World Academy of Science, Engineering and Technology
International Journal of Health and Medical Engineering
Vol:16, No:9, 2022

# Literature-Based Discoveries in Lupus Treatment

Oluwaseyi Jaiyeoba, Vetria Byrd



*Abstract*—Systemic lupus erythematosus (aka lupus) is a chronic disease known for its chameleon-like ability to mimic symptoms of other diseases rendering it hard to detect, diagnose and treat. The heterogeneous nature of the disease generates disparate data that are often multifaceted and multi-dimensional. Musculoskeletal manifestation of lupus is one of the most common clinical manifestations of lupus. This research links disparate literature on the treatment of lupus as it affects the musculoskeletal system using the discoveries from literature-based research articles available on the PubMed database. Several Natural Language Processing (NPL) tools exist to connect disjointed but related literature, such as Connected Papers, Bitola, and Gopalakrishnan. Literature-based discovery (LBD) has been used to bridge unconnected disciplines based on text mining procedures. The technical/medical literature consists of many technical/medical concepts, each having its sub-literature. This approach has been used to link Parkinson's, Raynaud, and Multiple Sclerosis treatment within works of literature. Literature-based discovery methods can connect two or more related but disjointed literature concepts to produce a novel and plausible approach to solving a research problem. Data visualization techniques with the help of natural language processing tools are used to visually represent the result of literature-based discoveries. Literature search results can be voluminous, but Data visualization processes can provide insight and detect subtle patterns in large data. These insights and patterns can lead to discoveries that would have otherwise been hidden from disjointed literature. In this research, literature data are mined and combined with visualization techniques for heterogeneous data to discover viable treatments reported in the literature for lupus expression in the musculoskeletal system. This research answers the question of using literature-based discovery to identify potential treatments for a multifaceted disease like lupus. A three-pronged methodology is used in this research: text mining, natural language processing, and data visualization. These three research-related fields are employed to identify patterns in lupus-related data that, when visually represented, could aid research in the treatment of lupus. This work introduces a method for visually representing interconnections of various lupus-related literature. The methodology outlined in this work is the first step toward literature-based research and treatment planning for the musculoskeletal manifestation of lupus. The results also outline the interconnection of complex, disparate data associated with the manifestation of lupus in the musculoskeletal system. The societal impact of this work is broad. Advances in this work will improve the quality of life for millions of persons in the workforce currently diagnosed and silently living with a musculoskeletal disease associated with lupus.

*Keywords*—Systemic lupus erythematosus, LBD, Data Visualization, musculoskeletal system, treatment.

## I. Introduction

Millions of people are diagnosed with lupus each year, however, breakthroughs in the study of lupus lag behind advances in other diseases [3]. A complex disease like lupus requires flexible analysis solutions that allow for various views of interactions and connections of disparate data [2]. The combination of such disparate but related data can create a view of the interconnected elements of the disease. This interconnection of literature is currently lacking in the study of lupus. Lupus is a result of friendly fire from the body's own immune system due to its inability to distinguish the body's system from foreign infections [2]. When such poor judgment occurs in the immune system, the results can be catastrophic for the organ involved. The possibility of lupus attacking a wide range of organs makes its possible range of symptoms very diverse. If the immune systems attack the pancreas, then there is a high chance of the patient having diabetes. Similarly, an attack on the circulatory system could lead to myocarditis, among other diseases [2], [5].

Black population in Europe and America make up the majority of people who are vulnerable to the lupus disease, with women particularly more vulnerable [4]. Lupus prevalence in women occurs more likely in early adulthood or after giving birth.

A diagnosis of lupus can be devastating news to a patient, but it is even worse when the disease is present but undiagnosed. Diagnosing lupus is tricky because it does not have definitive symptoms; it mimics symptoms of other diseases giving it a wide range of misleading diagnoses [6]. Physicians' criteria for lupus diagnosis have evolved over the years. The use of the 1997 American College of Rheumatology's (ACR) "Eleven Criteria of Lupus" was among the earlier methods to diagnose lupus. Using the ACR 1997, if a patient is diagnosed with four out of these eleven lupus symptoms, then it qualifies as a lupus diagnosis. This method of lupus diagnosis was updated by the European League Against Rheumatism (EULAR) and the American College of Rheumatology (ACR) in 2019, known as the EULAR/ACR 2019 criteria. The new EULAR/ACR 2019 criteria involve rigorous methodology with increased sensitivity and specificity [7].

The sensitivity of a diagnosis refers to the ability of a diagnosis to designate a person more accurately with a disease as positive. On the other hand, the specificity of a diagnosis refers to the percentage of people who recorded a negative test for a specific disease among a collection of people without the disease [8].

Lupus may attack various body organs and systems, including the skin, musculoskeletal, cardiovascular, respiratory, renal, hematologic, and central nervous systems. The possibility of attacking multiple body systems could foster a heterogenous clinical presentation of lupus with complex pathogenesis of the diseases. Musculoskeletal systems manifestation is among the most common features of systemic lupus erythematosus, usually requiring long-term management from the first diagnosis. The musculoskeletal system comprises bones, muscles, cartilage, tendons,

Oluwaseyi Jaiyeoba is a PhD student and Vetria Byrd is an Associate Professor with the Department of Computer Graphics Technology, Purdue University, West Lafayette, IN, 47907 USA (e-mail: ojaiyeob@purdue.edu, vbyrd@purdue.edu).

ligaments, joints, and other connective tissues that support and bind different tissues and organs together. The musculoskeletal system runs through most parts of the entire human body. Therefore musculoskeletal attack of lupus overlaps with some other lupus organ manifestation attacks such as the cardiovascular system and the cutaneous organ [9]. Musculoskeletal manifestation of lupus is usually presented in up to 50% of lupus patients and up to 95% during the clinical course [1]. Significant disabilities are associated with the long-term impact of the musculoskeletal manifestation of lupus. Conventional treatments of SLE using azathioprine, mycophenolate, or cyclophosphamide are less effective on musculoskeletal lupus manifestations like arthritis [1]. A lupus attack on the musculoskeletal system produces a wide range of medical conditions and symptoms in different parts of the body [1]. The musculoskeletal system comprises bones, muscles, cartilage, tendons, ligaments, joints, and other connective tissues that support and bind different tissues and organs together.

According to the International Association of Scientific, Technical, and Medical Publishers, there are about 2.5 million new scientific papers published each year by an average of 4.16 million unique authors, where each article is having an average of 4.2 authors. There are estimated to be of the order of 10,000 journal publishers globally [11]. This report shows a lot of disconnection between researchers as each researcher or group of researchers often work in isolation producing individual findings that may have subtle relationships with each other. Such relationships could lead to new innovations if discovered. However, due to the large volume of published information yearly, it may be difficult to discover these hidden innovations. These hidden innovations in literature have led to an area of research known as the Literature-Based-Discovery (LBD). LBD involves seeking new knowledge from existing literature using machine learning and data analytics tools to enhance the process of finding subtle connections and discoveries within the vast amount of literature out there. LBD builds bridges across the numerous islands of isolated interdisciplinary research [12].

## II. Research Question

LBD has been used for research discoveries and is more pronounced in medical discovery [13]. The use of LBD in disease research has mainly involved investigating medical conditions with non-multifaceted symptoms such as kidney diseases and cardiovascular diseases. However, some challenges are presented for multifaceted diseases like lupus, which has various possible medical conditions and symptoms. Thus, LBD in lupus would require a more robust LBD approach. This research seeks to answer the following research questions.

1) How can LBD be applied to multifaceted diseases like lupus?
2) What is the best method of visualizing a large lupus LBD database?

Using the scientific literature database on lupus from PubMed, this research looks to develop an LBD approach to discover potential treatment options for a multifaceted symptom disease like lupus using the musculoskeletal manifestations of lupus as a case study. The goal is to aid the discovery of viable treatments for major lupus expressions in the musculoskeletal system.

## III. Background on Literature-Based Discovery

LBD has led to numerous discoveries in the biomedical field since its first introduction by Swanson (1986), which led to the discovery of a previously unknown link between Raynaud's disease and fish oil [14]. Since this discovery, LBD research has been applied to discovering innovations and knowledge in numerous diseases like cancer [15], Parkinson's Disease [16], and Cardiac arrest [14] to mention a few. These studies often involve the design and implementation of analytical tools to automate or semi-automate the process of going through the vast amount of literature data [17]. Popular tools for computer-aided literature searching or information retrievals like PubMed and Google Scholar are insufficient in recognizing connections between fragments in literature. LBD systems are separate and distinct from the popular information retrieval tools and Biomedical text-mining applications [18]. An LBD technique will solve this problem by directly addressing the problem of knowledge overspecialization.

LBD is an information science technique that combines existing knowledge from scientific literature in an attempt to discover new knowledge. While LBD technique employs information retrieval as a step in its process, information retrieval is not enough for a robust relationship in defining the implicit associations and relations between different terms in literature. For instance, querying PubMed for a new therapeutic drug on a disease returns thousands of articles based on the disease. The majority of the search result lacks important information regarding the implicit relationship to a therapeutic drug.

Natural language processing (NLP) is a technique that plays a significant role in the robustness of LBD in identifying relationships between literature. There are two important goals to be achieved by LBD. First is identifying the existence of a relationship between two or more disjointed literature, and secondly, exploring the depth of the connection between these works of literature. The LBD process starts by collecting literature from an information retrieval system such as PubMed, then seeking relationships within the unstructured literature collection [19].

The first LBD research performed by Swanson (1986) involved an extensive manual search of the literature database [13]. The process involved reading the main titles and abstracts of scientific publications. Swanson and Smalheiser (1996) then went on to develop a system called Arrowsmith [20], which uses the Medline database and goes through the title of articles using co-occurrence of words or phrases to generate hidden connections between articles. The Arrowsmith approach, known as the Statistical or Probabilistic Model, relies on a statistical measure to determine the relationship between publications. It follows a frequency-based metric of word occurrence to obtain targets. The Statistical Model is

one of the earliest and basic methods employed in LBD research. The model is limited because it does not consider the semantics aspect of the terms found in each article. An updated version of the statistical method included the use of scores such as token, frequency, record frequency, term frequency-inverse document frequency (TF-IDF), and relative frequency. However, the drawback to these approaches is that they fail to identify important associations that exist with words that do not frequently occur, even though they are easy to compute and implement.

Since the first LDB tool, several LBD tools have been developed using NLP algorithms leading to many testable hypotheses for which some have been validated experimentally [20]. Modern-day approaches involve more advanced and automated methods and techniques for LBD, utilizing PubMed as the literature database while evaluating the work on Swanson's results. Using LBD, [21] discovered that the drug thalidomide had four new therapeutic applications in myasthenia gravis, chronic hepatitis C, Helicobacter pylori-induced gastritis, and acute pancreatitis. The use of curcuma longa, a dietary substance, was found to have therapeutic benefits on diseases and disorders related to the spine [22]. More recent use of LBD research has been used to identify a potential candidate gene for the interaction between myocardial infraction and depression [23] and also gain insights into the metabolomic processes of cardiac arrest [14].

The co-occurrence method has been mostly replaced by other models such as semantic and distributional models. The semantic model is similar to the co-occurrence model but differs only in defining the relationship that exists between articles. While the co-occurrence model uses the frequency of words or concepts, the semantic model uses a sentence's meaning. The semantic model involves the conversion of natural language words to a logical machine-learning representation of their meaning [12]. This approach enables the system to determine what constitutes a relationship. The semantic model is more popularly used in tools for LBD technique. The use of semantics model eliminates many false positive relationships and ensures that the extracted relationship is more accurate.

Semantic MEDLINE Database (SemMedDB) is an example of a repository that employs semantic predictions using SemRep, a semantic interpreter for biomedical text where relationships are extracted in a discovery pattern. The distributional models are more complex and use the association of words to imply the relationship between words in a sentence [24]. With distributional models (or distributional semantic models), information is derived from the words around each other. With this information, articles within similar word associations are established. The co-occurrence information is used to construct a vector representation of a particular word. When similar vector construction with similar co-occurrence of words is established between articles, then a relationship can be explored. These vector constructions have theoretical backgrounds in cognitive representations of words and approximate the idea of conceptual spaces [25]. Several variations of the different Language processing models exist for the co-occurrence model, semantic model, and distribution model [12]. In line with Swanson's (1986) work, there are two models of LBD technique:

### A. The Open Discovery Model

The open discovery approach begins with a start point problem and finds all possible articles associated with that problem. Open-based discovery starts with concept A, and this concept is used to query the online database such as PubMed to get intermediate-related literature B. The intermediate concept B is then used to consult works of literature to generate a new set of relationships C. the process goes on until the researcher determines the target concept, as shown in Fig. 1. Using Swanson's work as an example, Fish Oil was discovered to treat Raynaud's Disease because it lowers blood viscosity. The hypothesis is then checked to ensure there is no overlap between the start and target literature [26].

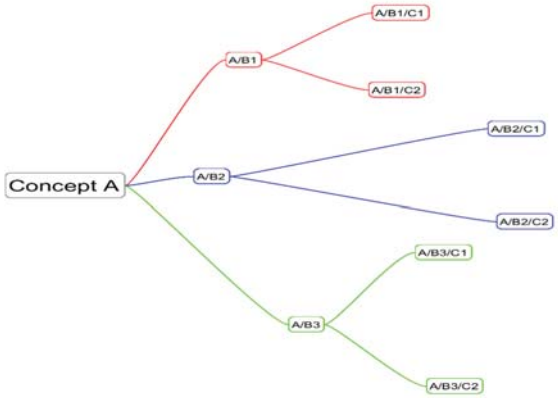


Fig. 1 Open discovery system [15]

### B. The Closed Discovery Model

The closed discovery approach begins with a starting point and an already established endpoint, finding all possible associated articles linking both points [15]. In a closed discovery, relationships are defined using an explicate connection between literature. This begins with the assumption of an undefined relationship between two concepts, A and D, as demonstrated in Fig. 2. The research begins by finding all the existing paths that exist between two concepts A and D. An example of a closed-based discovery is finding a link between hypogonadism and diminished sleep quality in aging men [27]. Hypogonadism result from little production of testosterone in men, while deprivation is often the cause of low testosterone in men. The research investigated what relationship exists between testosterone and age-related sleeping behaviors in men. The research used the semantic MEDLINE tool to study the characteristics of testosterone and sleep individually before using the semantic MEDLINE tool to generate links between testosterone and sleep. Cortisol, a primary stress hormone, was found through LBD as the link connected to declining testosterone levels [27]. LBD would help find literature discoveries that would aid research into treatment or relieve solutions to medical conditions arising due to lupus.

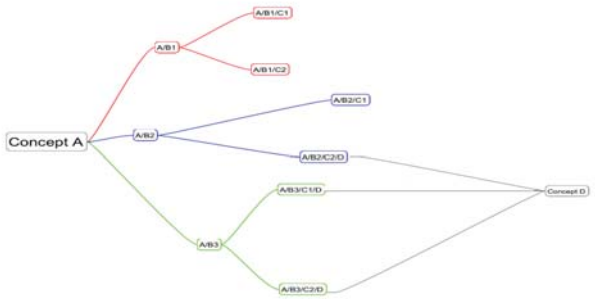


Fig. 2 Closed Discovery System [15]

## IV. Methodology

The general methodology for LBD involves two major processes: Article retrieval and article selection process [28]. Both processes can either be automated using natural processing language tools like semantic Medline or manually by painstaking going through the individual articles. Usually, both processes are involved for speed and efficiency [28]. This research methodology implements a new design approach for LBD on the musculoskeletal manifestation of lupus. Fig. 4 shows the methodology for this research. This research begins by investigating lupus and literature-based discoveries, and the first major step is Identifying the medical conditions associated with lupus manifestation in the musculoskeletal system through text mining of literature in PubMed. This research identifies the various symptoms and underlining diseases associated with lupus based on publications in PubMed. The following search terms were used to identify publications related to musculoskeletal lupus disease using Medical Subject Headings (MeSH) search in PubMed.

- ((”Lupus Erythematosus, Systemic”[Mesh]) AND ”Musculoskeletal System”[Mesh]) AND ( ”Diagnosis”[Mesh] OR ”Signs and Symptoms”[Mesh] )

MeSH terms are official words or phrases selected to represent particular biomedical concepts. Articles are labeled based on these MeSH terms to give better search results in PubMed [29].

908 PubMed publications from different journals were identified in October 2021. These publications included research, reports, clinical trials, journal articles, case reports, etc. After carefully analyzing these publications, 109 musculoskeletal medical conditions were identified following the methodology approach in Fig. 4. Table I shows the major underlining disease and symptoms identified from PubMed search and their occurrence based on MeSH term literature search.

The word cloud in Fig. 3 shows the resulting medical conditions identified in the MeSh term search. The major medical conditions identified from the MeSH term are osteonecrosis, arthropathy, myopathy, myocardial infarction, and shrinking lung syndrome, among others, as shown with larger text in the word cloud of Fig. 3. The word cloud is based on the publications count from the MeSh term search that identified the medical conditions of lupus manifestation

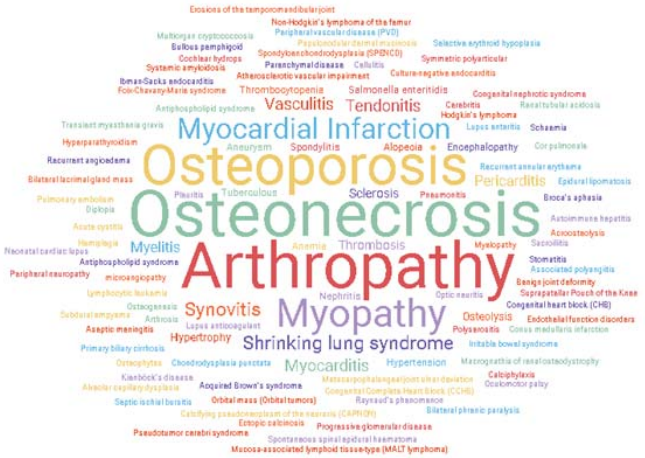


Fig. 3 Word cloud of conditions associated with musculoskeletal lupus disease identified in the Mesh search literature

in the musculoskeletal system. Osteonecrosis was found to have the most count occurrence in the MeSH term of literature studies. Osteonecrosis involves the death of cells in an organ or tissue caused by a lack of blood and oxygen in the tissues of a particular part. Arthropathy is another very common medical condition associated with lupus on the musculoskeletal system, which usually manifests as arthritis.

## V. Results of LBD Multifaceted Approach on the Musculoskeletal Manifestation of Lupus

SemRep software application is used as the LBD technique outlined in Fig. 4 of this paper. SemRep software application predicts publications extracted from PubMed titles. Semantic MEDLINE is a web-based application that visualizes SemRep predictions extracted from MEDLINE citations. SemRep works by it combine syntactic and semantic principles with structured biomedical domain knowledge contained in the Unified Medical Language System (UMLS) to extract semantic relations. SemRep Semantic model extracts meanings from article topics and uses the meaning to predict results for user queries [30]. Fig. 5 shows an example of how the semantic model takes an article topic and translates it to the semantic meaning, thus producing relevant results for a search on the treatment of arthritis. Other relationships can be derived such as disrupts, affects, interacts_with, simulates, etc. This visualization is displayed as an interactive graph with colored nodes and edges. The color symbolizes the node and the relationship between nodes shown by edges [27].

An open LBD search using semantic Medline is done on the major musculoskeletal manifestation of lupus identified in works of literature, as shown in Table I. An example is shown for some parts of Arthropathy medical condition in Table II. Each sentence from the abstract and title that reveals a treatment is shown in the first column of Table II, along with the PubMed publication number. The classification is extracted manually from the first column using the proposed treatment from the first column in Table II. Medical treatment classification is divided into Pharmacology, Naturopathic Medicine, Procedure, or Physiotherapy. This

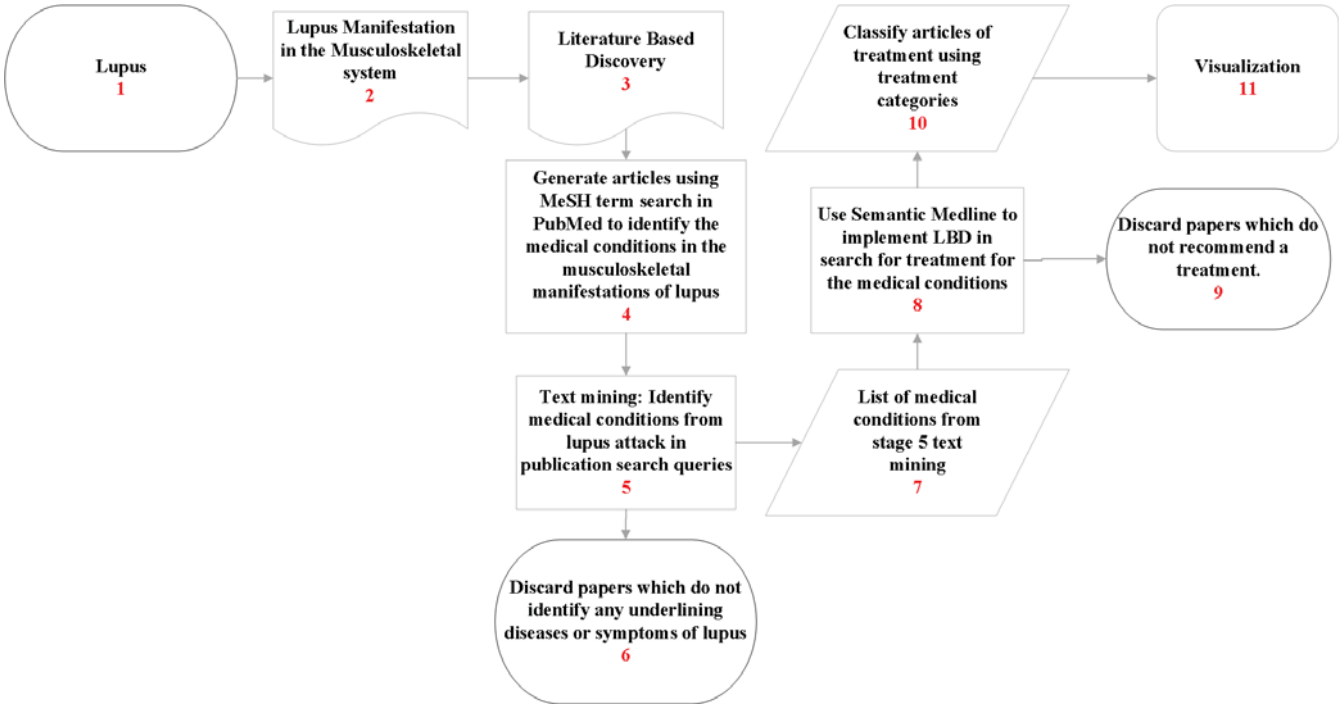


Fig. 4 Methodology Flowchart

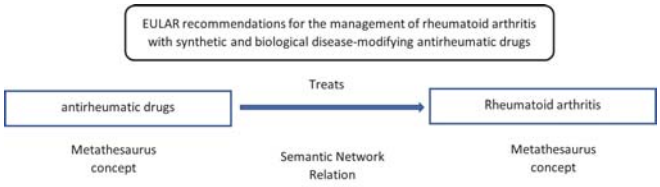


Fig. 5 Semantic interpretation

classification gives an idea of treatment solutions indicated in the publications.

Table III shows the summary for the number of publications found using Semantic MEDLINE LBD search and the treatment classification identified from each publication (Pharmacotherapy, Procedure, Naturopathy, Physiotherapy). Unsurprisingly, the result shows numerous treatment options involving pharmacotherapy and procedure (surgical and non-surgical procedures). The Shrinking lung syndrome (SLS) had the lowest publication result for LBD treatment options.

Historically, an interactive node graph has represented literature-based discovery visualization architecture. An interactive node graph has been used in applications such as LION LBD and Semantic Medline [15]. These visualizations are adequate when the literature data are few. However, they become occluded, and very hard to get insights when the literature data gets large into hundreds of data points. This research implements interactive zoomable circle packing graph visualization to solve this issue. Circle packing can visualize a large amount of data in a hierarchy. Circle packing gives a good overview of a large dataset with a clear representation of grouping and structural relations, as seen in Fig. 6. The circles' size indicates the number of statements in a publication that indicate treatment of the medical condition. Thus, the circle's size would signify the paper's importance regarding treatment possibilities. However, treatment ideas could come from the real world's faintest possibilities. The colors indicate the color calcifications, as shown in the legend. Fig. 6 shows an example of LBD treatment visualization for myopathy medical conditions based on the treatment classification.

## VI. Discussion

LBD techniques have been used on various chronic kidney and cancer diseases [15], [31]. To the knowledge of authors, no prior publication exists using LBD technique to find hidden treatments for lupus flare or any similar multifaceted diseases like lupus. Diseases such as kidney disease and cancer show few specific symptoms, which are usually not ambiguous. On the other hand, lupus might exhibit any number of different medical conditions called flares, which in turn have their symptoms. These intricate symptoms make lupus a complex disease that may require multiple treatment options. From the literature investigation in this research, 109 musculoskeletal lupus flares were identified. A total of 627 publications were identified as potential LBD treatments for the top 10 medical conditions identified as lupus musculoskeletal manifestation. These LBD results address the research question of how literature-based discovery can be used to discover potentially hidden treatment options for lupus or other multifaceted diseases. The approach outlined in this research presents a tedious but rigorous approach to identify hidden research that could potentially be a source of further treatment for lupus flares. The LBD approach of this research gives various treatment options to common musculoskeletal lupus flares. Lupus researchers could use this

TABLE I
MAJOR CONDITIONS ASSOCIATED WITH LUPUS MUSCULOSKELETAL MANIFESTATION AS THEY OCCUR IN THE MESH SEARCH IN PUBMED LITERATURE

| Medical conditions | Variations of medical conditions identified | No of Occurrences |
|---|---|---|
| Osteonecrosis (Bone Necrosis) | corticosteroid-associated osteonecrosis, multifocal osteonecrosis, symptomatic knee osteonecrosis (KON), asymptomatic osteonecrosis, Bilateral osteonecrosis of the femoral head, osteonecrosis of the femoral head, Bilateral calcaneal osteonecrosis, osteonecrosis of bone, asymptomatic avascular osteonecrosis (AVN), corticosteroid-induced osteonecrosis, osteonecrosis in distal tibia, avascular osteonecrosis, multiple osteonecrosis, noncollapsed and asymptomatic osteonecrosis of the femoral head associated with corticosteroids, Steroid-induced osteonecrosis, Bilateral distal tibial osteonecrosis, Atraumatic avascular necrosis (AVN), hip necrosis, avascular necrosis (AVN), aseptic necrosis of the temporomandibular joint, aseptic necrosis of femur heads, Multiple aseptic bone necrosis, Aseptic necrosis of the temporomandibular joint, avascular necrosis, Avascular necrosis of the mandibular, Perfusion of the femoral head, aseptic necrosis, Avascular necrosis of the hip, Aseptic necrosis of bone, Avascular necrosis of the talus, Avascular necrosis of the scaphoid, Avascular necrosis of the metatarsal head, Ischemic necrosis of bone | 60 |
| Arthropathy | hand arthropathy, arthritis , Enthesitis, arthralgia, hand and wrist arthralgia (HA), Jaccoud's arthropathy, erosive, arthropathy, unilateral erosive arthritis, hand osteoarthritis (OA), Retro-odontoid Pseudotumor, inflammatory arthritis, rheumatoid arthritis, Jaccoud's deformity, peripheral arthritis, Mycobacterium kansasii arthritis, Arthropathy of the Sternoclavicular Joint, erosive arthritis with extracapsular cysts, Cricoarytenoid arthritis, arthritis of the finger joints, Pseudoseptic arthritis, inflammatory arthroplasties, erosive arthritis, Total knee arthroplasty with corticosteroid use Hip arthroplasty, Enterococcal arthritis, Deforming arthropathy, ephemeral migrating asymmetrical arthralgia, staphylococcal septic arthritis of the joint, Salmonella septic arthritis, Gonococcal arthritis, Polyarthritis, polyarthralgia, polyarthritis, Gonarthritis | 59 |
| Osteoporosis | bone mineral density (BMD) loss, Secondary osteoporosis deterioration of bone mineral density (BMD), low bone density, glucocorticoid-induced Osteoporosis, osteoporotic compression fractures (steroid use), Osteoporosis in murine loss of trabecular bone (BMD loss), Osteoporosis due to corticosteroids | 51 |
| Myopathy | subclinical cardiomyopathy, Smooth muscle myopathy, Lupus myopathy, vacuolar myopathy, diaphragm myopathy, Vacuolar myopathy associated induced by chloroquine, inflammatory myopathy, cardiomyopathy, Restrictive cardiomyopathy, valvular cardiomyopathy, Hypertrophic cardiomyopathy, hypertrophic cardiomyopathy due to Chloroquine use, Chloroquine-induced neuromyopathy Joints Involvement, Myositis ossificans (also known as traumatic myositis ossificans or myositis ossificans circumscripta), dermatomyositis, polymyositis, Periorbital heliotrope edema,Primary Tuberculous Pyomyositis | 34 |
| Myocardial Infarction | myocardial inflammation, myocardial edema, accelerated atherosclerotic vascular disease (ASVD), Heart attack (Myocardial injury), myocardial dysfunction and heart failure, myocardial perfusion defects, myocardial perfusion abnormalities, myocardial derangements, myocardial dystrophy, focal myocardial fibrosis, myocardial fibrosis | 24 |
| Shrinking lung syndrome (SLS) | | 14 |
| Synovitis | Synovial effusion, bone erosions, metatarsophalangeal (MTP) joint, subclinical synovitis in hand or wrist joints, joint synovitis, Synovial Chondromatosis (osteochondromatosis), Synovial proliferation, synovial hypertrophy, synovial hypertrophy, metacarpophalangeal (MCP) joint synovial hypertrophy, subclinical synovial hypertrophy | 11 |
| Vasculitis | cerebral neutrophilic vasculitis, Mesenteric vasculitis, mesenteric vasculitis, systemic vasculitis, lymphocytic vasculitis, Urticarial vasculitis, Skeletal muscle lymphocytic vasculitis | 11 |
| Myocarditis | Lupus Myocarditis | 10 |
| Tendonitis | tendons involvements, tendon dislocation, tendon tear, tendon thinning | 10 |



approach to test the discoveries and validate the efficiency on lupus flares, particularly the options not widely used in conventional medicine.

Zoomable interactive circle packing visualization proposed in this research gives better navigation, as shown in Fig. 6, compared to the widely used interactive node graph visualization. The node graph visualization shows the relationship between different concepts, but it becomes difficult to understand as the size and complexity of the graph increase. The zoomable interactive circle packing visualization clearly distinguishes concepts while still showing vivid relationships between them using circle hierarchies addressing visualization issues from a large-scale LBD database. While lupus flares can be treated and help to bring relief to patients, there is no known cure to rid a patient of lupus diseases completely. This research is limited in addressing treatment options for the root cause of lupus. To expand on this research, LBD treatment technique should be applied to the symptoms of each identified lupus flare to provide a next-level layer of hidden literature treatment discoveries.

TABLE II
CLASSIFICATION EXAMPLE OF TREATMENTS IDENTIFIED FROM LBD

| Arthropathy | Classification |
|---|---|
| 32561964—Arthroplasty, Replacement, Partial Knee—topp—TREATS—Degenerative polyarthritis—dsyn—Outcomes of lateral unicompartmental knee arthroplasty in post-traumatic osteoarthritis, a retrospective comparative study.PURPOSE: We asked whether the clinical and radiographic outcomes and survivorship after unicompartmental knee arthroplasty (UKA) for osteoarthritis (OA) consequent to lateral tibial plateau fracture were comparable with those obtained after lateral UKA for primary OA. | Procedure |
| 32583972—Analgesics, Opioid—orch—TREATS—Chronic disease—dsyn—A Systematic Review and Meta-Analysis.OBJECTIVE: Opioids have long been prescribed for chronic pain conditions, including osteoarthritis (OA). | Pharmacotherapy |
| 32569719—Medicine, Folk—topp—TREATS—Inflammation—patf—Nees, is one of the traditional plant used as a folk medicine for the management of inflammation, arthritis, viral-bacterial infections and other ailments in India, China, Malaysia and other South-East Asian countries. | Naturopathy |
| 32588487—Physical therapy—topp—PREVENTS—Pain—sosy—A qualitative study exploring barriers and facilitators to physical therapy utilization for knee osteoarthritis.BACKGROUND: Physical therapy (PT) is recommended to reduce pain and improve function. | Physiotherapy |

TABLE III
NUMBER OF WORKS OF THE LITERATURE IDENTIFIED FOR EACH MEDICAL CONDITION

| Medical Condition | Naturopathy | Pharmacotherapy | Physiotherapy | Procedure |
|---|---|---|---|---|
| Myopathy | 2 | 5 | 12 | 11 |
| Tendonitis/Peritendinitis | 2 | 23 | 43 | 15 |
| Osteonecrosis | 5 | 47 | 2 | 75 |
| Vasculitis | 0 | 20 | 0 | 1 |
| Arthropathy | 4 | 15 | 11 | 18 |
| Myocardial Infarction | 2 | 31 | 0 | 51 |
| Myocarditis | 0 | 48 | 0 | 9 |
| Osteopenia and Osteoporosis | 13 | 75 | 3 | 11 |
| Shrinking lung syndrome (SLS) | 0 | 4 | 0 | 1 |
| Synovitis | 2 | 43 | 0 | 23 |

## VII. CONCLUSION

This research highlights the medical conditions associated with the musculoskeletal manifestations of lupus reported in works of literature. Using a LBD tool in semantic Medline, a potential treatment for the various medical conditions can be identified and visualized using interactive circle packing. LBD is a field with the immense potential to provide plausible hypotheses to solve various issues. It is particularly useful in the medical field, where numerous medical conditions are connected in ways that have not been explored. The major contribution of this research is outlining a robust approach to applying LBD to lupus, a multifaceted disease. The result of this LBD approach provides publications in PubMed, which could be explored as therapeutic options for the musculoskeletal manifestations of lupus. An insightful visualization technique using interactive circle packing is employed to provide better navigation through the bulky literature result from the application of LBD. This approach can be applied to other manifestations of lupus, such as renal or hematological manifestations. The approach outlined in this paper also highlights the different treatment classifications in which treatment research can be explored. The limitation of this research is that this LBD approach was nOt designed to treat the fundamental cause of lupus but rather to relieve the resulting medical conditions resulting from lupus flares on the musculoskeletal system. The proposed approach of this paper successfully identifies the medical conditions associated with lupus musculoskeletal manifestations before using the Medline LBD tool to identify treatment connections available in the literature.

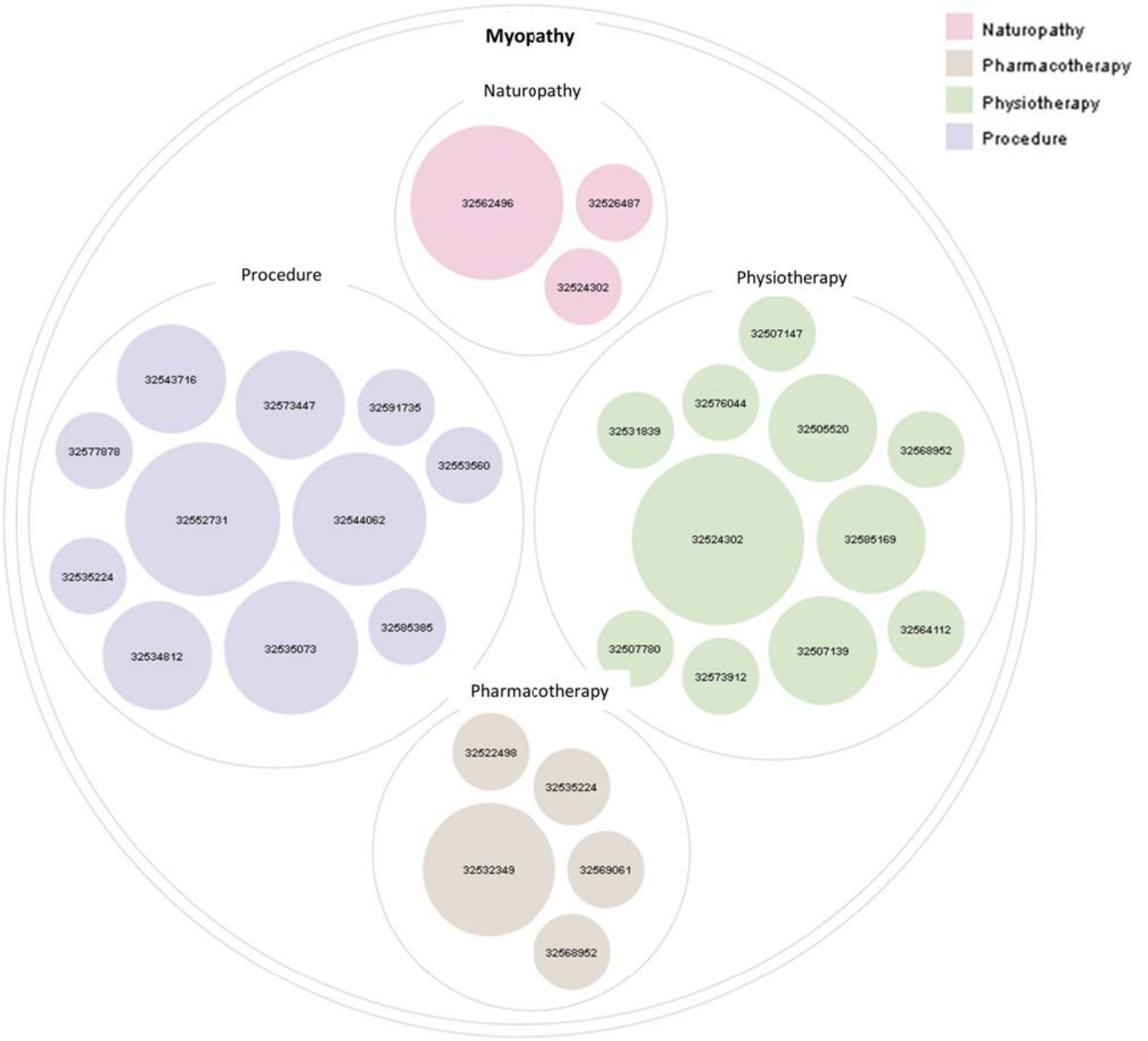


Fig. 6 Interactive Circle Packing visualization for Arthropathy LBD search in Semantic Medline

## REFERENCES

[1] K. Mahmoud, A. Zayat, and E. M. Vital, "Musculoskeletal manifestations of systemic lupus erythmatosus," Current opinion in rheumatology, vol. 29, pp. 486-492, 2017.
[2] T. Dörner and R. Furie, "Novel paradigms in systemic lupus erythematosus," The Lancet, vol. 393, pp. 2344-2358, 2019.
[3] S. Yavuz and P. E. Lipsky, "Current Status of the Evaluation and Management of Lupus Patients and Future Prospects," Frontiers in medicine, vol. 8, 2021.
[4] E. Heiman, S. S. Lim, G. Bao, and C. Drenkard, "Depressive symptoms are associated with low treatment adherence in African American individuals with systemic lupus erythematosus," Journal of clinical rheumatology: practical reports on rheumatic & musculoskeletal diseases, vol. 24, p. 368, 2018.
[5] G. R. V. Hughes, Lupus: The Facts: Oxford University Press, 2000.
[6] A. Stockl, "Complex syndromes, ambivalent diagnosis, and existential uncertainty: the case of Systemic Lupus Erythematosus (SLE)," Social science & medicine, vol. 65, pp. 1549-1559, 2007.
[7] S. R. Johnson, R. Brinks, K. H. Costenbader, D. Daikh, M. Mosca, R. Ramsey-Goldman, et al., "Performance of the 2019 EULAR/ACR classification criteria for systemic lupus erythematosus in early disease, across sexes and ethnicities," Annals of the rheumatic diseases, vol. 79, pp. 1333-1339, 2020.
[8] M. E. Orme, A. Voreck, R. Aksouh, R. Ramsey-Goldman, and M. W. Schreurs, "Systematic review of anti-dsDNA testing for systemic lupus erythematosus: a meta-analysis of the diagnostic test specificity of an anti-dsDNA fluorescence enzyme immunoassay," Autoimmunity Reviews, vol. 20, p. 102943, 2021.
[9] C. Arrigoni, S. Lopa, C. Candrian, and M. Moretti, "Organs-on-a-chip as model systems for multifactorial musculoskeletal diseases," Current opinion in biotechnology, vol. 63, pp. 79-88, 2020.
[10] K. Mahmoud, A. Zayat, and E. M. Vital, "Musculoskeletal manifestations of systemic lupus erythmatosus," Current opinion in rheumatology, vol. 29, pp. 486-492, 2017.
[11] R. Johnson, A. Watkinson, and M. Mabe, "The STM report: An overview of scientific and scholarly publishing," International Association of Scientific, Technical and Medical Publishers, pp. 1-214, 2018.
[12] S. Henry and B. T. McInnes, "Literature based discovery: models, methods, and trends," Journal of biomedical informatics, vol. 74, pp. 20-32, 2017.
[13] H.-T. Yang, J.-H. Ju, Y.-T. Wong, I. Shmulevich, and J.-H. Chiang,

”Literature-based discovery of new candidates for drug repurposing,” Briefings in bioinformatics, vol. 18, pp. 488-497, 2017.
[14] S. Henry, D. S. Wijesinghe, A. Myers, and B. T. McInnes, ”Using literature based discovery to gain insights into the metabolomic processes of cardiac arrest,” Frontiers in Research Metrics and Analytics, vol. 6, p. 32, 2021.
[15] S. Pyysalo, S. Baker, I. Ali, S. Haselwimmer, T. Shah, A. Young, et al., ”LION LBD: a literature-based discovery system for cancer biology,” Bioinformatics, vol. 35, pp. 1553-1561, 2019.
[16] Y. Zhu, W. Jung, F. Wang, and C. Che, ”Drug repurposing against Parkinson’s disease by text mining the scientific literature,” Library Hi Tech, 2020.
[17] G. Crichton, S. Baker, Y. Guo, and A. Korhonen, ”Neural networks for open and closed Literature-based Discovery,” Plos one, vol. 15, p. e0232891, 2020.
[18] A. Usai, M. Pironti, M. Mital, and C. A. Mejri, ”Knowledge discovery out of text data: a systematic review via text mining,” Journal of knowledge management, 2018.
[19] V. Gopalakrishnan, K. Jha, W. Jin, and A. Zhang, ”A survey on literature based discovery approaches in biomedical domain,” Journal of biomedical informatics, vol. 93, p. 103141, 2019.
[20] M. Weeber, J. A. Kors, and B. Mons, ”Online tools to support literature-based discovery in the life sciences,” Briefings in bioinformatics, vol. 6, pp. 277-286, 2005.
[21] M. Weeber, R. Vos, H. Klein, L. T. de Jong-van den Berg, A. R. Aronson, and G. Molema, ”Generating hypotheses by discovering implicit associations in the literature: a case report of a search for new potential therapeutic uses for thalidomide,” Journal of the American Medical Informatics Association, vol. 10, pp. 252-259, 2003.
[22] P. Srinivasan and B. Libbus, ”Mining MEDLINE for implicit links between dietary substances and diseases,” Bioinformatics, vol. 20, pp. i290-i296, 2004.
[23] Z. Dai, Q. Li, G. Yang, Y. Wang, Y. Liu, Z. Zheng, et al., ”Using literature-based discovery to identify candidate genes for the interaction between myocardial infarction and depression,” BMC Medical Genetics, vol. 20, pp. 1-10, 2019.
[24] A. Lenci, ”Distributional models of word meaning,” Annual review of Linguistics, vol. 4, pp. 151-171, 2018.
[25] P. Bruza, R. Cole, D. Song, and Z. Bari, ”Towards operational abduction from a cognitive perspective,” Logic Journal of IGPL, vol. 14, pp. 161-177, 2006.
[26] M. C. Ganiz, W. M. Pottenger, and C. D. Janneck, ”Recent advances in literature based discovery,” Journal of the American Society for Information Science and Technology, JASIST (Submitted), 2005.
[27] C. M. Miller, T. C. Rindflesch, M. Fiszman, D. Hristovski, D. Shin, G. Rosemblat, et al., ”A closed literature-based discovery technique finds a mechanistic link between hypogonadism and diminished sleep quality in aging men,” Sleep, vol. 35, pp. 279-285, 2012.
[28] M. Thilakaratne, K. Falkner, and T. Atapattu, ”A systematic review on literature-based discovery: general overview, methodology, & statistical analysis,” ACM Computing Surveys (CSUR), vol. 52, pp. 1-34, 2019.
[29] C. E. Lipscomb, ”Medical subject headings (MeSH),” Bulletin of the Medical Library Association, vol. 88, p. 265, 2000.
[30] H. Kilicoglu, G. Rosemblat, M. Fiszman, and D. Shin, ”Broad-coverage biomedical relation extraction with SemRep,” BMC bioinformatics, vol. 21, pp. 1-28, 2020.
[31] R. N. Kostoff and U. Patel, ”Literature-related discovery and innovation: Chronic kidney disease,” Technological Forecasting and Social Change, vol. 91, pp. 341-351, 2015.