Original paper

Reumatologia 2026; 64, 3: 166–173
DOI: https://doi.org/10.5114/reum/220541

Machine learning analysis of treatment and complication patterns in rheumatoid arthritis case reports (1995–2025)

Naruaki Ogasawara

The Japanese Society of Internal Medicine, Bunkyo-Ward, Tokyo, Japan

Abstract

Introduction: This study aimed to identify research trends and thematic structures in case reports related to rheumatoid arthritis (RA) and complications published between 1995 and September 2025. By applying unsupervised machine learning (ML), the study sought to uncover longitudinal patterns and cross-domain relationships that may not be fully captured in traditional reviews or large-scale epidemiological studies.

Material and methods: Bibliographic data were collected from the Web of Science Core Collection using the search term "rheumatoid arthritis and complications," limited to case reports published between 1995 and 2025. Titles, key words, and abstracts were combined into unified text documents and processed with natural language processing (NLP) techniques. Term Frequency–Inverse Document Frequency was applied for text vectorization. Topic extraction employed non-negative matrix factorization, and clustering was performed using K-means. The optimal number of clusters was determined based on the highest Silhouette Score (0.636). Analyses were conducted in Python (Version 3.10.5) within the PyCharm environment (Version 2022.1.3).

Results: The unsupervised ML framework identified two major and stable clusters among 1,200 case reports: a pharmacological and immunological management cluster and a surgical and postoperative complications cluster. Cluster 1 included 890 reports characterized by pharmacological and immunological terms such as "TNF," "therapy," and "anti," while cluster 2 contained 310 reports dominated by orthopedic terms such as "arthroplasty," "knee," and "revision." Reproducibility tests across five runs demonstrated consistent clustering patterns, supporting methodological reliability.

Conclusions: This NLP and ML-based analysis revealed a dual structure in RA complication literature, reflecting immunological treatment-related complications and surgical outcomes. Beyond confirming known domains, the study provides a data-driven characterization of their evolution and overlap, offering insights into the evolution of research themes in RA management. This approach demonstrates the potential of unsupervised ML for longitudinal mapping of clinical research domains.

Key words: rheumatoid arthritis, complications, case reports, natural language processing, machine learning, topic modeling, pharmacovigilance.

Introduction

Rheumatoid arthritis (RA) is a typical autoimmune disease characterized by chronic inflammation of the synovial membrane of joints. Delayed early diagnosis and treatment can lead to joint destruction and functional impairment, significantly impairing patients' quality of life [1, 2]. In recent years, significant progress has been made in controlling RA with the introduction of conventional synthetic disease-modifying antirheumatic drugs (csDMARDs) such as methotrexate (MTX), biologic disease-modifying antirheumatic drugs (bDMARDs) including tumor necrosis factor (TNF) inhibitors, and targeted synthetic disease-modifying antirheumatic drugs (tsDMARDs) such as Janus kinase (JAK) inhibitors and other agents with specific molecular targets [3, 4].

Address for correspondence

Naruaki Ogasawara, The Japanese Society of Internal Medicine, 3-28-8 Hongo Bunkyo-Ward, Tokyo 113-8433, Japan,
e-mail: n-ogasawara@naika.or.jp

Submitted: 16.12.2025; **Accepted:** 10.04.2026; **Published online:** 30.06.2026

However, approximately 30–50% of patients do not respond adequately to csDMARDs such as MTX [5, 6]. Furthermore, up to 40% of patients may develop difficult-to-treat RA, defined by the European Alliance of Associations for Rheumatology (EULAR) as failure of ≥ 2 bDMARDs or tsDMARDs with different mechanisms of action, persistent disease activity, and problematic disease management as perceived by the clinician or patient. These patients experience intolerable adverse events or require complex treatment strategies [7, 8]. In drug-resistant or advanced cases, surgical interventions, including synovectomy, joint-preserving surgery, and total joint arthroplasty, have become an important treatment option [9, 10]. One Japanese study recorded 1,569 cases between 2004 and 2019 [11].

Treatment strategies for RA are broadly divided into medical management (pharmacological treatment with disease-modifying antirheumatic drugs [DMARDs] and adjunctive therapies) and surgical management (including synovectomy, joint-preserving surgery, and total joint replacement). Case reports provide detailed descriptions of the treatment course and decision-making for individual cases, providing a wealth of clinical information. Additionally, case reports capture rare complications, unexpected treatment responses, and nuanced clinical decision-making processes that are often absent in large-scale cohort studies or randomized controlled trials. These reports serve as valuable resources for hypothesis generation, personalized medicine, and clinical education [12–15].

This study used natural language processing (NLP) and unsupervised machine learning (ML) to analyze bibliographic data on case reports related to "rheumatoid arthritis and complications" listed in the Web of Science (WoS) from 1995 to September 2025. Unlike traditional reviews or meta-analyses that rely on manual selection and thematic coding, NLP/ML methods enable scalable, reproducible, and data-driven identification of hidden patterns across thousands of reports. This approach can capture both prevailing pharmacological and surgical management trends and emerging clinical challenges, providing a more nuanced picture of RA management. However, to date, few studies have specifically applied NLP or ML techniques to RA case reports. By using NLP and ML, this study aims not only to confirm known complication domains but also to quantify their relative prominence, temporal evolution, and inter-cluster relationships, providing a more data-driven and longitudinal understanding of RA management.

Study purpose and scope

The purpose of this study is to systematically analyze case reports listed in the WoS database regarding RA treatment and complications, aiming to extract clinical characteristics and challenges that are not fully captured by existing large-scale epidemiological studies. It is expected that this study will provide new insights into treatment selection and complication management for RA patients.

Significance and contribution

This study represents a large-scale analysis of RA and its complications using case report data from WoS, applying NLP and ML techniques. By systematically extracting and categorizing clinical topics, this research identifies prevailing trends in RA-related complications and treatment strategies while highlighting previously underexplored areas such as the distinct role of arthroplasty in surgical management. The findings may support clinicians, researchers, and policymakers in understanding evolving research priorities, identifying safety-related concerns, and guiding future investigations.

Material and methods

Data collection and preprocessing

This study collected case reports related to "rheumatoid arthritis and complications" from the WoS database in September 2025. Each case report contained an article title, key words, and abstract. The study selection procedure consisted of the following steps:

1. Initial search: A total of 8,435 records were retrieved using the topic search query "rheumatoid arthritis and complications".
2. Document type filtering: The dataset was further restricted to include case reports, by adding the filter "case reports," yielding a dataset of 1,283 records.
3. Year filtering: The records were refined to include only those published between 1995 and 2025, resulting in 1,200 records.

These three fields were concatenated into a single text string per case report. The preprocessing pipeline handled missing values by replacing them with empty strings. All text data underwent conversion to lowercase and tokenization for subsequent analysis.

Text vectorization using Term Frequency–Inverse Document Frequency

The study applied Term Frequency–Inverse Document Frequency (TF-IDF) vectorization to transform textual data into numerical representations using the TfidfVectorizer from the scikit-learn Python library [16], a widely used open-source ML tool in Python. The vectorization process removed common English stop words and excluded terms appearing in more than 95% of documents or fewer than 2 documents (Max_df = 0.95,

Min_df = 2). This approach generated a sparse matrix that captured the relative importance of each term across the corpus [17] and has been effectively applied in emotional text classification tasks [18].

Topic modeling with non-negative matrix factorization

The analysis implemented topic modeling employing non-negative matrix factorization (NMF) [19] with an optimal number of components (topics) determined by silhouette score evaluation ranging from 2 to 15. The optimal number of components (topics) was selected based on the average silhouette score across 10 independent runs for each topic count ranging from 2 to 15.

Clustering with K-means

The study applied K-means clustering [20], a widely used unsupervised learning algorithm based on iterative centroid optimization, to the NMF feature matrix to categorize case reports based on their topic distributions. The number of clusters was set equal to the optimal number of topics identified through NMF. The clustering was performed multiple times with random initialization to assess stability. Each report received cluster label assignment, and visualization displayed the distribution of reports across clusters. The clustering algorithm used random initialization (i.e., the random state was not fixed); therefore, the results may vary across runs.

Statistical characterization of clusters

The analysis calculated the number of reports, mean, and standard deviation of the NMF feature values for each cluster. These metrics provided insights into the internal consistency and density of individual clusters. The study compiled all results into summary tables and exported them as Excel files for further examination.

Validation of clustering quality

Clustering quality was evaluated using three established metrics: the Silhouette Score [21], the Calinski-Harabasz Index [22], and the Davies-Bouldin Index [23].

Reproducibility assessment of topic modeling

The analysis conducted five independent runs of the NMF topic modeling process using different random seeds to evaluate model stability. For each topic, the study compared the top 10 key words across all runs and calculated their overlap with the original key word set. The average key word overlap across runs served as a reproducibility score, indicating the consistency of topic identification. Note that this study used the simple number of key word matches (overlap count) to evaluate NMF reproducibility and did not use ratio-based metrics such as the Jaccard Index or percentage overlap.

Bioethical standards

Due to the nature of the study, the consent of the bioethics committee was not required.

Results

Overview of clustering outcomes

Using TF-IDF vectorization, NMF, and K-means clustering, two major thematic clusters were identified among the case reports related to "rheumatoid arthritis and complications" published between 1995 and September 2025.

This study used a method to quantify the importance of words in a text (TF-IDF), a method to identify hidden features and trends within the text (NMF), and a method to group papers with similar content (K-means clustering). As a result, the optimal number of clusters was determined as 2, based on the Silhouette Score (0.636), along with the Calinski-Harabasz Index (2733.643) and the Davies-Bouldin Index (0.531), indicating that the clusters are well separated and internally cohesive, as reflected by the high Silhouette Score and Calinski-Harabasz Index and the low Davies-Bouldin Index (Table I).

Cluster 1 contained 890 case reports (n = 890), while cluster 2 contained 310 case reports (n = 310). The relatively low standard deviations of the average TF-IDF scores across both clusters (cluster 1 = 0.0364, cluster 2 = 0.0645) indicate consistent internal text similarity and term importance (Table II). The temporal evolution over the 30-year literature is characterized by the rapid growth of cluster 1 after 2000, with an explosive increase especially after 2020, while cluster 2 has only shown a gradual increase (Fig. 1).

Topic-specific key words and weight distribution

The top-ranked key words and their corresponding TF-IDF scores within each cluster are presented in Table III.

Table I. Clustering validity

Optimal topic count	Silhouette Score	Calinski-Harabasz Index	Davies-Bouldin Index
2	0.63638533	2,733.643464	0.531957014

Table II. Cluster statistics

Cluster	Topic label	Number of reports	Mean TF-IDF score	TF-IDF standard deviation
1	RA, arthritis, rheumatoid	890	0.043621009	0.03643496
2	Patients, arthroplasty, total	310	0.067337745	0.06456454

RA – rheumatoid arthritis, TF-IDF – Term Frequency–Inverse Document Frequency.

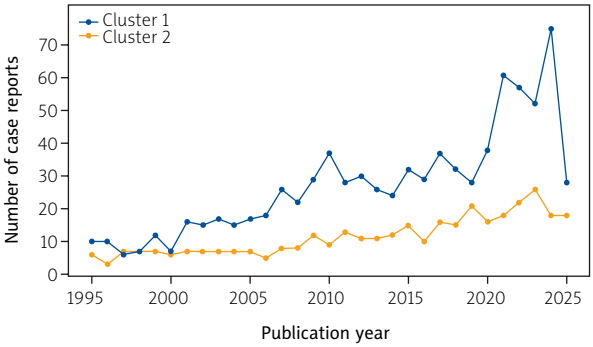

Fig. 1. Temporal evolution of two clusters (1995–2025).

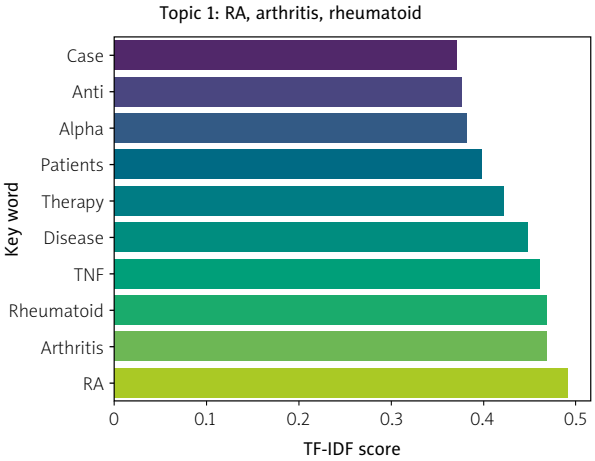

Fig. 2. Analysis of rheumatoid arthritis treatment and complications using Term Frequency–Inverse Document Frequency scores cluster 1.

RA – rheumatoid arthritis, TF-IDF – Term Frequency–Inverse Document Frequency, TNF – tumor necrosis factor.

The analysis of case reports related to "rheumatoid arthritis and complications" published between 1995 and September 2025 revealed the following clusters: the first cluster (cluster 1) consisted of case reports related to "rheumatoid arthritis disease and treatment" (890 cases), and the second cluster (cluster 2) consisted of case reports related to "rheumatoid arthritis surgery and follow-up" (310 cases).

Cluster 1 was characterized by immunological and pharmacological terms such as "TNF," "therapy," "alpha," and "anti", suggesting a focus on disease pathogenesis and biologic treatment (Fig. 2, Table III). By contrast, cluster 2 centered around "arthroplasty," "knee," "elbow," and "revision," indicating a distinct focus on surgical complications and orthopedic outcomes (Fig. 3, Table III).

These findings indicate an initial clear separation between medical-therapeutic and surgical-mechanical domains of RA complications, particularly in earlier publication periods. Within each cluster, the words used were internally consistent, supporting stable cluster assignment despite increasing inter-cluster overlap over time (Figure 4).

Clustering validity and reproducibility

To evaluate the quality of the clustering results, this study employed three metrics: the Silhouette Score, the Calinski-Harabasz Index, and the Davies-Bouldin In-

Table III. Topic key word scores

Topic	Key word	TF-IDF score
1	RA	0.491735426
1	Arthritis	0.468112257
1	Rheumatoid	0.468084563
1	TNF	0.460721937
1	Disease	0.447705025
1	Therapy	0.422323735
1	Patients	0.397992105
1	Alpha	0.382348399
1	Anti	0.37630643
1	Case	0.37118593
2	Patients	0.47450481
2	Arthroplasty	0.432641023
2	Total	0.363353083
2	Knee	0.351021459
2	Elbow	0.318907035
2	Surgery	0.266615052
2	Revision	0.263702504
2	Years	0.249892609
2	Follow	0.236676265
2	Fusion	0.232738428

RA – rheumatoid arthritis, TF-IDF – Term Frequency–Inverse Document Frequency, TNF – tumor necrosis factor.

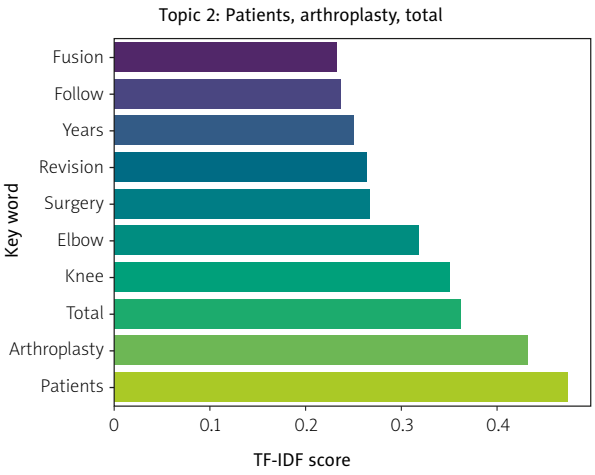

Fig. 3. Analysis of rheumatoid arthritis treatment and complications using Term Frequency–Inverse Document Frequency scores, cluster 2.

RA – rheumatoid arthritis, TF-IDF – Term Frequency–Inverse Document Frequency.

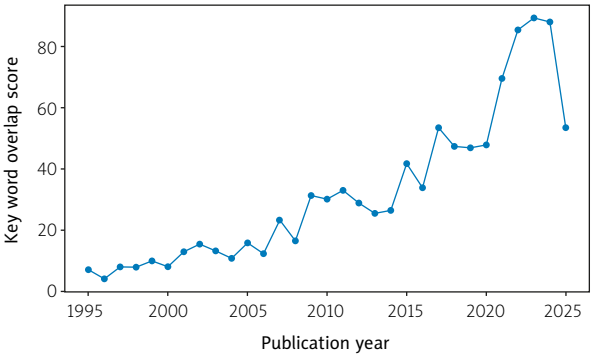

Fig. 4. Interrelationship between two clusters over time.

dex. The optimal topic number of two clusters achieved the best clustering validity with the indicators (Table I). Validation metrics indicated that the model produced well-separated and compact clusters, supporting the robustness of topic differentiation.

Reproducibility tests were conducted across 5 independent runs, each yielding highly consistent topic assignments with full overlap (average overlap count = 10 for both clusters). These results suggest high stability and reproducibility of the clustering model (Table IV).

Interpretation of topic structures

Cluster 1 ("RA, arthritis, rheumatoid") primarily included reports describing systemic inflammatory mechanisms, immunotherapy responses, and biologic drug-related complications such as TNF inhibitor-induced infections or autoimmune exacerbations. The presence of high-weight key words such as "TNF," "therapy," and "anti" further suggests a dominant focus on pharmacological management and immunological outcomes.

Cluster 2 ("patients, arthroplasty, total") represented literature dealing with orthopedic complications following joint replacement procedures in RA patients, particularly knee and elbow arthroplasties. Terms such as "revision," "fusion," and "follow" reflected the clinical and surgical follow-up nature of these case reports. This cluster therefore captures the mechanical and postoperative dimension of RA-related complications.

Table IV. Reproducibility summary

Topic	Average overlap count
1	10
2	10

Summary

The unsupervised ML framework successfully revealed two distinct and stable thematic structures within 30 years of RA complication literature:

– pharmacological/medical management cluster: focusing on pathophysiology, TNF-related therapy, and biologic drug–induced complications;
– surgical and postoperative complication cluster: concentrating on postoperative complications, prosthesis outcomes, and joint revision surgery.

These findings highlight the dual landscape of RA complications research, encompassing both conservative treatment and surgical treatment, and suggest that future interdisciplinary approaches should integrate both aspects to achieve comprehensive patient management. The strong reproducibility of the clustering results further supports the robustness of this ML-based framework for long-term bibliometric trend analysis.

Discussion

This ML-based analysis of case reports on "rheumatoid arthritis and complications" from 1995 to September 2025 revealed two major thematic clusters that correspond to distinct clinical domains: pharmacological/medical complications and surgical and postoperative complications. The former was characterized by pharmacological and immunological terms such as "TNF," "therapy," "alpha," and "anti," highlighting the dominance of biologic treatments and their associated adverse effects, such as infection and secondary autoimmune responses. The latter cluster was centered around surgical and orthopedic terminology such as "arthro-

plasty," "knee," "elbow," and "revision," indicating the ongoing challenges of joint replacement, prosthesis management, and surgical complications in patients with advanced RA.

These findings indicate that research on RA complications has evolved into two parallel streams over the past three decades: one driven by advances in pharmacological and immunological management, and another by surgical and postoperative interventions. While these domains are clinically recognized, this study provides a novel contribution by systematically quantifying their representation, temporal evolution, and inter-cluster relationships across 30 years of literature. This longitudinal mapping offers a data-driven perspective that complements traditional expert reviews.

The increasing co-occurrence of terms related to immunosuppressive therapy and surgical revision suggests a growing intersection between pharmacological failure and orthopedic intervention – an emerging clinical pattern not systematically captured in prior reviews. This insight may inform future research on treatment sequencing and risk stratification in difficult-to-treat RA.

Clinical and research implications

The results of this study revealed that research and treatment trends related to RA complications can be broadly divided into two clusters. Each cluster has important implications for medical practice and research.

The first cluster is related to drug therapy. Since the 2000s, new drugs such as TNF inhibitors and JAK inhibitors have emerged, bringing significant advances in RA treatment [24, 25]. However, these drugs are associated with safety concerns, such as the risk of infection (particularly herpes zoster reactivation) and thrombosis [26, 27]. This study applied unsupervised ML techniques to quantitatively identify how these side effect risks are manifested in the literature.

The second cluster is related to surgical procedures. Although the introduction of new drugs such as TNF and JAK inhibitors has led to significant advances in the treatment of RA, patients with severe RA may develop joint damage that requires orthopedic intervention (joint replacement) [28, 29]. In this second cluster, terms such as "revision," "fusion," and "prosthesis" frequently appear, revealing that issues include post-operative complications, the lifespan of artificial joints, infection, and mechanical problems.

Significance as a research method

In this study, unsupervised ML was used to identify meaningful groups (clusters) from a large amount of WoS research data without human intervention. The clusters were also highly reproducible (Silhouette Score = 0.636, number of overlaps in repeated runs = 10), confirming the reliability of this method.

This type of ML-based analysis can help discover new trends and rare complications that would be difficult to find using traditional human reviews alone and is expected to serve as a starting point for precision medicine and new research [30]. While the complication categories identified in this study are clinically established, the reproducible application of data-driven temporal quantification and automated clustering is a novel contribution that may be useful for future longitudinal and pharmacovigilance analyses. Moreover, the reproducibility test across five independent runs confirmed the model's stability, which is important for ensuring the reliability and validity of unsupervised ML applications in medical literature analysis.

Study limitations

Despite its contributions, several limitations should be acknowledged. First, the study relies exclusively on the WoS database, which, while comprehensive, may introduce selection bias. Certain journals, case series, or non-English publications not indexed in WoS could have been excluded, potentially leading to incomplete representation of the global literature. Second, case reports inherently reflect publication bias, favoring unusual, severe, or novel complications rather than typical or negative outcomes. This may overemphasize rare phenomena at the expense of everyday clinical experience. Third, the study's text-based modeling approach does not incorporate full-text content or structured clinical data such as patient demographics, laboratory results, or treatment duration. Consequently, the extracted topics reflect textual emphasis rather than detailed clinical causality. Lastly, as unsupervised ML involves algorithmic abstraction without expert supervision, topic labeling was inferred from high-frequency terms rather than domain-curated validation, which may limit semantic precision.

Future research should aim to integrate data from multiple bibliographic sources (e.g., PubMed, Scopus), incorporate multilingual text mining, and apply hybrid supervised–unsupervised approaches to enhance both the accuracy and interpretability of topic identification. The inclusion of full-text and metadata analysis would further deepen understanding of complication mechanisms and risk stratification in RA.

Conclusions

This study used NLP and unsupervised ML to identify two distinct and stable thematic structures (pharmaco-

therapy cluster and surgical cluster) in case reports on "rheumatoid arthritis and its complications" published between 1995 and September 2025. The former focuses on complications related to drug therapy, driven by advances in biological and molecular-targeted drug therapies such as TNF and JAK inhibitors. The latter focuses on postoperative complications following arthroplasty and arthrodesis, reflecting the ongoing challenges of surgical management in patients with advanced RA.

This study applies ML-based analytic methods, reflecting the recent expansion of ML-related research in the medical field [30]. Methodologically, this study demonstrates the potential of unsupervised ML for extracting clinically meaningful insights from large-scale literature data.

The high reproducibility of the clustering results supports the robustness of this analytical method. By uncovering longitudinal trends, cross-domain overlaps, and emerging intersections between pharmacological and surgical complications, this study demonstrates how ML-based text analysis can complement expert-driven reviews and support future applications in precision medicine, pharmacovigilance, and automated literature surveillance. This framework provides a scalable and reproducible approach that may be extended to track evolving treatment paradigms and safety-related trends in rheumatology.

Disclosures

Conflicts of interest: The author declares no conflicts of interest.
Funding: No external funding.
Ethics approval: Due to the nature of the study, consent is not required.
Availability of data and material: The data that support the findings of this study are available on request from the author (N.O.).

References

1. Zhu M, Ding Q, Lin Z, et al. New targets and strategies for rheumatoid arthritis: from signal transduction to epigenetic aspect. Biomolecules 2023; 13: 766, DOI: 10.3390/biom13050766.
2. Athanasiou A, Papazachou O, Rovina N, et al. The effects of exercise training on functional capacity and quality of life in patients with rheumatoid arthritis: a systematic review. J Cardiovasc Dev Dis 2024; 11: 161, DOI: 10.3390/jcdd11060161.
3. Alivernini S, Firestein GS, McInnes IB. The pathogenesis of rheumatoid arthritis. Immunity 2022; 55: 2255–2270, DOI: 10.1016/j.immuni.2022.11.009.
4. Jain A, Bishnoi M, Prajapati SK, et al. Targeting rheumatoid arthritis: a molecular perspective on biologic therapies and clinical progress. J Biol Eng 2025; 19: 67, DOI: 10.1186/s13036-025-00534-8.
5. Gupta V, Katiyar S, Singh A, et al. CD39 positive regulatory T cell frequency as a biomarker of treatment response to methotrexate in rheumatoid arthritis. Int J Rheum Dis 2018; 21: 1548–1556, DOI: 10.1111/1756-185X.13333.
6. Yu J, Zhou P. The advances of methotrexate resistance in rheumatoid arthritis. Inflammopharmacology 2020; 28: 1183–1193, DOI: 10.1007/s10787-020-00741-3.
7. Qiao Y, Li F, Zhang L, et al. A systematic review and meta-analysis comparing outcomes following total knee arthroplasty for rheumatoid arthritis versus for osteoarthritis. BMC Musculoskelet Disord 2023; 24: 484, DOI: 10.1186/s12891-023-06601-9.
8. Hawker G, Ravi B, Goodman S. Outcomes in rheumatoid arthritis patients undergoing total joint arthroplasty. Int J Clin Rheumatol 2014; 9: 585–593, DOI: 10.2217/ijr.14.47.
9. Sigmund A, Russell LA. Optimizing rheumatoid arthritis patients for surgery. Curr Rheumatol Rep 2018; 20: 48, DOI: 10.1007/s11926-018-0757-x.
10. Taylor-Williams O, Nossent J, Inderjeeth CA. Incidence and complication rates for total hip arthroplasty in rheumatoid arthritis: a systematic review and meta-analysis across four decades. Rheumatol Ther 2020; 7: 685–702, DOI: 10.1007/s40744-020-00238-z.
11. Hotta Y, Nasu Y, Nishida K, et al. FRI0523. The recent trend of orthopaedic surgeries for rheumatoid arthritis: an analysis of 1569 cases from 2004 to 2019. Ann Rheum Dis 2020; 79 (Suppl 1): 860–861, DOI: 10.1136/annrheumdis-2020-eular.1974.
12. Jackson D, Daly J, Saltman DC. Aggregating case reports: a way for the future of evidence-based health care? Clin Case Rep 2014; 2: 23–24, DOI: 10.1002/ccr3.58.
13. Cheung A, O'Sullivan F, Walker S, et al. MSR77. Making a case for case reports: a scoping review of the use of and challenges with case report data in quantitative synthesis. Value Health 2023; 26 (Suppl): S291.
14. Lowenfels AB, Mamtani R, Solomon LW, et al. The value of case reports for graduate medical education. J Grad Med Educ 2022; 14: 529–532, DOI: 10.4300/JGME-D-21-01115.1.
15. Ogasawara N, Matsunaga K, Isomoto H, Shimizu W. Internal medicine year in review 2022. Intern Med 2023; 62: 3431–3435, DOI: 10.2169/internalmedicine.2266-23.
16. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. J Mach Learn Res 2011; 12: 2825–2830.
17. Manning CD, Raghavan P, Schütze H. Introduction to information retrieval. Cambridge University Press, Cambridge 2008.
18. Alfarizi MI, Syafaah L, Lestandy M. Emotional text classification using TF-IDF (Term Frequency–Inverse Document Frequency) and LSTM (Long Short-Term Memory). JUITA J Inform 2022; 10: 225–234, DOI: 10.30595/juita.v10i2.13262.
19. Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. Nature 1999; 401: 788–791, DOI: 10.1038/44565.
20. Lloyd S. Least squares quantization in PCM. IEEE Trans Inf Theory 1982; 28: 129–137, DOI: 10.1109/TIT.1982.1056489.
21. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 1987; 20: 53–65, DOI: 10.1016/0377-0427(87)90125-7.
22. Caliński T, Harabasz J. A dendrite method for cluster analysis. Commun Stat 1974; 3: 1–27, DOI: 10.1080/03610927408827109.

23. Davies DL, Bouldin DW. A cluster separation measure. IEEE Trans Pattern Anal Mach Intell 1979; PAMI-1: 224–227, DOI: 10.1109/TPAMI.1979.4766909.
24. Angelini J, Talotta R, Roncato R, et al. JAK-Inhibitors for the treatment of rheumatoid arthritis: a focus on the present and an outlook on the future. Biomolecules 2020; 10: 1002, DOI: 10.3390/biom10071002.
25. Domańska-Poboża J, Wisłowska M. Evolving strategies in the treatment of rheumatoid arthritis: a historical perspective. Reumatologia 2025; 63: 116–130, DOI: 10.5114/reum/195012.
26. Riley TR, George MD. Risk for infections with glucocorticoids and DMARDs in patients with rheumatoid arthritis. RMD Open 2021; 7: e001235, DOI: 10.1136/rmdopen-2020-001235.
27. Venetsanopoulou AI, Voulgari PV, Drosos AA. Janus kinase versus TNF inhibitors: where we stand today in rheumatoid arthritis. Expert Rev Clin Immunol 2022; 18: 485–493, DOI: 10.1080/1744666X.2022.2064275.
28. Goodman SM, George MD. Should we stop or continue conventional synthetic (including glucocorticoids) and targeted DMARDs before surgery in patients with inflammatory rheumatic diseases? RMD Open 2020; 6: e001214, DOI: 10.1136/rmdopen-2020-001214.
29. Girdler SJ, Ye I, Tang R, Kirschner N. Altering the natural history of rheumatoid arthritis: the role of immunotherapy and biologics in orthopaedic care. J Orthop 2019; 17: 17–21, DOI: 10.1016/j.jor.2019.06.030.
30. Ogasawara N, Matsunaga K, Isomoto H, Shimizu W. Internal medicine year in review 2023. Intern Med 2024; 63: 3137–3140, DOI: 10.2169/internalmedicine.4396-24.