Article

Not peer-reviewed version

# Development of a Medication-Related Osteonecrosis of the Jaw Prediction Model Using the FDA Adverse Event Reporting System Database and Machine Learning

Shinya Toriumi *, Komei Shimokawa, Munehiro Yamamoto, Yoshihiro Uesawa *

Posted Date: 5 February 2025

doi: 10.20944/preprints202502.0246.v1

Keywords: medication-related osteonecrosis of the jaw; bisphosphonates; epidemiological research; disproportionality analysis; spontaneous report database; in silico analysis; quantitative structure-activity relationship; machine learning; artificial neural network; molecular structure descriptor

*Article*

# Development of a Medication-Related Osteonecrosis of the Jaw Prediction Model Using the FDA Adverse Event Reporting System Database and Machine Learning

**Shinya Toriumi [1,2,*], Komei Shimokawa [2], Munehiro Yamamoto [3] and Yoshihiro Uesawa [1,*]**

[1] Department of Medical Molecular Informatics, Meiji Pharmaceutical University, Kiyose 204-8588, Japan
[2] Department of Pharmacy, National Hospital Organization Kanagawa Hospital, Hadano 257-8585, Japan
[3] Department of Orthopedic Surgery, National Hospital Organization Kanagawa Hospital, Hadano 257-8585, Japan
* Correspondence: sn.toriumi@gmail.com (S.T.); uesawa@my-pharm.ac.jp (Y.U.); Tel.: +81-42-495-8983 (Y.U.)

**Abstract:** Medication-related osteonecrosis of the jaw (MRONJ) is a rare adverse event associated with antiresorptive and antiangiogenic drugs that can significantly impact quality of life. In this study, we conducted a quantitative structure-activity relationship (QSAR) analysis using the U.S. Food and Drug Administration Adverse Drug Reaction Database System (FAERS) and machine learning to construct a drug prediction model for MRONJ induction based solely on chemical structure information. We analyzed 4,815 drugs from FAERS to assess their association with MRONJ and predict MRONJ-positive and MRONJ-negative drugs. By incorporating 326 chemical structure descriptors, we performed QSAR analysis on 60 positive and 108 negative drugs. Three machine learning algorithms were evaluated along with the number of chemical structure descriptors for QSAR analysis. The optimal MRONJ induction drug prediction model was established using an artificial neural network algorithm and eight chemical structure descriptors (area under the receiver operating characteristic curve = 0.778). Among these descriptors, drugs with polar surface area characteristics were identified as potentially linked to MRONJ. This study demonstrates a promising approach for predicting MRONJ risk, which could enhance drug safety assessment and streamline drug screening in clinical and preclinical settings.

**Keywords:** medication-related osteonecrosis of the jaw; bisphosphonates; epidemiological research; disproportionality analysis; spontaneous report database; in silico analysis; quantitative structure-activity relationship; machine learning; artificial neural network; molecular structure descriptor

## 1. Introduction

Medication-related osteonecrosis of the jaw (MRONJ) is a rare adverse event associated with long-term administration of bisphosphonates (BPs) and denosumab [1,2]. However, MRONJ has also been linked to drugs with mechanisms of action distinct from those of bone resorption inhibitors, such as the angiogenesis inhibitors bevacizumab and sunitinib, as well as the immunosuppressants methotrexate and everolimus [1–4]. This suggests that various drugs may contribute to the development of MRONJ through different pathways. Although MRONJ significantly reduces quality of life, it is recommended to continue treatment with MRONJ-related drugs while implementing strategies to minimize its risk, as the therapeutic benefits of these drugs often outweigh the risks [5–8]. Therefore, the ability to predict and evaluate the risk of MRONJ in advance would be valuable for managing such adverse events.

Spontaneous reporting systems, which collect data on adverse events in clinical settings over an extended period, play an important role in epidemiological studies, particularly in drug safety

 doi:10.20944/preprints202502.0246.v1



evaluations [9–14]. A spontaneous reporting system is a drug adverse event database that gathers spontaneous reports from patients, medical professionals, pharmaceutical companies, and other sources. These databases accumulate a huge amount of adverse event reports that are often challenging to obtain at a single institution. In addition, they include data on patients with diverse backgrounds, such as those with renal or hepatic disorders, making spontaneous reporting system an excellent tool for inductively understanding drug-related adverse events. This reflects not only unique pharmacological and pharmacokinetic characteristics but also prescription and usage conditions [15,16]. In spontaneous reporting system databases, a signal detection approach can be used to identify potential causal relationships between adverse events and drugs, even when such relationships were previously unknown [17–19]. Many studies have utilized these databases to explore the association between drugs and adverse events [12,20]. The U.S. Food and Drug Administration (FDA) Adverse Event Reporting System (FAERS) is one of the largest spontaneous reporting databases globally [21].

In recent years, there has been a growing interest in using in silico analysis to evaluate drug toxicity (adverse event evaluation) [22,23]. The physiological activity and physical properties of a drug are typically determined by its chemical structure, which can be analyzed through structural similarity. Quantitative structure-activity relationship (QSAR) analysis is a method that models the relationship between chemical structure and drug efficacy based on this principle [24–27]. This method involves converting the chemical structure of a compound into computationally analyzable features and constructing mathematical models to relate structure to activity. Recently, advances in machine learning algorithms have been reported to enhance the prediction accuracy of such models [28,29]. Therefore, by leveraging machine learning to develop an MRONJ-inducing drug prediction model, it becomes possible to assess the risk of MRONJ for various drugs based solely on their chemical structure information. This approach is highly beneficial for screening new compounds and drugs prior to clinical use.

In this study, we integrated data from a drug adverse event database with machine learning techniques to construct an MRONJ-inducing drug prediction model. Initially, drugs associated with MRONJ were extracted from the FAERS drug adverse event database. Subsequently, molecular descriptors representing the structural information of the extracted drugs were calculated, and a classification model for MRONJ-inducing drugs was developed using machine learning.

## 2. Results

### *2.1. Creation of the FAERS Analysis Data Tables*

The FAERS analysis data table was created by combining information from the FAERS drug table (drug information), Reaction table (adverse event information), Demographic table (basic case information), and Therapy table (treatment period information). Duplicate records were eliminated from the four tables. Following deduplication, the Drug, Reaction, Demographic, and Therapy tables contained 103,252,306, 44,286,680, 14,836,487, and 53,686,946 records, respectively. These tables were merged, and data cleaning procedures were applied to create the FAERS analysis data table. The table contained adverse event data from 12,468,455 records involving 4,815 drugs. Among these, 3,427 cases (0.027%) were related to MRONJ.

### *2.2. Positive and Negative Drugs for MRONJ*

In the FAERS analysis data table, out of 4,815 drugs, 70 were identified as MRONJ-positive and 139 as MRONJ-negative (Supplementary Table S1). A volcano plot was generated to visually represent the relationship between the drugs reported in FAERS and MRONJ (Figure 1). Each point on the scatter plot represents a drug, with MRONJ-positive drugs located in the upper right quadrant and MRONJ-negative drugs in the upper left quadrant. The color of each point represents the total number of reported adverse events for each drug, with more red points and fewer blue points indicating higher numbers of adverse events. Among the 70 MRONJ-positive drugs, 11 were

 doi:10.20944/preprints202502.0246.v1



classified as malignant tumor drug protein kinase inhibitors (ATC code: L01E), and 8 were drugs affecting bone structure and mineralization (ATC code: M05B), such as BPs (Supplementary Table S1). The number of reported cases of MRONJ was 907 for denosumab, 702 for zoledronic acid, 264 for alendronate, 92 for ibandronate, 65 for sunitinib, and 60 for dexamethasone (Table 1).

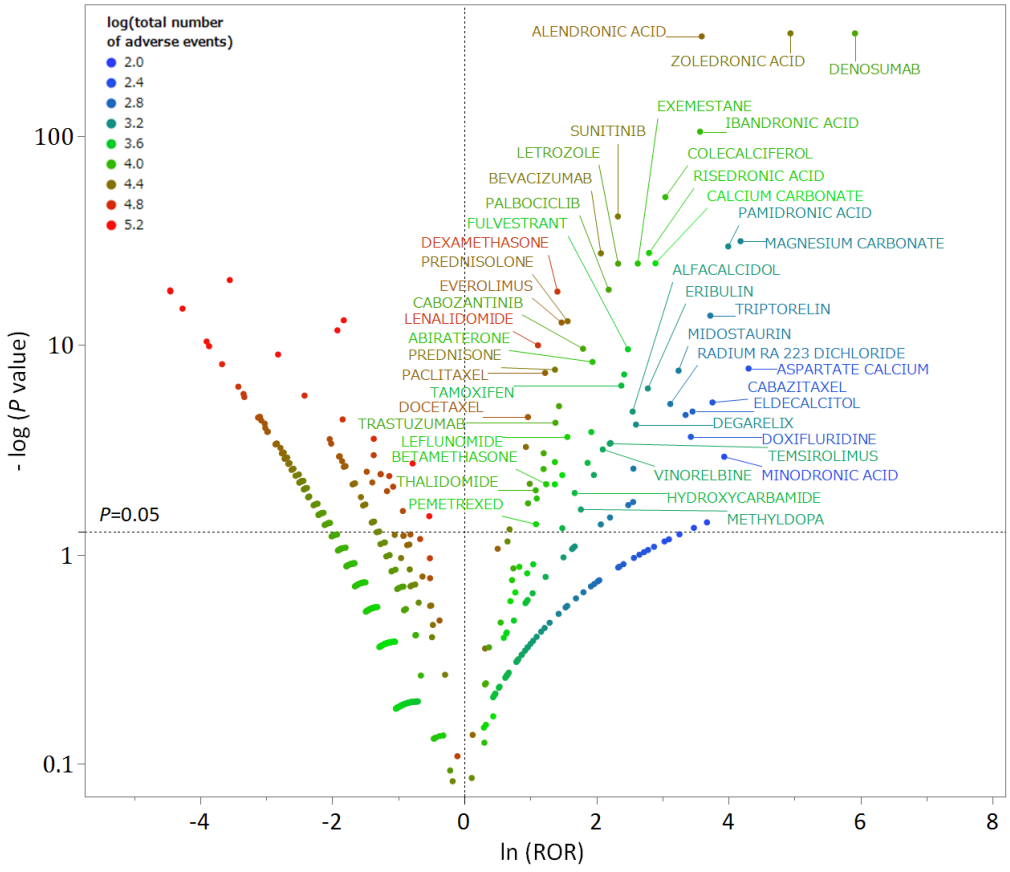


**Figure 1.** A volcano plot of drugs associated with medication-related osteonecrosis of the jaw (MRONJ). The x-axis represents the natural logarithm of the odds ratios (ln (ROR)), while the y-axis represents the common logarithm of the inverse *P*-value (−log10 [*P*]) from Fisher's exact test. The dotted line on the y-axis represents *P* = 0.05. The color of the plot represents the total number of adverse events reported for each drug. Drugs associated with MRONJ (MRONJ-positive drugs) are shown in the upper right part of the plot, while drugs not associated with MRONJ (MRONJ-negative drugs) are displayed in the upper left part.

**Table 1.** Twenty most frequently reported drugs as medication-related osteonecrosis of the jaw (MRONJ)-positive drugs.

| Drug name | Drug group | Number of MRONJ reports | *P*-value | ROR |
|---|---|---|---|---|
| Denosumab | Anti-RANKL antibody | 907 | <.0001 | 373.78 |
| Zoledronic acid | Bisphosphonates | 702 | <.0001 | 140.70 |
| Alendronic acid | Bisphosphonate | 264 | <.0001 | 36.69 |
| Ibandronic acid | Bisphosphonate | 92 | <.0001 | 35.77 |
| Sunitinib | Anticancer drugs | 65 | <.0001 | 10.30 |
| Dexamethasone | Corticosteroids | 60 | <.0001 | 4.12 |
| Cholecalciferol | Vitamin D | 55 | <.0001 | 21.13 |
| Bevacizumab | Anticancer drugs | 51 | <.0001 | 7.96 |
| Lenalidomide | Anticancer drugs | 47 | <.0001 | 3.07 |
| Everolimus | Anticancer drugs | 39 | <.0001 | 4.39 |
| Letrozole | Anticancer drugs | 38 | <.0001 | 10.34 |

 doi:10.20944/preprints202502.0246.v1



| | | | | |
|---|---|---|---|---|
| Prednisolone | Corticosteroids | 36 | <.0001 | 4.81 |
| Risedronic acid | Bisphosphonates | 33 | <.0001 | 16.50 |
| Exemestane | Anticancer drugs | 32 | <.0001 | 13.92 |
| Palbociclib | Anticancer drugs | 31 | <.0001 | 8.96 |
| Paclitaxel | Anticancer drugs | 29 | <.0001 | 3.42 |
| Calcium carbonate | Calcium | 28 | <.0001 | 18.19 |
| Docetaxel | Anticancer drugs | 25 | <.0001 | 2.64 |
| Prednisone | Corticosteroids | 25 | <.0001 | 3.97 |
| Pamidronic acid | Bisphosphonates | 22 | <.0001 | 54.81 |

*P*-value; Fisher's exact test; ROR, reported odds ratio.

### *2.3. QSAR Analysis Data Table*

A QSAR Analysis Data Table was created by incorporating 326 chemical structure descriptors for MRONJ-positive and -negative candidate drugs identified from the FAERS analysis data table (Supplementary Table S2). The Simplified Molecular Input Line-Entry System (SMILES) of 70 MRONJ-positive drugs and 139 MRONJ-negative drugs in FAERS was verified, and 326 types of chemical structure descriptors calculated using the Molecular Operating Environment, a chemical calculation environment, were included. The QSAR Analysis Data Table comprised 60 MRONJ-positive drugs and 108 MRONJ-negative drugs for which descriptors were available. Among the drugs affecting bone structure and mineralization (ATC code: M05B), all six drugs, including BPs, were classified as MRONJ-positive drugs (zoledronic acid, alendronic acid, ibandronic acid, risedronic acid, pamidronic acid, and minodronic acid).

### *2.4. QSAR Analysis Using Machine Learning (Construction of MRONJ-Induced Drug Prediction Model)*

In this study, QSAR analysis was conducted to evaluate the machine learning algorithm and the number of chemical structure descriptors to be incorporated into the prediction model. The analysis utilized all chemical structure descriptors for the machine learning algorithms random forest, gradient boosting, and artificial neural network to construct MRONJ-induced drug prediction models (Table 2). Default hyperparameter values of JMP analysis software were used for the three machine learning algorithms. For random forest, the hyperparameter conditions included 100 trees, 81 terms per branch, and a minimum branch size of 5, resulting in an area under the receiver operating characteristic curve (AUROC) of 0.726 for model validation. Gradient boosting utilized two branches per tree, 48 layers, and a learning rate of 0.02, with an AUROC of 0.714. The artificial neural network employed the Tan H activation function, 3 layers, and a learning rate of 0.1, achieving an AUROC of 0.741. Therefore, among the algorithms, the artificial neural network demonstrated the highest prediction accuracy (validation AUROC = 0.741; Table 2).

Furthermore, to enhance computational efficiency and prediction accuracy, we investigated the optimal number of chemical structure descriptors to be incorporated into the artificial neural network prediction model with the highest AUROC (Table 3). Due to the challenge of assessing the importance of each chemical structure descriptor in the artificial neural network prediction model, descriptors with the highest contribution rates from the random forest model were selected (Supplementary Table S3). Using the top 5, 6, 7, 8, 9, 10, 20, and 30 chemical structure descriptors with the highest contribution rates, the validation AUROCs were 0.699, 0.724, 0.719, 0.778, 0.761, 0.748, 0.724, and 0.716, respectively (Table 3). The model incorporating the top 8 chemical structure descriptors exhibited the highest prediction accuracy (validation AUROC = 0.778). The validation AUROC value of 0.778 indicates the success of our prediction model in identifying MRONJ-inducing drugs [30].

The eight key descriptors of the top-performing artificial neural network algorithm included ASA_P (total polar surface area), PEOE_VSA_FHYD (fractional hydrophobic dw surface area), PEOE_VSA-5 (total negative 5 dw surface area), h_pavgQ (average total charge), lip_acc (Lipinski Acceptor Count), vsa_acc (VDW acceptor surface area [A**2]), vsa_pol (VDW polar surface area [A**2]), and CASA- (charge-weighted negative surface area) (Table 4). Among these descriptors,

 doi:10.20944/preprints202502.0246.v1



ASA_P (total polar surface area) was higher in the MRONJ-positive drugs group than in the MRONJ-negative drugs group (Figure 2). Specifically, BPs and anticancer drugs exhibited higher values for ASA_P (total polar surface area) in the MRONJ-positive drug group (Table 5).

The accuracy rates of the 168 drugs incorporated into the MRONJ prediction model constructed in this study, categorized by drug efficacy group, are presented in Table 6 (top 13 drug classes) and Supplementary Table S4 (all drug classes). Notably, M05B Drugs affecting bone structure and mineralization and L04A Immunosuppressants achieved accuracy rates exceeding 80%. Conversely, "J05A Direct acting antivirals" and "N06A Antidepressants" had accuracy rates below 0.5.

The predictive model's performance was assessed by excluding drugs near the cutoff value in the ROC curve to define the applicability domain (Table 7). By setting the applicability domain, the model's reliability within a specific data range can be determined. Excluding drugs within ±0%, ±10%, and ±20% of the cutoff value resulted in 42, 32, and 17 drugs falling within the applicability domain, respectively, with balanced accuracies of 0.693, 0.750, and 0.800, F values of 0.593, 0.645, and 0.778, and Matthews correlation coefficients of 0.409, 0.488, and 0.618, respectively. Therefore, the performance of the MRONJ predictive model was improved by narrowing the applicability domain.

**Table 2.** Examination of the machine learning algorithms.

| Machine Learning Algorithms | AUROC of the training data | AUROC of the validation data | Cutoff value | Accuracy | Precision/positive predictive value | Negative predictive value | Recall/Sensitivity | Specificity | Balanced accuracy | F1-score | Matthews correlation coefficient |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Random Forest | 0.996 | 0.726 | 0.533 | 0.714 | 0.600 | 0.778 | 0.600 | 0.778 | 0.689 | 0.600 | 0.378 |
| Gradient Boosting | 0.956 | 0.714 | 0.484 | 0.714 | 0.636 | 0.742 | 0.467 | 0.852 | 0.659 | 0.538 | 0.347 |
| Artificial Neural Networks | 0.849 | 0.741 | 0.526 | 0.714 | 0.579 | 0.826 | 0.733 | 0.704 | 0.719 | 0.647 | 0.421 |

**Table 3.** Examination of the number of chemical structure descriptors in the artificial neural networks.

| Number of chemical structure descriptors * | AUROC of the training data | AUROC of the validation data | Cutoff value | Accuracy | Precision/positive predictive value | Negative predictive value | Recall/Sensitivity | Specificity | Balanced accuracy | F1-score | Matthews correlation coefficient |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 Descriptors | 0.713 | 0.699 | 0.363 | 0.667 | 0.600 | 0.676 | 0.200 | 0.926 | 0.563 | 0.300 | 0.186 |
| 6 Descriptors | 0.837 | 0.724 | 0.479 | 0.714 | 0.600 | 0.778 | 0.600 | 0.778 | 0.689 | 0.600 | 0.378 |
| 7 Descriptors | 0.703 | 0.719 | 0.554 | 0.714 | 0.615 | 0.759 | 0.533 | 0.815 | 0.674 | 0.571 | 0.361 |
| 8 Descriptors | 0.871 | 0.778 | 0.291 | 0.738 | 0.667 | 0.767 | 0.533 | 0.852 | 0.693 | 0.593 | 0.409 |
| 9 Descriptors | 0.871 | 0.761 | 0.383 | 0.714 | 0.600 | 0.778 | 0.600 | 0.778 | 0.689 | 0.600 | 0.378 |
| 10 Descriptors | 0.877 | 0.748 | 0.265 | 0.667 | 0.533 | 0.741 | 0.533 | 0.741 | 0.637 | 0.533 | 0.274 |
| 20 Descriptors | 0.786 | 0.724 | 0.463 | 0.762 | 0.778 | 0.758 | 0.467 | 0.926 | 0.696 | 0.583 | 0.458 |
| 30 Descriptors | 0.777 | 0.716 | 0.274 | 0.738 | 0.667 | 0.767 | 0.533 | 0.852 | 0.693 | 0.593 | 0.409 |

*Chemical structure descriptors incorporated into the artificial neural network were used with the top contributions in the random forest.

 doi:10.20944/preprints202502.0246.v1



**Table 4.** Eight chemical structure descriptors contributed to the MRONJ prediction model.

| Descriptor | Descriptor description | Number of branches* |
|---|---|---|
| ASA_P | Total polar surface area | 7 |
| PEOE_VSA_FHYD | Fractional hydrophobic dw surface area | 3 |
| PEOE_VSA-5 | Total negative 5 dw surface area | 3 |
| h_pavgQ | Total average charge (pH = 7) | 3 |
| lip_acc | Lipinski Acceptor Count | 3 |
| vsa_acc | VDW acceptor surface area (A**2) | 2 |
| vsa_pol | VDW polar surface area (A**2) | 2 |
| CASA- | Charge-weighted negative surface area | 2 |

*Number of splits in the random forest.


**Figure 2.** Comparison of the descriptor ASA_P values between MRONJ-positive and -negative drugs.

**Table 5.** Top 20 MRONJ-positive drugs by descriptor ASA_P values.

| Drug name | ATC code | Drug group | ASA_P |
|---|---|---|---|
| Detirelix | L02BX02 | Anticancer drug (hormone-related drugs) | 533.2 |
| Triptorelin | L02AE04 | Anticancer drug (hormone-related drugs) | 509.8 |
| Leuprorelin | L02AE02 | Anticancer drug (hormone-related drugs) | 412.1 |
| Cefcapene | J01DD17 | Antibiotics | 322.2 |
| Pamidronic acid | M05BA03 | Bisphosphonates | 305.2 |
| Alendronic acid | M05BA04 | Bisphosphonates | 302.7 |
| Pemetrexed | L01BA04 | Anticancer drug (metabolic antagonists) | 294.9 |
| Docetaxel | L01CD02 | Anticancer drug (taxanes) | 284.6 |
| Melphalan | L01AA03 | Anticancer drug (alkylating agents) | 283.3 |
| Bicalutamide | L02BB03 | Anticancer drug (hormone-related drugs) | 267.5 |
| Epacadostat | L01XX58 | Anticancer drug (others) | 267.1 |
| Paclitaxel | L01CD01 | Anticancer drug (taxanes) | 265.3 |
| Zoledronic acid | M05BA08 | Bisphosphonates | 263.6 |
| Temsirolimus | L01EG01 | Anticancer drug (protein kinase inhibitors) | 260.5 |
| Allelism | L01EM03 | Anticancer drug (protein kinase inhibitors) | 259.5 |
| Anastrozole | L02BG03 | Anticancer drug (hormone-related drugs) | 256.4 |
| Fulvestrant | L02BA03 | Anticancer drug (hormone-related drugs) | 254.4 |
| Ibandronic acid | M05BA06 | Bisphosphonates | 252.1 |
| Risedronic acid | M05BA07 | Bisphosphonates | 244.2 |
| Capecitabine | L01BC06 | Anticancer drug (metabolic antagonists) | 240.3 |

 doi:10.20944/preprints202502.0246.v1



*Number of splits in the random forest.

**Table 6.** Accuracy rate by therapeutic class of drugs incorporated into the MRONJ prediction model.

| Drug classes in the ATC classification | FAERS analysis data table | Classification results for the MRONJ prediction model | | |
|---|---|---|---|---|
| | Number of drugs (positive/negative) | Positive | Negative | Accuracy |
| L01E PROTEIN KINASE INHIBITORS | 14 (11/3) | 13 | 1 | 0.75 |
| L02B HORMONE ANTAGONISTS AND RELATED AGENTS | 7 (7/0) | 6 | 1 | 0.75 |
| L01X OTHER ANTINEOPLASTIC AGENTS | 6 (6/0) | 5 | 1 | 0.71 |
| M05B DRUGS AFFECT BONE STRUCTURE AND MINERALIZATION | 6 (6/0) | 6 | 0 | 1.00 |
| L04A IMMUNOSUPPRESSANTS | 9 (5/4) | 4 | 5 | 0.80 |
| A11C VITAMIN A AND D, INCL. COMBINATIONS OF THE TWO | 4 (4/0) | 4 | 0 | 1.00 |
| H02A CORTICOSTEROIDS FOR SYSTEMIC USE, PLAIN | 4 (4/0) | 4 | 0 | 1.00 |
| L01C PLANT ALKALOIDS AND OTHER NATURAL PRODUCTS | 4 (4/0) | 4 | 0 | 1.00 |
| R01A DECONGESTANTS AND OTHER NASAL PREPARATIONS FOR TOPICAL USE | 7 (3/4) | 5 | 2 | 0.56 |
| D07A CORTICOSTEROIDS, PLAIN | 5 (3/2) | 5 | 0 | 0.43 |
| A07E INTESTINAL ANTIINFLAMMATORY AGENTS | 4 (3/1) | 4 | 0 | 0.60 |
| C05A AGENTS FOR TREATMENT OF HEMORRHOIDS AND ANAL FISSURES FOR TOPICAL USE | 4 (3/1) | 3 | 1 | 1.00 |
| S01B ANTIINFLAMMATORY AGENTS | 4 (3/1) | 3 | 1 | 1.00 |

**Table 7.** Predictive performance of the MRONJ predictive models for compounds within the application domain.

| Applicability Domain | Number of drugs in applicability domain | Accuracy | Precision/positive predictive value | Negative predictive value | Recall/sensitivity | Specificity | Balanced accuracy | F1-score | Matthews correlation coefficient |
|---|---|---|---|---|---|---|---|---|---|
| Exclusion: Cutoff value 0.5 ± 0 (No exclusion) | 42 | 0.738 | 0.667 | 0.767 | 0.533 | 0.852 | 0.693 | 0.593 | 0.409 |
| Exclusion: Cutoff value 0.5 ± 0.1 (Applicability: 40–60% exclusion) | 32 | 0.656 | 0.476 | 1.000 | 1.000 | 0.500 | 0.750 | 0.645 | 0.488 |
| Exclusion: Cutoff value 0.5 ± 0.2 (Applicability: 30–70% exclusion) | 17 | 0.765 | 0.636 | 1.000 | 1.000 | 0.600 | 0.800 | 0.778 | 0.618 |

## 3. Discussion

### *3.1. Analysis of the Adverse Drug Reaction Database FAERS*

In this study, we developed a prediction model to classify MRONJ-inducing drugs based solely on their structural information using the FAERS adverse drug reaction database and a machine learning algorithm. To the best of our knowledge, this is the first study to build an MRONJ-inducing drug prediction model utilizing an adverse drug reaction database. Due to the challenge of accurately determining MRONJ risk from the FAERS database, we assessed it using three indicators: reported odds ratio (ROR), Fisher's exact test, and the total number of reports for each drug. ROR is widely used for signal detection of adverse events in adverse drug event databases. However, ROR is susceptible to inflation and false signal detection when the number of reports is limited. Therefore, in this study, in addition to ROR, we comprehensively evaluated MRONJ risk by incorporating Fisher's exact test and the total number of reports.

 doi:10.20944/preprints202502.0246.v1



From the FAERS analysis data table, 80 drugs were identified as MRONJ-positive and 139 as MRONJ-negative. Among the MRONJ-positive drugs, frequently reported drugs included BPs such as zoledronic acid, alendronic acid, and ibandronic acid, the anti-receptor activator of nuclear factor kappa B ligand (RANKL) antibody denosumab, anticancer drugs like sunitinib, bevacizumab, everolimus, and letrozole, as well as corticosteroids such as dexamethasone and prednisolone (Table 1). BPs (ATC code: M05B) have a strong affinity for bone hydroxyapatite, inhibit osteoclast activity, reduce bone resorption, and are used to treat osteoporosis and malignant tumors [31]. BPs are closely associated with MRONJ [2]. Monitoring the use of BPs for both malignant tumors and osteoporosis is crucial. Denosumab, an anti-RANKL antibody, has also been linked to MRONJ [32,33]. However, denosumab was not included in the QSAR analysis due to its nature as an antibody preparation. Several protein kinase inhibitors (ATC code: L01E) used in anticancer therapy were also found to be associated with MRONJ, with sunitinib exacerbating MRONJ in renal cell carcinoma [34]. Antiangiogenic drugs also contribute to MRONJ development [35], with varying effects depending on the drug's mechanism of action [36]. Corticosteroids such as dexamethasone and prednisolone (ATC code: D07A) have been shown to increase the risk of developing MRONJ [37] by delaying wound healing through immunosuppression and altering oral microbiota, increasing the risk of oral infections and MRONJ [32,38]. Selective estrogen receptor modulators, oral contraceptives, and sex hormone preparations may also influence MRONJ development. Estrogen, a sex hormone, has been shown to impact bone remodeling, potentially affecting jaw bone remodeling [39]. This study highlights the potential association between various drugs and MRONJ. While MRONJ is commonly linked to bone resorption inhibitors and antiangiogenic drugs, other medications have also been reported to induce this condition [14,40]. In this study, all drugs registered in the adverse event database were comprehensively examined under the same analysis conditions. The fact that some bone resorption inhibitors and antiangiogenic drugs were detected as MRONJ-positive drugs using this analysis method ensures the reliability of other detected drugs.

### *3.2. Construction of the MRONJ-Induced Drug Prediction Model*

In this study, we constructed an MRONJ-induced drug prediction model using an artificial neural network machine learning algorithm and eight chemical structure descriptors, achieving a validation AUROC of 0.871 (Table 3). Our goal was to improve the prediction model's accuracy by comparing three different machine learning algorithms and determining the optimal number of chemical structure descriptors. To address the challenge of assessing the individual contribution of each descriptor in the artificial neural network, a prediction model was constructed using chemical structure descriptors with a large contribution rate in a random forest. The best MRONJ prediction model demonstrated a negative predictive value of 0.767, specificity of 0.852, and minimal false negatives (Table 3). Notably, the prediction model exhibited high accuracy rates for drug categories affecting bone structure and mineralization (ATC code: M05B) and immunosuppressants (ATC code: L04A), including BPs, known to be associated with MRONJ development (Table 6). Our findings suggest that the developed MRONJ-inducing drug prediction model can effectively identify important MRONJ-positive drugs while excluding MRONJ-negative drugs, thereby reducing the risk of overlooking critical medications.

The MRONJ-inducing drug prediction model selected eight molecular descriptors, many of which were related to the polar surface area of the compound (Table 4). The MRONJ-positive drug group exhibited higher values for ASA_P (total polar surface area), a descriptor incorporated into the MRONJ prediction model constructed in this study, than the MRONJ-negative drug group (Figure 2). Specifically, the MRONJ-positive drug group, which included BPs and anticancer drugs, showed elevated levels of the descriptor (ASA_P: total polar surface area; Table 5). This suggests that drugs with strong polar surface area characteristics, such as BPs and anticancer drugs, may contribute to MRONJ development. BPs, known to be closely associated with MRONJ, are highly polar drugs due to the presence of hydroxyl groups of BPs that have a strong affinity for hydroxyapatite, the main component of bone tissue, leading to their specific adsorption to bone tissue [41]. BPs bound to bone

 doi:10.20944/preprints202502.0246.v1



tissue are taken up by osteoclasts, inhibiting their function and reducing bone resorption [42,43]. It has been reported that excessive bone resorption inhibition may be related to MRONJ [44]. The presence of hydroxyl groups contributes to a larger polar surface area in BPs, as these groups are known to have a high affinity for bone hydroxyapatite. Similarly, other drugs containing hydroxyl groups may also impact MRONJ by adsorbing the hydroxyl groups to bone hydroxyapatite. Anticancer drugs, which are designed to target specific molecules, often have a large polar surface area [45]. Among the 14 anticancer drugs analyzed, 11 protein kinase inhibitors (ATC code: L01E) were found to be MRONJ-positive drugs (Supplementary Table S1).

The applicability domain of the MRONJ-induced drug prediction model was confirmed to exclude probabilities close to the cutoff value of the ROC curve. By setting the applicability domain within a chemical space, drugs falling within this region can yield reliable prediction outcomes. Consequently, narrowing the application region enhanced the model's performance.

### *3.3. Limitations*

Drug adverse reaction databases such as FAERS have biases and limitations in information collection. FAERS relies on self-reported adverse drug reaction information, which introduces human biases such as reporting bias [46], making it challenging to accurately assess drug risks. Moreover, the detection of false signals of adverse drug reactions may occur when multiple drugs are used [47]. Therefore, in this study, we carefully identified MRONJ-inducing drugs by employing time-series data cleaning methods that consider the characteristics of FAERS and three evaluation indicators (ROR, Fisher's exact test, and total number of reports) for signal detection.

The quality of a machine learning-based prediction model depends on the quality of the input data [48]. Ideally, reliable data encompassing both positive and negative MRONJ drugs should be included in the learning dataset for QSAR analysis. However, the FAERS data utilized in this study may contain inappropriate reports, leading to limitations in the prediction model's accuracy. Additionally, due to the rarity of MRONJ, the number of reports in FAERS was limited, resulting in a constrained number of MRONJ-positive drugs in the QSAR analysis dataset. Consequently, the prediction model's applicability domain in this study may be limited [49,50]. Furthermore, the lack of patient information, such as the genetic background of patients, poses a challenge in explaining individual differences in the onset of adverse events.

## 4. Materials and Methods

### *4.1. Creation of the FAERS Analysis Data Table*

For the analysis in this study, data reported to FAERS from January 2004 to March 2022 were utilized to create data tables for FAERS analysis (Figure 3). The adverse event data reported to FAERS are stored in seven data tables. In this study, four data tables were utilized: the Drug table (containing drug information), the Reaction table (providing adverse event information), the Demographic table (offering basic case information), and the Therapy table (presenting treatment duration information), with duplicate reports removed [26,28]. The drugs in the Drug table were categorized into first and second suspected drugs, concomitant drugs and interactions, with only the first and second suspected drugs being considered in this study. World Health Organization drug classification ATC codes were assigned to each drug to facilitate drug effect tabulation [51,52]. The Reaction table documented adverse events according to the ICH International Glossary of Pharmaceutical Terms (Medical Dictionary for Regulatory Activities version 25.0; MedDRA ver. 25.0) based on the preferred term [53,54]. In this study, the adverse event "osteonecrosis of the jaw" in the Reaction table was defined as "medication-related osteonecrosis of the jaw," with a column added to indicate whether it was MRONJ or not. The Drug and Therapy tables were initially joined using "Primary ID" and "Drug sequence," followed by the joining of the Reaction and Demographic tables using "Primary ID." Additionally, for data cleaning purposes, only data from the Demographic table with adverse

 doi:10.20944/preprints202502.0246.v1



event onset dates falling within the Therapy table treatment start and end dates were extracted to create a consistent time-series data table for FAERS analysis.

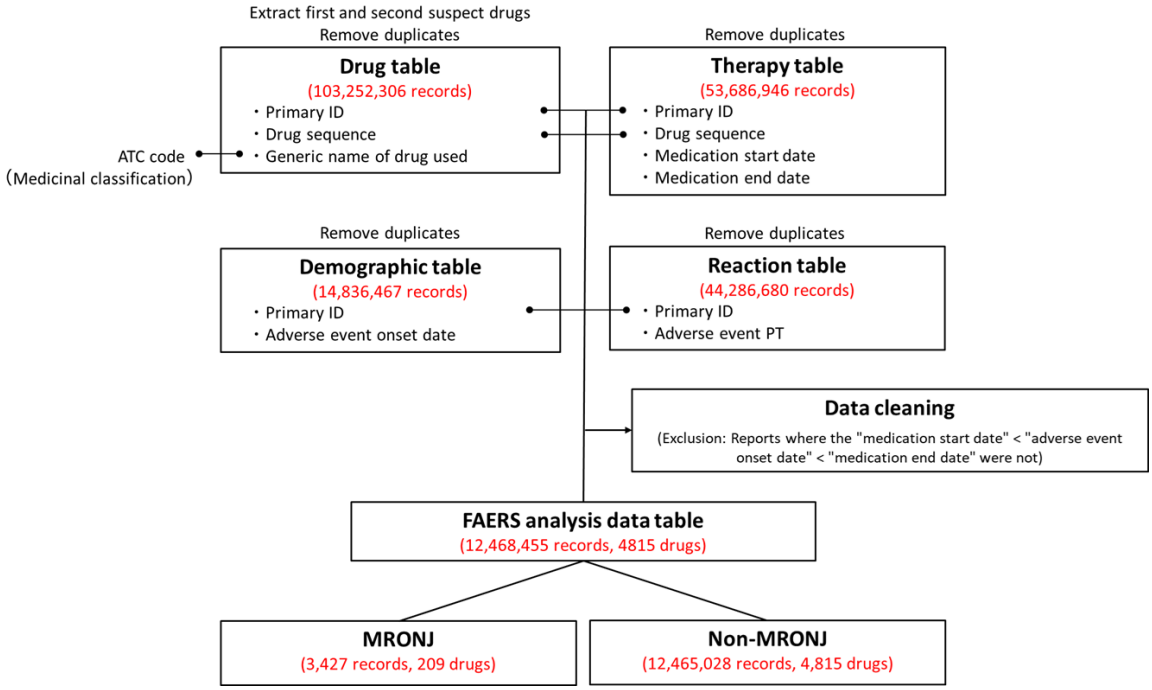


**Figure 3.** Procedure for creating the U.S. Food and Drug Administration Adverse Drug Reaction Database System (FAERS) Analysis Data Table. Duplicate data were removed from the Drug, Therapy, Demographic, and Reaction tables. Only the "first suspected drug" and "second suspected drug" were extracted from the DRUG table. Initially, the Drug and Therapy tables were merged using the Primary ID and Drug sequence. Subsequently, the Demographic and Reaction tables were joined using the Primary ID. To ensure data accuracy, reports that did not adhere to the order of treatment start date, adverse event onset date, and treatment end date were excluded. Out of the 12,468,455 reports in the FAERS analysis data table, 3,427 were related to the MRONJ.

### *4.2. Examination of the FAERS Analysis Data Tables (Extraction of Positive and Negative MRONJ Drugs)*

The drugs in the FAERS Analysis Data Tables were assessed using three indices: the Reporting Odds Ratio and Fisher's exact test, along with the total number of reports for each drug. Initially, a 2 × 2 contingency table for MRONJ was created for each of the 4,815 drugs in the FAERS analysis data table, and the *P*-values for ROR and Fisher's exact test were calculated (Figure 4). To stabilize the estimate, a correction was applied by adding 0.5 to all cells (Haldane Anscombe 1/2 correction) [55,56].

ROR is a key indicator in nonproportional analysis methods utilized for detecting adverse drug signals in pharmacovigilance [57]. It offers high sensitivity and low bias, enabling the estimation of the association between the drug and the adverse event [19]. However, classical signal detection indicators such as ROR may overestimate the signal and lead to unstable statistical estimates in cases of low reporting [58,59]. To address this, Eudra Vigilance guidelines recommend a minimum number of reports to ensure a stable signal [60]. In the present study, a threshold of 100 reports (Figure 4; a+b ≥ 100) was set for the total number of reports for each drug (Figure 4; a+b) to prevent the oversight of commonly used drugs [61]. In addition, Fisher's exact test was used to assess the independence of drugs and MRONJ in the 2 × 2 contingency table in Figure 4. Consequently, the criteria for identifying MRONJ-positive drugs included ROR > 1, Fisher's exact test *P*-values <0.05, and total adverse event reports ≥100, while MRONJ-negative drugs met the criteria of ROR <1, Fisher's exact test *P*-values <0.05, and total adverse event reports ≥100.

In addition, this study utilized a scatter plot (volcano plot) to visualize the MRONJ-positive and -negative drug candidates from the FAERS analysis data table. Volcano plots, commonly used in bioinformatics to analyze gene expression trends, were employed in this study [61–63]. In the volcano

 doi:10.20944/preprints202502.0246.v1



plot, the x-axis represents the natural logarithm of the reported odds ratio (lnROR), while the y-axis represents the ordinary logarithm [−log (*P*-value)] of Fisher's exact test *P*-value. The x-axis indicates the risk of MRONJ development when lnROR > 0 (ROR > 1), while the y-axis indicates a −log *P*-value >1.3 ($P < 0.05$), indicating a significant difference in the 2 × 2 contingency table shown in Figure 4. Each point on the plot represents a drug, with the color of the point indicating the total number of adverse event reports (a+b in Figure 4), where drugs with a high number of reports are depicted in red and those with a low number in blue. Only drugs with 100 or more adverse event reports were included in the analysis. Therefore, MRONJ-positive drugs are located in the upper right-hand corner of the plot, while MRONJ-negative drugs are in the upper left-hand corner.

| | Reports with a suspected adverse event | Reports without a suspected adverse event |
|---|---|---|
| Report with a suspected drug | a | b |
| All other reports | c | d |

ROR (reporting odds ratio) = a × d / b × c

**Figure 4.** Cross-tabulation and formula used to calculate the ROR for an adverse event. The table is organized with reports for the suspected drug, all other reports, reports with an adverse event, and reports without an adverse event (a–d represent the number of reports).

*4.3. Creation of QSAR Analysis Data Tables (Addition of Chemical Structure Descriptors)*

Data tables for the QSAR analysis were created by incorporating chemical structure descriptors for MRONJ-positive and -negative drugs (Figure 5, Supplementary Table S2). The chemical structures of the drugs were obtained from the PubChem compound database in the form of SMILES, a linear representation of molecular structures [64]. Chemical structure descriptors were calculated using the Molecular Operating Environment version 2022.02 (Chemical Computing Group, Inc, Montreal, Canada) [65], a specialized chemical computing platform. Prior to descriptor calculations, water molecules and counter ions were eliminated through desalting. Each drug was converted to a three-dimensional structure, assessed for partial charge, and optimized using force field calculations (Amber 10 EHT). A total of 326 chemical structure descriptors were calculated for each drug. Descriptor variables with missing values or perfect collinearity ($r^2 = 1$) were excluded. Mixtures, large peptides, bacterial preparations, inorganic compounds, organometallic compounds, and drugs with unspecified names or abbreviations were also removed. Enoxaparin was excluded due to duplication in the dataset. Consequently, the data table for the QSAR analysis comprised 326 chemical structure descriptors for 60 MRONJ-positive and 108 MRONJ-negative drugs. To validate the model, the data table was randomly divided into a 3:1 ratio for training and validation purposes.

 doi:10.20944/preprints202502.0246.v1



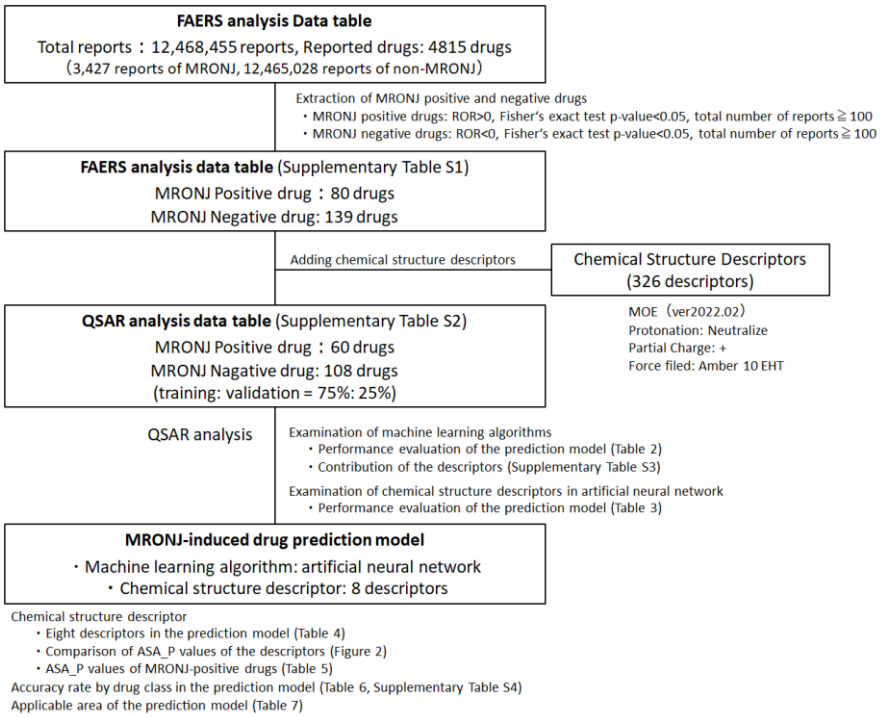


**Figure 5.** Procedure for creating the Quantitative Structure-Activity Relationship (QSAR) Analysis Data Table. Positive and negative drugs for MRONJ were estimated based on the drugs listed in the FAERS analysis data table. A total of 326 chemical structure descriptions, representing structural features, were added to the positive and negative drugs for MRONJ. In the QSAR analysis, models were constructed and compared using three machine learning algorithms and varying numbers of descriptors. The selection of chemical structure for the artificial neural network was guided by their contribution in the random forest model. The model with the highest performance was checked for accuracy, the significance of the descriptors used, accuracy rates across pharmacological groups of the incorporated drugs, and the applicability domain.

### *4.4. QSAR Analysis Using Machine Learning Algorithms (Construction of MRONJ-Induced Drug Prediction Model)*

A MRONJ-induced drug prediction model was constructed through QSAR analysis using machine learning algorithms (Figure 5). The algorithms considered were random forests, gradient boosting, and artificial neural networks, all available in the JMP analysis software. Each algorithm has distinct approaches and characteristics. Random Forest [66] and gradient boosting [67] are ensemble learning methods that combine several weak learners, such as decision trees. Random Forest is known for its stability and improved accuracy through bagging [66], while gradient boosting achieves high prediction performance through sequential error correction by boosting [67]. On the other hand, artificial neural networks consist of multi-layered structures with input, hidden, and output layers containing multiple neurons, enabling them to learn complex nonlinear relationships [68]. In this study, the artificial neural network was constructed using a multilayer perceptron neural network with a back-propagation algorithm for nonlinear regression. Boosting was employed as the ensemble method for the artificial neural network. It is crucial to utilize these algorithms differently due to their unique approaches and characteristics. In this study, the prediction models were built using default values for chemical structure descriptors and hyperparameters for each algorithm.

Furthermore, the study investigated the optimal number of chemical structure descriptors in the artificial neural network. It is beneficial to develop predictive models with fewer descriptors for computational efficiency and explainability. Artificial neural networks are effective in capturing nonlinear relationships, but determining the significance of each descriptor in the model can be challenging. On the other hand, random forests utilize decision trees to assess feature importance. In this study, the top chemical structure descriptors with the largest contribution in the random forest

 doi:10.20944/preprints202502.0246.v1



model were selected and integrated into the artificial neural network algorithm [69,70]. In addition, we analyzed the descriptors used in constructing the MRONJ predictive model and interpreted the drug characteristics associated with MRONJ.

The prediction performance of the MRONJ-induced drug prediction model was evaluated using metrics such as AUROC, accuracy, precision (positive predictive value), negative predictive value, recall-sensitivity, specificity, balanced accuracy, F1-score, and Matthews correlation coefficient. The hold-out method was employed for validation.

The applicability domain was assessed by determining the cutoff value on the ROC curve of the artificial neural network prediction model [71]. Defining the scope of application helps to establish the reliable prediction range of the constructed model. In this study, the cutoff value was determined using Youden's index [72] and normalized to 0.5. Additionally, the model's performance within the applicability domain was assessed by building a predictive model that excluded drugs with deviations of ±0%, ±10%, and ±20% from the cutoff value.

### *4.5. Statistical Analysis*

All analyses were conducted using JMP Pro 16.2.0 (SAS Institute Inc., NC, USA), and a *P*-value less than 0.05 was considered significant.

## 5. Conclusions

In this study, an MRONJ-induced drug prediction model was constructed using chemical structure information, the FEARS database of drug adverse events, and machine learning. The model, based on an artificial neural network algorithm and eight chemical structure descriptors, identified drugs with polar surface area characteristics as potential contributors to MRONJ. These findings could enhance risk assessment in clinical trials and postmarketing surveillance, as well as streamline screening in new drug development.

**Supplementary Materials:** The following supporting information can be downloaded at the website of this paper posted on Preprints.org., Table S1: 70 MRONJ-positive drugs and 139 MRONJ-negative drugs. Table S2: QSAR analysis data table. Table S3: Contribution of chemical structure descriptors in random forests. Table S4: Accuracy rate of the MRONJ prediction model by ATC classification.

**Author Contributions:** Conceptualization, S.T. and Y.U.; methodology, Y.U.; software, Y.U.; validation, S.T., K.S., M.Y. and Y.U.; formal analysis, S.T. and Y.U.; investigation, S.T. and Y.U.; resources, S.T. and Y.U.; data curation, S.T. and Y.U.; writing—original draft preparation, S.T., K.S., M.Y. and Y.U.; writing—review and editing, S.T., K.S., M.Y. and Y.U.; visualization, S.T. and Y.U.; supervision, S.T. and Y.U.; project administration, Y.U.; funding acquisition, Y.U. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Ethical review and approval were waived for this study by the Ethics Committee of Meiji Pharmaceutical University as the study used anonymized data from an open-access database.

**Informed Consent Statement:** Patient consent was waived by the Ethics Committee of Meiji Pharmaceutical University as the study used anonymized data from an open-access database.

**Data Availability Statement:** Data are contained within the article. Data from the FAERS database were downloaded from the website of the U.S. Food and Drug Administration (FDA) (https://www.fda.gov/drugs/drug-approvals-and-databases/fda-adverse-event-reporting-system-faers-database) in June 2022.

 doi:10.20944/preprints202502.0246.v1



**Acknowledgments:** The authors thank the staff of the National Hospital Organization Kanagawa Hospital and the Department of Medical Molecular informatics at Meiji Pharmaceutical University as well as our families for their support. The authors would like to thank Enago (www.enago.jp) for the English language review.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| AUROC | Area under the receiver operating characteristic curve |
| FDA | Food and Drug Administration |
| MRONJ | Medication-related osteonecrosis of the jaw |
| QSAR | Quantitative structure-activity relationship |
| RANKL | Receptor activator of nuclear factor kappa B ligand |
| ROR | Reported odds ratio |
| SMILES | The Simplified Molecular Input Line-Entry System |

## References

1. Khan, A.A.; Morrison, A.; Hanley, D.A.; Felsenberg, D.; McCauley, L.K.; O'Ryan, F.; Reid, I.R.; Ruggiero, S.L.; Taguchi, A.; Tetradis, S.; et al. International Task Force on Osteonecrosis of the Jaw. Diagnosis and management of osteonecrosis of the jaw: a systematic review and international consensus. *J Bone Miner Res.* **2015,** *30*, 3-23. doi: 10.1002/jbmr.2405
2. Ruggiero, S.L.; Dodson, T.B.; Aghaloo, T.; Carlson, E.R.; Ward, B.B, Kademani, D. American Association of Oral and Maxillofacial Surgeons' Position Paper on Medication-Related Osteonecrosis of the Jaws-2022 Update. *J Oral Maxillofac Surg.* **2022,** *80*, 920-943. doi: 10.1016/j.joms.2022.02.008
3. Ishimaru, M.; Ono, S.; Morita, K.; Matsui, H.; Hagiwara, Y.; Yasunaga, H. Prevalence, Incidence Rate, and Risk Factors of Medication-Related Osteonecrosis of the Jaw in Patients with Osteoporosis and Cancer: A Nationwide Population-Based Study in Japan. *J Oral Maxillofac Surg.* **2022**, *80*, 714-727. doi: 10.1016/j.joms.2021.12.007
4. King, R.; Tanna, N.; Patel, V. Medication-related osteonecrosis of the jaw unrelated to bisphosphonates and denosumab-a review. *Oral Surg Oral Med Oral Pathol Oral Radiol.* **2019**, *127*, 289-299. doi: 10.1016/j.oooo.2018.11.012
5. Anastasilakis, A.D.; Pepe, J.; Napoli, N.; Palermo, A.; Magopoulos, C.; Khan, A.A.; Zillikens, MC.; Body, J.J. Osteonecrosis of the Jaw and Antiresorptive Agents in Benign and Malignant Diseases: A Critical Review Organized by the ECTS. *J Clin Endocrinol Metab.* **2022**, *107*, 1441-1460. doi: 10.1210/clinem/dgab888
6. Japanese Allied Committee on Osteonecrosis of the Jaw; Yoneda, T.; Hagino, H.; Sugimoto, T.; Ohta, H.; Takahashi, S.; Soen, S.; Taguchi, A.; Nagata, T.; Urade, M.; Shibahara, T.; et al. Antiresorptive agent-related osteonecrosis of the jaw: Position Paper 2017 of the Japanese Allied Committee on Osteonecrosis of the Jaw. *J Bone Miner Metab.* **2017**, *35*, 6-19. doi: 10.1007/s00774-016-0810-7
7. Schiodt, M.; Otto, S.; Fedele, S.; Bedogni, A.; Nicolatou-Galitis, O.; Guggenberger, R.; Herlofson, B.B.; Ristow, O.; Kofod, T. Workshop of European task force on medication-related osteonecrosis of the jaw-Current challenges. *Oral Dis.* **2019**, *25*, 1815–1821. doi: 10.1111/odi.13160
8. Campisi, G.; Mauceri, R.; Bertoldo, F.; Bettini, G.; Biasotto, M.; Colella, G.; Consolo, U.; Di Fede, O.; Favia, G.; Fusco, V.; et al. Medication-Related Osteonecrosis of Jaws (MRONJ) Prevention and Diagnosis: Italian Consensus Update 2020. *Int J Environ Res Public Health.* **2020**, *17*, 5998. doi: 10.3390/ijerph17165998
9. Edwards, I.R.; Lindquist, M.; Wiholm, B.E.; Napke, E. Quality criteria for early signals of possible adverse drug reactions. *Lancet.* **1990**, *336*, 156-158. doi: 10.1016/0140-6736(90)91669-2
10. Kimura, K.; Kikegawa, M.; Kan, Y.; Uesawa, Y. Identifying Crude Drugs in Kampo Medicines Associated with Drug-Induced Liver Injury Using the Japanese Adverse Drug Event Report Database: A Comprehensive Survey. *Pharmaceuticals (Basel).* **2023**, *16*, 678. doi: 10.3390/ph16050678
11. Toriumi, S.; Mimori, R.; Sakamoto, H.; Sueki, H.; Yamamoto, M.; Uesawa, Y. Examination of Risk Factors and Expression Patterns of Atypical Femoral Fractures Using the Japanese Adverse Drug Event Report

Database: A Retrospective Pharmacovigilance Study. *Pharmaceuticals (Basel).* **2023**, *16*, 626. doi: 10.3390/ph16040626

12. Fang, H.; Su, Z.; Wang, Y.; Miller, A.; Liu, Z.; Howard, P.C.; Tong, W.; Lin, SM. Exploring the FDA adverse event reporting system to generate hypotheses for monitoring of disease characteristics. *Clin Pharmacol Ther.* **2014**, *95*, 496-8. doi: 10.1038/clpt.2014.17
13. Hosoya, R.; Ishii-Nozawa, R.; Terajima, T.; Kagaya, H.; Uesawa, Y. The Association between Molecular Initiating Events and Drug-Induced Hiccups. *Pharmaceuticals (Basel).* **2024**, *17*, 379. doi: 10.3390/ph17030379
14. Ahdi, H.S.; Wichelmann, T.A.; Pandravada, S.; Ehrenpreis, E.D. Medication-induced osteonecrosis of the jaw: a review of cases from the Food and Drug Administration Adverse Event Reporting System (FAERS). *BMC Pharmacol Toxicol.* **2023**, *24*, 15. doi: 10.1186/s40360-023-00657-y
15. Nathan, K.T.; Conn, K.M.; van Manen, R.P.; Brown, J.E. Signal detection for bleeding associated with the use of direct oral anticoagulants. *Am J Health Syst Pharm.* **2018**, *75*, 973-977. doi: 10.2146/ajhp170529
16. Sanagawa, A.; Hotta, Y.; Kondo, M.; Nishikawa, R.; Tohkin, M.; Kimura, K. Tumor lysis syndrome associated with bortezomib: A post-hoc analysis after signal detection using the US Food and Drug Administration Adverse Event Reporting System. *Anticancer Drugs.* **2020**, *2*, 183-189. doi: 10.1097/CAD.0000000000000862
17. Evans, S.J.; Waller, P.C.; Davis, S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiol Drug Saf.* **2001**, *6*, 483-6. doi: 10.1002/pds.677
18. Sakaeda, T.; Tamon, A.; Kadoyama, K.; Okuno, Y. Data mining of the public version of the FDA Adverse Event Reporting System. *Int J Med Sci.* **2013**, *7*, 796-803. doi: 10.7150/ijms.6048
19. van Puijenbroek, E.P.; Bate, A.; Leufkens, H.G.; Lindquist, M.; Orre, R.; Egberts, A.C. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. *Pharmacoepidemiol Drug Saf.* **2002**, *11*, 3-10. doi: 10.1002/pds.668
20. Toriumi, S.; Kobayashi, A.; Sueki, H.; Yamamoto, M.; Uesawa, Y. Exploring the Mechanisms Underlying Drug-Induced Fractures Using the Japanese Adverse Drug Event Reporting Database. *Pharmaceuticals (Basel).* **2021**, *14*, 1299. doi: 10.3390/ph14121299
21. FDA Adverse Event Reporting System (FAERS). Available online: https://www.fda.gov/drugs/drug-approvals-and-databases/fda-adverse-event-reporting-system-faers (accessed on 5 June 2022).
22. Zhang, S. Computer-aided drug discovery and development. *Methods Mol Biol.* **2011**, *716*, 23-38. doi: 10.1007/978-1-61779-012-6_2
23. Kunimoto, R.; Bajorath, J.; Aoki, K. From traditional to data-driven medicinal chemistry: A case study. *Drug Discov Today.* **2022**, *8*, 2065-2070. doi: 10.1016/j.drudis.2022.04.017
24. Hansch, C.; Maloney, P.; Fujita, T.; Muir, R.M. Correlation of Biological Activity of Phenoxyacetic Acids with Hammett Substituent Constants and Partition Coefficients. *Nature* **1962**, *194*, 178–180. doi: 10.1038/194178b0
25. Hansch, C.; Fujita, T. p-σ-π Analysis. A Method for the Correlation of Biological Activity and Chemical Structure. *Journal of the American Chemical Society*, **1964**, *86*, 1616-1626. doi: 10.1021/ja01062a035
26. Hansen, K.; Mika, S.; Schroeter, T.; Sutter, A.; ter Laak, A.; Steger-Hartmann, T.; Heinrich, N.; Müller, K.R. Benchmark data set for in silico prediction of Ames mutagenicity. *J Chem Inf Model.* **2009**, *49*, 2077-2081. doi: 10.1021/ci900161g
27. Ambe, K.; Ohya, K.; Takada, W.; Suzuki, M.; Tohkin, M. In Silico Approach to Predict Severe Cutaneous Adverse Reactions Using the Japanese Adverse Drug Event Report Database. *Clin Transl Sci.* **2021**, *2*, 756-763. doi: 10.1111/cts.12944
28. Uesawa, Y. Quantitative structure-activity relationship analysis using deep learning based on a novel molecular image input technique. *Bioorg Med Chem Lett.* **2018**, *28*, 3400-3403. doi: 10.1016/j.bmcl.2018.08.032
29. Zhang, J.; Mucs, D.; Norinder, U.; Svensson, F. LightGBM: An Effective and Scalable Algorithm for Prediction of Chemical Toxicity-Application to the Tox21 and Mutagenicity Data Sets. *J Chem Inf Model.* **2019**, *59*, 4150-4158. doi: 10.1021/acs.jcim.9b00633
30. Polo, T.C.F.; Miot, H.A. Use of ROC curves in clinical and experimental studies. *J Vasc Bras.* **2020**, *19*, e20200186. doi: 10.1590/1677-5449.200186

31. Graharm, R.; Russell, G. Bisphosphonates: The first 40 years. *Bone* **2011**, *49*, 2–19. doi: 10.1016/j.bone.2011.04.022
32. Saad, F.; Brown, S.F.; Poznak, C.V.; Ibrahim, T.; Stemmer, S.M.; Stopeck, A.T.; Diel, I.J.; Takahashi, S.; Shore, N.; Henry, D.H.; et al. Incidence, risk factors, and outcomes of osteonecrosis of the jaw: Integrated analysis from three blinded active-controlled phase III trials in cancer patients with bone metastases. *Ann. Oncol.* **2012**, *23*, 1341–1347. doi: 10.1093/annonc/mdr435
33. Hasegawa, S.; Ikesue, H.; Satake, R.; Inoue, M.; Yoshida, Y.; Tanaka, M.; Matsumoto, K.; Wakabayashi, W.; Oura, K.; Muroi, N.; et al. Osteonecrosis of the Jaw Caused by Denosumab in Treatment-Naïve and Pre-Treatment with Zoledronic Acid Groups: A Time-to-Onset Study Using the Japanese Adverse Drug Event Report (JADER) Database. *Drugs Real World Outcomes.* **2022**, *9*, 659-665. doi: 10.1007/s40801-022-00324-4
34. Brunello, A.; Saia, G.; Bedogni, A.; Scaglione, D.; Basso, U. Worsening of osteonecrosis of the jaw during treatment with sunitinib in a patient with metastatic renal cell carcinoma. *Bone* **2009**, *44*, 173–175. doi: 10.1016/j.bone.2008.08.132
35. Pimolbutr, K.; Porter, S.; Fedele, S. Osteonecrosis of the Jaw Associated with Antiangiogenics in Antiresorptive-Naïve Patient: A Comprehensive Review of the Literature. *Biomed Res. Int.* **2018**, 8071579. doi: 10.1155/2018/8071579
36. Eguia, A.; Bagán-Debón, L.; Cardona, F. Review and update on drugs related to the development of osteonecrosis of the jaw. *Med. Oral Patol. Oral Cir. Bucal.* **2020**, *25*, e71–e83. doi: 10.4317/medoral.23191
37. Di Fede, O.; Bedogni, A.; Giancola, F.; Saia, G.; Bettini, G.; Toia, F.; D'Alessandro, N.; Firenze, A.; Matranga, D.; Fedele, S.; et al. BRONJ in patients with rheumatoid arthritis: A multicenter case series. *Oral Dis.* **2016**, *22*, 543–548. doi: 10.1111/odi.12490
38. Tsao, C.; Darby, I.; Ebeling, P.R.; Walsh, K.; O'Brien-Simpson, N.; Reynolds, E.; Borromeo, G. Oral health risk factors for bisphosphonate-associated jaw osteonecrosis. *J. Oral Maxillofac. Surg.* **2013**, *71*, 1360–1366. doi: 10.1016/j.joms.2013.02.016
39. Khosla, S.K.; Oursler, M.J.; Monroe, D.G. Estrogen and the skeleton. *Trends Endocrinol. Metab.* **2012**, *23*, 576–581. doi: 10.1016/j.tem.2012.03.008
40. Toriumi, S.; Kobayashi, A.; Uesawa, Y. Comprehensive Study of the Risk Factors for Medication-Related Osteonecrosis of the Jaw Based on the Japanese Adverse Drug Event Report Database. *Pharmaceuticals (Basel).* **2020**, *13*, 467. doi: 10.3390/ph13120467
41. Puljula, E.; Turhanen, P.; Vepsäläinen, J.; Monteil, M.; Lecouvey, M.; Weisell, J. Structural requirements for bisphosphonate binding on hydroxyapatite: NMR study of bisphosphonate partial esters. *ACS Med Chem Lett.* **2015**, *21*, 6, 397-401. doi: 10.1021/ml5004603
42. Hughes, D.E.; MacDonald, B.R.; Russell, R.G.; Gowen, M. Inhibition of osteoclast-like cell formation by bisphosphonates in long-term cultures of human bone marrow. *J Clin Invest.* **1989,** *83*, 1930-1935. doi: 10.1172/JCI114100
43. Chavassieux, P.M.; Arlot, M.E.; Reda, C.; Wei, L.; Yates, A.J.; Meunier, P.J. Histomorphometric assessment of the long-term effects of alendronate on bone quality and remodeling in patients with osteoporosis. *J Clin Invest.* **1997**, *100*, 1475-1480. doi: 10.1172/JCI119668
44. Lombard, T.; Neirinckx, V.; Rogister, B.; Gilon, Y.; Wislet, S. Medication-Related Osteonecrosis of the Jaw: New Insights into Molecular Mechanisms and Cellular Therapeutic Approaches. *Stem Cells Int.* **2016**, 8768162. doi: 10.1155/2016/8768162
45. Roskoski, R.Jr. Properties of FDA-approved small molecule protein kinase inhibitors: A 2024 update. *Pharmacol Res.* **2024**, *200*, 107059. doi: 10.1016/j.phrs.2024.107059
46. Weiss-Smith, S.; Deshpande, G.; Chung, S.; Gogolak, V. The FDA drug safety surveillance program: adverse event reporting trends. *Arch Intern Med.* **2011**, *171*, 591-3. doi: 10.1001/archinternmed.2011.89
47. Neha, R.; Beulah, E.; Anusha, B.; Vasista, S.; Stephy, C.; Subeesh, V. Aromatase inhibitors associated osteonecrosis of jaw: signal refining to identify pseudo safety signals. *Int J Clin Pharm.* **2020**, *42*, 721-727. doi: 10.1007/s11096-020-01018-z
48. Gudivada, V.N.; Apon, A.; Ding, J. Data Quality Considerations for Big Data and Machine Learning: Going Beyond Data Cleaning and Transformations. *International Journal on Advances in Software* **2017**, *10*, 1-20.

 doi:10.20944/preprints202502.0246.v1




49. Joonho, G.; Hyunjoong, K. RHSBoost: Improving classification performance in imbalance data. *Comput. Stat. Data Anal.* **2017**, *111*, 1-13. doi: 10.1016/j.csda.2017.01.005
50. Ezzat, A.; Wu, M.; Li, X.L.; Kwoh, C.K. Drug-target interaction prediction via class imbalance-aware ensemble learning. *BMC Bioinformatics.* **2016**, *17*, 509. doi: 10.1186/s12859-016-1377-y
51. ATC/DDD Index. Available online: https://atcddd.fhi.no/atc_ddd_index/ (accessed on 3 July 2022)
52. Lumini, A.; Nanni, L. Convolutional Neural Networks for ATC Classification. *Curr Pharm Des.* **2018**, *34*, 4007-4012. doi: 10.2174/1381612824666181112113438
53. ICH harmonisation for berrer health. Available online: https://www.ich.org/ (accessed on 1 July 2022)
54. MedDRA; Medical Dictionary for Regulatory Activities. Available online: https://www.meddra.org/how-to-use/support-documentation/japanese/welcome (accessed on 3 July 2022)
55. Kan, Y.; Nagai, J.; Uesawa, Y. Evaluation of antibiotic-induced taste and smell disorders using the FDA adverse event reporting system database. *Sci Rep.* **2021**, *11*, 9625. doi: 10.1038/s41598-021-88958-2
56. Okunaka, M.; Kano, D.; Uesawa, Y. Nuclear receptor and stress response pathways associated with antineoplastic agent-induced diarrhea. *Int J Mol Sci.* **2022**, *23*, 12407. doi: 10.3390/ijms232012407
57. Oshima, Y.; Tanimoto, T.; Yuji, K.; Tojo, A. Association between GvHD and nivolumab in the FDA adverse event reporting system. *Bone Marrow Transplant.* **2017**, *52*, 1463-1464. doi: 10.1038/bmt.2017.158
58. Rothman, K.J.; Lanes, S.; Sacks, S.T. The reporting odds ratio and its advantages over the proportional reporting ratio. *Pharmacoepidemiol Drug Saf.* **2004**, *13*, 519-23. doi: 10.1002/pds.1001
59. Harpaz, R.; DuMouchel, W.; LePendu, P.; Bauer-Mehren, A.; Ryan, P.; Shah, N.H. Performance of pharmacovigilance signal-detection algorithms for the FDA adverse event reporting system. *Clin Pharmacol Ther.* **2013**, *93*, 539-46. doi: 10.1038/clpt.2013.24
60. European Medicines Agency. "GUIDELINE ON THE USE OF STATISTICAL SIGNAL DETECTION METHODS IN THE EUDRAVIGILANCEDATAANALYSISSYSTEM". Available online: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/draft-guideline-use-statistical-signal-detection-methods-eudravigilance-data-analysis-system_en.pdf (accessed on 5 July 2022).
61. Kurosaki, K.; Uesawa, Y. Molecular Initiating Events Associated with Drug-Induced Liver Malignant Tumors: An Integrated Study of the FDA Adverse Event Reporting System and Toxicity Predictions. *Biomolecules.* **2021**, *11*, 944. doi: 10.3390/biom11070944
62. Chen, J.J.; Wang, S.J.; Tsai, C.A.; Lin, C.J. Selection of differentially expressed genes in microarray data analysis. *Pharmacogenomics J.* **2007**, *7*, 212-220. doi: 10.1038/sj.tpj.6500412
63. Cui, X.; Churchill, G.A. Statistical tests for differential expression in cDNA microarray experiments. *Genome Biol,* **2003**, *4*, 210. doi: 10.1186/gb-2003-4-4-210
64. PubChem. Available online: https://pubchem.ncbi.nlm.nih.gov/ (accessed on 1 July 2022).
65. MOE. Available online: https://www.chemcomp.com/Products.htm (accessed on 18 July 2022).
66. Breiman, L. Random Forests. *Machine Learning.* **2001**, *45*, 5–32. doi: 10.1023/A:1010933404324
67. Jerome, H. Friedman. "Greedy function approximation: A gradient boosting machine." *Ann. Statist.* **2001**, *29*, 1189–1232. doi: 10.1214/aos/1013203451
68. Hornik, K.; Stinchcombe, M.; White, H. Multilayer feedforward networks are universal approximators. *Neural networks*, **1989**, *2*, 359-366. doi: 10.1016/0893-6080(89)90020-8
69. Mamada, H.; Iwamoto, K.; Nomura, Y.; Uesawa, Y. Predicting blood-to-plasma concentration ratios of drugs from chemical structures and volumes of distribution in humans. *Mol Divers.* **2021**, *3*, 1261-1270. doi: 10.1007/s11030-021-10186-7
70. Nishikiori, K.; Tanaka, K.; Uesawa, Y. Construction of a prediction model for drug removal rate in hemodialysis based on chemical structures. *Mol Divers.* **2022**, *5*, 2647-2657. doi: 10.1007/s11030-021-10348-7
71. Sahigara, F.; Mansouri, K.; Ballabio, D.; Mauri, A.; Consonni, V.; Todeschini, R. Comparison of different approaches to define the applicability domain of QSAR models. *Molecules.* **2012**, *17*, 4791-4810. doi: 10.3390/molecules17054791
72. Akobeng, A.K. Understanding diagnostic tests 3: Receiver operating characteristic curves. *Acta Paediatr.* **2007**, *96*, 644-647. doi: 10.1111/j.1651-2227.2006.00178.x

 doi:10.20944/preprints202502.0246.v1