Surveillance of atypical femoral fractures in a nationwide fracture register

Hans Peter BÖGL [1,2], Georg ZDOLSEK [2], Lukas BARNISIN [3], Michael MÖLLER [3,5] and Jörg SCHILCHER [2,4]

[1] Department of Orthopedic Surgery, Gävle Hospital, Gävle; [2] Department of Orthopedic Surgery and Department of Biomedical and Clinical Sciences, Linköping University, Linköping and Center for Medical Image Science and Visualization, Linköping University, Linköping; [3] Department of Orthopedics, Sahlgrenska University Hospital, Gothenburg; [4] Wallenberg Centre for Molecular Medicine, Linköping University, Linköping; [5] Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
Correspondence: hans.peter.bogl@regiongavleborg.se
Submitted 2021-02-09. Accepted 2021-12-07.

Background and purpose — To continuously assess the incidence of atypical femoral fractures (AFFs) in the population is important, to allow the evaluation of the risks and benefits associated with osteoporosis treatment. Therefore, we investigated the possibility to use the Swedish Fracture Register (SFR) as a surveillance tool for AFFs in the population and to explore means of improvement.

Patients and methods — All AFF registrations in the SFR from January 1, 2015 to December 31, 2018 were enrolled in the study. For these patients, radiographs were obtained and combined with radiographs from 176 patients with normal femoral fractures, to form the study cohort. All images were reviewed and classified into AFFs or normal femur fractures by 2 experts in the field (gold-standard classification) and 1 orthopedic resident educated on the specific radiographic features of AFF (educated-user classification). Furthermore, we estimated the incidence rate of AFFs in the population captured by the register through comparison with a previous cohort and calculated the positive predictive value (PPV) and, where possible, the inter-observer agreement (Cohen's kappa) between the different classifications.

Results — Of the 178 available patients with AFF in the SFR, 104 patients were classified as AFF using the gold-standard classification, and 89 using the educated-user classification. The PPV increased from 0.58 in the SFR classification to 0.93 in the educated-user classification. The inter-observer agreement between the gold-standard classification and the educated-user classification was 0.81.

Interpretation — With a positive predictive value of 0.58 the Swedish Fracture Register outperforms radiology reports and reports to the Swedish Medical Products Agency on adverse drug reactions as a diagnostic tool to identify atypical femoral fractures.

Atypical femoral fractures (AFF) represent a new entity of insufficiency fractures that have a strong association with bisphosphonate treatment (1–5). These fractures are rare, and the radiographic pattern may be difficult to discriminate from those of other types of fractures of the femoral shaft (6). There is little published information on how AFFs are identified by radiologists and orthopedic surgeons in daily clinical practice. Some reports have suggested that the current American Society for Bone and Mineral Research (ASBMR) criteria (2) are poorly applied and very few AFFs are correctly identified as such (7) (Figure 1). Therefore, epidemiological surveillance in the population is difficult.
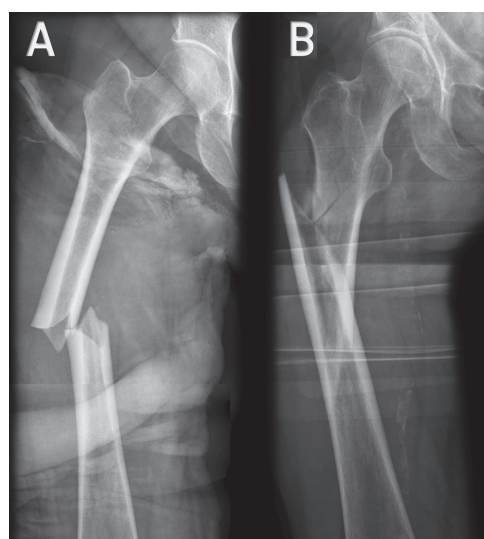

Figure 1. A. Atypical femoral fracture with radiographic features i.e., non-comminuted fracture with local cortical thickening and "beaking" at the fracture site, transverse orientation of the fracture line in the lateral cortex with a medial spike. B. Normal femoral fracture.

The Swedish Fracture Register (SFR) is a nationwide register in which trauma- and fracture-specific data are registered by the treating physicians, in a web-based interface, irrespective of the type of treatment administered. In 2015, we added the possibility to register AFF in the SFR.

The surveillance of AFF in the population through the SFR and other fracture registers should help patients and prescribers to arrive at a realistic picture of the risks and benefits of bisphosphonate treatment, even when treatment recommendations and clinical practice change over time. However, such a surveillance system relies on the correct classification of fractures.

Therefore, we investigated the Swedish Fracture Register (SFR) as a surveillance tool for atypical femoral fractures (AFFs) in the population and to explore means of improvement. Another secondary aim was to estimate the incidence of AFF in the population between 2015 and 2018 using SFR.

Patients and methods

Swedish Fracture Register

The Swedish Fracture Register was inaugurated in 2011 at Sahlgrenska University Hospital in Gothenburg, Sweden and has now reached a national coverage of 100% (53 hospitals). The completeness of the data in the SFR ranges widely depending on the fracture type and the reporting hospital. With currently more than 600,000 fractures included in its database, the SFR is one of the world's largest registers of fractures and related treatments.

The registration process is performed through a personal login on the register's homepage (https://sfr.registercentrum.se). Every registration includes the date of trauma, trauma mechanism, fracture classification (modified AO/OTA classification), treatment (non-operative and operative), type of surgical treatment, date of treatment, and any subsequent surgeries.

Atypical femoral fractures in the register

In 2015, we introduced atypical fractures as a subtype of subtrochanteric and diaphyseal femoral fractures (ICD S72.2 and S72.3; International Classification of Diseases, 10th revision). When a femoral fracture (S72.2 and S72.3) is registered on the homepage using the patient's national personal identification number, the user will be prompted with a pop-up window asking: "Is this an atypical femoral fracture, yes/no." The pop-up window also contains the following reminder of the features of atypical fractures: "Atypical fractures are stress fractures with a transverse fracture line originating from the lateral cortex on the frontal plane X-ray. A callus reaction is present (cortical thickening). The fracture may present as an incomplete thin fracture line or as a complete fracture. These fractures have a strong association with bisphosphonate treatment."

Detailed information can be registered regarding the trauma mechanism and whether a fracture is related to primary or secondary bone tumors or pre-existing implants or joint prostheses.
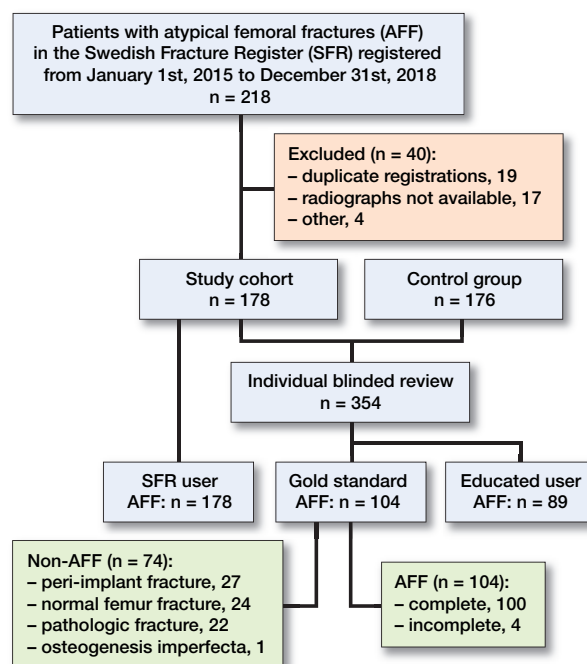

Figure 2. Flowchart showing the selection of radiographs and the results of the different radiographic reviews.

Study population

All the patients with AFF in the SFR were identified between January 1, 2015 and December 31, 2018. The personal identification number was used to retrieve all the radiographs of the ipsilateral femur for each patient from the relevant radiology department and stored in the PACS at Linköping University Hospital, Sweden. To provide a control group of patients with normal femur fractures, we randomly selected an additional 176 patients from our previous study cohort with normal femoral fractures (NFFs) from 2008 to 2010 (8) (mean age, 82 years; SD 9.6; 80% females in the initial study cohort), to form the current study cohort (Figure 2). Radiographs from these patients served as a negative control population to derive the gold-standard classification and the educated-user classification described below.

Fracture types

We classified fractures into 2 types: AFFs and NFFs. AFFs were characterized according to the following features: (1) located in the femoral diaphysis (between the tip of the lesser trochanter and the supracondylar flare); (2) the fracture originated in the lateral cortex and ran perpendicular to the longitudinal axis of the femur; (3) the fracture was not at all or only minimally comminuted, and (4) showed a medial spike, and (5) there was endosteal/periosteal thickening at the fracture site. According to the ASBMR criteria (2), 4 out of 5 major criteria need to be fulfilled for the classification of AFF. In our definition, the transverse fracture line and endosteal/periosteal thickening were compulsory features of AFF (9). NFFs were not trans-

Incidence (per 100,000 inhabitants) trends of AFF over time

Year	2008	2009	2010	2008–2010 mean	2018
AFF rate	23	18	29	23	13
Population [a]	1,036,859	1,046,832	1,057,143	1,046,945	1,142,417
Incidence	2.2	1.7	2.7	2.2	1.1

[a] Accrued inhabitants and permanent residents, 55 years and older in the counties Västra Götaland, Kalmar, Gävleborg, Dalarna, Blekinge, Uppsala, Värmland, and Jämtland.

verse on the lateral side. Both fracture types showed no signs of previous surgery or malignancy. The number of patients receiving bisphosphonate treatment could not be determined.

Classification of fracture types

Radiographs were made available blinded from all background information to the 3 reviewers, who then classified the fractures individually into AFF or NFF. 2 of the reviewers (HPB and GZ) with experience in research on AFF were assigned as expert reviewers (gold-standard classification). The remaining reviewer (LB) was a final-year orthopedic resident without research background, regarded as an average SFR user. Before classification of the fracture types, LB was given a short, online introduction to the field of AFF, the ASBMR classification criteria, and our thoughts on the classification of AFFs based on their insufficiency-type fracture appearance (educated-user classification) (10).

The expert reviewers were in agreement on all but 9 cases (inter-rater kappa of 0.93), which were agreed upon via video conference, to establish a gold-standard classification. Eventually, there were 3 different classifications of the study cohort available for comparison: the original classification of AFF cases in the fracture register (SFR classification); the gold standard classification; and the educated-user classification.

Statistics

We tested the diagnostic performances of the SFR classification and the educated-user classification compared with the gold-standard classification using positive predictive values (PPVs) with corresponding 95% confidence intervals (CI). Cohen's kappa coefficient for inter-observer agreement was calculated for the gold-standard classification and the educated-user classification to evaluate the extent of agreement.

Furthermore, we calculated the sensitivity of the SFR to detect AFF by comparing the number of AFFs in the gold-standard classification with the number of expected AFFs in the population in 2018. To calculate the expected number of AFFs, we used historical data on AFF in patients aged > 55 years (10) from counties with a coverage for hip fractures of > 80% in the SFR in 2018 (Table). The year 2018 was chosen because the extent of coverage of the register is steadily increasing and 2018 was the most recent year with complete

register data at the time of data extraction. We used a well-defined, nationwide cohort of patients with AFF for the period 2008–2010 as a reference population (8), to extrapolate the expected incidence rate of AFF in the whole population of women and men aged > 55 years in 2018 (Table), using population data from Statistics Sweden (https://www.scb.se/en/; the governmental agency for official statistics in Sweden).

Ethics, funding, and potential conflicts of interest

Results

Between January 1, 2015 and December 31, 2018 there were 5,247 registrations of low-energy subtrochanteric and diaphyseal femoral fractures in the SFR of which 218 (178 females) were AFFs. We excluded 40 of the 218 AFF cases from the analysis owing to multiple registrations (n = 19), missing radiographs (n = 17), and other reasons (n = 4).

Of the remaining 178 patients registered with AFF in the SFR (mean age, 75 [SD 11] years; 147 [83%] females), 104 were classified as having AFF in the gold-standard classification. The remaining 74 cases from the SFR cohort were not classified as AFF by the 2 expert reviewers (27 peri-implant fractures, 24 normal fractures, 22 fractures with malignant changes in the ipsilateral femur, and 1 fracture with features typical of osteogenesis imperfecta). The SFR classification had a diagnostic accuracy of 0.58 (CI 0.51–0.66) when compared with the gold-standard classification.

The educated-user classification identified 89 patients as AFF. Of these, 83 patients were also identified as AFF in the gold-standard classification (inter-observer agreement of 0.81 (CI 0.74–0.88), Cohen's kappa coefficient) (11). The remaining 265 cases were classified as non-AFFs. The PPV for the educated-user classification compared with the gold-standard classification was 0.93 (CI 0.85–0.97). Of the 104 patients with AFF in the gold-standard classification, 97 were ≥ 55 years of age.

We identified 4,781 low-energy subtrochanteric and diaphyseal femoral fractures (including peri-implant fractures but excluding insufficiency and pathological fractures) registered in patients aged 55 years or older in the SFR for the period 2015–2018. For the year 2018 (fractures of the subtrochan-

teric or diaphyseal region of the femur, n = 1,164), we calculated an incidence rate of 1.1 AFF per 100,000 women and men aged ≥ 55 years and at risk in the selected counties (n = 1,142,417). The mean incidence rate of AFF in the selected counties for the period 2008–2010 was 2.2 per 100,000 per year (range 1.7–2.7). Comparing this number of AFF cases in the population with the number of AFF cases in the gold-standard classification, the SFR allowed detection of roughly half of the expected patients with AFF in the population, assuming an unchanged incidence rate of AFF over time.

Discussion

We studied the diagnostic accuracy of AFF diagnoses registered in the Swedish Fracture Register to evaluate the potential of the register as a surveillance tool for these rare fractures. Approximately 58% (104/178) of all AFFs registered in the SFR were confirmed as such in the gold-standard classification. Even if this percentage appears low, the register is more successful in this respect than radiology reports alone (< 7%) [7] and reports on drug adverse reactions to the Swedish Medical Products Agency. For the period 2015–2018, only 20 patients (18 women) were reported to the drug adverse reaction database (BiSi; extraction date: November 12, 2020) for the whole country. Thus, improvement is both desirable and urgent.

We chose to exclude AFF cases from the gold-standard classification if their radiographs showed pre-existing implants in the ipsilateral femur. This was in line with our previous research [3,9,10,12–23] and according to interpretations [24] of the internationally accepted case definition of AFF proposed by ASBMR [2]. However, recent studies have questioned the exclusion of peri-implant fractures as AFFs [24]. Inclusion of these fractures in a surveillance tool would be reasonable in terms of allowing a thorough assessment in the later stages. In the present study, 27 of the patients who were registered as AFF in the SFR showed pre-existing implants in the ipsilateral femur and 14 of these were classified as AFF by the 2 experts but excluded due to the pre-existing implants. Including these peri-implant fractures in the accuracy assessment would increase the agreement between the SFR classification and the gold-standard classification to 66% and would be in line with recent findings and thoughts on whether to exclude AFF in the presence of orthopedic implants [25,26]. Of the 22 AFF cases in the SFR classification that were identified as having pathological fractures in the gold-standard classification, 10 patients were also identified through the pathological fracture classification tool in the SFR.

We observed a difference in the registration accuracy between the educated-user classification and the SFR classification with an increase in the PPV from 0.58 to 0.93. This observed improvement using 1 single examiner does not allow any conclusions on how to improve accuracy in the register but has prompted us to publish a movie clip along with the regis-

tration process of AFF in the SFR (https://stratum.blob.core.windows.net/sfr/Movies/AtypFemureFract.mp4). The importance of accurate diagnosis of AFF based on established diagnostic criteria has been emphasized in previous studies, especially in terms of whether the fracture line runs transverse and whether a minimal periosteal callus reaction can be seen or not [6,27]. Our findings indicate that most classification failures are related to poor compliance with the ASBMR criteria. Registrations of treatment choices in the register are typically done by the operating surgeon. Most fracture classifications on the other hand are done by junior doctors in the emergency department. Emergency departments are often staffed with doctors at the intern or resident level who might not have sufficient knowledge of AFFs. The improvement that we observed was based on the fracture classification of a volunteer, final-year orthopedic resident. The observed effect might be influenced by selection and expert bias, thus limiting generalizability of our findings to the standard SFR user responsible for the initial classification of fractures. Choosing a control group of familiar size gives a clear indication of the prevalence of AFF in the sample. This might have influenced the PPV of the educated-user classification as well as the kappa value, because we defined the outcome prevalence by our design. Furthermore, the higher PPV of the educated-user classification seems to come at the cost of overall diagnostic accuracy. Because the number of false-negative diagnoses was increased, the sensitivity of the educated user classification was lower. Another limitation is the choice of control radiographs that were previously reviewed by the expert reviewers [8]. This might have influenced the gold standard classification through recall bias. However, because reviews were several years apart, such an influence is very unlikely [28]. Furthermore, we are unaware of the number of false-negative AFFs in the register because such radiographs were not reviewed.

Based on extrapolated estimates on incidence rates of AFF in the population, the SFR enabled detection of roughly half of the expected cases. These estimates are highly speculative given that the true rate for patients with AFF in the population of Sweden for 2018, as well as the time trends of the incidence rates for the period 2008–2010 compared with 2018, are unknown. Considering the decline in bisphosphonate prescription rates [29] and the emphasis on regimens with individualized but limited treatment duration of bisphosphonates for osteoporosis [30], a decrease in the incidence of AFF would be expected.

We were not able to acquire information regarding the status of treatment with bisphosphonates or other anti-resorptive drugs because these variables are not captured in the SFR. Our results are highly relevant for healthcare systems with large registers that allow the registration of specific fracture types. Furthermore, this investigation was limited to complete AFFs that were treated surgically; no incomplete fractures were included.

Surveillance of AFFs in the population is hindered by difficulties related to the timely and correct diagnosis of patients.

Despite the currently moderate accuracy level of registered AFFs in the SFR, the register provides valuable possibilities for surveillance across individual hospitals with minimal effort. With further assistance provided by artificial intelligence algorithms to aid in the correct classification of initial radiographs (31,32), and the coupling of instructional materials concerning the ASBMR criteria to the registration process in the SFR, the accuracy of this and other registers can be further improved.

In conclusion, the SFR represents a useful tool to identify the rare event of AFF in the population. The diagnostic accuracy of AFF data in the SFR might be improved through minimal educational effort.

1. **Shane E, Burr D, Ebeling P R, Abrahamsen B, Adler R A, Brown T D, et al.** Atypical subtrochanteric and diaphyseal femoral fractures: report of a task force of the American Society for Bone and Mineral Research. J Bone Miner Res 2010; 25(11): 2267-94.

2. **Shane E, Burr D, Abrahamsen B, Adler R A, Brown T D, Cheung A M, et al.** Atypical subtrochanteric and diaphyseal femoral fractures: second report of a task force of the American Society for Bone and Mineral Research. J Bone Miner Res 2014; 29(1): 1-23.

3. **Schilcher J, Michaelsson K, Aspenberg P.** Bisphosphonate use and atypical fractures of the femoral shaft. N Engl J Med 2011; 364(18): 1728-37.

4. **Odvina C V, Zerwekh J E, Rao D S, Maalouf N, Gottschalk F A, Pak C Y C.** Severely suppressed bone turnover: a potential complication of alendronate therapy. J Clin Endocrinol Metab 2005; 90(3): 1294-301.

5. **Kwek E B K, Goh S K, Koh J S B, Png M A, Sen H T.** An emerging pattern of subtrochanteric stress fractures: a long-term complication of alendronate therapy? Injury 2008; 39(2): 224-31.

6. **Rydholm A.** Highly different risk estimates for atypical femoral fracture with use of bisphosphonates: debate must be allowed! Acta Orthop 2012; 83(4): 319-20.

7. **Harborne K, Hazlehurst J M, Shanmugaratnam H, Pearson S, Doyle A, Gittoes N J, et al.** Compliance with established guidelines for the radiological reporting of atypical femoral fractures. Br J Radiol 2016; 89(1057): 20150443.

8. **Bogl H P, Michaelsson K, Zdolsek G, Hoijer J, Schilcher J.** Increased rate of reoperation in atypical femoral fractures is related to patient characteristics and not fracture type: a nationwide cohort study. Osteoporos Int 2020; 31(5): 951-9.

9. **Schilcher J, Koeppen V, Ranstam J, Skripitz R, Michaelsson K, Aspenberg P.** Atypical femoral fractures are a separate entity, characterized by highly specific radiographic features: a comparison of 59 cases and 218 controls. Bone 2013; 52(1): 389-92.

10. **Schilcher J, Koeppen V, Aspenberg P, Michaelsson K.** Risk of atypical femoral fracture during and after bisphosphonate use. Acta Orthop 2015; 86(1): 100-7.

11. **Landis J R, Koch G G.** The measurement of observer agreement for categorical data. Biometrics 1977; 33(1): 159-74.

12. **Schilcher J, Aspenberg P.** Incidence of stress fractures of the femoral shaft in women treated with bisphosphonate. Acta Orthop 2009; 80(4): 413-15.

13. **Aspenberg P, Schilcher J, Fahlgren A.** Histology of an undisplaced femoral fatigue fracture in association with bisphosphonate treatment. Acta Orthop 2010; 81(4): 460-2.

14. **Schilcher J.** Epidemiology, radiology and histology of atypical femoral fractures. Acta Orthop Suppl. 2013; 84(352): 1-26.

15. **Aspenberg P, Schilcher J.** Atypical femoral fractures, bisphosphonates, and mechanical stress. Curr Osteoporos Rep 2014; 12(2): 189-93.

16. **Schilcher J, Koeppen V, Aspenberg P, Michaelsson K.** Risk of atypical femoral fracture during and after bisphosphonate use. N Engl J Med 2014; 371(10): 974-6.

17. **Schilcher J, Sandberg O, Isaksson H, Aspenberg P.** Histology of 8 atypical femoral fractures: remodeling but no healing. Acta Orthop 2014; 85(3): 280-6.

18. **Schilcher J.** High revision rate but good healing capacity of atypical femoral fractures: a comparison with common shaft fractures. Injury 2015; 46(12): 2468-73.

10. **Schilcher J, Howe T S, Png M A, Aspenberg P, Koh J S.** Atypical fractures are mainly subtrochanteric in Singapore and diaphyseal in Sweden: a cross-sectional study. J Bone Miner Res 2015; 30(11): 2127-32.

20. **Gustafsson A, Schilcher J, Grassi L, Aspenberg P, Isaksson H.** Strains caused by daily loading might be responsible for delayed healing of an incomplete atypical femoral fracture. Bone 2016; 88: 125-30.

21. **Kharazmi M, Michaelsson K, Hallberg P, Schilcher J.** Lateral fixation: an alternative surgical approach in the prevention of complete atypical femoral fractures. Eur J Orthop Surg Traumatol 2017; 28(2): 299-304.

22. **Bogl H P, Aspenberg P, Schilcher J.** Undisturbed local bone formation capacity in patients with atypical femoral fractures: a case series. Osteoporos Int 2017; 28(8): 2439-44.

23. **Bogl H P, Zdolsek G, Michaelsson K, Hoijer J, Schilcher J.** Reduced risk of reoperation using intramedullary nailing with femoral neck protection in low-energy femoral shaft fractures. J Bone Joint Surg Am 2020; 102(17): 1486-94.

24. **Lee J Y, Soh T, Howe T S, Koh J S, Kwek E B, Chua D T.** Bisphosphonate-associated peri-implant fractures: a new clinical entity? Acta Orthop 2015; 1-5.

25. **Robinson J de D, Leighton R K, Trask K, Bogdan Y, Tornetta P 3rd.** Periprosthetic atypical femoral fractures in patients on long-term bisphosphonates: a multicenter retrospective review. J Orthop Trauma 2016; 30(4): 170-6.

26. **MacKenzie S A, Ng R T, Snowden G, Powell-Bowns M F R, Duckworth A D, Scott C E H.** Periprosthetic atypical femoral fractures exist and are associated with duration of bisphosphonate therapy. Bone Joint J 2019; 101-b(10): 1285-91.

27. **Feldstein A C, Black D, Perrin N, Rosales A G, Friess D, Boardman D, et al.** Incidence and demography of femur fractures with and without atypical features. J Bone Miner Res 2012; 27(5): 977-86.

28. **Campbell W S, Talmon G A, Foster K W, Baker J J, Smith L M, Hinrichs S H.** Visual memory effects on intraoperator study design: determining a minimum time gap between case reviews to reduce recall bias. Am J Clin Pathol 2015; 143(3): 412-18.

29. **Jha S, Wang Z, Laucis N, Bhattacharyya T.** Trends in media reports, oral bisphosphonate prescriptions, and hip fractures 1996-2012: an ecological analysis. J Bone Miner Res 2015; 30(12): 2179-87.

30. **Whitaker M, Guo J, Kehoe T, Benson G.** Bisphosphonates for osteoporosis: where do we go from here? N Engl J Med 2012; 366(22): 2048-51.

31. **Olczak J, Fahlberg N, Maki A, Razavian A S, Jilert A, Stark A, et al.** Artificial intelligence for analyzing orthopedic trauma radiographs. Acta Orthop 2017; 88(6): 581-6.

32. **Groot O Q, Bongers M E R, Ogink P T, Senders J T, Karhade A V, Bramer J A M, et al.** Does artificial intelligence outperform natural intelligence in interpreting musculoskeletal radiological studies? A systematic review. Clin Orthop Relat Res 2020; 478(12): 2751-64.