. J Diabetes Metab Disord. 2023;22:775-85.
43. Van Everdingen A, Van Reesema DS, Jacobs J, Bijlsma J. Low-dose glucocorticoids in early rheumatoid arthritis: discordant effects on bone mineral density and fractures? Clin Exp Rheumatol. 2003;21:155-60.
44. Ishii S, Cauley JA, Greendale GA, Crandall CJ, Danielson ME, Ouchi Y, et al. C-reactive protein, bone strength, and nine-year fracture risk: data from the Study of Women's Health Across the Nation (SWAN). J Bone Miner Res. 2013;28:1688-98.
45. Laurent MR, Goemaere S, Verroken C, Bergmann P, Body JJ, Bruyère O, et al. Prevention and treatment of glucocorticoid-induced osteoporosis in adults: consensus recommendations from the Belgian Bone Club. Front Endocrinol (Lausanne). 2022;13:908727.
46. Van Staa T. The pathogenesis, epidemiology and management of glucocorticoid-induced osteoporosis. Calcif Tissue Int. 2006;79:129-37.
47. Compston J. Glucocorticoid-induced osteoporosis: an update. Endocrine. 2018;61:7-16.
48. Sargın G, Köse R, Şentürk T. Relationship between bone mineral density and anti-citrullinated protein antibody and rheumatoid factor in patients with rheumatoid arthritis. Eur J Rheumatol. 2018;6:29.
49. Gupta N, Kanwar N, Arora A, Khatri K, Kanwal A. The interplay of rheumatoid arthritis and osteoporosis: exploring the pathogenesis and pharmacological approaches. Clin Rheumatol. 2024;43:1421-33.

Original Investigation / Orijinal Araştırma

DOI: 10.4274/tod.galenos.2025.01488
Turk J Osteoporos 2025;31(3):162-8

A New Era in Digital Health: Evaluation of Artificial Intelligence Supported Chatbots (ChatGPT-4, BingAI, and Gemini) Responses on Osteoporosis

Dijital Sağlıkta Yeni Bir Dönem: Yapay Zeka Destekli Sohbet Robotlarının (ChatGPT-4, BingAI ve Gemini) Osteoporozla İlgili Yanıtlarının Değerlendirilmesi

Gamze Gül Güleç[1], Özge Özpolat Bulut[2], Fatih Bağıcer[3]

[1]Başkent University Adana Dr. Turgut Noyan Education and Research Center, Department of Physical Medicine and Rehabilitation, Adana, Türkiye
[2]Viranşehir State Hospital, Clinic of Physical Medicine and Rehabilitation, Şanlıurfa, Türkiye
[3]University of Health Sciences Türkiye, Başakşehir Çam and Sakura City Hospital, Department of Physical Medicine and Rehabilitation, İstanbul, Türkiye

Abstract

Objective: This study aimed to evaluate and compare the quality and readability of osteoporosis-related information generated by three artificial intelligence (AI) chatbots: ChatGPT-4, BingAI, and Gemini.
Materials and Methods: 25 frequently asked questions about osteoporosis (obtained via Google Trends through Gemini) were submitted to each chatbot on December 23, 2024. The first responses were evaluated for readability [Flesch-Kincaid Reading Ease (FKRE) and Flesch-Kincaid Grade Level (FKGL)] and quality (EQIP tool). Two experienced clinicians assessed the accuracy and completeness using Likert scales.
Results: The mean FKRE scores were 34.5±12.9 (ChatGPT-4), 33.8±14.3 (BingAI), and 36.1±10.9 (Gemini), indicating difficulty in reading the texts. The FKGL scores ranged from 11.2 to 12.5, suggesting that college-level reading ability was required. However, BingAI (EQIP: 55.4±7.9) and Gemini (54.4±8.8) outperformed ChatGPT-4 (48.6±6.3) in terms of quality (p=0.005). Accuracy and completeness were high across all models, with mean scores exceeding 4.3/5 for each.
Conclusion: While all three AI chatbots delivered accurate and complete answers on osteoporosis, their content readability remained suboptimal. BingAI and Gemini provide higher-quality information, possibly due to real-time web integration. Future chatbot development should focus on enhancing readability and real-time data access to support effective health communication, particularly in conditions such as osteoporosis, where patient understanding is crucial.
Keywords: Osteoporosis, artificial intelligence, chatbot, ChatGPT, readability, health communication

Öz

Amaç: Bu çalışmanın amacı, yapay zeka (AI) destekli sohbet robotlarının (ChatGPT-4, BingAI ve Gemini) osteoporoz ile ilgili verdiği bilgilerin kalitesini ve okunabilirliğini değerlendirmek ve karşılaştırmaktır.
Gereç ve Yöntem: Google Trends üzerinden osteoporoz hakkında en sık sorulan 25 soru belirlendi ve her bir sohbet robotuna ayrı ayrı soruldu. İlk verilen yanıtlar okunabilirlik [Flesch-Kincaid Okuma Kolaylığı (FKRE) ve Flesch-Kincaid Sınıf Düzeyi (FKGL)] ve bilgi ve yazım kalitesi (EQIP aracı) açısından değerlendirildi. Yanıtların doğruluğu ve yeterliliği iki deneyimli klinisyen tarafından Likert ölçeğiyle değerlendirdi.
Bulgular: Ortalama FKRE skorları ChatGPT-4, BingAI ve Gemini için sırasıyla 34,5, 33,8 ve 36,1 idi. FKGL puanları 11,2 ile 12,5 arasında değişmekteydi. Bu skorlar metinlerin okunmasının zor olduğunu ve üniversite düzeyinde okuma becerisi gerektirdiğini ortaya koydu. Kalite açısından BingAI (EQIP: 55,4±7,9) ve Gemini (54,4±8,8), ChatGPT-4'ten (48,6±6,3) anlamlı şekilde daha iyi performans gösterdi (p=0,005). Tüm modellerde doğruluk ve yeterlilik yüksek olup, ortalama puanlar 5 üzerinden 4,3'ün üzerindeydi.
Sonuç: Üç yapay zeka sohbet robotu da osteoporoz hakkında doğru ve yeterli yanıtlar üretse de içeriklerinin okunabilirliği hala istenilen seviyede değildir. BingAI ve Gemini, muhtemelen anlık veri kullandığından daha yüksek kaliteli bilgiler sunmaktadır. Sohbet robotlarının güncellemelerinde okunabilirliğin artırılması ve güncel veri erişiminin sağlanması, osteoporoz gibi anlaşılması önem arzeden konularda sağlık iletişimini güçlendirebilir.
Anahtar kelimeler: Osteoporoz, yapay zeka, sohbet robotu, ChatGPT, okunabilirlik, sağlık iletişimi

Corresponding Author/Sorumlu Yazar: Gamze Gül Güleç MD, Başkent University Adana Dr. Turgut Noyan Education and Research Center, Deparmet of Physical Medicine and Rehabilitation, Adana, Türkiye
E-mail: gamzegulgulec@gmail.com **ORCID ID:** orcid.org/0000-0003-2020-1507
Received/Geliş Tarihi: 12.04.2025 **Accepted/Kabul Tarihi:** 23.06.2025 **Epub:** 24.07.2025 **Publication Date/Yayınlanma Tarihi:** 05.12.2025

Cite this article as/Atıf: Güleç GG, Özpolat Bulut Ö, Bağcıer F. A new era in digital health: evaluation of artificial intelligence supported chatbots (ChatGPT-4, BingAI, and Gemini) responses on osteoporosis. Turk J Osteoporos. 2025;31(3):162-8

Introduction

Osteoporosis is a widespread metabolic bone disorder that has a profound impact on the health burden of the aging population and affects millions of individuals globally. This disease is characterized by a decrease in the mineral density of bone tissue, leading to brittle bones (1). This means that there is a risk of more serious fractures, particularly in the spine, hip and wrists (2). In the US, approximately 2 million osteoporotic fractures occur annually, impairing individuals' quality of life and imposing a high economic burden on the healthcare system (3). Osteoporosis is often under-recognized or diagnosed late (4). As a result, asignificant majority of patients remain undiagnosed and do not receive treatment until a fracture develops. Research has revealed that only 25% of patients with osteoporosis even know that they have it (1). This lack of awareness is especially evident in areas with limited healthcare access and among populations with limited health literacy (5).

Today, digital health solutions, particularly innovative technologies such as artificial intelligence (AI)-powered chatbots, have begun to play a crucial role in the management of chronic diseases, such as osteoporosis. There are different types of AI, such as machine learning and natural language processing, large language models (LLMs) (6). These technologies facilitate health management by offering services such as information dissemination, symptom tracking, and treatment recommendations. However, online health information often lacks adequate moderation, leading to significant variability in the quality and reliability of information (7).

In particular, LLMs have become a major focus of research and development because their ability to process and generate human-like text, given their training on large datasets, has generated significant interest. Among them are ChatGPT-4, BingAI, and Gemini, to name a few, each boasting particular characteristics and capabilities (8).

OpenAI's ChatGPT-4 is one of the Generative Pre-trained Transformer series and is recognized for its advanced natural language understanding and generation capabilities. To create a more well-behaved model, it was fine-tuned with both supervised learning and reinforcement learning, resulting in highly fluid and contextually appropriate answers on a wider array of subjects (9). For instance, BingAI, an LLM integrated with the Microsoft Bing Search Engine, supports a variant of the GPT model and provides a version optimized for real-time information retrieval along with research benefits, which improves the accuracy and relevance of the result segments. BingAI's integration with a search engine allows it to provide up-to-date information, making it a valuable tool for accessing current medical guidelines and studies (9). Gemini, developed by Google, is based on the Language Model for Dialogue Applications and is designed to create informative and conversational content, continuously updating its knowledge base with the latest web information to ensure that its responses are both current and contextually relevant (10).

The quality of health-related information on osteoporosis found through AI chatbots has been examined in the literature (11). Previous comparative studies of different chatbots have assessed the differences in readability and quality between responses to the same theme generated by different chatbots, but not for osteoporosis. The present study aimed to evaluate and compare the quality and readability of information provided by three different AI chatbots for the most common questions on osteoporosis.

Materials and Methods

The study was conducted on December 23, 2024, at the Clinic of Physical Therapy and Rehabilitation at Viranşehir State Hospital. This study did not involve any processes with live animals or human participants; therefore, institutional ethical approval was not required. To avoid any potential bias, all personal data from the browser were cleared before the searches. Additionally, all chat sessions were initiated in clean browsers with cleared cookies and no previous prompt history to eliminate prior interaction effects.

Three separate chatbots (ChatGPT-4, BingAI and Gemini) were posed 25 of the most frequently asked questions about osteoporosis on 23.12.2024. A prompt was submitted to Gemini to retrieve the 25 most frequently asked questions about osteoporosis based on Google Trends data. Gemini served solely as an interface to access publicly available search query data, without generating or altering the content. This approach was chosen to reflect real-world public interest in a neutral and reproducible manner and has been validated by similar studies in the literature (7,8,11). The prompt given to Gemini was, "Can you write the 25 most frequently asked questions about osteoporosis according to Google Trends?" Since LLMs can produce different answers to the same question, only the first answers were considered for each question. This is because the first answers tend to reflect what LLMs consider to be the most likely and correct responses. The word count was not limited, which allowed for extensive explanations. Each question was entered into the chatbots individually on a separate page.

The answers obtained from the chatbots about osteoporosis were obtained by a researcher. The responses were then evaluated by 2 different clinicians, each with at least 5 years of experience in the diagnosis and management of OP. If there were differences between the clinicians' evaluations, they were evaluated by a third independent clinician and a joint decision was made. If a response included a reference or DOI, it was manually verified via academic databases such as PubMed and CrossRef. Inter-rater agreement for the 5-point accuracy ratings was calculated using Cohen's kappa (κ=0.92), indicating excellent reliability. The clinicians were blind to which LLM the texts belonged to.

The accuracy and adequacy of the texts obtained from the LLMs were evaluated according to the Likert scale, based on previous studies (12). The accuracy of the texts was assessed according to a 5-point Likert scale (1: very poor accuracy or unacceptable inaccuracies with high risk of harm; 2: poor accuracy or potentially harmful errors; 3: negligible moderate inaccuracies; 4: good level of accuracy with minor inaccuracies; 5: very good level of accuracy, no risk of harm). The adequacy of the texts was assessed according to a 3-point Likert scale (1: incomplete presentation of important parts of information addressing some aspects of the problem; 2: adequate presentation of information addressing all aspects of the problem; 3: more information than expected addressing all aspects of the problem).

To evaluate the quality of the text generated by large language models, we used the ensuring quality information for patients (EQIP) tool. This assessment tool evaluates the content in thirty-two different ways, including whether the information is consistent and whether the writing is appropriate (13). The tool consists of 20 questions answered "yes," "somewhat," "no" or "not applicable." The scoring is done by multiplying the number of "yes" by 1 (so the more of these you have, the better), the number of "partially" by 0.5 and the number of "no" by 0. These are summed, with the total number of "does not apply" responses subtracted from 20 total items, then divided by the new total number of items. The final value is multiplied by 100, to obtain the EQIP score which is expressed as a percentage. EQIP are classified as follows: 76-100%: Well-written, great quality; 51-75%: Good quality, minor issues; 26-50%: Serious quality issues; 0-25%: Severe quality issues.

Flesch-Kincaid Reading Ease (FKRE) and Flesch-Kincaid Grade Level (FKGL) scores were used to evaluate the readability of the texts from the LLMs. The FKRE score, which ranges from 0 to 100, is a widely used readability score tool, and a higher score corresponds to improved readability. FKRE score =206.835-1.015 × (average sentence length) + 84.6 × (average word length) The FKGL score is a modified version of the FKRE score, which denotes the average US school grade level that is capable of understanding the text, with a lower score signifying an increase in readability (14).

Statistical Analysis

All statistical analyses were conducted with IBM SPSS version 22.0 software (IBM Corp., Armonk, NY, USA). Normality of the data distribution was assessed by Kurtosis-Skewness values and the Kolmogorov-Smirnov/Shapiro-Wilk test. Mean, standard deviation and median were calculated to describe the study variables. Group differences were evaluated by ANOVA or, as appropriate, the Kruskal-Wallis test. If significant differences were found, pairwise comparisons were performed using the t-test or Mann-Whitney U test. Statistical significance was defined as $p<0.05$.

Results

In this study, Gemini was used to retrieve a cumulative total of 25 frequently asked questions relating to osteoporosis based on Google Trends data. The questions cover various aspects of osteoporosis, including its meaning, its warning signs, prevention, risks, and treatment. The top five questions looked to learn more about the condition itself ("What is osteoporosis?"), determining its symptoms ("What are the symptoms of osteoporosis?"), and ways to prevent the disease ("What are the best ways to prevent osteoporosis?") (Table 1).

Geographic analysis showed higher search interest in osteoporosis in Puerto Rico, Ecuador, and Bolivia (Figure 1). The popularity of osteoporosis over time (the analysis was conducted on a time range starting from 2004) is presented according to the Google Trends analysis (Figure 2). Responses generated by all three AI chatbots—ChatGPT-4, BingAI and Gemini—were subjected to readability analysis using the FKRE and FKGL metrics. The readability scores for all models suggested that the outputted information was hard for the average population to read (all models $p>0.05$). Mean FKRE scores obtained with ChatGPT-4, BingAI, and Gemini were 34.5, 33.8, and 36.1, respectively, all classified as "difficult" (FKRE) scoring. In a similar manner, FKGL scores from 11.2 to 12.5 were noted, suggesting a need for a college-level education to comprehend the content (Table 2).

Responses were analyzed for quality using the EQIP tool. Chatbots produced significantly different quality scores ($p=0.005$). The mean EQIP score for ChatGPT-4 was a mere 48.6 points, significantly lower than that of either BingAI (55.4) or Gemini (54.4).

For both the completeness and correctness of responses, all three models performed similarly without any statistically significant differences among them ($p>0.05$). On a 5-point Likert scale of accuracy, the mean accuracy scores were 4.3, 4.4, and 4.3 for ChatGPT-4, BingAI, and Gemini, respectively, indicating that the model-generated responses were generally accurate but less so in some instances. All answers from the models had no major mistakes (1 or 2 points).

Discussion

In this study, we assessed the responses of three distinct Big Language Models: ChatGPT-4, BingAI, and Gemini. Responses were examined with regard to multiple dimensions, including readability, information quality, completeness and accuracy. Although no statistically significant differences were observed between the models for most parameters, statistically significant differences were observed between the models for the EQIP scores. These results show that although the general performance of the models was similar, the quality of information provided by ChatGPT was worse than that of the other models.

Table 1. Top 25 questions searched about osteoporosis across countries: 2004-2024 (based on Google Trends data)

Rank	Question	Category of the topic based on EQIP
1	What is osteoporosis?	Condition or illness
2	What are the symptoms of osteoporosis?	Condition or illness
3	What are the most effective methods to prevent osteoporosis?	Prevention or after care
4	Which age groups are at risk of osteoporosis?	Condition or illness
5	Which drugs are used in the treatment of osteoporosis?	Medication or product
6	What is the relationship between osteoporosis and nutrition?	Condition or illness
7	What is the role of physical activity in osteoporosis?	Miscellaneous
8	How is osteoporosis diagnosed?	Test, operation, investigation, or procedure
9	What are the genetic factors of osteoporosis?	Condition or illness
10	What factors increase the risk of osteoporosis in women?	Condition or illness
11	What are the best dietary recommendations for people with osteoporosis?	Miscellaneous
12	Which vitamins and minerals are effective in preventing osteoporosis?	Prevention or after care
13	What is the relationship between osteoporosis and menopause?	Condition or illness
14	What kind of exercise should be done to reduce the risk of osteoporosis?	Prevention or after care
15	What are the psychological effects of osteoporosis?	Miscellaneous
16	What is the relationship between osteoporosis and fractures?	Condition or illness
17	Are there natural methods to treat osteoporosis?	Prevention or after care
18	What are the most common misconceptions about osteoporosis?	Condition or illness
19	What are the latest technologies in the treatment of osteoporosis?	Medication or product
20	What is the impact of regular screenings on osteoporosis?	Miscellaneous
21	What are the differences between osteoporosis and other bone diseases?	Condition or illness
22	What role does physical therapy play in the treatment of osteoporosis?	Services
23	What tests should be done for osteoporosis?	Test, operation, investigation, or procedure
24	What should people with osteoporosis pay attention to in their daily life?	Prevention or after care
25	Common myths about osteoporosis	Miscellaneous

EQIP: Ensuring quality information for patients

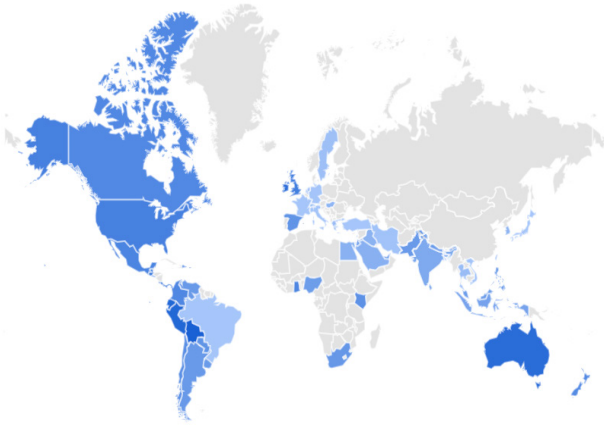

Figure 1. World map showing the relative search interest for the term "osteoporosis" by country based on Google Trends data. Darker shades of blue indicate higher levels of interest, whereas grey indicates regions with insufficient data

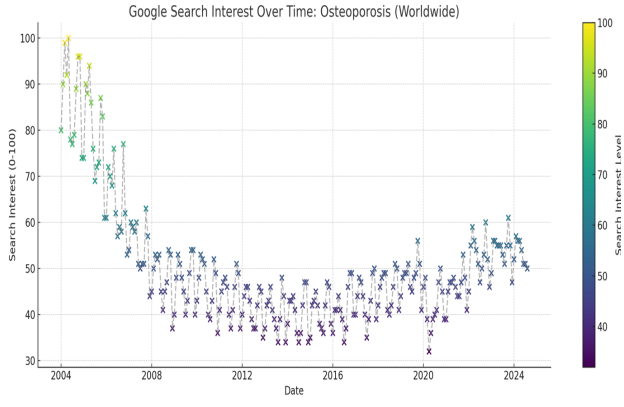

Figure 2. Timeline of global Google search interest for the term "osteoporosis" from January 2004 onwards. The color intensity of the data points corresponds to the level of search interest, with brighter colors indicating greater public attention. This trend reflects temporal changes in public awareness and potential influences such as awareness campaigns, research publications, or media coverage

Table 2. Comparison of large language models in terms of readability, quality, completeness, and accuracy

	ChatGPT	BingAI	Gemini	p
FKRE Mean ± SD Median (min-max)	 34.5±12.9 36.5 (16.0-61.2)	 33.8±14.3 34.8 (11.2-66.7)	 36.1±10.9 33.7 (23.0-67.3)	0.789
FKGL Mean ± SD Median (min-max)	 11.3±2.3 11.0 (7.1-15.6)	 12.5±2.7 12.1 (7.1-17.1)	 11.2±1.8 11.1 (8.1- 15.0)	0.127
EQIP Mean ± SD Medain (min-max)	 48.6±6.3*+ 50.0 (34.2-60.0)	 55.4±7.9 55.6 (40.7-69.4)	 54.4±8.8 55.5 (36.5-68.7)	**0.005**
Completeness Mean ± SD Median (min-max)	 2.1±0.2 2.0 (2.0-3.0)	 2.0±0.7 2.0 (1.0-3.0)	 2.0±0.2 2.0 (2.0-3.0)	0.553
Accuracy Mean ± SD Median (min-max)	 4.3±0.5 4.0 (3.0-5.0)	 4.4±0.5 4.0 (4.0-5.0)	 4.3±0.6 4.0 (3.0-5.0)	0.907

FKRE: Flesch-Kincaid Reading Ease, SD: Standard deviation, EQIP: Ensuring quality information for patients

There was no statistically significant difference between all three models in terms of FKRE and FKGL scores with all models producing content that required a higher education level to understand. These findings are consistent with previous studies reviewing AI-driven chatbots in health, for instance, those concerning retinopathy of prematurity or erectile dysfunction, which noted similar readability issues (9,15). Despite its long-established theory, the limited spread between the FKRE and FKGL scores invites novel questions about the generative mechanisms that govern these metrics across varying models. Another possible reason for the consistency in readability is the common use of large-scale datasets with a lot of technical and specialty medical information. These models are trained on large corpora of text, many of which are sourced from academic texts, clinical guidelines, and research papers, producing outputs that naturally mirror the complexity of their sources. This may lead to high and consistent FKGL scores, as rewriting complex medical acronyms in a simple manner while maintaining the same level of information is a complex task for LLMs.

Additionally, the small differences in readability scores could imply that current AI models prioritize writing accurate and complete information over writing accessible information. This is consistent with earlier research finding that LLMs can provide purportedly detailed and contextually accurate information without presenting it in an easily understandable manner for the general population. A study comparing the readability of online health information on stuttering has shown that even widely utilized resources frequently fall short of the recommended readability levels for medical literature (16). Therefore, it appears that readability, especially for medical information, remains a wider problem not limited to AI models.

Even if the FKRE and FKGL scores were similar, it might be worth speculating about the lack of a model that outperformed the others in terms of readability. Therefore, a possible explanation could be the training methodologies used in different LLMs. ChatGPT-4, BingAI, and Gemini are all from different companies, yet they may share similar pre-processing methods for medical terminology that normalize language complexity. Alternatively, all of the models may be limited by the trade-offs that come with readability versus accuracy—using basic language risks sacrificing the impact of the medical content, but the more detail you give, the more complicated it becomes.

Notably, despite no major differences in readability scores, the EQIP tool noted clear differences in the quality of information between the two models. ChatGPT-4 was greatly outperformed in quality by the production versions of BingAI and Gemini, both of which are powered by real-time information retrieval systems. This means that while all models might have difficulty with readability, having up-to-date information might result in more relevant and higher-quality content. In contrast, ChatGPT-4, which is more reliant on its pre-trained dataset with no real-time data, might be slower in providing the most up-to-date health information, which may negatively affect its EQIP score.

This finding contrasts with studies that reported more consistent EQIP scores across different AI models in other medical contexts, such as a study evaluating responses to spinal cord injury-related questions (17). One possible reason for this discrepancy is the nature of the conditions being discussed. Osteoporosis, as a chronic and evolving disease, requires up-to-date knowledge of recent clinical guidelines, medications, and prevention strategies. In conditions where real-time information plays a critical role, the advantages of models with real-time data retrieval, such as BingAI and Gemini, become more pronounced. This raises the possibility that the performance gap observed in this study could widen further in rapidly evolving fields of medicine, where new treatments and guidelines frequently emerge.

The other aspect of the question is how much their responses depend on the user interaction. BingAI and Gemini are likely

to be related to search engines and are trained on real-time data; therefore, there is a high chance that they were fine-tuned on real user queries and data. This real-time feedback loop may help improve response quality over time, unlike ChatGPT-4's static model, which would not see such incremental learning adaptively.

From a completeness and accuracy standpoint, there were no noteworthy differences between the models, which all performed well in answering the questions. This finding indicates that LLMs of all types and training data have comparable proficiency in processing all core medical concepts relevant to osteoporosis. Some responses contained small inaccuracies, highlighting the need for improvements, especially in more nuanced medical topics. The results are consistent with the existing literature, in that AI models have shown very high accuracy whenever presented with general medical knowledge but not with more specialized or context-driven information (18). Future advancements in LLMs may include improving both content readability and quality by implementing intelligent algorithms that adjust the complexity of the language to suit the individual reader's capacity to understand. Such technology could tailor the content in real time according to the user's previous encounters with content or health literacy. At the moment, human supervision is required for maintaining the readability and accuracy of AI-generated health information (19). Until now, AI and human experts have been working together in such a way that people can more easily get AI-driven health information in a hybrid style.

One limitation of this study is the relatively small number of questions used to evaluate the models. A larger and more diverse dataset could provide a more comprehensive comparison. Additionally, while readability and quality were measured, user satisfaction and engagement were not, which could be important metrics for evaluating the practical utility of these models.

Conclusion

This study underlines the strengths and weaknesses of ChatGPT-4, BingAI and Gemini for osteoporosis-related health information. While the output was similar in terms of readability levels across the models, BingAI and Gemini produced superior responses, in part due to their access to real-time data. As these AI tools become more prominent in health communication, future use should also focus on accessibility, the provision of real-time updates, and human oversight to mitigate their inaccurate use.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this manuscript, the authors used ChatGPT-4o to obtain background information and assist with language editing. All outputs generated by the AI tool were carefully reviewed, revised, and verified by the authors. The authors take full responsibility for the accuracy and integrity of this final manuscript. The use of AI was limited to information retrieval and language refinement and did not influence the study's data, analysis, or conclusions.

Ethics

Ethics Committee Approval: This study did not require an ethics committee certificate as it was not conducted on humans.
Informed Consent: Since this study was not conducted on human subjects, patient consent was not required.

Footnotes

Authorship Contributions

Concept: Ö.Ö.B., F.B., Design: F.B., Data Collection or Processing: Ö.Ö.B., Analysis or Interpretation: G.G.G., Ö.Ö.B., Literature Search: G.G.G., F.B., Writing: G.G.G., Ö.Ö.B., F.B.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Sözen T, Özışık L, Başaran NÇ. An overview and management of osteoporosis. Eur J Rheumatol. 2017;4:46-56.
2. Pouresmaeili F, Kamalidehghan B, Kamarehei M, Goh YM. A comprehensive overview on osteoporosis and its risk factors. Ther Clin Risk Manag. 2018;14:2029-49.
3. Rashki Kemmak A, Rezapour A, Jahangiri R, Nikjoo S, Farabi H, Soleimanpour S. Economic burden of osteoporosis in the world: a systematic review. Med J Islam Repub Iran. 2020;34:154.
4. LeBoff MS, Greenspan SL, Insogna KL, Lewiecki EM, Saag KG, Singer AJ, et al. The clinician's guide to prevention and treatment of osteoporosis. Osteoporos Int. 2022;33:2049-102. Erratum in: Osteoporos Int. 2022;33:2243.
5. Allen-Meares P, Lowry B, Estrella ML, Mansuri S. Health literacy barriers in the health care system: barriers and opportunities for the profession. Health Soc Work. 2020;45:62-4.
6. Singareddy S, Sn VP, Jaramillo AP, Yasir M, Iyer N, Hussein S, et al. Artificial intelligence and its role in the management of chronic medical conditions: a systematic review. Cureus. 2023;15:e46066.
7. Battineni G, Baldoni S, Chintalapudi N, Sagaro GG, Pallotta G, Nittari G, et al. Factors affecting the quality and reliability of online health information. Digit Health. 2020;6:2055207620948996.
8. De Angelis L, Baglivo F, Arzilli G, Privitera GP, Ferragina P, Tozzi AE, et al. ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health. Front Public Health. 2023;11:1166120.
9. Durmaz Engin C, Karatas E, Ozturk T. Exploring the role of ChatGPT-4, BingAI, and gemini as virtual consultants to educate families about retinopathy of prematurity. Children (Basel). 2024;11:750.
10. Yu P, Xu H, Hu X, Deng C. Leveraging generative AI and large language models: a comprehensive roadmap for healthcare integration. Healthcare (Basel). 2023;11:2776.
11. Erden Y, Temel MH, Bağcıer F. Artificial intelligence insights into osteoporosis: assessing ChatGPT's information quality and readability. Arch Osteoporos. 2024;19:17.
12. Huang AS, Hirabayashi K, Barna L, Parikh D, Pasquale LR. Assessment of a large language model's responses to questions and cases about glaucoma and retina management. JAMA

Ophthalmol. 2024;142:371-5. Erratum in: JAMA Ophthalmol. 2024;142:393.

13. Moult B, Franck LS, Brady H. Ensuring quality information for patients: development and preliminary validation of a new instrument to improve the quality of written health care information. Health Expect. 2004;7:165-75.
14. Alas AN, Bergman J, Dunivan GC, Rashid R, Morrisroe SN, Rogers RG, et al. Readability of common health-related quality-of-life instruments in female pelvic medicine. Female Pelvic Med Reconstr Surg. 2013;19:293-7.
15. Şahin MF, Ateş H, Keleş A, Özcan R, Doğan Ç, Akgül M, et al. Responses of five different artificial intelligence chatbots to the top searched queries about erectile dysfunction: a comparative analysis. J Med Syst. 2024;48:38.
16. Tsiamtsiouris J, Kollia B, Kamowski-Shakibai MT, Garcia P, Basch, CH. Applying tests of readability to online stuttering information. Adv Neurodev Disord. 2020;4:279-83.
17. Temel MH, Erden Y, Bağcıer F. Information quality and readability: ChatGPT's responses to the most common questions about spinal cord injury. World Neurosurg. 2024;181:e1138-44.
18. Samaan JS, Yeo YH, Rajeev N, Hawley L, Abel S, Ng WH, et al. Assessing the accuracy of responses by the language model ChatGPT to questions regarding bariatric surgery. Obes Surg. 2023;33:1790-6.
19. Akkara JD, Kuriakose A. Commentary: is human supervision needed for artificial intelligence? Indian J Ophthalmol. 2022;70:1138-9.

DOI: 10.4274/tod.galenos.2025.86094
Turk J Osteoporos 2025;31(3):169-77

Acute Impact of Spinal Manipulation on Pain and Muscle Mechanical Properties in Chronic Low Back Pain: An Assessor Blinded, Randomized Controlled Trial

Kronik Bel Ağrısında Spinal Manipülasyonun Ağrı ve Kasın Mekanik Özellikleri Üzerine Akut Etkisi: Değerlendiriciye Kör, Randomize Kontrollü Bir Çalışma

Ayça Aracı[1], Emine Eda Kurt[2], Serkan Taş[3]

[1]Alanya Alaaddin Keykubat University Faculty of Health Sciences, Department of Physiotherapy and Rehabilitation, Antalya, Türkiye
[2]Alanya University Faculty of Health Sciences, Department of Physiotherapy and Rehabilitation, Antalya, Türkiye
[3]Toros University Faculty of Health Sciences, Department of Physiotherapy and Rehabilitation, Mersin, Türkiye

Abstract

Objective: Chronic low back pain (CLBP) represents one of the most challenging and costly musculoskeletal conditions to manage. A variety of therapeutic approaches, including exercise training, pain management strategies, and spinal manipulation and mobilization, are employed in its treatment. Among these, clinical spinal manipulation and mobilization techniques are widely regarded as one of the most effective interventions for reducing pain and disability, offering both short- and long-term benefits. High-velocity spinal manipulation is commonly adopted for treating CLBP and has been associated with changes in muscle activity, but the evidence is controversial. The aim of this study was to analyze the immediate effects of two manual spinal techniques (MST) on pain, flexibility, and muscle mechanical properties in CLBP.

Materials and Methods: This single-blinded, randomized comparative trial used a pre- and post-test design. Participants were randomly assigned to two groups: Group 1 received Maitland's posterior-anterior central vertebral pressure mobilization technique, and group 2 underwent the lumbar roll technique. Assessments were conducted at baseline and immediately after the interventions. Muscle mechanical properties were measured using MyotonPro, pain intensity was evaluated using a visual analog scale (VAS), and flexibility was assessed through the sit and reach test and the modified Schober test.

Results: Following the interventions, significant improvements were observed in VAS scores and sit-and-reach test results in both groups. Analysis of Schober test data revealed a significant improvement in group 2 ($p<0.001$). Mechanical properties of the paravertebral muscles at the L3-L4 vertebral level were assessed using MyotonPro, showing statistically significant enhancements in elasticity (Hz) and dynamic stiffness (N/m) in both groups post-intervention. However, no statistically significant differences were identified between the groups.

Conclusion: Both MSTs demonstrated efficacy in alleviating pain, reducing muscle stiffness, and enhancing flexibility. In this study, manipulations were applied to the symptomatic side, which yielded positive outcomes in pain reduction and muscle properties. However, further research is needed to determine whether the symptomatic side is superior in terms of therapeutic efficacy.

Keywords: Chronic low back pain, spinal manipulation, mobilization, muscle mechanical properties, MyotonPro

Öz

Amaç: Kronik bel ağrısı (KBA), yönetimi en zorlayıcı ve en maliyetli kas-iskelet sistemi rahatsızlıklarından biri olarak kabul edilmektedir. Egzersiz eğitimi, ağrı yönetim stratejileri ve spinal manipülasyon gibi çeşitli terapötik yaklaşımlar bu durumun tedavisinde kullanılmaktadır. Bu yöntemler arasında klinik spinal manipülasyon, hem kısa hem de uzun vadeli faydalar sağlayarak ağrı ve fiziksel kısıtlılığı azaltmada en etkili müdahalelerden biri olarak yaygın şekilde kabul görmektedir. Yüksek hızda uygulanan spinal manipülasyon teknikleri, KBA tedavisinde yaygın olarak tercih edilmekte ve kas aktivitesinde değişikliklerle ilişkilendirilmektedir; ancak bu konuda mevcut kanıtlar çelişkilidir. Bu çalışmanın amacı, KBA'lı bireylerde iki farklı spinal manipülasyon tekniğinin klinik sonuçlar ve kas mekanik özellikleri üzerindeki anlık etkilerini analiz etmektir.

Gereç ve Yöntem: Bu tek-körlemeli, randomize karşılaştırmalı çalışma, ön test–son test desenine sahiptir. Katılımcılar rastgele iki gruba ayrılmıştır: Grup 1, Maitland'ın postero-anterior merkezi vertebral bası tekniğini alırken, Grup 2 ise lumbal roll tekniğine tabi tutulmuştur.

Corresponding Author/Sorumlu Yazar: Lec, Ayça Aracı, MD, Alanya Alaaddin Keykubat University Faculty of Health Sciences, Department of Physiotherapy and Rehabilitation, Antalya, Türkiye
E-mail: uyanayca@gmail.com **ORCID ID:** orcid.org/0000-0002-1089-3370
Received/Geliş Tarihi: 11.04.2025 **Accepted/Kabul Tarihi:** 01.07.2025 **Epub:** 21.08.2025 **Publication Date/Yayınlanma Tarihi:** 05.12.2025

Cite this article as/Atıf: Aracı A, Kurt EE, Taş S. Acute impact of spinal manipulation on pain and mechanical properties in chronic low back pain: an assessor blinded, randomized controlled trial. Turk J Osteoporos. 2025;31(3):169-77

Öz

Değerlendirmeler, müdahale öncesi ve hemen sonrasında gerçekleştirilmiştir. Kas mekanik özellikleri MyotonPro cihazı ile ölçülmüş; ağrı düzeyi görsel analog skala (VAS) ile değerlendirilmiş; esneklik ise otur-uzan testi ve modifiye Schober testi kullanılarak belirlenmiştir.

Bulgular: Uygulamalar sonrasında her iki grupta da VAS skorları ile otur-uzan testi sonuçlarında anlamlı iyileşmeler gözlenmiştir. Schober testi analizinde, yalnızca Grup 2'de istatistiksel olarak anlamlı bir gelişme saptanmıştır (p<0,001). L3-L4 vertebral düzeyindeki paravertebral kasların mekanik özellikleri, MyotonPro cihazı ile değerlendirilmiş olup, her iki grupta da elastisite (Hz) ve dinamik sertlik (N/m) parametrelerinde anlamlı artışlar görülmüştür. Bununla birlikte, gruplar arasında istatistiksel olarak anlamlı bir fark bulunmamıştır.

Sonuç: Her iki spinal manipülasyon tekniği de, ağrının hafifletilmesi, kas sertliğinin azaltılması ve esnekliğin artırılması açısından etkili bulunmuştur. Elde edilen bulgular doğrultusunda, terapötik etkinliğin en üst düzeye çıkarılması için manipülasyonların semptomatik tarafa uygulanması önerilmektedir.

Anahtar kelimeler: Kronik bel ağrısı, spinal manipülasyon, mobilizasyon, kas mekanik özellikleri, MyotonPro

Introduction

Worldwide, low back pain (LBP) has become the leading cause of disability (1). In most cases, a specific patho-anatomic cause of back pain cannot be identified (2). A range of treatment approaches, including exercise therapy, pain management strategies, and spinal manipulation, are commonly employed in the management of chronic LBP (3,4).

Stochkendahl et al. (1) proposed an evidence-based, stepped-care clinical guideline for the diagnosis and management of LBP, emphasizing non-pharmacologic and patient-centered strategies. Their approach highlights the importance of patient education, encouragement of physical activity, and, where appropriate, the use of manual therapies such as spinal manipulation. Advanced imaging and invasive interventions are reserved for cases presenting with "red flag" symptoms. Within the scope of such guidelines, spinal manipulative therapy (SMT) is widely recognized as a valuable tool for reducing pain and disability (5). Nevertheless, uncertainty remains regarding which SMT techniques are most effective, the ideal treatment frequency, and which patient profiles are likely to benefit. A central debate persists on whether SMT should be applied to regions with the most pronounced biomechanical dysfunction or to areas exhibiting the highest pain sensitivity (6).

Reduced spinal mobility is one of the causes of back pain. Specifically, the decrease in the isometric resistance and strength of the spinal extensor muscles leads to greater loading on passive structures (6). This can result in plastic deformation, possible strain, and consequently, a loss of flexibility in these structures. When flexibility is reduced, the mechanical efficiency of the joint is limited, and energy consumption increases (6). When spinal mobility is reduced, the likelihood of LBP increases. In 90% of patients with LBP, movement restrictions are observed in at least one of the following tests: fingertip-to-floor distance, trunk flexion/extension, lateral flexion, modified Schober test, and knee extension (6). Therefore, the use of SMT is necessary to increase the mobility of the lumbar region.

SMT, which encompasses both mobilization and manipulation techniques, exerts its therapeutic effects on pain and movement restriction through distinct yet complementary mechanisms (7-9). Mobilization involves the application of low-velocity, rhythmic passive forces targeting joint capsules and surrounding soft tissues to improve tissue flexibility and enhance range of motion. This technique also contributes to pain modulation by reducing muscle spindle hyperexcitability and activating endogenous inhibitory pathways within the central nervous system (10). In contrast, manipulation involves a high-velocity, low-amplitude (HVLA) thrust that often produces joint cavitation, leading to changes in intradiscal pressure and mechanical mobility, while simultaneously triggering a cascade of neurophysiologic responses that inhibit nociceptive transmission (8). Recent evidence indicates that SMT is particularly effective in managing acute and subacute musculoskeletal disorders, primarily due to its ability to rapidly induce peripheral and central neurophysiological responses (11-13). Activation of type I and II mechanoreceptors during SMT suppresses nociceptive input at the level of the spinal dorsal horn, resulting in localized hypoalgesia (11). Concurrently, descending pain modulatory systems such as the periaqueductal gray matter in the brainstem are engaged, producing widespread analgesic effects beyond the site of application. This bidirectional neurophysiologic mechanism also induces transient changes in corticospinal excitability and sensorimotor integration, potentially affecting not only pain perception but also motor outputs such as postural control and muscle tone (11-13). These effects typically emerge within 5 to 30 minutes post-intervention and diminish within a few hours, positioning SMT as a clinically relevant and effective option for short-term symptom relief, particularly in acute care settings (12).

Techniques such as the Maitland spinal mobilization used in this study involve central posteroanterior (PA) pressure applications to reduce muscle spasm and alleviate LBP. It has been reported to be particularly effective when pain is of equal intensity on both sides (14). On the other hand, the literature frequently mentions the use of side-lying manipulation technique in cases of chronic low back pain (CLBP) (15). However, studies evaluating the acute effects of mobilization and manipulation techniques on muscle mechanical properties are rare in the literature (6,16,17). The primary objective of this study was to investigate the immediate effects of Maitland's PA central vertebral pressure mobilization technique and the side-lying lumbar manipulation

technique on muscle mechanical properties, pressure pain threshold, and joint range of motion in patients with CLBP. A secondary objective was to compare the relative effectiveness of these two manual therapy techniques in improving the measured outcomes.

Materials and Methods

Study design and setting: This randomized, assessor-blinded comparative study aimed to evaluate the effects of the lateral lumbar spinal manipulation technique and Maitland's PA central vertebral mobilization technique (18) on the mechanical properties of muscles and their impact on the pressure-pain threshold.

Sample size: A power analysis was conducted to determine the required sample size for this study. To achieve 80% statistical power at a 5% level of significance, it was estimated that detecting a clinically meaningful difference of approximately 64 N/m (equivalent to 15%) in lumbar extensor muscle stiffness would require at least 10 participants per group. This calculation was based on an assumed baseline stiffness of 320 N/m and a standard deviation (SD) of 74 N/m. (19).

Ethics approval and consent to participate: This study was approved by the Non-Interventional Ethics Committee of Alanya Alaaddin Keykubat University (approval no: 10354421, date: 14/03/2021) and was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki (1964) and its subsequent revisions. The trial was registered on ClinicalTrials.gov (ID: NCT05057091; last verified: 09/2024). All procedures were performed in compliance with relevant institutional guidelines and regulatory standards. Written informed consent was obtained from all participants prior to their inclusion in the study.

Eligibility criteria and search strategy: The data were collected in the Department of Physical Medicine and Rehabilitation at Alanya Alaaddin Keykubat University Faculty of Medicine in between June 2021 and September 2021. The diagnosis of CLBP was made by a physiatrist. Volunteers aged between 18 and 65 years who were diagnosed as having CLBP by a physician were included in the study. Participants were excluded (1) if they had a risk of pregnancy, (2) were using pacemakers or had metal implants in their body, (3) had vertebral fractures, (4) cancer, (5) osteoporosis, (6) body mass index (BMI) above 30 kg/m^2, (7) neurologic diseases that might lead to muscle weakness and spasticity, and (8) primary or/and secondary degenerative vertebral diseases (18)

All interventions were administered by a physiotherapist (A.A) who had 15 years of clinical experience, specialized training in spinal manipulation therapy, and was a certifed osteopathic practioner. This practitioner was responsible for obtaining medical histories and conducting clinical assessments for all participants. Pre- and post-intervention measurements were performed by a second physiotherapist (S.T) who was blinded to group allocation to minimize assessment bias.

Randomization: Covariate adaptive randomization was performed using a computer program for the optimal allocation of patients to two groups (covariates: age, sex, and BMI).

Intervention

Group 1 – Maitland's posterior-anterior central vertebral pressure mobilization technique: Participants in this group were positioned in the prone position. The therapist identified the vertebral segments associated with the highest pain sensitivity through palpation. Using the pisiform bones of the hands, the therapist applied PA central vertebral mobilization directly to the spinous processes of the identified vertebrae (20). For the unilateral PA mobilization, both thumbs were positioned one over the other on the spinous process of the targeted vertebra, and an average discrete pressure of approximately 4 kg was applied (20).

Illustration of the PA central vertebral mobilization technique, where the therapist applies pressure on the spinous process of the identified painful vertebrae (Figure 1).

Group 2 – Side-lying lumbar spinal manipulation: Participants in this group were positioned in the side-lying posture with the painful side facing upward, while the therapist stood in front of them. Initially, the upper leg was flexed until movement was detected in the vertebral segment corresponding to the pain site. The interspinous space was then palpated, and the foot of the flexed leg was placed into the popliteal fossa of the lower leg for stabilization. In the second phase of positioning, the therapist grasped the lower shoulder to induce contralateral trunk rotation and hyperextension. This was done by pulling the shoulder until movement was again perceived at the interspinous level, thereby isolating the symptomatic vertebra between adjacent segments. With this alignment maintained, the patient was gently rolled toward the therapist. Finally, a HVLA thrust was delivered using the therapist's arm and body. The thrust was applied simultaneously in two opposing directions: anteriorly to the pelvis and posteriorly to the shoulder (21).

Demonstration of the side-lying lumbar spinal manipulation

Figure 1. Maitland's PA central vertebral mobilization technique
PA: Posteroanterior

technique, highlighting the steps of positioning, alignment, and application of high-speed, low-amplitude thrust for targeted spinal mobilization (Figure 2).

Outcomes measures: This was a randomized single-blinded comparative study, involving pre- and post-measurement tests. Assessments were performed at baseline and immediately after SMT interventions. The dependent variables were the mechanical properties of muscles, LBP level, and lumbar region flexibility.

Lumbar Region Flexibility Assessment

Modify Schober test: Lumbar region flexion flexibility was measured using a modified Schober test. For measurement, the examiner marked both posterior superior iliac spine and then drew a horizontal line at the center of both marks.

Participants were instructed to perform maximum trunk flexion by bending forward. The distance between the two vertical marks was re-measured during full flexion. The difference between the baseline and flexed position was recorded in centimeters. A value between 0-5 cm was interpreted as reduced lumbar flexibility, >10 cm indicated increased flexibility, and values between 5-10 cm were considered within the normal range (22).

Sit and reach test: The patients sat in a long sitting position with their knees straight and the soles of their feet flat against the bottom of the test board. The feet were positioned to be approximately shoulder width apart, and the patient extended forward from the waist and hip so that the elbows, wrists, and fingers were stretched. During the test, attention was paid to keeping the knees straight. The tested patient pushed the measuring board forward on the test stand with their fingers and waited 1-2 seconds at the last point. The place where the feet made contact with the test stand was taken as the starting point, point 0. The distance between the fingertip and the starting point was measured and recorded in cm as "-" if it was ahead of the 0 point, and "+" if it was behind (23).

Mechanical properties measurements: MyotonPro was described as valid (24) and dependable (25-27) for assessing muscle mechanical parameters. The information below is provided by MyotonPro: (3) oscillation frequency (an indicator of tone), (4) dynamic stiffness (an indicator of stiffness), and (5) logarithmic attenuation (related to elasticity). A physiotherapist with 3 years of myotonometric measurement experience who was blinded to the groups conducted the measurements.

Measurements for the lumbar extensor muscle were taken 3 cm from the midpoint of the L3/L4 intervertebral gap. The mean of three consecutive measurements was recorded (25,26).

Figure 2. Side-lying lumbar spinal manipulation

Visual analog scale (VAS): The severity of the patient's pain was assessed using the VAS. This scale typically consists of a 10 cm horizontal or vertical line, with "no pain" at one end and "unbearable pain" at the other. The patient marks a point on the line to indicate their level of pain. The distance between the starting point and the marked point is measured in centimeters (cm) and recorded. On this scale, "0" represents no pain, "5" indicates moderate pain, and "10" pain ever experienced (28).

Statistical Analysis

For statistical analysis, the IBM SPSS version 20.0 software (IBM Corporation, Armonk, NY) was used. For normal distribution, data are reported as mean ± SD. Data that are not normally distributed are shown as the median interquartile range. Categorical data are represented as a percentage (%). The Shapiro-Wilk test was used to determine the normality of the data distribution. When the data distribution was normal, Taleb's test was used to evaluate statistically significant differences. The Mann-Whitney U test was employed if the data were not regularly distributed. The chi-square test was used to perform qualitative comparisons of the groups. The paired t-test was used to compare repeated measurements in each group if the data were regularly distributed. The Wilcoxon test was used if the data distribution was not normal. The statistical significance threshold was set as $p<0.05$.

Results

A total of 29 (23 females, 6 males) diagnosed with CLBP were enrolled in the study, with a mean age of 38.48±11.35 years. Of these, 14 patients were allocated to group 1, which received Maitland PA central vertebral pressure and PA unilateral vertebral pressure mobilization, and 15 patients were assigned to group 2, which underwent side-lying lumbar spine manipulation (Figure 3). This figure illustrates the process of patient enrollment, group allocation, interventions, and analysis throughout the study.

The demographics of the patients are illustrated in Table 1. There were no statistically significant variations in sex, age, BMI, or pain duration between the groups (all $p>0.005$).

The study addressed the differences between the before and after values of both groups. Change, as well as intra-group and inter-group differences, were assessed (Table 2).

Pain Intensity (VAS)

Following the interventions, both groups' pain levels decreased significantly. However, no significant difference in VAS change

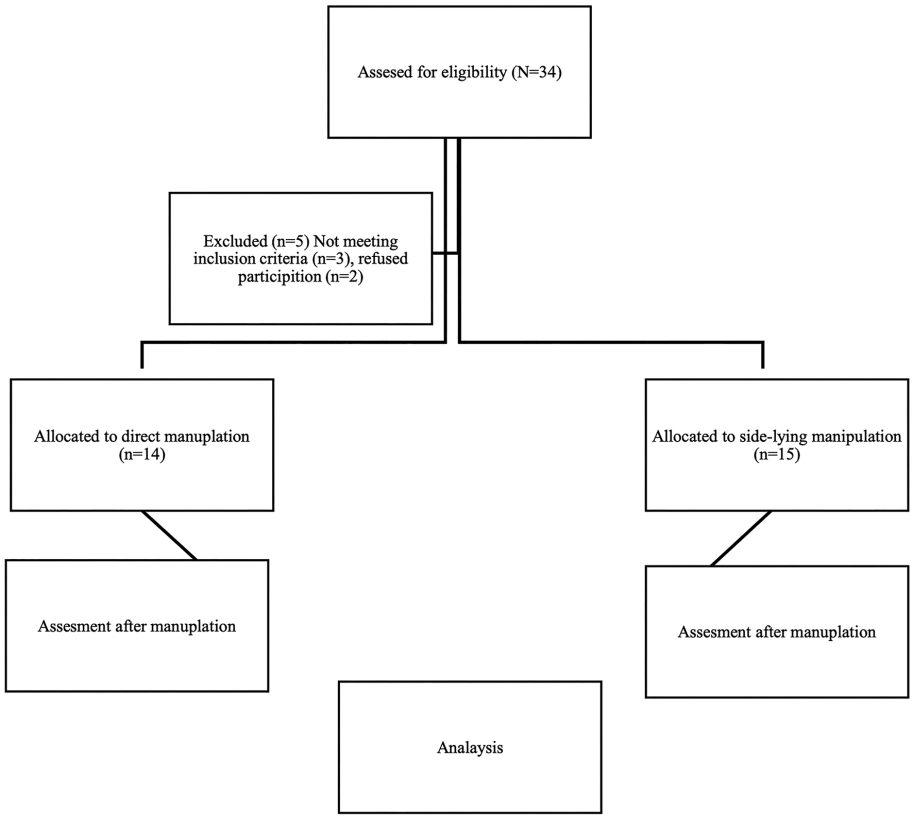

Figure 3. Flowchart of the study design

Table 1. Demographic data of the patients

	Maitland's posterior-anterior central vertebral pressure (group 1)	Lumbar roll techinique (group 2)	p
Age*	37.64±13.36	35.40±9.44	0.604
Sex (male/female)	3 (21.4%)/11 (78.6%)	3 (20%)/14 (80%)	0.924
BMI*	25.16±2.87	22.43±6.80	0.170
Pain duration* (month)	6.15±7.79	6.26±9.32	0.965

*: Mean ± standard deviation (normally distributed data), BMI: Body mass index

from baseline was seen between the groups. Group 1 reported a decrease in VAS from 5.71±1.77 to 3.64±2.27 (p=0.006), while Group 2 improved from 4.46±1.16 to 2.06±1.18 (p=0.001). However, there was no statistically significant difference between the groups in VAS change scores (p=0.533).

Lumbar Flexibility (Schober Test)

The Schober test revealed a significant increase only in Group 2 (from 4.57±3.49 to 5.33±3.37, p<0.001*), whereas no significant difference was detected in Group 1 (p=0.980). The intergroup comparison of change scores was not statistically significant (p=0.780).

Hamstring Flexibility (Sit and Reach Test)

When the sit and reach test data were compared before and after the intervention, both group 1 (p=0.004*b) and group 2 (p=0.003*b) showed a significant difference. However, the change scores between the groups were not significantly different (p=0.561).

Paravertebral Muscle Mechanical Properties at L3-L4 Level (MyotonPro)

When the mechanical changes in the paravertebral muscles in the L3-L4 vertebral space were assessed using a MyotonPro before and after the intervention, significant changes in (1) logarithmic decrement [related to elasticity (Hz)] (p=0.028*a) were found. In repeated tests in both groups, no difference in dynamic stiffness (2) [indicator of stiffness (N/m) (p=0.020*a)] or oscillation frequency [indicator of tone (dec)] was identified. In both groups, dynamic stiffness values [indicator of stiffness (N/m)] (p=0.020*a) and (3) oscillation frequency revealed no significant change.

Following the intervention, both groups' VAS scores and sit-reach tests improved significantly. The Schober test data were examined, and it was discovered that there was only a significant difference in group 2 (p=0.001*).

Discussion

The present study aimed to compare the acute effects of two spinal manipulation techniques Maitland's PA central vertebral pressure mobilization technique and side-lying lumbar spinal manipulation in individuals diagnosed with CLBP. The results demonstrated that both intervention methods were effective in reducing pain intensity, decreasing muscle stiffness, and enhancing flexibility.

Despite the positive outcomes observed in both groups, no statistically significant difference was identified between the techniques. This finding suggests that the two spinal manipulation approaches may produce comparable short-term clinical benefits in the management of CLBP. These results support the flexibility of clinical decision-making in choosing either technique based on practitioner expertise or patient preference, without compromising therapeutic efficacy.

The significant reduction in pain intensity observed in both groups aligns with previous literature (20,29), and is likely mediated through spinal and supraspinal pain modulation mechanisms. Although the systematic review by Coulter et al. (17) suggested spinal manipulation may be more effective than mobilization for pain reduction, it also noted limited effects on function. Similarly, Bussières et al. (30) emphasized that spinal manipulation should be integrated with education, exercise, and behavioral approaches to maximize its impact on range of motion and clinical outcomes.

Our results are consistent with previous research indicating that both mobilization and manipulation alleviate pain by modulating muscle spindle sensitivity and engaging endogenous analgesic pathways (29,31). In line with Cardinale et al. (32), even a single session of manipulation significantly improved flexibility outcomes such as the sit and reach test.

Table 2. Functional measurements at baseline and after manipulation; differences between two measurements

	Maitland's posterior-anterior central vertebral pressure (group 1)		**Lumbar roll techinique (group 2)**		**Group 1 vs. Group 2 intragroup**
	Mean ± SD	**p**	**Mean ± SD**	**p**	
VAS					
Baseline	5.71±1.77[a]	0.006*	4.46±1.16[a]	0.001*	
End of manuplation	3.642.27[a]		2.06±1.18[a]		
Change	1.50±2.07[b]		2.40±1.00[b]		0.533[b]
Schober test					
Baseline	6.00±4.00[b]	0.98	4.57±3.49[a]	**<0.001***	
End of manipulation	6.00±3.25[b]		5.33±3.37 [a]		
Change	1.00±1.00[b]		1.00±1.00[b]		0.780[b]
Sit and reach test					
Baseline	-4.00±12.25[b]	**0.004***[b]	-3.50±16.00[b]	**0.003***[b]	
End of manipulation	-1.00±13.75[b]		-1.00±13.00[b]		
Change	4.00±5.00[b]		2.00±4.00[b]		0.561[b]
L3-L4 frequency (Hz)					
Baseline	14.45±1.24[a]	0.546[a]	15.02±2.16[a]	**0.028***[a]	
End of manipulation	14.58±1.50[a]		14.58±1.83[a]		
Change	4.00±5.00[b]		2.00±4.00[b]		0.270[b]
L3-L4 stiffness (N/m)					
Baseline	268.50±54.04[a]	0.750[a]	292.33±79.72[a]	**0.020***[a]	0
End of manipulation	264.57±48.31[a]		268.20±66.40[a]		
Change	6.00±48.75[b]		14.00±34.00[b]		0.186[b]
L3-L4 dec					
Baseline	1.21±0.19[a]	0.840[a]	1.14±0.25[a]	0.412[a]	
End of manipulation	1.20±0.19[a]		1.16±0.26[a]		
Change	0.05±0.14[b]		- 0.02±0.07[b]		0.62[b]

[a]: Mean ± standard deviation (SD) (normally distributed data), [b]: Median ± interquartile range (non-normally distributed data), VAS: Visual analog scale, *: p<0.05

In the current study, both manipulation techniques led to significant improvements in sit and reach scores, indicating a favorable impact on hamstring flexibility and posterior chain mobility.

This finding highlights the potential of spinal interventions to modulate not only local spinal biomechanics but also distal muscle extensibility, possibly via reflex inhibition or improved neuromuscular coordination.

However, no significant improvement in the Schober test was observed in the Maitland group, mirroring the findings of Abe et al. (6), who reported limited acute effects of this technique on spinal flexion in young adults with LBP.

Notably, side-lying lumbar manipulation was more effective in improving lumbar flexion as assessed by the modified Schober test, suggesting this technique may offer advantages in enhancing segmental mobility. Improvement in fingertip-to-floor distance supports the role of manipulation in resolving mechanical restrictions and improving joint kinematics (33,34).

In terms of mechanical muscle properties, both groups showed improvements in elasticity and a reduction in stiffness, in contrast to Wu et al. (35), who reported increased stiffness post-manipulation. These discrepancies may be due to differences in population characteristics or measurement methods, and support the hypothesis that both biomechanical and neurophysiologic mechanisms including proprioceptive input and central pain modulation may be involved (36).

A recent meta-analysis by de Zoete et al. (37) confirmed the short-term efficacy of SMT for pain and function in patients with CLBP. However, although SMT showed superiority over sham interventions, it was only marginally more effective than other active interventions, highlighting the importance of incorporating SMT within a broader multimodal pain management strategy.

Study Limitations

There are some limitations to this study that should be noted. One of the main limitations is the relatively small sample size, which limits the generalizability of the findings and may reduce statistical power, especially for secondary outcomes. However, because this study was primarily designed to evaluate the effects SMT techniques on muscle architecture, power analysis was calculated based on muscle architecture measurements. A control or placebo group was not included in the study plan because it would be ethically inappropriate not to intervene in patients presenting with pain. The observed improvements may also be influenced by non-specific factors such as patient expectations, natural variability of symptoms, or repeated measurements, limiting the ability to isolate treatment-specific effects.

Despite these limitations, the study possesses several notable strengths. It addresses a significant gap in the literature by directly comparing the acute effects of two widely used SMT techniques—Maitland mobilization and side-lying lumbar manipulation—within a randomized, assessor-blinded design. Moreover, the use of clinically relevant outcome measures, including pain intensity (VAS), spinal flexibility (modified Schober and fingertip-to-floor tests), and muscle stiffness, enhances the applicability of the findings to clinical practice. Nevertheless, future studies incorporating control groups, long-term follow-up assessments, and stratified analyses are essential to deepen our understanding of the therapeutic mechanisms and optimize treatment strategies

Conclusion

In conclusion, both the Maitland mobilization technique and side-lying lumbar spinal manipulation appear to be effective interventions for the management of CLBP. Although no significant differences were found between the techniques in this study, their clinical utility may be optimized when selected based on individual patient profiles and incorporated into personalized treatment protocols.

Given the multifactorial nature of CLBP, the integration of these manual therapy techniques within multimodal rehabilitation approaches including education, exercise, and behavioral strategies may further enhance therapeutic outcomes. Future research should continue to explore how these methods can be effectively combined to maximize both short- and long-term benefits, guided by the growing body of evidence in the literature.

Ethics

Ethics Committee Approval: This study was approved by the Non-Interventional Ethics Committee of Alanya Alaaddin Keykubat University (approval no: 10354421, date: 14/03/2021) and was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki (1964) and its subsequent revisions.

Informed Consent: Informed consent for publication was obtained from all patients.

Acknowledgments

We express our deepest gratitude to all the participants who volunteered for our study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: A.A., S.T., Concept: A.A., E.E.K., S.T., Design: S.T., Data Collection or Processing: A.A., E.E.K., S.T., Analysis or Interpretation: A.A., E.E.K., S.T., Literature Search: A.A., E.E.K., S.T., Writing: A.A., E.E.K.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Stochkendahl MJ, Kjaer P, Hartvigsen J, Kongsted A, Aaboe J, Andersen M, et al. National Clinical Guidelines for non-surgical treatment of patients with recent onset low back pain or lumbar radiculopathy. Eur Spine J. 2018;27:60-75.

2. Hush J, Elkins M. Virtual special issue on low back pain. J Physiother. 2018;64:208-9.
3. Kreiner DS, Matz P, Bono CM, Cho CH, Easa JE, Ghiselli G, et al. Guideline summary review: an evidence-based clinical guideline for the diagnosis and treatment of low back pain. Spine J. 2020;20:998-1024. Erratum in: Spine J. 2021;21:726-7.
4. Wilhelm M, Cleland J, Carroll A, Marinch M, Imhoff M, Severini N, et al. The combined effects of manual therapy and exercise on pain and related disability for individuals with nonspecific neck pain: a systematic review with meta-analysis. J Man Manip Ther. 2023;31:393-407.
5. Rubinstein SM, de Zoete A, van Middelkoop M, Assendelft WJJ, de Boer MR, van Tulder MW. Benefits and harms of spinal manipulative therapy for the treatment of chronic low back pain: systematic review and meta-analysis of randomised controlled trials. BMJ. 2019;364:l689.
6. Abe KY, Tozim BM, Navega MT. Acute effects of Maitland's central posteroanterior mobilization on youth with low back pain. MTP&RehabJournal. 2015;13:1-5.
7. Chambers H. Physiotherapy and lumbar facet joint injections as a combination treatment for chronic low back pain. A narrative review of lumbar facet joint injections, lumbar spinal mobilizations, soft tissue massage and lower back mobility exercises. Musculoskeletal Care. 2013;11:106-20.
8. Maigne JY, Vautravers P. Mechanism of action of spinal manipulative therapy. Joint Bone Spine. 2003;70:336-41.
9. Snodgrass SJ, Rivett DA, Robertson VJ. Manual forces applied during posterior-to-anterior spinal mobilization: a review of the evidence. J Manipulative Physiol Ther. 2006;29:316-29.
10. Lascurain-Aguirrebeña I, Newham D, Critchley DJ. Mechanism of action of spinal mobilizations: a systematic review. Spine (Phila Pa 1976). 2016;41:159-72.
11. Aspinall SL, Leboeuf-Yde C, Etherington SJ, Walker BF. Manipulation-induced hypoalgesia in musculoskeletal pain populations: a systematic critical review and meta-analysis. Chiropr Man Therap. 2019;27:7.
12. Jupin C, Beltran Aibar V, Sarhan FR. Short-term effects of spinal manual therapy on the nervous system in managing musculoskeletal pain: a systematic review. J Clin Med. 2025;14:3830.
13. Nim C, Aspinall SL, Cook CE, Corrêa LA, Donaldson M, Downie AS, et al. The effectiveness of spinal manipulative therapy in treating spinal pain does not depend on the application procedures: a systematic review and network meta-analysis. J Orthop Sports Phys Ther. 2025;55:109-22.
14. Thomas JS, Clark BC, Russ DW, France CR, Ploutz-Snyder R, Corcos DM. Effect of spinal manipulative and mobilization therapies in young adults with mild to moderate chronic low back pain: a randomized clinical trial. JAMA Netw Open. 2020;3:e2012589.
15. McCarthy CJ, Potter L, Oldham JA. Comparing targeted thrust manipulation with general thrust manipulation in patients with low back pain. A general approach is as effective as a specific one. A randomised controlled trial. BMJ Open Sport Exerc Med. 2019;5:e000514.
16. Araujo FX, Ferreira GE, Angellos RF, Stieven FF, Plentz RDM, Silva MF. Autonomic effects of spinal manipulative therapy: systematic review of randomized controlled trials. J Manipulative Physiol Ther. 2019;42:623-34.
17. Coulter ID, Crawford C, Hurwitz EL, Vernon H, Khorsan R, Suttorp Booth M, et al. Manipulation and mobilization for treating chronic low back pain: a systematic review and meta-analysis. Spine J. 2018;18:866-79.
18. Bicalho E, Palma Setti JA, Macagnan J, Rivas Cano JL, Manffra EF. Immediate effects of a high-velocity spine manipulation in paraspinal muscles activity of nonspecific chronic low-back pain subjects. Man Ther. 2010;15:469-75.
19. Klich S, Ficek K, Krymski I, Klimek A, Kawczy ski A, Madeleine P, et al. Quadriceps and patellar tendon thickness and stiffness in elite track cyclists: an ultrasonographic and myotonometric evaluation. Front Physiol. 2020;11:607208.
20. de Oliveira RF, Costa LOP, Nascimento LP, Rissato LL. Directed vertebral manipulation is not better than generic vertebral manipulation in patients with chronic low back pain: a randomised trial. J Physiother. 2020;66:174-9.
21. Castro-Sánchez AM, Lara-Palomo IC, Matarán-Peñarrocha GA, Fernández-de-Las-Peñas C, Saavedra-Hernández M, Cleland J, et al. Short-term effectiveness of spinal manipulative therapy versus functional technique in patients with chronic nonspecific low back pain: a pragmatic randomized controlled trial. Spine J. 2016;16:302-12.
22. Ince G, Sarpel T, Durgun B, Erdogan S. Effects of a multimodal exercise program for people with ankylosing spondylitis. Phys Ther. 2006;86:924-35. Erratum in: Phys Ther. 2006;86:1452.
23. Flores L, Gatchel RJ, Polatin PB. Objectification of functional improvement after nonoperative care. Spine (Phila Pa 1976). 1997;22:1622-33.
24. Feng YN, Li YP, Liu CL, Zhang ZJ. Assessing the elastic properties of skeletal muscle and tendon using shearwave ultrasound elastography and MyotonPRO. Sci Rep. 2018;8:17064.
25. Taş S, Salkın Y. An investigation of the sex-related differences in the stiffness of the achilles tendon and gastrocnemius muscle: inter-observer reliability and inter-day repeatability and the effect of ankle joint motion. Foot (Edinb). 2019;41:44-50.
26. Taş S, Aktaş A, Tufek MT, Dağ F. MyotonPRO is a reliable and repeatable tool for measuring mechanical properties of the upper limb muscles in patients with chronic stroke. Physikalische Medizin Rehabilitationsmedizin Kurortmedizin. 2021;33:286-92.
27. Lettner J, Królikowska A, Ramadanov N, Oleksy Ł, Hakam HT, Becker R, et al. Evaluating the reliability of myotonpro in assessing muscle properties: a systematic review of diagnostic test accuracy. Medicina (Kaunas). 2024;60:851.
28. Senna MK, Machaly SA. Does maintained spinal manipulation therapy for chronic nonspecific low back pain result in better long-term outcome? Spine (Phila Pa 1976). 2011;36:1427-37.
29. Bialosky JE, Bishop MD, Price DD, Robinson ME, George SZ. The mechanisms of manual therapy in the treatment of musculoskeletal pain: a comprehensive model. Man Ther. 2009;14:531-8.
30. Bussières AE, Stewart G, Al-Zoubi F, Decina P, Descarreaux M, Haskett D, et al. Spinal manipulative therapy and other conservative treatments for low back pain: a guideline from the Canadian chiropractic guideline initiative. J Manipulative Physiol Ther. 2018;41:265-93.
31. Cross KM, Kuenze C, Grindstaff TL, Hertel J. Thoracic spine thrust manipulation improves pain, range of motion, and self-reported function in patients with mechanical neck pain: a systematic review. J Orthop Sports Phys Ther. 2011;41:633-42.
32. Cardinale M, Boccia G, Greenway T, Evans O, Rainoldi A. The acute effects of spinal manipulation on neuromuscular function in asymptomatic individuals: a preliminary study. Phys Ther Sport. 2015;16:121-6.
33. Miyamoto N, Hirata K, Miyamoto-Mikami E, Yasuda O, Kanehisa H. Associations of passive muscle stiffness, muscle stretch tolerance, and muscle slack angle with range of motion: individual and sex differences. Sci Rep. 2018;8:8274.
34. Hamilton RI, Garden CLP, Brown SJ. Immediate effect of a spinal mobilisation intervention on muscle stiffness, tone and elasticity in subjects with lower back pain - a randomized cross-over trial. J Bodyw Mov Ther. 2022;29:60-7.
35. Wu Z, Zhu Y, Xu W, Liang J, Guan Y, Xu X. Analysis of biomechanical properties of the lumbar extensor myofascia in elderly patients with chronic low back pain and that in healthy people. Biomed Res Int. 2020;2020:7649157.

36. Lederer AK, Maly C, Weinert T, Huber R. Tissue stiffness is not related to pain experience: an individually controlled study in patients with chronic neck and back pain. Evid Based Complement Alternat Med. 2019;2019:1907168.
37. de Zoete A, Rubinstein SM, de Boer MR, Ostelo R, Underwood M, Hayden JA, et al. The effect of spinal manipulative therapy on pain relief and function in patients with chronic low back pain: an individual participant data meta-analysis. Physiotherapy. 2021;112:121-34.

Original Investigation / Orijinal Araştırma

DOI: 10.4274/tod.galenos.2025.97720
Turk J Osteoporos 2025;31(3):178-88

Comparison of the Osteoporosis Knowledge Levels of Physiotherapy Students Pre-post Awareness Training: A Randomized Controlled Trial

Fizyoterapi Öğrencilerinin Farkındalık Eğitimi Öncesi ve Sonrası Osteoporoz Bilgi Düzeylerinin Karşılaştırılması: Randomize Kontrollü Çalışma

Bengisu Tüfekçi[1], Tuğba Birben Kurt[2], Ayşe Nur Şekeroğlu[3]

[1]Gaziantep İslam Science and Technology University Vocational School of Health Services, Department of Physiotherapy Program, Gaziantep, Türkiye
[2]Recep Tayyip Erdoğan University Güneysu Vocational School of Physical Therapy and Rehabilitation, Department of Women's Health and Rehabilitation, Rize, Türkiye
[3]Osmaniye Korkut Ata University Vocational School of Health Services, Department of Therapy and Rehabilitation, Osmaniye, Türkiye

Abstract

Objective: This research seeks to answer the question of whether education on osteoporosis is more effective when delivered via video conferencing or through informational brochures.
Materials and Methods: The sample of the study consisted of 223 physiotherapy and rehabilitation students. Students were divided into two groups (video group, brochure group) by stratified randomization method according to their achievement levels. Osteoporosis awareness training was conducted through an online video conferencing system and using informative brochures. The osteoporosis knowledge test-revised (R-OKT) was used to assess the level of osteoporosis knowledge and awareness.
Results: While there was no significant difference between the brochure and video training groups in terms of R-OKT scale nutrition, exercise and total knowledge scores ($p>0.05$), it was observed that the post-training scores increased significantly in both groups compared to the pre-education scores ($z>8.5$, $p<0.001$). Statistically significant differences were found between grade levels in post-education nutrition ($\chi^2=52.222$, $p<0.001$), exercise ($\chi^2=56.488$, $p<0.001$), and total knowledge scores ($\chi^2=62.863$, $p<0.001$). Pairwise comparisons revealed that 2nd-year students had significantly higher scores than all other grade levels ($p<0.05$). Positive and statistically significant relationships were found between the students' grade point average and the scores of the R-OKT subscales both before and after the training ($r=0.157$-0.302, $p<0.05$).
Conclusion: Current study showed that video education methods were as effective as traditional printed materials in increasing the level of knowledge and awareness about osteoporosis and that the effect of education may vary depending on class levels. The findings underscore the importance of providing early and well-structured osteoporosis education for future healthcare professionals, which is essential for the development of more effective preventive health strategies.
Keywords: Osteoporosis, awareness, health education strategies, video-based learning

Öz

Amaç: Bu araştırma, osteoporozla ilgili eğitimin video konferans yoluyla mı yoksa bilgilendirici broşürler aracılığıyla mı daha etkili olduğu sorusuna yanıt aramaktadır.
Gereç ve Yöntem: Çalışmanın örneklemi 223 fizyoterapi ve rehabilitasyon öğrencisinden oluşmuştur. Öğrenciler, başarı düzeylerine göre tabakalı randomizasyon yöntemi ile iki gruba (video grubu, broşür grubu) ayrılmıştır. Osteoporoz farkındalık eğitimi, çevrimiçi video konferans sistemi ve bilgilendirici broşürler kullanılarak gerçekleştirilmiştir. Osteoporoz bilgi testi-revize (R-OKT) ölçeği kullanılarak osteoporoz bilgi ve farkındalık düzeyi değerlendirilmiştir.
Bulgular: Broşür ve video eğitim grupları arasında R-OKT ölçeği beslenme, egzersiz ve toplam bilgi puanları açısından anlamlı bir fark bulunmazken ($p>0,05$), eğitim sonrası puanların her iki grupta da eğitim öncesi puanlara göre anlamlı olarak arttığı gözlemlenmiştir ($z>8,5$, $p<0,001$). Eğitim sonrası beslenme ($\chi^2=52,222$, $p<0,001$), egzersiz ($\chi^2=56,488$, $p<0,001$) ve toplam bilgi puanları ($\chi^2=62,863$, $p<0,001$) açısından sınıflar arasında anlamlı farklılık saptanmıştır. Yapılan ikili karşılaştırmalar sonucunda, 2. sınıf öğrencilerinin puanlarının diğer tüm sınıf

Corresponding Author/Sorumlu Yazar: Bengisu Tüfekçi PT, PhD Gaziantep İslam Science and Technology University Vocational School of Health Services, Department of Physiotherapy Program, Gaziantep, Türkiye
E-mail: tufekci.bengisu@gmail.com **ORCID ID:** orcid.org/0000-0002-3042-6083
Received/Geliş Tarihi: 20.07.2025 **Accepted/Kabul Tarihi:** 25.08.2025 **Epub:** 02.10.2025 **Publication Date/Yayınlanma Tarihi:** 05.12.2025

Cite this article as/Atıf: Tüfekçi B, Birben Kurt T, Şekeroğlu AN. Comparison of the osteoporosis knowledge levels of physiotherapy students pre-post awareness training: a randomized controlled trial. Turk J Osteoporos. 2025;31(3):178-88

seviyelerine göre anlamlı düzeyde daha yüksek olduğu belirlenmiştir (p<0,05). Öğrencilerin not ortalamaları ile eğitim öncesi ve sonrası R-OKT alt ölçeklerinin puanları arasında pozitif ve istatistiksel olarak anlamlı ilişkiler bulundu (r=0,157-0,302, p<0,05).
Sonuç: Çalışmamız, video eğitim yöntemlerinin osteoporoz hakkında bilgi ve farkındalık düzeyini artırmada geleneksel basılı materyaller kadar etkili olduğunu ve eğitimin etkisinin sınıf düzeyine göre değişebileceğini göstermiştir. Sonuçlar, gelecekteki sağlık profesyonelleri için erken ve yapılandırılmış osteoporoz eğitiminin önemini vurgulamakta ve bu eğitimin, daha etkili önleyici sağlık stratejilerinin geliştirilmesine katkı sağlayacak nitelikte olduğunu göstermektedir.
Anahtar kelimeler: Osteoporoz, farkındalık, sağlık eğitimi stratejileri, video tabanlı öğrenme

Introduction

Osteoporosis (OP) is a systemic disease characterized by low bone density and deteriorated microarchitecture of bone tissue, leading to an increased fragility of the skeletal system (1). Often referred to as the "silent epidemic of the 21st century", OP typically progresses without symptoms until it manifests through fractures (2). In 2019, it was estimated that 25.5 million women and 6.6 million men in the European Union, the United Kingdom, and Switzerland were affected by OP. Additionally, 4.3 million patients have a history of osteoporotic fractures (3). By 2046, it is predicted that the rate of hospital admissions due to osteoporotic fractures will increase by 150% (4). Osteoporotic fractures significantly reduce patients' quality of life, leading to dependence in daily activities (5). In addition to causing deformities, they also result in serious mortality and morbidity (6). In this context, OP represents a significant public health issue (7).

The etiology of OP involves both genetic and environmental factors. A family history of OP, low body mass index (BMI), calcium and vitamin D deficiencies, excessive caffeine consumption, insufficient sunlight exposure, alcohol and tobacco use, sedentary lifestyle, and menopause are among the risk factors (8). The increasing level of OP with advancing age diminishes quality of life. Physical dysfunctions due to OP have negative effects both sensorially and psychologically (9).

In order to prevent OP, bone tissue should be strengthened and bone health should be protected at an early age. Evaluation of risk factors in the early period and awareness-raising studies in this regard are very important for stopping the development of OP, preventing its progression, and avoiding fractures that may arise from OP (10).

There are studies in the international literature that assess the level of knowledge and awareness of OP in various segments of society (11-13). However, there has been no study in the literature examining the impact of different types of education on the awareness levels of physiotherapy and rehabilitation students. The fundamental aim of our study is to determine the knowledge level and awareness of physiotherapy and rehabilitation students regarding OP, who will play significant roles in every stage of OP treatment when they begin practicing their profession. This research specifically seeks to answer the question of whether education on OP is more effective when delivered via video conferencing or through informational brochures. This study aims to identify methods that will contribute to a more effective learning process for students regarding OP.

Materials and Methods

Type of Research and Ethical Considerations

Our study was randomized and experimental type, and ethical approval was obtained from Osmaniye Korkut Ata University Health Sciences Research Ethics Committee (no: E.177950, date: 23.05.2024). Participants who voluntarily agreed to participate and gave written informed consent were included in the study.

Population and Sample

The population of the study consisted of students studying at Recep Tayyip Erdoğan University, Güneysu School of Physical Therapy and Rehabilitation (n=270). The sample size was not calculated, and it was aimed to reach the entire population. Reaching 70% of the population was accepted as sufficient sample size for the study (14). The sample of the study consisted of 223 (82.6%) physical therapy and rehabilitation students who accepted to participate in the study and completed the scale form completely in May 2024 after ethical approval was obtained. Although the study aimed to reach the entire population, potential selection bias due to voluntary participation was minimized through stratified randomization and a high response rate (82.6%).

Data Collection Tools

Students were divided into groups by stratified randomization method according to their achievement levels [those with grade point average (GPA) >2, those with GPA between 2-3, and those with GPA >3]. Two groups of students were randomized using the sealed envelope method, with equal numbers from each stratum. Both groups received training on OP. OP education was given to the first group [video group (VG)] online via video conferencing system. The online training video used in our study was 15 minutes long. This video content was created by Dr. Physiotherapist, one of the researchers in our study, with verbal and visual expression accompanied by a power point presentation. The second group [brochure group (BG)] was educated by distributing a brochure containing OP content prepared by the same Dr. Physiotherapist who was also involved in our study. The content of the video and the brochure was prepared to be identical in terms of the scope of information. The "personal information form' was used to record the socio-demographic characteristics of the students, and OP knowledge test-revised (R-OKT)" was used to determine the level of OP knowledge and awareness. Before the training, the two groups answered the R-OKT at the same time but in different classes under the supervision of an instructor. Immediately after

answering the R-OKT, the VG students were made to watch a 15-minute OP education video. After the completion of the video conference, the R-OKT was answered by the students for the second time without being removed from the classroom. In the BG, the information brochure was distributed immediately after answering the R-OKT for the first time and they were asked to read it for 15 minutes. After the brochure training, the BG answered the R-OKT for the second time without being removed from the classroom. Participating students were unaware that they would answer the R-OKT scale for the second time at the end of the study. The answers obtained from the study were evaluated in a single blinded manner by another Dr. physiotherapist who took part in our study.

Data were collected online using Google forms. The reason why this online method was preferred in collecting the data was to ensure that the participants responded impartially and without being influenced because the researcher conducting the study was a Asst. Prof. at the university of the participants to whom the questionnaire would be applied, and in addition, to prevent paper waste (15). Each participant was allowed to respond only once; no personal information was collected, and all responses were recorded anonymously. Participants were informed to complete the form voluntarily and attentively. The pre- and post-tests were administered on the same day to minimize environmental influences, ensure the standardization of educational delivery, and prevent information transfer between groups. To enhance data quality and minimize methodological limitations, the Google Forms platform was configured to allow only one response per participant, and all questions were set as mandatory to prevent missing data and duplicate submissions. Participants were required to answer each question before proceeding to the next, in order to reduce inattentive or careless responses. The survey link was distributed exclusively to the target population through restricted access. Participants were clearly informed about the scientific purpose of the study and were explicitly told that participation was entirely voluntary and that they would face no consequences should they choose not to participate. The importance of providing accurate and thoughtful responses for scientific integrity was also emphasized.

Personal Information Form

It consists of 9 questions in which gender, age, class, smoking, presence of chronic diseases, knowledge of regular physical activity, height, body weight, BMI of physical characteristics, and GPA to determine the academic achievement level of the students were questioned.

Osteoporosis Knowledge Test (R-OKT)

The OKT scale was first developed by Kim et al. (16) in 1991 to include 24 multiple-choice questions to measure the level of knowledge about OP. The scale includes questions related to exercise, activity level and diet to prevent OP. The Turkish validity and reliability of the revised form of the OKT (R-OKT), which was revised by Gendler et al. (17) and in which the number of questions was increased to 32, was performed by Şimşir Atalay et al. (18). R-OKT was used in our study. In the R-OKT, questions 1-11 ask about OP risk factors. The answers consist of "high probability of OP", "low probability of OP", "not related to the development of OP" and "don't know". The answers "It is not related to the development of OP" or "I do not know" are considered incorrect and 0 points are given, while the answers "It is likely to have OP" or "it is unlikely to have OP" are considered correct and 1 point is given. The other questions contain 4 optional answers, and 1 point is awarded when the correct answer is marked. The R-OKT has 2 subgroups: The nutrition subgroup contains 26 questions (1-11 and 18-32), and the exercise subgroup contains 20 questions (1-17 and 30-32). These two subgroups have 14 questions in common (1-11 and 30-32). This is taken into account in the total score, and the total score is between 0-32. A high score indicates a good level of OP knowledge (17).

Statistical Analysis

Shapiro-Wilk test was used to evaluate whether the variables in the study conformed to normal distribution. Mean ± standard deviation and median (minimum-maximum) values were given in the descriptive statistics of the variables.

Mann-Whitney U test was used to compare pre-education (pre-e) R-OKT score, post-education (post-e) R-OKT score, age, height, weight, BMI, and GPA values according to brochure-video grouping.

Wilcoxon signed-rank test was used to examine whether the R-OKT scores in the study differed at the measurement times (pre-post education).

Cross tabulations were created; number (n), percentage (%) and chi-square (χ^2) test statistics were given for the comparison of categorical variables according to brochure-video grouping.

Kruskal-Wallis non-parametric analysis of variance was used to compare the R-OKT scores of individuals according to the class of education. Bonferroni correction was made for pairwise comparisons, and the results of the analyses were presented.

Spearman non-parametric correlation coefficient was used in the correlation analysis between R-OKT scores and grade and GPA.

For statistical analyses and calculations, IBM SPSS Statistics 21.0 (IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.) and MS-Excel 2007 programs were used. Statistical significance level was accepted as $p<0.05$.

Results

It was determined that the gender, grade level, smoking status, and physical activity levels of students in the Brochure and VGs were similar ($p>0.05$). The mean age of students in the BG was significantly higher than that of students in the VG ($z=2.067$, $p=0.039$). However, the groups were similar in terms of height, weight, BMI, and GPA ($p>0.05$) (Table 1).

No significant difference was found between the pre-e and post-e R-OKT nutrition scores of students in the Brochure and VGs ($p>0.05$). However, in both groups, post-e nutrition scores

Table 1. Distribution of socio-demographic and physical characteristics by group				
	Brochure group (n=107)	**Video group (n=116)**	**Test statistic**	
Socio-demographic characteristics	**n (%)**	**n (%)**	**χ^2**	**p**
Gender				
Female	88 (82.2)	97 (83.6)	χ^2=0.075	0.785
Male	19 (17.8)	19 (16.4)		
Academic year				
1st year	23 (21.5)	31 (26.7)	χ^2=2.239	0.524
2nd year	31 (29.0)	39 (33.6)		
3rd year	28 (26.2)	24 (20.7)		
4th year	25 (23.3)	22 (19.0)		
Smoking				
Yes	23 (21.5)	25 (21.6)	χ^2=0.001	0.992
No	84 (78.5)	91 (78.4)		
Physical activity				
Yes	30 (28.0)	30 (25.9)	χ^2=0.134	0.714
No	77 (72.0)	86 (74.1)		
Physical characteristics	**Mean ± SD median (min-max)**	**Mean ± SD median (min-max)**	**Z**	**p**
Age (years)	21.40±1.77	20.96±1.49	z=2.067	0.039*
	21.0 (18-29)	21.0 (18-25)		
Height (m)	1.65±0.09	1.66±0.08	z=0.310	0.756
	1.64 (1.45-1.88)	1.65 (1.53-2.01)		
Weight (kg)	64.63±14.14	64.57±16.12	z=0.444	0.657
	62.0 (34-130)	61.0 (44-140)		
BMI (kg/m²)	23.57±4.31	23.34±4.36	z=0.476	0.634
	22.9 (15.2-39.2)	22.6 (17.3-38.8)		
GPA	2.53±0.42	2.45±0.45	z=1.641	0.101
	2.56 (1.0-3.47)	2.50 (1.0-3.43)		

n: Number of individuals, %: Percentage; χ^2: Chi-square test, m: Meter, kg: Kilogram, SD: Standard deviation, BMI: Body mass index, GPA: Grade point average, Min: Minimum, Max: Maximum, z: Mann-Whitney U test, p<0.05

increased significantly compared to pre-e scores (*z>8.9, p<0.001). This finding indicates that both educational methods positively influenced students' nutrition knowledge (Table 2, Figure 1).

Similarly, the pre-e and post-e R-OKT exercise scores of students in the Brochure and VGs were statistically similar (p>0.05). However, in both groups, post-e scores increased significantly compared to pre-e scores (*z>8.5, p<0.001). This result suggests that both educational methods were effective in improving students' exercise knowledge, with no significant difference in the degree of improvement between the groups (Table 2, Figure 1). A similar pattern was observed for total R-OKT scores. Although no statistically significant difference was found between the groups (p>0.05), post-e scores in both groups increased significantly compared to pre-e scores (*z>8.9, p<0.001) (Table 2, Figure 1).

Students' pre-e nutrition scores were similar across grade levels (p>0.05). However, post-e nutrition scores showed statistically significant differences between grade levels (χ^2=52.222, p<0.001). Pairwise comparisons revealed a significant difference between the post-e data of 2nd-year students and all other grade levels (p<0.05). This finding indicates that 2nd-year students achieved higher nutrition scores post-education, demonstrating greater improvement in nutrition knowledge. Additionally, pairwise comparisons of post-e data between 1st and 4th years showed that upper-grade students had significantly higher knowledge (p<0.05) (Table 3).

A significant difference was found in pre-e exercise scores across grade levels (χ^2=21.985, p<0.001). Pairwise comparisons indicated that 4th-year students' exercise scores were significantly higher than those of 1st- and 2nd-year students (p<0.05). Post-e exercise scores also showed significant differences between

Table 2. Comparison of pre-education (PT) and post-education (PT) R-OKT scores between and within groups

	Mean ± SD	Median (min-max)	Mean ± SD	Median (min-max)	p (group)	
R-OKT nutrition						
Pre-e	15.00±2.83	15.0 (7-21)	14.12±3.23	14.0 (3-21)	z**=1.765	0.077
Post-e	20.77±2.57	21.0 (11-25)	20.34±2.64	20.0 (13-26)	z**=1.458	0.145
P (time)	*z=8.926; p<0.001		*z=9.135; p<0.001			
R-OKT exercise						
Pre-e	11.77±2.45	12.0 (6-17)	11.05±2.68	11.0 (3-16)	z**=1.608	0.108
Post-e	16.16±2.33	16.0 (8-20)	15.61±2.38	16.0 (9-20)	z**=1.940	0.052
P (time)	*z=8.508; p<0.001		*z=8.896; p<0.001			
R-OKT total score						
Pre-e	17.88±3.19	18.0 (9-25)	16.90±3.63	17.0 (6-25)	z**=1.714	0.086
Post-e	25.03±3.34	25.0 (14-31)	24.30±3.63	24.0 (14-32)	z**=1.712	0.087
P (time)	*z=8.907; p<0.001		*z=9.099; p<0.001			

R-OKT: Osteoporosis knowledge test-revised, Pre-e: Pre-education, Post-e: Post-education, SD: Standard deviation, Min: Minimum, Max: Maximum, *: Wilcoxon signed-rank test, **: Mann-Whitney U test

grade levels (χ^2=56.488, p<0.001). Pairwise comparisons revealed that 2nd-year students' exercise scores were significantly higher than those of all other grade levels (p<0.05). This suggests that 2nd-year students had higher exercise knowledge both pre-education (except compared to 4th years) and post-education compared to other grades (Table 3).

No statistically significant difference was found in pre-e total scores across grade levels (p>0.05). However, post-e total scores showed significant differences between grade levels (χ^2=62.863, p<0.001). Post-hoc pairwise comparisons (Mann-Whitney U test with Bonferroni correction) revealed that 2nd-year students' total scores were significantly higher than those of all other grade levels (p<0.05). This finding suggests that 2nd-year students exhibited higher overall performance post-education compared to other grades (Table 3).

Analysis of pre-e data revealed a positive and significant correlation between grade level and students' overall academic achievement (GPA) and exercise subscale scores (r=0.387, p<0.001 and r=0.297, p<0.001, respectively). This suggests that as grade level increases, both academic achievement and exercise knowledge improve. Similarly, a positive but weaker correlation was found between grade level and nutrition and total scores (r=0.177, p=0.008 and r=0.173, p=0.010, respectively), indicating a positive relationship between grade level and students' nutrition knowledge and total R-OKT scores (Table 4).

A low-level, positive correlation was observed between GPA and nutrition subscale scores (r=0.172, p=0.010), suggesting that as academic achievement increases, nutrition knowledge also improves. Similarly, moderate (r=0.302, p<0.001) and low-level (r=0.220, p=0.001) positive correlations were found between GPA and exercise and total scores, respectively. These results indicate that as GPA increases, students' nutrition and exercise knowledge, as well as their overall OP knowledge, also improve (Table 4).

Strong, positive linear relationships were identified between students' nutrition and exercise scores and their total scores. A high-level, statistically significant positive correlation was found between nutrition and exercise scores (r=0.748, p<0.001), indicating a strong relationship between nutrition knowledge and physical activity knowledge. Similarly, a very high-level, statistically significant positive correlation was observed between nutrition scores and total scores (r=0.939, p<0.001), suggesting that nutrition knowledge significantly impacts the total OP knowledge score. A very high-level, statistically significant positive correlation was also found between exercise scores and total scores (r=0.827, p<0.001), supporting the significant role of exercise knowledge in increasing the total OP knowledge score (Table 4).

A very weak, positive, and statistically significant relationship was detected between GPA and both nutrition and total scores (r=0.158, p=0.018; r=0.157, p=0.019).

Post-e scores showed strong, positive linear relationships. As GPA increased, the R-OKT total score, i.e., OP knowledge level, also increased. A very high-level, statistically significant positive correlation was found between nutrition scores and both exercise scores (r=0.823, p<0.001) and total scores (r=0.954, p<0.001), indicating that post-e nutrition knowledge directly affects both exercise knowledge and overall OP knowledge. Similarly, a very high-level, statistically significant positive correlation was observed between exercise scores and total scores (r=0.915, p<0.001), suggesting that exercise knowledge significantly impacts the total OP knowledge level and interacts synergistically with nutrition knowledge (Table 4).

Figure 1. Osteoporosis knowledge test-revised

Table 3. Evaluation of R-OKT scores by grade level

Grade	Nutrition subscale		Exercise subscale		Total score	
	Pre-e Mean ± SD median (min-max)	Post-e Mean ± SD median (min-max)	Pre-e Mean ± SD median (min-max)	Post-e Mean ± SD median (min-max)	Pre-e Mean ± SD median (min-max)	Post-e Mean ± SD median (min-max)
1st year	13.74±3.23	19.20±2.21	10.67±2.68	14.91±2.02	16.43±3.64	22.59±2.74
	14.0 (3-21)	19.0 (13-24)	10.5 (5-17)	15.0 (9-19)	17.0 (6-25)	22.0 (14-28)
2nd year	14.37±3.14	22.24±2.54	10.80±2.61	17.50±2.31	17.20±3.64	27.23±3.48
	15.0 (6-19)	23.0 (15-26)	11.0 (3-16)	18.0 (10-20)	18.0 (8-24)	28.0 (16-32)
3rd year	14.73±2.99	19.65±2.71	11.75±2.14	14.81±2.33	17.73±3.19	23.59±3.37
	15.0 (7-21)	20.0 (11-25)	12.0 (6-16)	15.0 (8-19)	18.0 (10-25)	24.0 (14-30)
4th year	15.53±2.62	20.55±1.43	12.72±2.39	15.74±1.37	18.30±3.01	24.34±1.87
	15.0 (10-20)	21.0 (17-23)	13.0 (6-17)	16.0 (12-19)	18.0 (12-24)	25.0 (20-28)
Test statistic	χ^2=7.690 p=0.053	χ^2=52.222 p<0.001*	χ^2=21.895 p<0.001*	χ^2=56.488 p<0.001*	χ^2=7.110 p=0.068	χ^2=62.863 p<0.001*

R-OKT: Osteoporosis knowledge test-revised, Pre-e: Pre-education; Post-e: Post-education, SD: Standard deviation, Min: Minimum, Max: Maximum, χ^2: Kruskal-Wallis test; p<0.05

Table 4. Correlation between grade level, GPA, and pre-education and post-education R-OKT scores (n=223)

Pre-e						
		Grade	**GPA**	**Nutrition**	**Exercise**	**Total**
Grade	r					
	p					
GPA	r	0.387				
	p	<0.001*				
Nutrition	r	0.177	0.172			
	p	0.008*	0.010*			
Exercise	r	0.297	0.302	0.748		
	p	<0.001*	<0.001*	<0.001*		
Total	r	0.173	0.220	0.939	0.827	
	p	0.010*	0.001*	<0.001*	<0.001*	
Post-e						
Grade	r					
	p					
GPA	r	0.387				
	p	<0.001*				
Nutrition	r	0.057	0.158			
	p	0.396	0.018*			
Exercise	r	-0.029	0.127	0.823		
	p	0.664	0.058	<0.001*		
Total	r	0.054	0.157	0.954	0.915	
	p	0.419	0.019*	<0.001*	<0.001*	

R-OKT: Osteoporosis knowledge test-revised, GPA: Grade point average, Pre-e: Pre-education; Post-e: Post-education, r: Spearman correlation coefficient, *: Statistically significant correlation (p<0.05)

Discussion

Since OP is an important public health problem that is preventable and manageable, recent studies have emphasized the importance of increasing individuals' awareness and knowledge levels about OP (19-21). However, when the literature is analyzed, it is seen that studies conducted to raise awareness in OP are focused on postmenopausal women and geriatric individuals (22-24). Although there is a need to raise awareness about OP in every age period, raising awareness at young ages and increasing the level of knowledge on this subject will provide more effective results in improving bone health and preventing OP (25). Therefore, in our study, we aimed to increase the awareness and knowledge levels of physiotherapy and rehabilitation students who will have an important role in health services in the fight against OP. In this context, we compared the OP knowledge levels of the students before and after the awareness training given thorough video and brochure methods. In the studies presented in the literature on OP knowledge and awareness level, it is seen that scales prepared by researchers are mostly used as assessment tools (26-28). In our study, we preferred to use the R-OKT scale, whose validity and reliability have been proven nationally and internationally (17,18).

When the current literature on OP awareness level is examined, it is seen that both video conferencing methods and traditional education methods given with printed materials provide a significant increase in the level of OP knowledge. In this regard, Lopez-Olivo et al. (29) compared the effectiveness of written booklet and video-based education in a randomized controlled study with 225 patients with OP. As a result, they found that the level of OP knowledge increased statistically significantly in both education groups (29). In the randomized controlled study of Şahin et al. (30) investigating the level of OP knowledge in patients with Parkinson's disease, a comparison was made between informative brochure and face-to-face traditional education model given in addition to this. As a result of the study, it was reported that the level of OP knowledge increased significantly in both groups in the early period ($p<0.001$), but there was no significant difference between the two groups in terms of knowledge level after the training (30). Chotiyarnwong et al. (31) conducted a randomized controlled study with 413 participants who applied to an orthopaedic clinic for similar purposes. In this study, two different groups were formed as traditional lecture and video lecture, and the effectiveness levels of these types of education on OP knowledge levels were compared. It was reported that there was no statistically significant difference between the two groups before and after the training and that video-based training was as effective as traditional face-to-face training (31). In the present study, the effectiveness levels of two different training methods, brochure and video, were compared in order to increase the OP knowledge level of physiotherapy and rehabilitation students. Similar to the literature, it was seen that the OP knowledge level of both the BG and the VG increased significantly after the training ($p<0.001$). However, there was no significant difference between the two groups in terms of OP knowledge level after the training. We think that the fact that the brochure material contains visual expression and important information is emphasized with bold fonts is effective on the similar results with video training. Unlike the findings of our study, in the study conducted by Therdyothin and Amphansap (32) which compared training through Video Podcasts with traditional face-to-face education using reading materials in orthopedic residents' OP education, the VG initially had a lower OP knowledge level. However, after the training, the test score average of the VG was significantly higher than that of the traditional group ($p<0.001$). It was reported that the increase in knowledge was more pronounced among younger participants in the VG. Interestingly, however, participants stated that they preferred live lectures over Video Podcasts (32). Sunthornsup et al. (33) compared brochure and video methods for family education of children with juvenile idiopathic arthritis. In this study, which had a similar chronic disease education with OP, it was reported that the VG was significantly superior to the BG in terms of knowledge score in the evaluation performed immediately after the education ($p=0.003$). In addition, it was reported that video education (83.5%) increased knowledge more than brochure (69.1%) in parents who were found to have a lower level of education. However, in the follow-up evaluation performed four weeks later, it was reported that the scores of both groups were significantly higher than the baseline level, but the difference between the two groups decreased. In this study, the researchers emphasized that the video method provided ease of comprehension especially in participants with low education level, but both training methods created a permanent effect (33). In the light of this information, it has been observed that OP education by videoconferencing method can be as effective as traditional printed materials such as brochures, and it can even provide higher short-term learning in some conditions such as young age and high education level in the literature. However, in many studies conducted on this subject, it was observed that there was no statistically significant superiority between the two methods used for learning. Video-containing trainings may be more interesting and more memorable for students who have a high use of technology and are prone to audiovisual learning. As a matter of fact, in the study conducted with orthopaedic assistants, it was observed that the VG provided much higher success. Therefore, considering that the most effective learning method is the one that is sensitive to the needs and preferences of the target audience, we think that it would be important for educators to use traditional and video conferencing methods together by using hybrid methods. Vallée et al. (34) reported that hybrid education consistently showed superior effects on knowledge outcomes compared to traditional education methods in health education in their study in which they compared education methods in medical students, which supports this idea.

In the present study, when the OP knowledge level and awareness of the participating students were evaluated using

the R-OKT scale, it was shown that it increased significantly after the training ($p<0.005$). In addition, while the knowledge level of the classes was similar before the training, it was observed that there was a statistically significant difference after the training ($\chi^2=52.222$, $p<0.001$); especially 2nd grade students had higher scores in the nutrition and exercise subheadings compared to other classes. This suggests that the knowledge about OP was reinforced within the scope of the current study and that the previous academic background contributed to the learning process. Similar to the results of our study, a positive correlation was observed between academic achievement (GPA) before and after the training and OP knowledge scores ($r=0.220$, $p=0.001$). This result supports that participants with high academic achievement have more knowledge about OP and that the training shows more effective results in these students. In the current study, when the total R-OKT scores of the first year students before the training were analyzed, it was observed that they had lower scores than the fourth year students. Erçalık et al. (35) examined the OP knowledge levels of students studying in the departments of nutrition and dietetics and physiotherapy and rehabilitation, whose participant population closely resembled that of our study. Using the R-OKT scale, as in our study, they similarly found that fourth-year students had significantly higher total scores than first-year students (35). On the other hand, when analyzing the post-education results of the R-OKT nutrition and exercise subscales in our study, a more pronounced increase in scores was observed in first-year students compared to their pre-education levels, whereas, despite having higher baseline knowledge levels, the increase in fourth-year students was more limited. We believe that this may be due to the fact that as the class level increases, knowledge levels reach a certain threshold, thereby reducing the additional contribution of the provided education. The findings of Aybala Koçak et al. (36) in their study on physical therapy and rehabilitation students in which they investigated the level of knowledge about calcium and vitamin D deficiency and OP are supportive of our results as they reported that the level of knowledge of students about OP, calcium and vitamin D gradually increased with education during school education. In this context, the inclusion of education on OP awareness in the curriculum from the early stages and its comprehensive handling will contribute to the physical therapy and rehabilitation students who will work with both the risk group and osteoporotic patients after graduation to start clinical practice better equipped in this context. These results reveal the necessity of periodically repeating the trainings with different methods in order to increase the long-term permanence of the education offered in the early period, especially since the rate of increase in the knowledge of the students in the advanced classes is slower.

Study Limitations

The limitations of our study are that only short-term knowledge assessment was performed after the training. The reason for this preference is that the outcome assessment performed immediately after the training can be performed without loss of sample. In addition, it was predicted that large sample losses may occur due to the presence of students approaching graduation status among the participating students. For this reason, further studies are needed to evaluate the long-term permanence of the effects of the training. Although the study aimed to reach the entire population, the risk of selection bias due to voluntary participation could not be entirely eliminated. However, stratified sampling based on academic achievement and a high participation rate were intended to mitigate this risk. In addition, standardisation and quality assessment of different training materials should be considered in future studies to be conducted in this context. There is a need for studies evaluating the evidence-based, up-to-date, comprehensible and socio-cultural appropriateness of video and brochure training contents. Additionally, methods for collecting participants' feedback on the clarity and effectiveness of educational materials can be incorporated into the study, thereby expanding the scope of the findings.

Conclusion

In the present study, we examined the effectiveness of two different training methods to increase OP awareness in physiotherapy and rehabilitation students. We concluded that both brochure and video conference methods statistically significant increased the level of OP knowledge. Especially the fact that second-year students had higher post-education knowledge levels in the nutrition, exercise and total subcategories compared to other grades suggests that the effect of education may vary depending on class levels. Furthermore, the fact that the increase in post-training knowledge of fourth-grade students was more limited compared to the first-grade students indicates that the effect of the trainings given in the early years of education is stronger and the rate of knowledge acquisition decreases in the following grades. In addition, the positive correlation between students' grade point averages revealing their academic achievement and their OP knowledge scores suggests that academic achievement may be an important factor in increasing OP awareness.

The results of our study show that both traditional printed materials and video education methods are effective in increasing the level of knowledge and awareness of OP. Especially considering that the early effects of the education given in the early grades are more pronounced, it will be important to integrate different educational methods into the curriculum to be repeated periodically from the beginning in terms of the permanence of the information. In the future, we recommend that studies should be conducted to evaluate the quality and reliability of different types of educational materials such as videos and brochures used for OP education in terms of content, and to investigate long-term knowledge gains.

Ethics

Ethics Committee Approval: Our study was randomized and experimental type, and ethical approval was obtained from Osmaniye Korkut Ata University Health Sciences Research Ethics Committee (no: E.177950, date: 23.05.2024).

Informed Consent: Participants who voluntarily agreed to participate and gave written informed consent were included in the study.

Footnotes

Authorship Contributions

Concept: B.T., Design: B.T., Data Collection or Processing: B.T., T.B.K., A.N.Ş., Analysis or Interpretation: B.T., T.B.K., A.N.Ş., Literature Search: B.T., T.B.K., A.N.Ş., Writing: B.T., T.B.K., A.N.Ş.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Clynes MA, Harvey NC, Curtis EM, Fuggle NR, Dennison EM, Cooper C. The epidemiology of osteoporosis. Br Med Bull. 2020;133:105-7.
2. Aibar-Almazán A, Voltes-Martínez A, Castellote-Caballero Y, Afanador-Restrepo DF, Carcelén-Fraile MDC, López-Ruiz E. Current status of the diagnosis and management of osteoporosis. Int J Mol Sci. 2022;23:9465.
3. Willers C, Norton N, Harvey NC, Jacobson T, Johansson H, Lorentzon M, et al. Osteoporosis in Europe: a compendium of country-specific reports. Arch Osteoporos. 2022;17:23.
4. McCabe E, Ibrahim A, Singh R, Kelly M, Armstrong C, Heaney F, et al. A systematic review of the Irish osteoporotic vertebral fracture literature. Arch Osteoporos. 2020;15:34.
5. Spångeus A, Willerton C, Enthoven P, Grahn Kronhed AC. Patient education improves pain and health-related quality of life in patients with established spinal osteoporosis in primary care-a pilot study of short- and long-term effects. Int J Environ Res Public Health. 2023;20:4933.
6. Yang M, Yin H, Zhen D, Ding Y, Wang Y, Sun L, et al. Exposure to famine in every stage of life and the risk of osteoporosis and fractures later in life: a cross-sectional study. Bone. 2023;168:116644.
7. Khan AA, Slart RHJA, Ali DS, Bock O, Carey JJ, Camacho P, et al. Osteoporotic fractures: diagnosis, evaluation, and significance from the international working group on DXA best practices. Mayo Clin Proc. 2024;99:1127-41.
8. Carvalho ACL, Gomes FA, Bernardo Silva AV, Araújo MS, Barbosa GM, Avila MA, et al. Self-care during the COVID-19 pandemic: development of a virtual educational booklet for postmenopausal women with osteoporosis. J Bodyw Mov Ther. 2023;34:74-80.
9. Hopman WM, Berger C, Joseph L, Morin SN, Towheed T, Anastassiades T, et al. Longitudinal assessment of health-related quality of life in osteoporosis: data from the population-based Canadian multicentre osteoporosis study. Osteoporos Int. 2019;30:1635-44.
10. Ayyash M, Jaber K, Daghash R, Abu-Farha R, Alefishat E. Perception and awareness of osteoporosis and its related risk factors among women: a cross-sectional study. Electron J Gen Med. 2023;20:em475.
11. Kadam N, Chiplonkar S, Khadilkar A, Khadilkar V. Low knowledge of osteoporosis and its risk factors in urban Indian adults from Pune city, India. Public Health Nutr. 2019;22:1292-9.
12. Chelf S, Davis RE, Bass MA, Ford MA, Firouzabadi AD, Leo JT, et al. Osteoporosis knowledge and health beliefs among middle-aged men and women in the Southern United States. J Osteopath Med. 2022;122:453-9.
13. Karakaş A, Erdinç Gündüz N, Özçelik S, Limoncu H, Dilek B, Gülbahar S. Awareness and knowledge levels of osteoporosis in patients with multiple sclerosis. Turk J Osteoporos. 2022;28:131-6.
14. Sumbuloglu K, Sumbuloglu V. Biyoistatistik. 9. Baskı. Hatipoğlu Basım ve Yayın San. 2000 Tic. Ltd. Sti., Ankara.
15. Ball H. Conducting online surveys. J Hum Lact Consultant. 2019;35:413-7.
16. Kim KK, Horan ML, Gendler P. Osteoporosis knowledge test. Grand Valley State University; 1991.
17. Gendler PE, Coviak CP, Martin JT, Kim KK, Dankers JK, Barclay JM, et al. Revision of the osteoporosis knowledge test: reliability and validity. West J Nurs Res. 2015;37:1623-43.
18. Şimşir Atalay N, Akkaya N, Şahin F. The psychometric properties of the turkish version of revised 2011-osteoporosis knowledge test. Turk J Osteoporos. 2015;21:127-31.
19. Chen X, Jia X, Lan J, Wu W, Ni X, Wei Y, et al. Association between risk factors and bone mineral density and the development of a self-assessment tool for early osteoporosis screening in postmenopausal women with type 2 diabetes. PeerJ. 2024;12:e18283.
20. Li MH, Chen IC, Yang HW, Yen HC, Ke YY, Chen YM, et al. Association of methylenetetrahydrofolate reductase rs1801133 polymorphism with osteoporosis and fracture risk in Taiwan. Int J Med Sci. 2024;21:2261-71.
21. Minalyan A, Li T, D'Anna K, Alfraji N, Gabrielyan L, Downey C. Suboptimal osteoporosis care in hospitalized patients: a retrospective analysis of vertebral compression fractures detected on computed tomography. Rheumatol Int. 2024;44:2599-605.
22. Tabor E, Grodzki A, Pluskiewicz W. Higher education and better knowledge of osteoporosis improve bone health in Polish postmenopausal women. Endokrynol Pol. 2022;73:831-6.
23. Senthilraja M, Cherian KE, Jebasingh FK, Kapoor N, Paul TV, Asha HS. Osteoporosis knowledge and beliefs among postmenopausal women: a cross-sectional study from a teaching hospital in southern India. J Family Med Prim Care. 2019;8:1374-8.
24. Chan CY, Subramaniam S, Chin KY, Ima-Nirwana S, Muhammad N, Fairus A, et al. Knowledge, beliefs, dietary, and lifestyle practices related to bone health among middle-aged and elderly Chinese in Klang Valley, Malaysia. Int J Environ Res Public Health. 2019;16:1787.
25. Chan CY, Mohamed N, Ima-Nirwana S, Chin KY. A review of knowledge, belief and practice regarding osteoporosis among adolescents and young adults. Int J Environ Res Public Health. 2018;15:1727.
26. D'Silva F, Pinto CA. Knowledge level of pre and post menopausal women on osteoporosis: a cross-sectional study. IOSR Journal of Nursing and Health Science (IOSR-JNHS). 2017;6:70-5.
27. Selçuk EB, Tetik BK, Sönmez B, Tekindal MA. Assesment of the knowledge, attitudes and behaviours of postmenopausal women about osteoporosis. Ankara Med J. 2015;15:114-9.
28. Hassan AOA, Alshammari KF, Bakrshoom YF, Alhamazani AF, Alsadun AS. Awareness of osteoporosis among general population in Ha'il city, Saudi Arabia. Medical Science. 2021;25:689-95.
29. Lopez-Olivo MA, des Bordes JKA, Lin H, Rizvi T, Volk RJ, Suarez-Almazor ME. Comparison of multimedia and printed patient education tools for patients with osteoporosis: a 6-month randomized controlled trial. Osteoporos Int. 2020;31:857-66.
30. Şahin MA, Aydemir MD, Dönmez Çolakoğlu B, Çakmur R, Ünal B, Gülbahar S. The effect of osteoporosis education on osteoporosis knowledge level and daily life in Parkinsonıs disease patients: a 12-week, randomized-controlled trial. Turk J Phys Med Rehabil. 2024;70:379-89.
31. Chotiyarnwong P, Boonnasa W, Chotiyarnwong C, Unnanuntana A. Video-based learning versus traditional lecture-based learning

for osteoporosis education: a randomized controlled trial. Aging Clin Exp Res. 2021;33:125-31.

32. Therdyothin A, Amphansap T. Comparison between video podcasts and traditional learning in osteoporosis for orthopedic resident physicians during the COVID-19 pandemic in Thailand. J Southeast Asian Orthop. 2023;47:3-10.
33. Sunthornsup W, Vilaiyuk S, Soponkanaporn S. Effect of educational brochure compared with video on disease-related knowledge in patients with juvenile idiopathic arthritis: a randomized controlled trial. Front Pediatr. 2022;10:1048949.
34. Vallée A, Blacher J, Cariou A, Sorbets E. Blended learning compared to traditional learning in medical education: systematic review and meta-analysis. J Med Internet Res. 2020;22:e16504.
35. Erçalık C, Adar S, Başkaya MÇ, Persil Özkan Ö, Şener S, Uğur E. Examination and comparison of osteoporosis knowledge level of students studying in nutrition and dietetics and physiotherapy and rehabilitation departments. Turk J Osteoporos. 2019;25:83-7.
36. Aybala Koçak F, Barut Ö, Kurt EE, Şaş S, Öztürk Durmaz H, Tuncay F. Knowledge level and awareness about calcium deficiency, vitamin D deficiency and osteoporosis among physical therapy and rehabilitation school students. Turk J Osteoporos. 2020;26:23-9.

DOI: 10.4274/tod.galenos.2025.04875
Turk J Osteoporos 2025;31(3):189-94

Performance Comparison of Gemini, DeepSeek, and ChatGPT-4o on American Board of Physical Medicine and Rehabilitation Board Exam Practice Questions

Amerikan Fiziksel Tıp ve Rehabilitasyon Kurulu Sınavı Deneme Sorularında Gemini, DeepSeek ve ChatGPT-4o'nun Performans Karşılaştırması

Gonca Sağlam Akkaya, Hanife Baykal Şahin

Karadeniz Technical University Faculty of Medicine, Department of Physical Medicine and Rehabilitation, Trabzon, Türkiye

Abstract

Objective: The rapid advancement of large language models (LLMs) has demonstrated their important potential in medical education and assessment. This study aimed to evaluate the performance of three prominent LLMs (Gemini, DeepSeek, and ChatGPT-4o) on practice questions designed to be representative of the American Board of Physical Medicine and Rehabilitation (ABPMR) certification examination. By comparing their accuracy across various medical domains, we sought to understand their current capabilities as supplementary tools for medical trainees.

Materials and Methods: We used a comprehensive set of 100 publicly available ABPMR practice questions from 2015, ensuring a consistent benchmark for comparison. These questions, which cover a wide range of topics and clinical scenarios, were systematically fed into Gemini, DeepSeek, and ChatGPT-4o via their web interfaces. The responses were then independently analyzed by a blinded physical medicine and rehabilitation specialist to ensure an unbiased evaluation.

Results: DeepSeek achieved the highest overall accuracy at 88%, significantly outperforming Gemini (81%, p=0.022) but not showing a statistically significant difference compared to ChatGPT-4o (86%, p=0.238). The models displayed varying strengths across different specialty areas. ChatGPT-4o performed best in neurologic disorders (90%) and electrodiagnosis (87%). In contrast, DeepSeek led in musculoskeletal medicine (88%), patient management (97%), and amputation (100%). Gemini performed comparably to DeepSeek in equipment/assistive technology (90%). No significant inter-model differences were found in domains such as rehabilitation problems (93%), basic sciences (80%), and applied sciences (83%).

Conclusion: Our findings suggest that while DeepSeek demonstrated superior aggregate performance, all three LLMs possess unique, complementary strengths across different domains of physical medicine and rehabilitation. The lack of significant differences in domain-stratified analyses points to the task-specific nature of LLM efficacy. These results indicate that LLMs are promising supplementary educational tools, but their persistent limitations in complex clinical reasoning necessitate continued human oversight and validation.

Keywords: Large language models, physical medicine and rehabilitation, medical education

Öz

Amaç: Büyük dil modellerinin (BDM) hızlı gelişimi, tıp eğitimi ve değerlendirmesinde önemli bir potansiyel göstermiştir. Bu çalışmanın amacı, önde gelen üç BDM olan Gemini, DeepSeek ve ChatGPT-4o'nun, Amerikan Fiziksel Tıp ve Rehabilitasyon Kurulu (ABPMR) sertifika sınavını temsil eden deneme sorularını yanıtlama performansını değerlendirmekti. Bu modellerin tıp öğrencileri için yardımcı araçlar olarak mevcut yeteneklerini anlamak için farklı tıbbi alanlardaki doğruluklarını karşılaştırma hedeflendi.

Gereç ve Yöntem: 2015 yılında erişime sunulmuş olan 100 adet ABPMR deneme sorusundan oluşan kapsamlı bir set kullandıldı. Bu sorular, geniş konu çeşitliliği ve klinik senaryoları kapsamakta olup, Gemini, DeepSeek ve ChatGPT-4o'nun web arayüzlerine sistematik bir şekilde girildi. Yanıtlar, tarafsız bir değerlendirme sağlamak amacıyla, hangi BDM tarafından üretildiği bilinmeyen (körleme yöntemi) bağımsız bir fiziksel tıp ve rehabilitasyon uzmanı tarafından analiz edildi.

Corresponding Author/Sorumlu Yazar: Assoc. Prof. Gonca Sağlam Akkaya MD, Karadeniz Technical University Faculty of Medicine, Department of Physical Medicine and Rehabilitation, Trabzon, Türkiye
E-mail: goncasaglam@hotmail.com **ORCID ID:** orcid.org/0000-0001-7713-4435
Received/Geliş Tarihi: 14.08.2025 **Accepted/Kabul Tarihi:** 22.09.2025 **Epub:** 08.10.2025 **Publication Date/Yayınlanma Tarihi:** 05.12.2025

Cite this article as/Atıf: Sağlam Akkaya G, Baykal Şahin H. Performance comparison of Gemini, DeepSeek, and ChatGPT-4o on American board of physical medicine and rehabilitation board exam practice questions. Turk J Osteoporos. 2025;31(3):189-94

Öz

Bulgular: DeepSeek, %88 ile en yüksek genel doğruluğa ulaştı. Gemini'den (%81, p=0,022) önemli ölçüde daha iyi performans göstermiş, ancak ChatGPT-4o'dan (%86, p=0,238) istatistiksel olarak anlamlı bir farkla ayrılmamıştı. Modeller, farklı uzmanlık alanlarında değişen güçlü yönler sergiledi. ChatGPT-4o, nörolojik bozukluklar (%90) ve elektrodiyagnoz (%87) alanlarında en yüksek performansı gösterdi. Buna karşılık, DeepSeek kas-iskelet tıbbı (%88), hasta yönetimi (%97) ve ampütasyon (%100) alanlarında lider oldu. Gemini ise ekipman/yardımcı teknoloji (%90) alanında DeepSeek ile benzer bir performans sergiledi. Rehabilitasyon sorunları (%93), temel bilimler (%80) ve uygulamalı bilimler (%83) gibi alanlarda ise modeller arasında anlamlı bir fark bulunmadı.

Sonuç: Bulgularımız, DeepSeek'in genel performansta üstünlük gösterse de, her üç BDM'nin de fiziksel tıp ve rehabilitasyonun farklı alanlarında benzersiz ve tamamlayıcı güçlü yönlere sahip olduğunu düşündürmektedir. Alana göre yapılan analizlerde istatistiksel olarak anlamlı farklılıkların bulunmaması, BDM etkinliğinin göreve özgü değişkenliğini vurgulamaktadır. Bu sonuçlar, BDM'lerin tıp eğitiminde umut verici ek araçlar olduğunu göstermekle birlikte, karmaşık klinik muhakemedeki kalıcı sınırlamaları nedeniyle insan gözetiminin ve doğrulamasının kritik önemini koruduğunu vurgulamaktadır.

Anahtar kelimeler: Büyük dil modelleri, fiziksel tıp ve rehabilitasyon, tıp eğitimi

Introduction

The field of artificial intelligence (AI) has seen rapid advancements in recent years, particularly in large language models (LLMs). They have demonstrated their potential to revolutionize various fields, including healthcare and medical education (1,2). These models, trained on vast datasets of text and code, can generate human-like text, translate languages, write different kinds of creative content, and answer questions in an informative way (3). In the medical domain, LLMs are being explored for applications ranging from assisting with clinical decision-making (4) and summarizing medical records (5) to generating patient education materials (6) and potentially aiding in exam preparation.

The American Board of Physical Medicine and Rehabilitation (ABPMR) examination serves as a crucial benchmark for physicians specializing in this field, assessing their knowledge and clinical reasoning skills (7). Success on this exam is essential for board certification and signifies competency in the specialty. This study aimed to evaluate the performance of three prominent LLMs; Gemini, DeepSeek, and ChatGPT-4, in their ability to answer questions representative of ABPMR practice questions.

Materials and Methods

Data Collection

We obtained a comprehensive set of 100 practice questions from the ABPMR. These questions were sourced from publicly available practice materials and previous examination sets, ensuring a representative sample of the board's assessment style and content. They were released by the ABPMR in June 2015 as a study tool and have been permanently removed from their active examination item banks. They present a wide range of topics relevant to physical medicine and rehabilitation (PMR). These questions spanned all core domains of PMR, distributed as follows: Neurologic disorders (30 questions), musculoskeletal medicine (32 questions), electrodiagnosis (15 questions), amputation (5 questions), rehabilitation problems (15 questions), basic sciences (15 questions), and applied sciences (15 questions). Additionally, questions were categorized by focus: Patient management (32 questions) and equipment/assistive technology (10 questions). The questions are in a multiple-choice format, often presenting clinical scenarios, designed to assess foundational knowledge and clinical reasoning relevant to the certification examination. The original document includes an answer key. The static nature and prior public release of these questions ensure a consistent and accessible benchmark for comparing the performance of LLM.

AI Models and Question Processing

Three state-of-the-art LLMs were selected for this study: Gemini, DeepSeek, and ChatGPT-4. Each model was accessed through its web interface, using the most recent available versions. The ABPMR questions were formatted and input into each AI model without modification. We ensured that the input format was consistent across all three models to maintain fairness in the comparison.

Response Generation

Each AI model was prompted with the ABPMR questions individually. The models were instructed to provide their best attempt at answering each question without any additional context or information beyond what was provided in the question itself. The analysis was performed after a total of three attemps. A PMR specialist, blinded to which AI model generated each response, independently scored the answers. The performance was assessed by another specialist who calculated the overall accuracy, percentage and number of correctly answered questions, for each LLM. Additionally, the performances were assessed across different question topics to identify potential strengths and weaknesses of each model in specific areas of PMR.

This study was conducted in compliance with ethical guidelines for AI research. No patient data or confidential examination materials were used. The study focused solely on the AI models' performance on publicly available practice questions. This study did not require ethical approval as it was based on publicly

available data and did not involve human participants or private medical information.

Statistical Analysis

The statistical analysis was performed using SPSS version 29 (IBM, Armonk, NY). Chi-squared tests of independence were used for categorical comparisons of correct versus incorrect responses within each question category. Fisher's exact test was used in categories with small sample sizes, such as "amputation". To compare the mean accuracy percentages across the three models, a One-Way ANOVA was conducted for overall performance, as well as for questions categorized by organ system and by question focus. When a significant result was found in the ANOVA for overall performance, a Tukey HSD post-hoc analysis was performed to identify specific pairwise differences between the models, with adjustments made to control for multiple comparisons. Key assumptions for the ANOVA, including homogeneity of variances (verified by Levene's test) and normality of residuals (confirmed by Shapiro-Wilk tests), were checked and met.

Results

All models showed incremental improvement from first to third attempts, with DeepSeek maintaining the highest accuracy at each stage (86%, 86%, 88%). DeepSeek significantly outperformed Gemini (p=0.022), though no significant differences existed between DeepSeek and ChatGPT-4o (p=0.238) or ChatGPT-4o and Gemini (p=0.100) (Table 1). Table 2 presents the summary of key findings.

When stratified by organ system, model performance varied by specialty. In neurologic disorders (30 questions), ChatGPT-4o achieved the highest accuracy (90%), exceeding Gemini (80%) and DeepSeek (87%). For musculoskeletal medicine (32 questions), DeepSeek led (88%) over Gemini (81%) and ChatGPT-4o (84%). DeepSeek achieved perfect accuracy in amputation (5 questions, 100%), while Gemini and ChatGPT-4o both scored 80%. All models performed equally in rehabilitation problems (93%) and basic sciences (80%). No significant differences across models for organ system-based performance was found (F=1.12, p=0.350), consistent with individual categories (all p>0.05) (Figure 1).

Analysis by question focus revealed additional task-specific strengths. In electrodiagnosis (15 questions), ChatGPT-4o excelled (87%) over DeepSeek (80%) and Gemini (67%). DeepSeek dominated patient management (32 questions, 97%) compared to ChatGPT-4o (91%) and Gemini (88%). For equipment/assistive technology (10 questions), Gemini and DeepSeek tied (90%), outperforming ChatGPT-4o (80%). No model differences emerged in applied sciences (all 83%). No significant difference was observed for inter-model variation (F=0.63, p=0.548), aligning with non-significant chi-squared tests for all focus categories (Figure 2).

Discussion

LLMs are creating a major change in medical education, allowing for new uses in knowledge sharing, customized self-assessment tools, and simulated practice of clinical reasoning. As these systems are increasingly used for high-stakes tasks like preparing for board exams and assisting with clinical decisions, it is crucial to rigorously evaluate their accuracy, limitations, and potential biases to ensure safe implementation (8). This imperative is underscored by documented instances of LLMs generating plausible yet incorrect medical information, highlighting critical gaps in reliability (1,9).

Table 1. Comparison of overall performances of AI models (Gemini, DeepSeek and ChatGPT-4o) for ABPMR board practice questions

		Gemini (%)	Deepseek (%)	ChatGPT-4o (%)	p
1st attemtp	Incorrect	20	14	16	0.21
	Correct	80	86	84	
2nd attemp	Incorrect	19	14	16	0.41
	Correct	81	86	84	
3rd attempt	Incorrect	16	12	13	0.17
	Correct	84	88	87	

AI: Artificial intelligence, ABPMR: American Board of Physical Medicine and Rehabilitation

Table 2. Summary of findings

Metric	Gemini	DeepSeek	ChatGPT-4o
Highest overall accuracy	81.7%	86.7%	85.0%
Top specialty area	Equipment and assistive tecnology, medical rehabilitation	Musculoskeletal medicine, patient management, amputation	Neurologic disorders, electrodiagnosis
Statistical advantage	None	Outperformed Gemini (p=0.022)	None

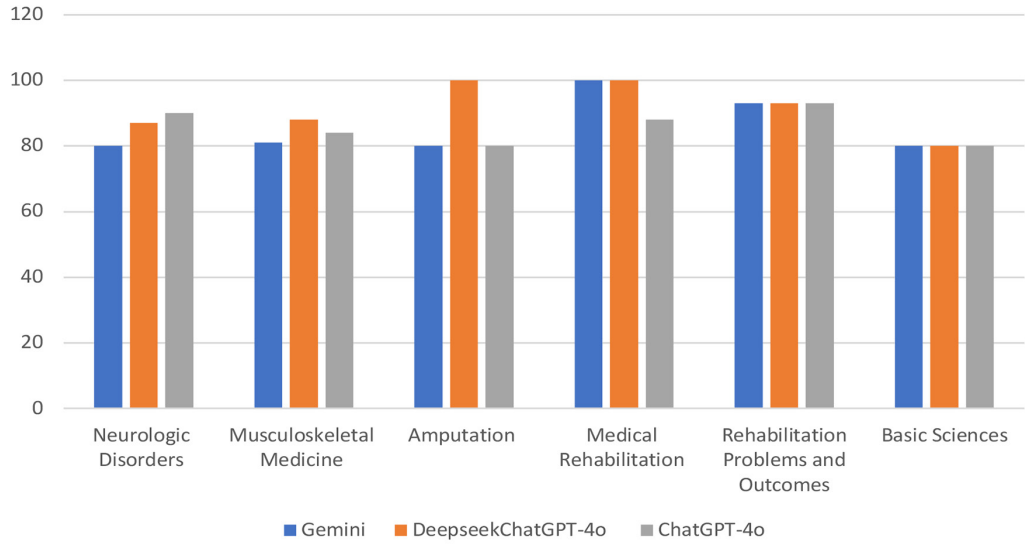

Figure 1. Comparative performance percentages of AI models (Gemini, DeepSeek, and ChatGPT-4o) for "type of problem/organ system" practice questions
AI: Artificial intelligence

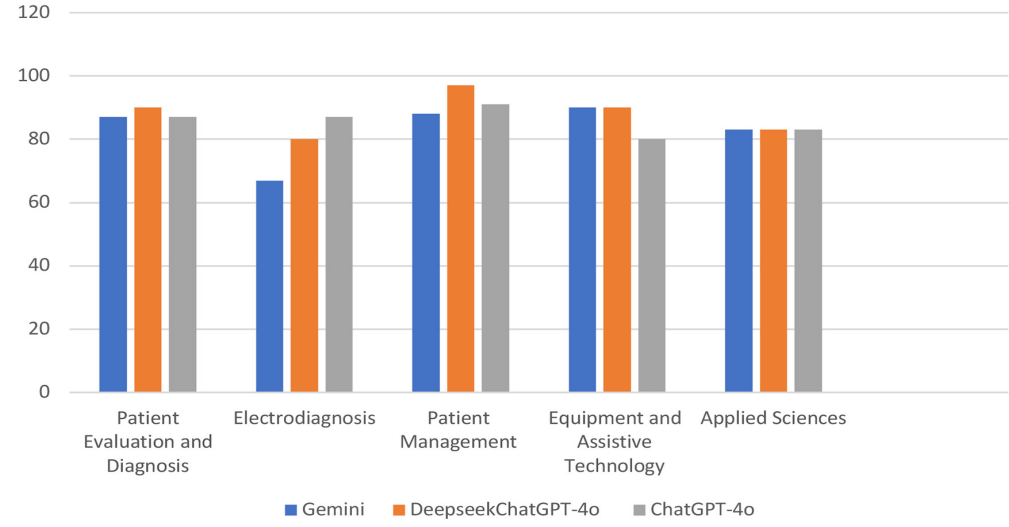

Figure 2. Comparative performance percentages of AI models (Gemini, DeepSeek, and ChatGPT-4o) for "focus of question and patient management" practice questions
AI: Artificial intelligence

Our analysis indicates that while all three LLMs demonstrated a certain level of proficiency in answering ABPMR-style questions, their accuracy varied significantly across different domains. DeepSeek exhibited the highest overall accuracy, suggesting that its training data and model architecture may be better optimized for medical reasoning tasks. This aligns with prior research demonstrating that advanced LLMs can achieve near-expert performance in certain medical domains. Gemini and DeepSeek, while also performing well, showed slightly lower accuracy rates, potentially due to differences in training methodologies and dataset composition (2).

While our study focused on the field of PMR, the utility of LLMs in other medical specialties has also been explored, with promising results in rheumatology (10). Recent studies have specifically evaluated the performance of various LLMs on rheumatology board-level questions, finding that they can achieve high accuracy. A comparative study using questions from the American College of Rheumatology's CARE-2022 Question Bank found that GPT-4 demonstrated a 78% accuracy rate, outperforming Claude 3: Opus (63%) and Gemini Advanced (53%) (11). Another study, which assessed the performance of LLMs on the Spanish access exam to specialized medical training (MIR), reported that GPT-4 achieved a remarkable accuracy of 93.71% on rheumatology questions, significantly higher than ChatGPT's 66.43% (12).

Similar to its utility in other medical specialties, AI and LLMs show significant promise for orthopedic study and board exam preparation. Recent studies have assessed the performance of various LLMs, including GPT-4 and Google Gemini, on standardized tests like the orthopaedic in-training examination (OITE) and the Turkish orthopedics and traumatology board examination (13,14). One study found that GPT-4 performed at the level of a third-year resident on the 2021 OITE (15). Another study using the 2022 OITE found that Google Gemini was the most accurate model, correctly answering 69.9% of questions, a performance level approaching that of fourth- and fifth-year residents (16).

While LLMs generally perform well on text-based, knowledge-recall questions, their accuracy can be significantly lower on questions that include images (17). Moreover, a key limitation highlighted by these studies is the tendency for LLMs to provide inaccurate in response to complex or fact-based questions (14). This suggests that while these tools are becoming valuable for their ability to provide explanations and enhance learning, they are not yet a substitute for human clinical judgment and must be used with caution and careful verification of their output.

LLMs are also showing considerable potential in neurology, with studies indicating their effectiveness in answering board-style questions and assisting with clinical reasoning (18,19). A study evaluating several LLMs on questions from the self-assessment in neurological surgery American Board of Neurological Surgery Primary Board Examination Review found that all models exceeded the passing threshold. The highest accuracy was achieved by OpenAI o1 (87.6%), followed by Claude 3.5 Sonnet (83.2%) and Gemini 2.0 (81.0%) (20). Another study, which specifically evaluated GPT-4 on neurology board-style questions, reported an accuracy rate of 75.0%, outperforming the average human test-taker score of 69% and the passing score of 70%. This study also highlighted that GPT-4 performed particularly well in subspecialties like neuromuscular disorders, pharmacology, and cognitive and behavioral disorders (21). While these results are promising for medical education and exam preparation, the studies also underscore limitations, such as lower performance on questions involving images and a need for continued physician supervision to ensure accuracy and reliability.

According to our results, DeepSeek excelled in musculoskeletal medicine and patient management, ChatGPT-4o led in neurologic disorders and electrodiagnosis, and Gemini performed strongly in equipment/assistive technology and medical rehabilitation (100%). All models achieved parity in rehabilitation problems and applied sciences. These findings suggest that while DeepSeek holds an aggregate advantage, task-specific expertise varies across models.

Our observed higher accuracy (e.g., DeepSeek: 88%, ChatGPT-4o: 86%) compared to prior studies in specialties like rheumatology (GPT-4: 78%) or orthopedics (Gemini: 69.9%) likely reflects rapid advancements in LLM capabilities rather than methodological differences alone. Key drivers include iterative model evolution (e.g., optimizations from GPT-4 to GPT-4o), architectural refinements improving clinical reasoning, and potential task-specific fine-tuning enhancing performance on medical benchmarks. This trajectory of technical progress suggests LLMs are steadily narrowing the accuracy gap with human expertise across medical domains, though persistent limitations in handling novel scenarios warrant ongoing validation.

Study Limitations

Despite these promising results, our study also underscores several limitations in the current generation of LLMs. Notably, none of the models achieved perfect accuracy, indicating that they are still prone to errors in medical reasoning and knowledge retrieval. Previous studies have also pointed out that LLMs can occasionally generate incorrect or misleading information, particularly when dealing with complex clinical scenarios (4). The ability of these models to handle complex and specialized knowledge makes them valuable tools for clinicians and trainees, potentially serving as a supplementary aid for exam preparation and continuing medical education (22).

Another key limitation is the evolving nature of AI models. As LLMs continue to be updated and refined, their performance on medical assessments may improve, necessitating ongoing evaluation and benchmarking. Moreover, while our study utilized publicly available ABPMR-style questions, it is possible that these questions do not fully capture the depth and breadth of the actual board examination. Further studies incorporating a larger and more diverse set of questions, as well as real-world clinical case evaluations, would provide a more comprehensive understanding of these models' capabilities.

Understanding the capabilities and limitations of these LLMs in this context is crucial for determining their potential role in medical education and assessment, while also highlighting areas where human oversight and expertise remain essential. We acknowledge potential limitations, including the evolving nature of AI models and the possibility that the questions used may not fully represent the current ABPMR examination. Additionally, we recognize that AI models' performance may not directly translate to clinical competence or decision-making ability.

Conclusion

This study highlights the growing potential of LLMs such as Gemini, DeepSeek, and ChatGPT-4 in medical education and assessment. While these models demonstrate impressive accuracy in answering board-style questions, their limitations emphasize the need for human oversight and further refinements to ensure reliability in clinical decision-making contexts. Future research should focus on improving LLM interpretability, optimizing training datasets, and developing hybrid AI-human systems to enhance the effectiveness of AI-assisted medical education.

Ethics

Ethics Committee Approval: This study did not require ethical approval as it was based on publicly available data.

Informed Consent: Informed consent was not required for this study as it did not involve human participants or private medical information.

Footnotes

Authorship Contributions

Concept: G.S.A., Design: G.S.A., Data Collection or Processing: G.S.A, H.B.Ş., Analysis or Interpretation: G.S.A, H.B.Ş., Literature Search: G.S.A., Writing: G.S.A.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol. 2017;2:230-43.
2. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542:115-8. Erratum in: Nature. 2017;546:686.
3. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. Advances in Neural Information Processing Systems. 2017;30:1-11.
4. Topol E. Deep medicine: how artificial intelligence can make healthcare human again. New York: Basic Books; 2019.
5. Singhal K, Azizi S, Tu T, Mahdavi SS, Wei J, Chung HW, et al. Large language models encode clinical knowledge. Nature. 2023;620:172-80. Erratum in: Nature. 2023;620:E19.
6. Aydin S, Karabacak M, Vlachos V, Margetis K. Large language models in patient education: a scoping review of applications in medicine. Front Med (Lausanne). 2024;11:1477898.
7. American Board of Physical Medicine and Rehabilitation. About the ABPMR. Available from: https://www.abpmr.org/About
8. Lee P, Bubeck S, Petro J. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. N Engl J Med. 2023;388:1233-9.
9. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare (Basel). 2023;11:887.
10. Is EE, Menekseoglu AK. Comparative performance of artificial intelligence models in rheumatology board-level questions: evaluating Google Gemini and ChatGPT-4o. Clin Rheumatol. 2024;43:3507-13.
11. Flores-Gouyonnet J, Cuéllar-Gutiérrez MC, Figueroa-Parra G, Kimbrough B, Joerns EK, Navarro-Mendoza E, et al. Performance of large language models in rheumatology board-like questions: accuracy, quality, and safety. Lancet Rheumatol. 2025;7:e152-4.
12. Madrid-García A, Rosales-Rosado Z, Freites-Nuñez D, Pérez-Sancristóbal I, Pato-Cour E, Plasencia-Rodríguez C, et al. Harnessing ChatGPT and GPT-4 for evaluating the rheumatology questions of the Spanish access exam to specialized medical training. Sci Rep. 2023;13:22129.
13. Ghanem D, Covarrubias O, Raad M, LaPorte D, Shafiq B. ChatGPT Performs at the level of a third-year orthopaedic surgery resident on the orthopaedic in-training examination. JB JS Open Access. 2023;8:e23.00103.
14. Yaş S, Ahmadov A, Baymurat AC, Tokgöz MA, Coşkun Yaş S, Odluyurt M, et al. ChatGPT vs. orthopedic residents! Who is the winner? GMJ. 2024;35:185-91.
15. Kung JE, Marshall C, Gauthier C, Gonzalez TA, Jackson JB. Evaluating ChatGPT performance on the orthopaedic in-training examination. JB JS Open Access. 2023;8:e23.00056.
16. Pamuk Ç, Uyanık AF, Kuyucu E, Uğurlar M. Can ChatGPT pass the Turkish Orthopedics and Traumatology Board Examination? Turkish orthopedic surgeons versus artificial intelligence. Ulus Travma Acil Cerrahi Derg. 2025;31:310-5.
17. Nawari A, Zahir J, Kumar S, Ocampo L, Opara O, Ahmad H, et al. Artificial intelligence large language models are nearly equivalent to fourth-year orthopaedic residents on the orthopaedic in-training examination: a cause for concern or excitement? J Orthop Educ Inst. 2025;6.
18. Schubert MC, Wick W, Venkataramani V. Performance of large language models on a neurology board-style examination. JAMA Netw Open. 2023;6:e2346721. Erratum in: JAMA Netw Open. 2024;7:e240194.
19. Romano MF, Shih LC, Paschalidis IC, Au R, Kolachalama VB. Large language models in neurology research and future practice. Neurology. 2023;101:1058-67.
20. Andrade NS, Donty S. Comparison of large language models' performance on neurosurgical board examination questions. medRxiv. 2025. [Preprint].
21. Shu E, Sharma A, Nanda P, Kothari S, Wang T. Large language model performance in neurology board questions (S33.001). Neurology. 2024;102(Suppl 17):S33.001.
22. Tarabanis C, Zahid S, Mamalis M, Zhang K, Kalampokis E, Jankelson L. Performance of publicly available large language models on internal medicine board-style questions. PLOS Digit Health. 2024;3:e0000604.

DOI: 10.4274/tod.galenos.2025.75010
Turk J Osteoporos 2025;31(3):195-200

Combined Surgical and Concentrated Growth Factor Therapy in the Management of Stage II and III MRONJ Associated with Denosumab: A Report of Two Cases

Denosumab ile İlişkili Evre II ve III MRONJ'un Cerrahi ile Birlikte Konsantre Büyüme Faktörü Uygulamasıyla Tedavisi: İki Olgu Sunumu

Mert Kırdemir, Muharrem Ergün Dudak, Burak Borlu

Ege University Faculty of Dentistry, Department of Oral and Maxillofacial Surgery, İzmir, Türkiye

Abstract

Medication-related osteonecrosis of the jaw (MRONJ) is a significant complication associated with antiresorptive agents such as denosumab, particularly in oncology patients. This report describes two patients with prostate cancer who developed MRONJ during denosumab treatment. One patient exhibited stage 3 MRONJ in the mandible, while the other had stage 3 MRONJ in the maxilla. Following clinical and radiological evaluations, both patients underwent surgical intervention, and concentrated growth factor (CGF) was applied locally to promote postoperative healing. In the first case, mandibular sequestrectomy was performed with reconstruction using a pre-shaped plate. In the second case, premaxillary resection was carried out after infection control. Both patients experienced favorable outcomes with successful, complication-free wound healing. These cases highlight the potential benefits of CGF-assisted surgical treatment in denosumab-related MRONJ and underscore the importance of individualized treatment planning in the absence of standardized protocols.

Keywords: Osteonecrosis, denosumab, MRONJ, concentrated growth factor

Öz

İlaç ilişkili çene osteonekrozu (MRONJ), özellikle onkoloji hastalarında denosumab gibi anti-rezorptif ajanlarla ilişkili önemli bir komplikasyondur. Bu raporda, denosumab tedavisi sırasında MRONJ gelişen prostat kanseri tanısı olan iki hasta sunulmaktadır. Bir hastada alt çenede evre 3 MRONJ gözlenirken, diğerinde üst çenede evre 3 MRONJ mevcuttu. Klinik ve radyolojik değerlendirmelerin ardından her iki hasta da opere edildi ve ameliyat sonrası iyileşmeyi desteklemek amacıyla lokal olarak konsantre büyüme faktörü (CGF) uygulandı. İlk olguda mandibular sequestrektomi yapılarak önceden şekillendirilmiş plak ile rekonstrüksiyon gerçekleştirildi. İkinci olguda ise enfeksiyon kontrolü sonrası premaksiller rezeksiyon yapıldı. Her iki hastada komplikasyonsuz başarılı yara iyileşmesi ile olumlu sonuçlar elde edildi. Bu olgular, denosumab ilişkili MRONJ'de CGF destekli cerrahi tedavinin potansiyel faydalarını vurgulamakta ve standart bir protokol olmaması durumunda bireyselleştirilmiş tedavi planlamasının önemini ortaya koymaktadır.

Anahtar kelimeler: Osteonekroz, denosumab, MRONJ, konsantre büyüme faktörü

Introduction

Medication-induced osteonecrosis of the jaw was first described by Marx (1). After it was established that bisphosphonates and other antineoplastic and antiangiogenic agents could cause maxillofacial osteonecrosis, the American Association of Oral and Maxillofacial Surgeons (AAOMS) named this condition "medication-related osteonecrosis of the jaw" (MRONJ) in 2014 (2).

For individuals taking antiresorptive or antiangiogenic drugs and with no history of radiotherapy, open bone for more than eight weeks or lesions with fistulas suggest MRONJ (2). The most common risk factor is dentoalveolar surgery, with the majority of cases occurring after tooth extraction (3). Other risk factors include implant operations, oral infections, poor hygiene, and ill-fitting dentures (2).

Treatment options range from conservative methods to extensive surgery. Conservative options include maintaining oral hygiene,

Corresponding Author/Sorumlu Yazar: Asst. Mert Kırdemir, Ege University Faculty of Dentistry, Department of Oral and Maxillofacial Surgery, İzmir, Türkiye
E-mail: mertkrdemir@gmail.com **ORCID ID:** orcid.org/0000-0003-0685-9062
Received/Geliş Tarihi: 06.06.2025 **Accepted/Kabul Tarihi:** 26.06.2025 **Epub:** 13.08.2025 **Publication Date/Yayınlanma Tarihi:** 05.12.2025

Cite this article as/Atıf: Kırdemir M, Dudak ME, Borlu B. Combined surgical and concentrated growth factor therapy in the management of stage II and III MRONJ associated with denosumab: a report of two cases. Turk J Osteoporos. 2025;31(3):195-200

antibiotic therapy and minimally invasive sequestrectomy (4). Surgical excision and curettage are also commonly used methods (5). Additionally, the regenerative effects of locally applied platelet concentrates have been reported in various studies to support tissue regeneration. Concentrated growth factor (CGF) is a leukocyte-rich platelet concentrate that supports angiogenesis and tissue regeneration, offering potential benefits in immunocompromised individuals (6-10).

These case reports describe the surgical treatment of extensive osteonecrosis of the maxilla and mandible in two patients who were receiving denosumab for prostate cancer alongside CGF. Although our cases are related to oncologic patients, denosumab is also widely used in osteoporotic patients and may similarly lead to osteonecrosis of the jaws. What distinguishes this report is the combined use of surgical intervention and CGF in stage 2 and 3 MRONJ cases associated with denosumab in patients with prostate cancer. In this respect, our study may contribute to the management of MRONJ in osteoporotic patients.

In this case series, all procedures involving human subjects were conducted in accordance with our institution's ethical standards and the Helsinki Declaration. All patient identifiers were meticulously safeguarded, and a comprehensive informed consent process was duly implemented.

Case Reports

Case 1

A 78-year-old male patient presented at our clinic, complaining of pain and exposed bone in right mandible (Figure 1). Radiological examination revealed radiolucent areas with irregular borders in the parasymphyseal region of the mandible (Figure 2). The patient's medical history indicated a diagnosis of prostate cancer, for which he had been receiving monthly 120 mg denosumab injections for a total of 8 months. There was no history of systemic disease, metastasis, or other oncologic treatments. Denosumab therapy had been discontinued three months prior to surgery, based on consultation with the oncology department. Preoperative cone beam computed tomography (CBCT) showed necrotic bone margins. Prior to the planned sequestrectomy, during which MRONJ was initially suspected, a 3D-printed model of the mandible was prepared to assist surgical planning. To reinforce mandibular integrity and prevent intra-operative or post-operative fracture, a pre-shaped reconstruction plate was applied. Surgical debridement was performed under local anesthesia. Following this, CGF was prepared by centrifuging the patient's own venous blood using a specialized device (Medifuge®, Silfradent, Italy). For this procedure, six tubes of

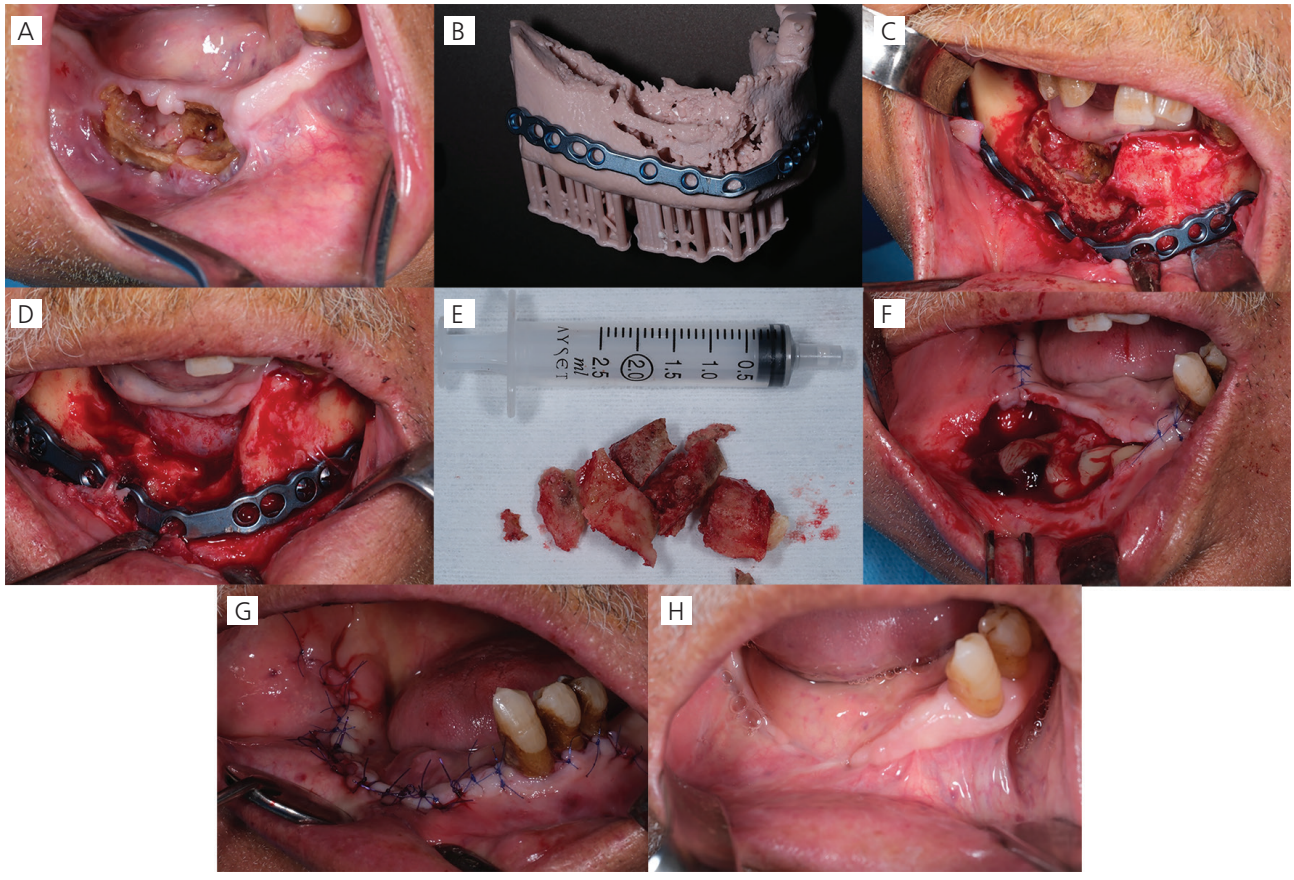

Figure 1. Clinical presentation of Case 1. **A)** Preoperative intraoral view, **B)** Pre-shaped reconstruction plate on 3D model, **C)** Intraoral view reconstruction plate and necrotic bone, mental nerve (white arrow), **D)** After sequestrectomy, mental nerve (white arrow), **E)** Necrotic bone fragments, **F)** Applied concentrated growth factor, **G)** Primary closure, **H)** Healing after 3 month

blood were collected. The obtained CGF was processed into membrane form and placed directly into the surgical site prior to primary closure. After adequate flap mobilization, the wound was closed primarily. The healing process was uneventful, with satisfactory soft tissue recovery observed at 8 weeks.

Case 2

A 64-year-old male patient attended our clinic, complaining of pain in the maxilla and halitosis following a tooth extraction at an external centre. On examination, non-healing tissue and purulent discharge were present at the extraction site. The patient had been receiving monthly 120 mg denosumab for six months due to prostate cancer. As in the first case, there was no history of systemic illness, metastasis, or other cancer therapies. Following oncology consultation, denosumab was discontinued, and a three-month drug holiday was observed prior to surgery. A pre-diagnosis of stage 2 MRONJ was made and amoxicillin and chlorhexidine mouthwash were prescribed to control the infection (Figure 3). Following oncology consultation, the patient's denosumab therapy was discontinued. A CBCT scan, performed 2 months after discontinuation of denosumab, revealed a broad demarcation line involving the nasal floor and the maxillary sinus (Figure 4). With a diagnosis of stage 3 MRONJ, premaxillary resection was performed under local anesthesia. Followed by CGF application using the same protocol as in Case 1. Six tubes of venous blood were collected, and the resulting CGF was processed into membrane form and applied to the surgical site before primary closure. After an uneventful eight-week healing period, a removable prosthesis was delivered to the patient.

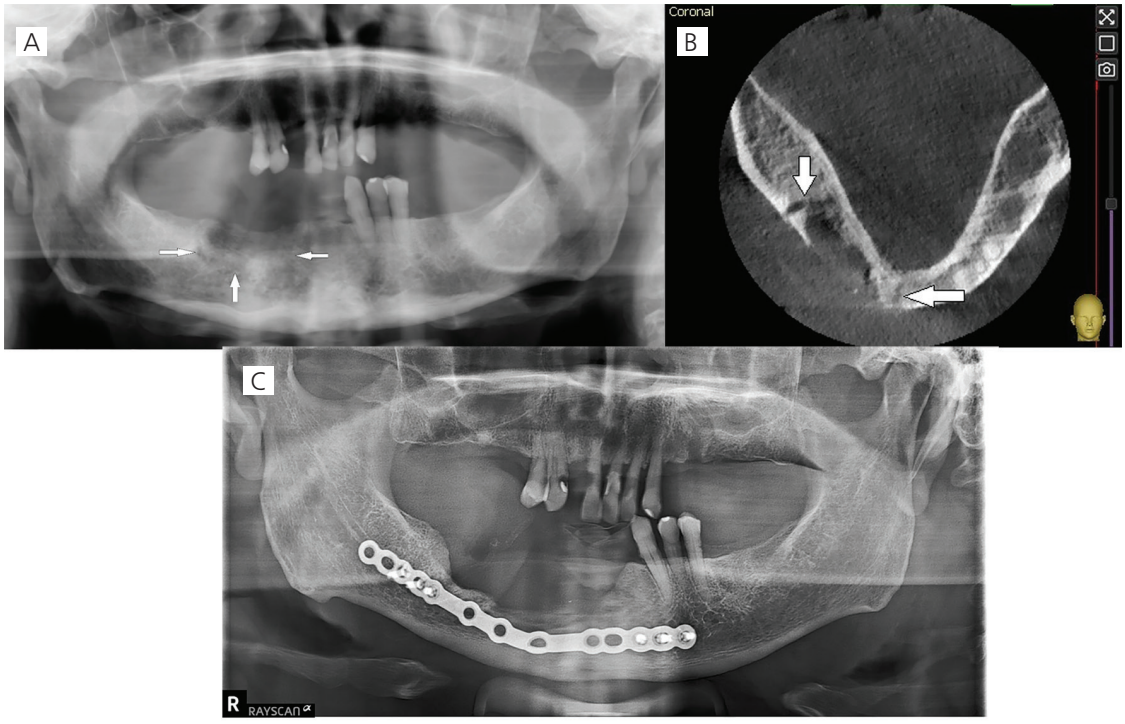

Figure 2. Radiological presentation of Case 1, **A)** Preoperative orthopantomography, **B)** Dental volumetric tomography axial seciton, **C)** Postoperative orthopantomography

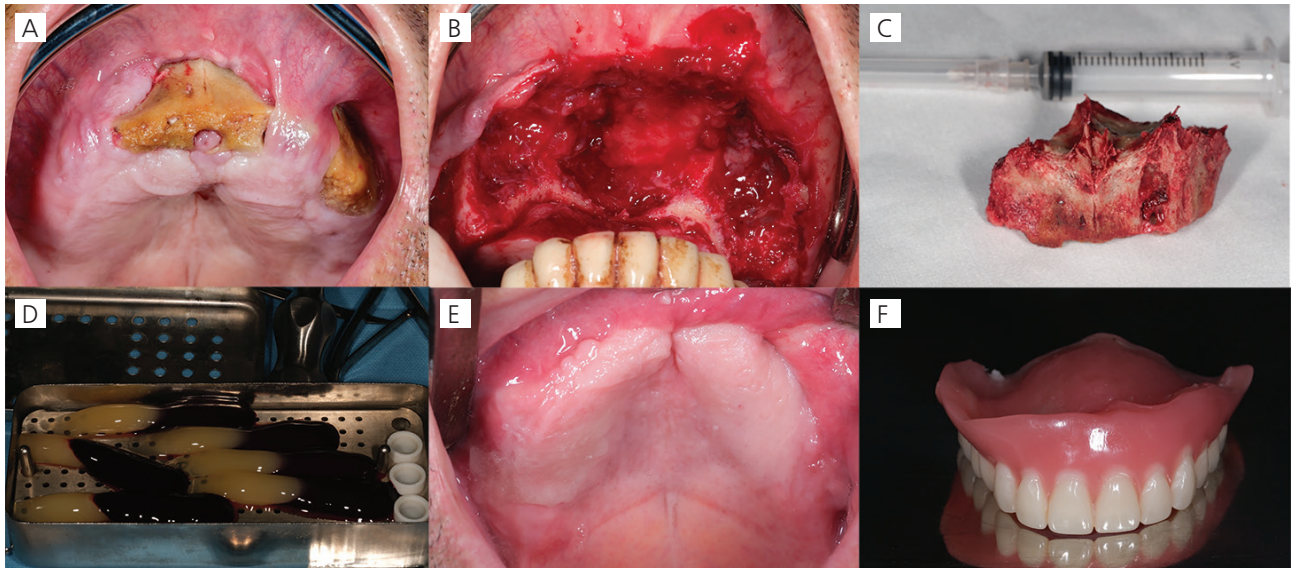

Figure 3. Clinical presentation of Case 2, **A)** Preoperative intraoral view, **B)** Intraoral view after lambo removal, **C)** Necrotic bone fragment, **D)** Concentrated growth factor, **E)** Healing after 8 weeks, **F)** Final prosthesis

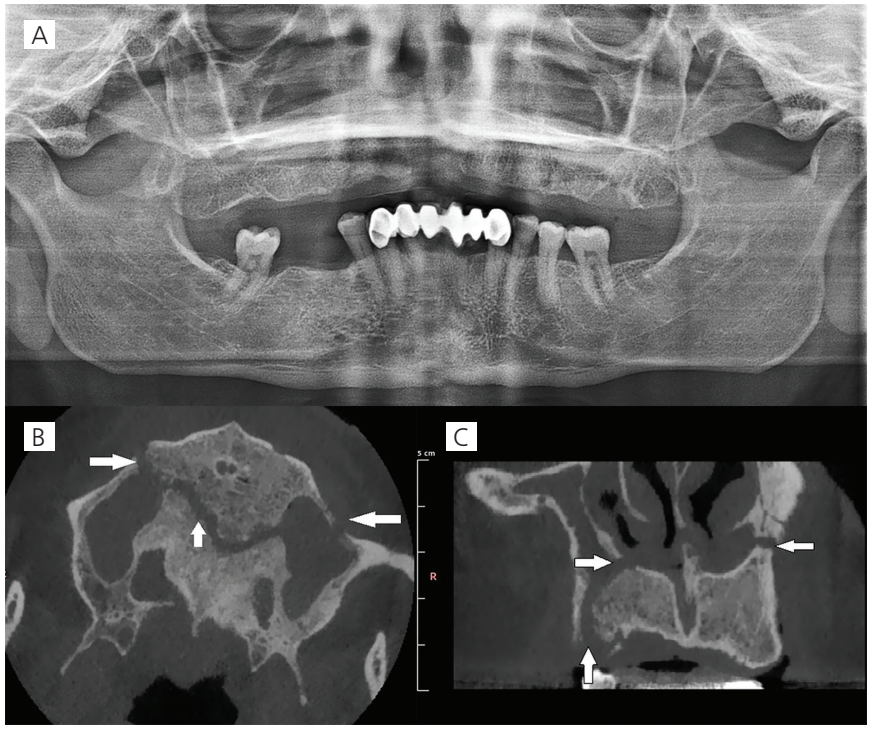

Figure 4. Radiological presentation of Case 2, **A)** Preoperative orthopantomography, **B)** DVT axial section, **C)** DVT sagittal section
DVT: Dental volumetric tomography

Discussion

MRONJ is a serious complication associated with the use of antiresorptive and antiangiogenic drugs. MRONJ is frequently observed in elderly female patients, particularly those using bisphosphonates or denosumab for osteoporosis or cancer (2). MRONJ is relatively uncommon in patients taking antiresorptive drugs for osteoporosis, with reported incidence rates ranging from 0.1% to 0.01%. An incidence rate of 0.052% has been reported in patients using denosumab. IV bisphosphonates used in oncology patients have been reported to cause MRONJ at a higher rate than denosumab (2). Another study reported a prevalence of MRONJ ranging from 0% to 27.5% after bisphosphonate use, whereas the risk of developing denosumab-induced MRONJ was up to 2% (7). Although denosumab is often considered to have fewer adverse effects, some recent studies have reported that it may pose a comparable or even higher risk of MRONJ than zoledronic acid (5).

While dentoalveolar surgery has been identified as the main risk factor, some studies have found that the impact of minor surgery is low. Chemotherapy, corticosteroid use and smoking are also risk factors (3). For individuals at risk of MRONJ, a pre-treatment clinical and radiological evaluation, elimination of infection, good oral hygiene and regular dental check-ups are recommended. This approach has been shown to significantly reduce the risk of MRONJ (7).

The AAOMS recommends interrupting drug treatment three months before and after invasive procedures, but this decision should be made jointly with the prescribing physician. In 2012, both the Italian Society of Maxillofacial Surgery and the Italian Society of Oral Pathology and Medicine recommended conservative surgery for early-stage lesions, as defined by the two associations. Surgical excision is recommended for advanced-stage lesions. Treatment protocols remain heterogeneous, and a universally accepted approach has yet to be established (5). While surgical treatment is generally considered more effective, it is not without limitations, as some patients may experience recurrence requiring additional interventions (8). In addition to surgery, other methods have been proposed, including low-dose laser therapy, fluorescent staining, ozone therapy, hyperbaric oxygen therapy, pentoxifylline and alpha-tocopherol (9).

Plasma-rich fibrin (PRF) stimulates the migration and proliferation of fibroblasts and osteoblasts, while L-PRF stimulates angiogenesis (7,8). CGF enables the formation of a denser fibrin matrix with a high concentration of growth factors (10). Studies have reported that the local application of CGF improves tissue regeneration in the surgical treatment of MRONJ (6).

Current literature suggests that CGF application may be beneficial in osteoporotic patients who develop MRONJ. In the treatment of early-stage MRONJ in osteoporotic individuals, the use of CGF in combination with conservative surgery has resulted in significant improvements in wound healing and bone regeneration (6). Borsani et al. (11) reported that CGF accelerates postoperative healing, while Yüce et al. (6) found that it enhances tissue integrity following sequestrectomy. However, these studies were mostly conducted in osteoporotic patients or early-stage MRONJ cases. In contrast, our study differs in that it demonstrates the supportive effect of CGF in the surgical treatment of stage 2 and 3 MRONJ in prostate cancer patients receiving denosumab. In this respect, our report represents one of the rare examples highlighting the effectiveness of CGF in advanced oncologic MRONJ cases.

Conclusion

These case reports demonstrate that stage 2 and stage 3 MRONJ cases resulting from denosumab use for cancer treatment can be managed surgically, and that the application of CGF may support healing. Although there is no standardized approach for the treatment of advanced-stage MRONJ, individualized, case-specific planning may enhance clinical success. Further advanced clinical and scientific studies are needed to better understand the pathophysiology of MRONJ and to determine the most effective treatment strategies.

Ethics

Informed Consent: All patient identifiers were meticulously safeguarded, and a comprehensive informed consent process was duly implemented.

Footnotes

Authorship Contributions

Surgical and Medical Practices: M.K., M.E.D., B.B., Concept: M.K., M.E.D., Design: M.K., M.E.D., Data Collection or Processing: M.K., M.E.D., B.B., Analysis or Interpretation: M.K., M.E.D., B.B., Literature Search: M.K., M.E.D., B.B., Writing: M.K., M.E.D., B.B.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Marx RE. Pamidronate (Aredia) and zoledronate (Zometa) induced avascular necrosis of the jaws: a growing epidemic. J Oral Maxillofac Surg. 2003;61:1115-7.
2. Ruggiero SL, Dodson TB, Fantasia J, Goodday R, Aghaloo T, Mehrotra B, et al. American Association of Oral and maxillofacial surgeons position paper on medication-related osteonecrosis of the jaw–2014 update. J Oral Maxillofac Surg. 2014;72:1938-56.
3. Soutome S, Otsuru M, Hayashida S, Murata M, Yanamoto S, Sawada S, et al. Relationship between tooth extraction and development of medication-related osteonecrosis of the jaw in cancer patients. Sci Rep. 2021;11:17226.
4. Torul D, Bereket MC. Bisphosphonate induced osteonecrosis of the jaws and current therapies. Cumhuriyet Dental Journal. 2017;20:122-31.
5. Kün-Darbois JD, Fauvel F. Medication-related osteonecrosis and osteoradionecrosis of the jaws: update and current management. Morphologie. 2021;105:170-87.
6. Yüce MO, Adalı E, Işık G. The effect of concentrated growth factor (CGF) in the surgical treatment of medication-related osteonecrosis of the jaw (MRONJ) in osteoporosis patients: a randomized controlled study. Clin Oral Investig. 2021;25:4529-41.

7. Nica DF, Riviş M, Roi CI, Todea CD, Duma V-F, Sinescu C. Complementarity of photo-biomodulation, surgical treatment, and antibiotherapy for medication-related osteonecrosis of the jaws (MRONJ). Medicina (Kaunas). 2021;57:145.
8. Tenore G, Zimbalatti A, Rocchetti F, Graniero F, Gaglioti D, Mohsen A, et al. Management of medication-related osteonecrosis of the jaw (MRONJ) using leukocyte- and platelet-rich fibrin (L-PRF) and photobiomodulation: a retrospective study. J Clin Med. 2020;9:3505.
9. Steel BJ. Management of medication-related osteonecrosis of the jaw (MRONJ) risk in patients due to commence anti-resorptive/anti-angiogenic drugs-how should pre-drug-treatment dental preventive care be organised? Community Dent Health. 2019;36:244-54.
10. Dudak ME, Calis A, Koca H. Effectiveness of concentrated growth factor on improving postoperative comfort in patients undergoing coronectomy for impacted lower third molars: a double-blinded split-mouth randomized controlled clinical study. J Oral Maxillofac Surg. 2025;83:345-52.
11. Borsani E, Bonazza V, Buffoli B, Nocini PF, Albanese M, Zotti F, et al. Beneficial effects of concentrated growth factors and resveratrol on human osteoblasts *in vitro* treated with bisphosphonates. Biomed Res Int. 2018;2018:4597321.

DOI: 10.4274/tod.galenos.2025.25986
Turk J Osteoporos 2025;31(3):201-4

Osteopoikilosis as an Uncommon Cause of Extremity Pain: A Case Report

Ekstremite Ağrısının Nadir Bir Nedeni Olarak Osteopoikilozis: Olgu Sunumu

Emre Fidan[1], Adil Can Karaoğlu[2], Ahmet Özdemir[2]

[1]Balıkesir University Faculty of Medicine, Balıkesir, Türkiye
[2]Balıkesir University Faculty of Medicine, Department of Neurosurgery, Balıkesir, Türkiye

Abstract

Osteopoikilosis, also known as spotted bone disease, is a rare, benign, and inherited sclerosing bone dysplasia characterized by multiple small, well-defined sclerotic foci, predominantly involving the appendicular skeleton. While the etiology remains unclear, it is believed to result from inherited defects in trabecular bone formation along stress lines. Lesions occur symmetrically around the knee joints, pelvic bones, tarsal bones, shoulder girdle, and carpal bones, while involvement of the skull and spine is considered extremely rare. The disease follows an autosomal dominant inheritance pattern and can often be confused with bone dysplasias, bone metastases and other sclerotic diseases on radiological examinations. The diagnosis is usually made incidentally on radiologic examinations. We report the case of a 40-year-old man who presented to our clinic after trauma, with pain in both wrists and the left knee. Radiologic imaging revealed widespread sclerotic lesions, notably involving the cervical, thoracic, lumbar, and sacral vertebrae-an uncommon pattern for osteopoikilosis. In addition to classical appendicular lesions, the extensive axial skeleton involvement distinguished this case from typical presentations.
Keywords: Osteopoikilosis, bone dsyplasia, radiographic sclerosis

Öz

Osteopoikilozis, benekli kemik hastalığı olarak da bilinen, nadir, benign ve kalıtsal bir sklerozan kemik displazisidir. Bu hastalık, çoğunlukla apendiküler iskeleti tutan, çok sayıda küçük, iyi sınırlı sklerotik odaklarla karakterizedir. Etiyolojisi tam olarak aydınlatılamamış olmakla birlikte, stres hatları boyunca trabeküler kemik oluşumundaki kalıtsal defektlerden kaynaklandığı düşünülmektedir. Lezyonlar simetrik olarak diz eklemleri, pelvis kemikleri, tarsal kemikler, omuz kuşağı ve karpal kemikler çevresinde görülürken, kafa ve omurga tutulumunun son derece nadir olduğu kabul edilmektedir. Hastalık, otozomal dominant bir kalıtım modeli izler ve radyolojik incelemelerde sıklıkla kemik displazileri, kemik metastazları ve diğer sklerotik hastalıklarla karıştırılabilir. Tanı genellikle radyolojik tetkiklerde tesadüfen konur. Bu olgu sunumunda, travma sonrası her iki el bileğinde ve sol dizinde ağrı şikayetiyle kliniğimize başvuran 40 yaşında bir erkek hastayı bildirmekteyiz. Radyolojik görüntülemede, servikal, torakal, lomber ve sakral vertebraları da içeren yaygın sklerotik lezyonlar saptanmış olup, bu durum osteopoikilozis için alışılmadık bir tutulum paternidir. Klasik apendiküler lezyonlara ek olarak, aksiyel iskeletteki bu geniş tutulum, olgumuzu tipik osteopoikilozis sunumlarından ayıran en belirgin özelliktir.
Anahtar kelimeler: Osteopoikliosis, kemik displazisi, radyolojik sklerozis

Introduction

Osteopoikilosis (OPK) (OMIM #166700) is an exceptionally uncommon hereditary bone disorder first described by Albers Schönberg in 1915 (1,2). OPK is an osteochondrodysplasia characterized by an increase in bone density without any accompanying changes in bone structure. OPK often does not cause any symptoms; however, in rare cases, it may manifest with dermatological signs. Dermatological manifestations may appear as small, yellowish papules or subcutaneous nodules, which can develop into keloid-like lesions resembling scleroderma. These sores tend to get keloid formations that look like scleroderma. Occasionally, these symptoms may cause pain and effusion in larger joints, although they do not result in any deformities (3). Since most patients don't show symptoms, we can establish the diagnosis using an imaging method for an unrelated cause (4). Radiological imaging reveals the presence of several uniformly distributed sclerotic lesions with distinct boundaries and higher

Corresponding Author/Sorumlu Yazar: Emre Fidan MD, Balıkesir University Faculty of Medicine, Balıkesir, Türkiye
E-mail: emreefidan12@gmail.com **ORCID ID:** orcid.org/0000-0002-7871-771X
Received/Geliş Tarihi: 22.02.2025 **Accepted/Kabul Tarihi:** 21.05.2025 **Epub:** 14.08.2025 **Publication Date/Yayınlanma Tarihi:** 05.12.2025

Cite this article as/Atıf: Fidan E, Karaoğlu AC, Özdemir A. Osteopoikilosis as an uncommon cause of extremity pain: a case report. Turk J Osteoporos. 2025;31(3):201-4

radiodensity (5). The main objective of diagnosis is to exclude OPK as a potential cause, particularly when considering more severe cases such as osteoblastic metastases (6).

Case Report

A 40-year-old man was admitted to our clinic with complaints of pain in both wrists and the left knee after a fall two days ago. He stated that the pain was aggravated by movement and partially relieved by rest and that there was no increase in his complaints at night. No systemic symptoms such as fever, weight loss, or night sweats were reported by the patient. Physical examination revealed tenderness on palpation in both wrists and metacarpophalangeal joints, especially in the left one. Swelling was noted in the left wrist with a positive distal radioulnar joint (DRUJ) test. Range of motion at the wrists and knees was slightly limited due to pain, but no deformity or muscle weakness was noted.

Imaging of the patient showed a cervical spine radiograph showing multiple osteosclerotic lesions involving vertebral bodies and spinous processes, consistent with OPK (Figure 1A). Thoracic and lumbar spine radiograph showing multiple sclerotic foci in the vertebral bodies, indicating axial skeletal involvement (Figure 1B), radiographs of the left hand and wrist showing sclerotic lesions in the carpal bones, metacarpals, and phalanges (Figure 1C). Bilateral knee radiographs show diffuse osteosclerotic areas in the bones surrounding both knee joints, typical of OPK (Figure 1D). A radiograph of the lumbar spine shows prominent sclerotic lesions in the vertebral bodies, more prominent in the thoracic region (Figure 1E). The patient's clinical history, physical examination, and radiologic imaging studies were evaluated together, and a diagnosis of OPK, which is not one of the musculoskeletal diseases we frequently encounter in clinical practice, was made. No surgical procedure was considered; painkillers were prescribed, and rest was recommended.

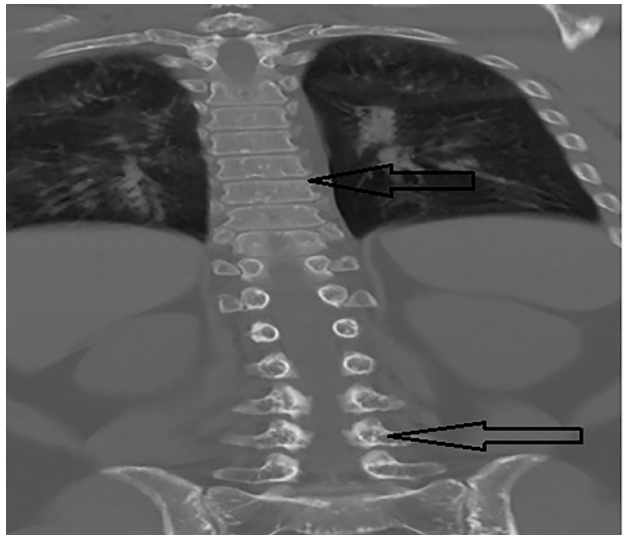

Figure 1B. Thoracic and lumbar spine radiograph showing multiple sclerotic foci in the vertebral bodies, indicating axial skeleton involvement

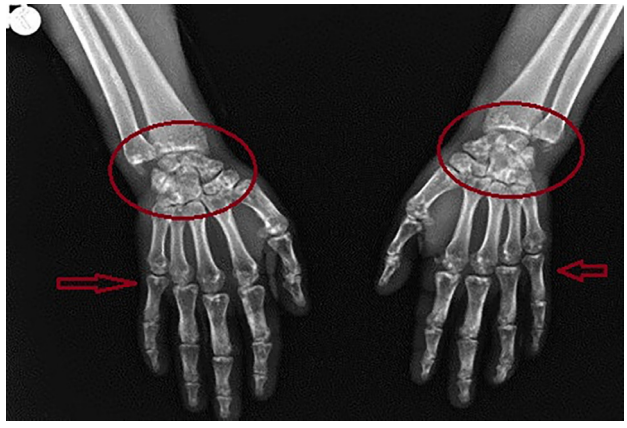

Figure 1C. Radiograph of the left hand and wrist revealing sclerotic lesions in the carpal bones, metacarpals, and phalanges. A positive distal radioulnar joint test was noted on the left side

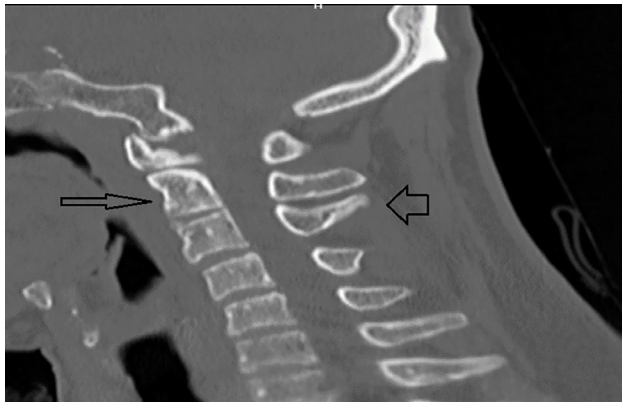

Figure 1A. Cervical spine radiograph demonstrating multiple osteosclerotic lesions involving the vertebral bodies and spinous processes, consistent with osteopoikilosis

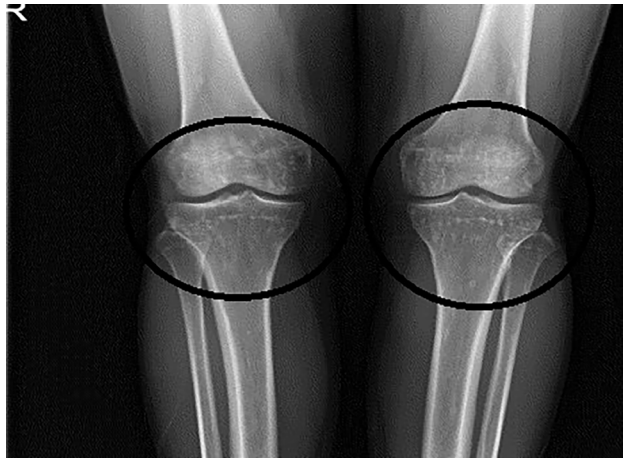

Figure 1D. Bilateral knee radiographs showing diffuse osteosclerotic areas in the bones surrounding both knee joints, typical of osteopoikilosis

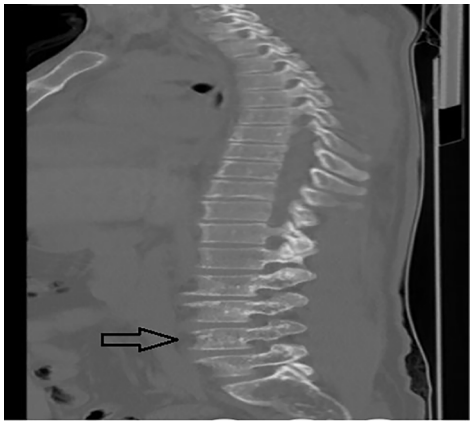

Figure 1E. Lumbar spine radiograph with prominent sclerotic lesions in the vertebral bodies, more pronounced than in the thoracic region

Discussion

When someone has OPK, they will have dense radiolucent spots that are spread out evenly across the epiphyses and metaphyses of long bones. It has a prevalence of roughly 1 in 50,000. Patients typically receive a diagnosis of this condition between the ages of 15 and 60. The lesions are characterized by several circular or oval-shaped patches that are 1-10 mm in diameter and have a hardened texture. These areas can appear in anyone, regardless of age or gender (3,7). Osteopoikilosis is characterized by the presence of a concentrated area of bone that resembles cortical bone and contains Haversian canals. The spongiosa, located directly below the cortex, typically displays this area (1). A lot of the time, osteopoikilosis, osteopathia striata, and melorheostosis happen together. This makes it possible that they are all different symptoms of the same disorder, which is called mixed sclerosing bone dysplasia (8). Researchers have demonstrated that loss-of-function mutations in the *LEM domain containing 3* (*LEMD3*) gene cause conditions characterized by elevated bone density, such as osteopoikilosis, Buschke-Ollendorff syndrome, and melorheostosis (9). A mutation in the *LEMD3* gene genetically links osteopoikilosis (3). While the inheritance of this condition is well-established, there have also been rare instances recorded. Plain radiography and computed tomography scans typically show bone islands near the joints in osteopoikilosis. The orientation of these islands aligns with the surrounding trabecula. Visualization on magnetic resonance imaging. The lesions appear as tiny, dark areas on both T1- and T2-weighted imaging due to their composition of mature, thick bone (4,8). While OPK is a rare condition, it is important to consider other disorders such as tuberous sclerosis, osteopathy striata, osteoblastic metastases, and Ollier disease when making a differential diagnosis. The most crucial concern for OPK is the presence of bone metastases, which are the most prevalent form of malignant bone tumors (5). The epiphyseal and metaphyseal areas of long bones are the most often affected anatomical sites in osteoporosis. The most frequently affected locations include the phalanges (100%), carpal bones (97.4%), metacarpals (92.3%), foot phalanges (87.2%), metatarsal (84.4%), tarsal (84.6%), pelvis (74.4%), femur (74.4%), radius (66.7%), ulna (66.7%), and tibia (20.5%); the ribs are less commonly affected (2,3). The key diagnostic markers include asymmetry, axial skeletal involvement, bone degeneration, and a positive bone scan. Typically, individuals with OPK do not show any symptoms. However, in certain circumstances, patients may experience pain and joint swelling without any specific deformity or malfunction. It does not require any special handling (3,10). Minor joint pain, which may or may not be associated with effusion, is the most commonly reported clinical finding in OPK. In some cases, it may be the sole symptom present. We use non-steroidal anti-inflammatory medicines to alleviate pain (4). Although complications of OPK are uncommon, they may include the development of osteosarcoma, chondrosarcoma, and giant cell tumors. However, it's crucial to acknowledge the lack of a definitive association (2,3).

Conclusion

OPK, also known as spotted bone disease, is an exceptionally uncommon genetic disorder affecting the bones. Several patches of excessive bone growth, typically found near the joints, characterize this condition. The inheritance pattern for this condition is autosomal dominant (6).

However, cutaneous manifestations of OPK may be associated with other conditions like rheumatic or skeletal problems (2). "Although OPK is generally asymptomatic, its radiologic appearance may mimic more serious conditions such as osteoblastic metastases".

In our case, the unusual involvement of the entire axial skeleton, including thoracic, cervical, lumbar, and sacral vertebrae, distinguished it from typical cases of OPK. However, following trauma, finding no other findings except for a fracture of the ulnar styloid, injury to the left radius DRUJ, and pain in the wrist upon palpation. In a patient presenting with pain over the extremities, you can keep in mind the possibility of osteopetrosis and diagnose this patient by direct radiography.

Ethics

Informed Consent: Informed consent was obtained.

Footnotes

Authorship Contributions

Surgical and Medical Practices: A.C.K., Concept: A.Ö., Design: E.F., Data Collection or Processing: E.F., Analysis or Interpretation: E.F., Literature Search: E.F., Writing: E.F.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Mohsenpour B, Ahmadi A. Osteopoikilosis: a case report. J Med Case Rep. 2023;17:292.
2. Paraskevas G, Raikos A, Stavrakas M, Spanidou S, Papaziogas B. Osteopoikilosis: a case report of a symptomatic patient. J Radiol Case Rep. 2009;3:38-43.
3. Mohammad N, Pouya H, Leila A, Ali MM, Mazyar B, Pouya K. Osteopoikilosis: a rare case with interesting imaging. Archives of Case Reports. 2023;7:12-4.
4. Özçiçek A, Özçiçek F, Karavaş E. Osteopoikilosis: a rare hereditary sclerotic bone dysplasia. Arch Basic Clin Res. 2019;1:78-80.
5. Hill CE, McKee L. Osteopoikilosis: an important incidental finding. Injury. 2015;46:1403-5.
6. D Di Primio G. Benign spotted bones: a diagnostic dilemma. CMAJ. 2011;183:456-9.
7. Deshayes S, Ly KH, Rieu V, Maigné G, Martin Silva N, Manrique A, et al. Steroid-sparing effect of anakinra in giant-cell arteritis: a case series with clinical, biological and iconographic long-term assessments. Rheumatology (Oxford). 2021;61:400-6.
8. Gaillard F, Kearns C, Baba Y, Smith M, Jones R, Evans D. Osteopoikilosis. Radiopaedia.org. 2024.
9. Hellemans J, Preobrazhenska O, Willaert A, Debeer P, Verdonk PC, Costa T, et al. Loss-of-function mutations in LEMD3 result in osteopoikilosis, Buschke-Ollendorff syndrome and melorheostosis. Nat Genet. 2004;36:1213-8.
10. Gaudio A, Xourafa A, Rapisarda R, Gorgone C, Gnoli M, Pedrini E, et al. Familiar osteopoikilosis: case report with differential diagnosis and review of the literature. Clin Case Rep. 2020;9:922-6.

DOI: 10.4274/tod.galenos.2025.16870
Turk J Osteoporos 2025;31(3):205-6

Re-evaluating Cardiovascular Risk in Osteoporotic Patients: A Case for Individualized Romosozumab Therapy

Osteoporozlu Hastalarda Kardiyovasküler Riskin Yeniden Değerlendirilmesi: Kişiselleştirilmiş Romosozumab Tedavisinin Önemi

Ömer Faruk Bucak[1], Rümeysa Çalışkan[1], Mustafa Hüseyin Temel[2], Fatih Bağcıer[1], Evrim Coşkun[1]

[1]University of Health Sciences Türkiye, Başakşehir Çam and Sakura City Hospital, Department of Physical Medicine and Rehabilitation, İstanbul, Türkiye
[2]University of Health Sciences Türkiye, Sultan 2. Abdülhamid Han Training and Research Hospital, Department of Physical Medicine and Rehabilitation, İstanbul, Türkiye

Keywords: Cardiovascular risk, osteoporosis, romosozumab
Anahtar kelimeler: Kardiyovasküler risk, osteoporoz, romosozumab

Dear Editor,

The cardiovascular safety profile of romosozumab is contentious issue, especially in light of findings from the ARCH trial suggesting a higher risk of the major adverse cardiovascular events (MACE) with romosozumab versus alendronate. Indeed, despite the impressive anti-fracture efficacy for romosozumab, its cardiovascular impact has led regulatory bodies to suggest cautious prescriptions or no recommendations at all. However, the cardiovascular risk may not be attributable to romosozumab per se but rather to baseline patient characteristics and study design, several real-world evidence and observational studies indicate (1).

Our patient is an 80-year-old lady with advanced osteoporosis who was previously treated for a year with ibandronate followed by a year of denosumab prior to being switched to romosozumab.

Notably, during denosumab treatment, the patient did not experience any adverse effects or new fractures. However, given the severity of her osteoporosis—including multiple vertebral compression fractures (T7, T9, T11, and L2) that required spinal stabilization surgery—her overall fracture risk remained substantially high. Therefore, a decision was made to escalate therapy by transitioning to romosozumab—an agent with both anabolic and antiresorptive effects—after prior use of antiresorptive agents including both bisphosphonates and denosumab. This switch reflected an individualized treatment approach tailored to the patient's persistently high fracture risk and therapeutic history.

She also had a history of hypertension and cardiac arrhythmia and had had cardiac ablation 1 year earlier. She was receiving apixaban for anticoagulation. Although she had a history of cardiovascular disease, cardiology clearance was obtained to initiate romosozumab. Her kidney function was normal (glomerular filtration rate: 65), as were her liver enzymes (alanine transaminase: 17, aspartate transaminase: 22), vitamin D (36 ng/mL) and calcium levels (9.64 mg/dL). During treatment with romosozumab, after 4 doses, no major cardiovascular events occurred on study.

After four doses of romosozumab (approximately six months of treatment), bone mineral density (BMD) improvements were observed. In the lumbar spine (L1-L4), the T-score improved from -5.15 to -4.16, and BMD increased from 0.445 g/cm² to 0.543 g/cm², reflecting a 22% relative gain. At the femoral neck, the T-score remained stable (-4.18 to -4.16), with a slight BMD increase from 0.376 g/cm² to 0.398 g/cm² (approximately 6% relative gain). These findings indicate a skeletal response, particularly in the spine, even after a relatively short course of romosozumab. No adverse events or new fractures were reported during this period.

This case highlights important points:

1. Romosozumab may still be an option in select high-risk cardiovascular patients if appropriate cardiology consultation and monitoring are performed.

Corresponding Author/Sorumlu Yazar: Asst. Rümeysa Çalışkan, University of Health Sciences Türkiye, Başakşehir Çam and Sakura City Hospital, Department of Physical Medicine and Rehabilitation, İstanbul, Türkiye
E-mail: rumeysa98clskn@hotmail.com **ORCID ID:** orcid.org/0009-0001-7035-7000
Received/Geliş Tarihi: 09.04.2025 **Accepted/Kabul Tarihi:** 30.04.2025 **Publication Date/Yayınlanma Tarihi:** 05.12.2025

Cite this article as/Atıf: Bucak ÖF, Çalışkan R, Temel MH, Bağcıer F, Coşkun E. Re-evaluating cardiovascular risk in osteoporotic patients: a case for ındividualized romosozumab therapy. Turk J Osteoporos. 2025;31(3):205-6

2. Real-world data on romosozumab use in high-cardiovascular-risk patients is needed, as controlled trials often exclude such individuals.

The discordant cardiovascular safety profile of romosozumab relates primarily to differences between clinical trials. The ARCH trial showed a statistically significant increase in MACE (hazard ratio: 1.87; 95% confidence intercal: 1.11-3.17; p=0.02) compared to alendronate given for 12 months, raising concerns regarding whether romosozumab may be pro-atherosclerotic (1). The FRAME trial, which compared romosozumab with placebo, found no difference in cardiovascular events (MACE 0.8% in both groups), suggesting that the potential cardioprotective effect of alendronate may have confounded the ARCH trial results (1). Moreover, the BRIDGE trial, which involved men with osteoporosis, showed a higher rate for cardiovascular events in users of romosozumab (4.9%) compared to placebo (2.5%), but did not achieve statistical significance (2).

Recent real-world data from a large propensity-score-matched cohort study (TriNetX, covering 136 million patients) provides additional insights. This study compared romosozumab with parathyroid hormone (PTH) analogues (teriparatide/abaloparatide) and found that romosozumab was associated with fewer major cardiovascular events than PTH analogues, challenging the assumption that romosozumab inherently increases cardiovascular risk (3). Also, an unusual frequency of cardiovascular events was noted as reported in Food and Drug Administration (FDA) adverse event reporting system pharmacovigilance studies in Japan, possibly attributable to higher baseline cardiovascular risk in patients in Japan, rather than being directly attributable to romosozumab (3).

Mechanistic hypotheses exist for sclerostin inhibition influencing vascular calcification and progression of atherosclerotic cardiovascular disease given that it is expressed in vascular smooth muscle cells. Yet definitive cardiovascular dysfunction with romosozumab has not been demonstrated in preclinical studies, and Mendelian randomization studies have been conflicting, showing that lower sclerostin associates with an increased risk of atherosclerosis but failing to show definitive causation (4).

Given these findings, a one-size-fits-all approach to cardiovascular risk and romosozumab may not be appropriate. Regulatory agencies have taken a cautious stance—the FDA recommends avoiding romosozumab in patients with a history of MI or stroke within the past year, while the EMA contraindicates its use in all patients with prior cardiovascular events. But increasing real-world evidence suggests that careful patient selection and multidisciplinary oversight may allow high-fracture-risk patients with stable cardiovascular conditions to take romosozumab with no undue risk (1).

In summary, this case highlights the importance of providing thoughtful consideration of risk-to-benefit ratio in selecting osteoporosis therapy in patients with concomitant cardiovascular disease. While regulatory guidelines err on the side of caution, clinicians should integrate real-world data, patient-specific cardiovascular profiles, and multidisciplinary decision-making when considering romosozumab therapy. Large-scale observational studies and prospective trials will be needed to further optimize patient selection and treatment safety.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Ö.F.B., E.C., Concept: F.B., Design: R.Ç., F.B., Data Collection or Processing: F.B., Analysis or Interpretation: M.H.T., E.C., Literature Search: R.Ç., M.H.T., Writing: F.B.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Reid IR. What is the risk of cardiovascular events in osteoporotic patients treated with romosozumab. Expert Opin Drug Saf. 2022;21:1441-3.
2. Stokar J, Szalat A. Cardiovascular safety of romosozumab vs PTH analogues for osteoporosis treatment: a propensity-score-matched cohort study. J Clin Endocrinol Metab. 2025;110:e861-7. Erratum in: J Clin Endocrinol Metab. 2025;110:e915.
3. Vestergaard Kvist A, Faruque J, Vallejo-Yagüe E, Weiler S, Winter EM, Burden AM. Cardiovascular safety profile of romosozumab: a pharmacovigilance analysis of the US food and drug administration adverse event reporting system (FAERS). J Clin Med. 2021;10:1660.
4. Alnaqbi KA, Al Zeyoudi J, Aljaberi AK. Cardiac arrhythmia and heart failure shortly after starting romosozumab for osteoporosis: a case-based review. Cureus. 2023;15:e50303. Erratum in: Cureus. 2023;15:c148.

2025 Hakem Dizini - 2025 Referee Index

Ahmet Kıvanç Cengiz
Ayhan Aşkın
Ayşe Yalıman
Coşkun Zateri
Çiğdem Bircan
Didem Erdem Gürsoy
Ece Aydoğ
Ece Çınar
Esra Giray
Ezgi Yıldız Güvercin
F Yeşim Gökçe Kutsal
Funda Çalış
Göksel Tanıgör
Gülseren Akyüz
Hatice Bodur
Hatice Uğurlu
Haydar Gök
Hilal Yeşil
İlknur Aktaş
İlknur Albayrak
Kazım Çapacı
Meliha Kasapoğlu Aksoy
Musa Baklacı
Nilgün Mesci
Özge Keniş Coşkun
Raikan Büyükavcı
Rana Terlemez
Rengin Güzel
Sevil Okan
Sibel Eyigör
Ülkü Zeynep Akarırmak
Zafer Günendi