Prediction of Response to Intra-Articular Injections of Hyaluronic Acid for Knee Osteoarthritis

Eva K. Lee[1,2,3][a], Fan Yuan[2], Barton J. Mann[4] and Marlene DeMaio[4,5]
[1]*Center for Operations Research in Medicine and Healthcare, The Data and Analytics Innovation Institute, Atlanta, U.S.A.*
[2]*Georgia Institute of Technology, Atlanta, U.S.A.*
[3]*AccuHealth Technologies, Atlanta, Georgia, U.S.A.*
[4]*The American Orthopedic Society for Sports Medicine, Chicago, U.S.A.*
[5]*Medical Corps, United States Navy, U.S.A.*

Keywords: Knee Osteoarthritis, Injections of Hyaluronic Acid, Machine Learning for Evidence-Based Practice, Branch-and-Bound, Particle Swarm Optimization.

Abstract: Osteoarthritis (OA) is a degenerative joint disease, with the knee the most frequently affected joint. Fifty percent of knee OA patients eventually undergo surgical procedures such as knee replacement to address pain and functional limitations. A significant number of these surgeries may be unnecessary, with intra-articular injections of hyaluronic acid (HA) serving as a non-invasive, cost-effective alternative. Although research studies have clearly demonstrated that HA improves knee function, the efficacy of this treatment remains controversial. Many physicians have observed that effects depend on several patient characteristics such as age, weight, gender, severity of the OA, and technical issues such as injection site and placement. In this study, a multi-stage, multi-group machine learning model is utilized to uncover discriminatory features that can predict the response status of knee OA patients to different types of HA treatment. The algorithm can identify certain subgroups of knee OA patients who respond well to HA therapy. The baseline results, based on factors such as patients' weight, smoking status and frequency, identifies the patients most suitable for HA injection. The model can achieve more than 89% blind prediction accuracy. The data derived from this study allows physicians to administer HA products more selectively, resulting in a higher therapy success rate. Information on the predicted responses could also be shared with patients beforehand to incorporate their values and preferences into treatment selection. The model's decision support tools also allow physicians to quickly determine whether a patient is exhibiting at least the expected treatment response, and if not, to potentially take corrective action. To the best of our knowledge, this work represents the first machine learning approach that predicts patient responses to HA injections for knee osteoarthritis. The model is generalizable and can be used to predict patient responses to other treatments and conditions.

1 INTRODUCTION

Osteoarthritis (OA) is a degenerative joint disease that can affect the many tissues of the joint. It is one of the most prevalent and costly chronic medical conditions. affecting more than 32.5 million adults in the United States (United States Bone and Joint Initiative 2018). During 2019–2021, 21.2% of U.S. adults (53.2 million) reported an arthritis diagnosis. (Elgaddal, et al., 2022; Fallon, et. al., 2023) and by 2040, it is projected to increase to 78.4 million Americans.

Arthritis increasingly is reported as the main cause of disability among U.S. adults (Theis, K.A. et al., 2019). Annual direct medical care expenditures for osteoarthritis in the U.S. is estimated to exceed $495.5 billion (United States Bone and Joint Initiative, 2019; Lo, et al., 2020). Worldwide, about 528 million people were living with osteoarthritis in 2019 (WHO 2023, GBD 2019). It is estimated that those with OA pain lost 31% of productive time at work due to presenteeism and 8% due to absenteeism, compared to 16% and 4%, respectively, for those who did not report OA pain (Leifer et al., 2022).

[a] https://orcid.org/0000-0003-0415-4640

Lee, E., Yuan, F., Mann, B. and DeMaio, M.
Prediction of Response to Intra-Articular Injections of Hyaluronic Acid for Knee Osteoarthritis.
DOI: 10.5220/0013071800003838

In *Proceedings of the 16th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2024) - Volume 1: KDIR*, pages 497-508
ISBN: 978-989-758-716-0; ISSN: 2184-3228

There is no known cure for OA. Instead, treatments aim to reduce pain, maintain or improve joint mobility, and limit functional impairment. Treatments are usually non-operative, such as physical therapy, rest, modification of daily activities, analgesics, and anti-inflammatory medication. For individuals who desire or require a high level of physical activity, rest and activity reduction are not viable treatment options. Oral non-steroidal anti-inflammatory drugs (NSAIDs) are often recommended, although frequent and serious adverse effects of NSAIDs have been reported (Zhang et al., 2010, Salis and Sainsbury, 2024). Over the past 25 years, intra-articular injection of hyaluronic acid (and similar hyaluronan preparations) has emerged as an additional tool for managing the symptoms of OA for patients who fail to respond to other conservative treatments. However, controversies exist regarding its safety and efficacy, the number of injections and courses, type of preparation, duration of its effects, and combining it with other drugs or molecules (Chavda et al., 2022). Other factors include patient characteristics such as age, weight, gender, and severity of the OA.

Knee OA happens when the cartilage in the knee joint breaks down, enabling the bones to rub together. The friction makes the knees hurt, become stiff, and sometimes swell. Knee OA is a leading cause of arthritis disability (Cui et al., 2020). Of significance for sport medicine, heavy physical activity, participation in high intensity contact sports, participation in certain elite level sports, and knee injury have all been linked to the development of knee OA (Chan, et al., 2020; Driban, et al., 2017; Lohmander, et al., 2007; McAlindon et al., 1999; Sharma, 2001; Spector et al., 1996; Turner, et al., 2000). Although it cannot be cured, treatments are available to slow its progression and ease the symptoms. Knee OA alone results in the loss of an average of 13 days of work per year (versus 3 days for those without Knee OA (Ayis & Dieppe, 2009).

Knee osteoarthritis affects more than 14 million Americans, and its symptoms often lead to physical inabilities, disabilities, and all sorts of inconveniences for patients. It is estimated that knee osteoarthritis is associated with approximately $27 billion in total healthcare costs every year, with about 800,000 knee surgeries performed annually. Specifically, 99% of these knee replacements are done to address pain and functional limitations (Barbour et al., 2017). In a multicenter longitudinal cohort study, it was reported that about one-third of knee replacements may be unnecessary (Riddle et al., 2014).

The management of knee pain depends on the diagnosis, inciting activity, underlying medical conditions, body mass, and chronicity. In general, non-operative management is the mainstay of initial treatment and includes rehabilitation, activity modification, weight loss when indicated, shoe orthoses, local modalities, and medication. The oral medication often prescribed is an analgesic, usually with anti-inflammatory properties. Supplements, such as chondroitin sulfate and glucosamine, have been shown to have a role. Since 1997, the regimen has expanded to include viscosupplementation. These agents are preparations of hyaluronic acid or their derivatives (HA) which are sterilely injected into the knee. Although research studies have clearly demonstrated that HA improves knee function, the efficacy of this treatment remains controversial. Many physicians have observed that effects seem to depend on several patient characteristics, such as age, weight, gender, severity of the OA and technical issues such as injection site and placement (Mora et al., 2018).

This study aims to answer an important question: whether different types of patients may respond differently to HA treatment. Is it possible to identify certain subgroups of knee OA patients who respond well (or those who don't) to HA therapy? Further, we question whether it is possible prior to treatment to predict a patient's response to HA injections based on patient and treatment characteristics. Physicians could then make empirically informed decisions about whether to treat a particular patient with HA and perhaps which type of HA preparation is most likely to produce the best treatment response for that individual patient.

The goal of this study is to evaluate which patient population, or patient characteristics, would benefit most from HA injection. Since at least 18% of out-patient visits to military treatment facilities by active-duty personnel are attributed to painful knee disorders, our study focuses on these patients. The study uses a prospective, double-blinded clinical trial. A multi-stage, multi-group machine learning model (Lee et al., 2016b; Lee, 2017; Lee & Egan, 2022; Lee et al., 2021, 2023a, 2023b) described in Section 2.3 is used to uncover discriminatory patterns that can predict suitability of treatment and outcomes. The resulting predictive rule can be implemented as part of a clinical practice guideline for evidence-based intervention. The model enables physicians to administer HA products more selectively and effectively to the targeted population to maximize cost effectiveness and the percentage of patients who experience a successful HA injection.

2 METHODS AND STUDY DESIGN

2.1 Patient Cohort, Treatment, and Outcome Measures

2.1.1 Patient Data

Three group of patients (active-duty military personnel, military retirees, and their families) through the Department of Orthopaedics at the Naval Medical Center Portsmouth were included. The cohort includes those between 18 and 65 who sought treatment for symptomatic osteoarthritis of the knee. All patients were evaluated by a board-certified orthopaedic surgeon. Each patient has had radiographic evidence of knee OA with a minimum Kellgren-Lawrence score of 1, has experienced symptoms for more than three months, has failed a minimum of three months of non-operative treatment, including, but not limited to, analgesic and anti-inflammatory medication, cortisone injection, physical therapy, bracing, and/or heel wedge. The cohort excludes patients with precautions or contraindications for viscosupplementation, those who had a cortisone injection within the past three months, those who had prior HA injections at any point, those with a history of deep knee infection, those currently experiencing peripheral neuropathy, chondrocalcinosis, or knee ligament instability, and those who were candidates for knee surgery.

Patients were randomly assigned to receive either Hylan G-F 20 (Synvisc®) [Sanofi Biosurgery, Cambridge, MA, USA], a high molecular weight (MW = 6000 kDa) cross-linked HA product derived from an avian source, or EUFLEXXA® [bioengineered 1% sodium hyaluronate (IA-BioHA); Ferring Pharmaceuticals, Inc., Parsippany, NJ], a medium weight (MW = 2400 - 2600 kDa) HA product derived from bacterial fermentation.

Treatment allocations were randomly assigned by the study pharmacist using the RANDBETWEEN(0,1) function in Microsoft Excel. Physicians, physicians performing the injections, patients, and research personnel were blinded to treatment assignment. To maintain blinding, the pharmacy removed the original manufacturer's label prior to dispensing and relabelled with the protocol title, subject identifier and expiration date. The two HA products had the same volume and color, so there was no ability to discern one from the other at the time of injection.

During a baseline evaluation before the first injection, the following data were collected:

- patient demographic data: age, sex, height, weight, BMI (as calculated from height and weight), and smoking history.
- the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC; Bellamy, 2002) as a measure of knee OA symptoms and functioning.
- the RAND-36 (Hays et al., 1993) as a measure of general health status.
- the MARX Knee Activity Rating Scale (Marx et al., 2001) to assess activity level (running, deceleration, cutting (changing directions while running) and pivoting.
- patient-rated health conditions (a) using a comorbidity questionnaire (Sangha et al., 2003) and (b) quality of life as measured by the EuroQOL EQ-5D (Brooks, 1996).
- a patient-completed Arthritis Self-Efficacy Scale (Lorig et al., 1989), an eight-item instrument that assesses patient's perceived ability to manage arthritis symptoms.

Specific patient treatment expectations (e.g., "Improve ability to go up and down stairs") and the importance of these expectations were evaluated with the scale developed by Mancuso (Mancuso et al., 2001). Patients were also asked to rate their global expectation for their response to the HA injections on a seven-point scale ranging from "No improvement. I don't have much hope that this treatment will help my symptoms at all" to "Excellent improvement. I expect complete or nearly complete relief from knee symptoms." Patients with bilateral OA were instructed to rate only the knee they perceived to be more severe in terms of pain and functional impairment on all instruments and to rate the same knee at baseline and follow-ups.

Prior to the first injection, a physician assessed quadriceps atrophy, presence of antalgic gait, knee effusion, pain on palpation of the knee, range of motion and alignment, and use of medication. Patients also received four baseline radiographs. These included (a) a standing anteroposterior (AP) of the knee weight-bearing view; (b) weight-bearing flexed view 400 posterior-anterior (PA) Rosenberg view; (c) a lateral x-ray at 300; and (d) a Merchant view. Digitized radiographs were evaluated for osteoarthritis severity and for alignment by a board-certified musculoskeletal radiologist and an orthopaedic surgeon blinded to assigned treatment or other patient characteristics. OA severity was rated using the Kellgren-Lawrence Grading System which incorporates joint space narrowing, osteophyte formation, sclerosis and bony deformation observed

on x-rays. Scores range from 0 (no radiographic features of OA) to 4 (large osteophytes, marked joint space narrowing, severe sclerosis, and definite bony deformity). Alignment was determined by measuring the following angles from x-rays: (a) condylar-hip angle of the femoral condylar tangent with respect to the mechanical axis of the femur expressed as degrees of deviation from 90°, negative for varus and positive for valgus; (b) plateau-ankle angle between the tibial margin tangent and the mechanical axis of the tibia expressed as degrees of deviation from 90°, negative for varus and positive for valgus; (c) condylar-plateau angle between the femoral and tibial joint surface tangents; and (d) hip-knee-ankle angle between a line drawn from the center of the femoral head to the midpoint of the tibial eminential spine and another line from this midpoint to the center of the talus surface of the ankle joint. The medial angle between the lines is the HKA angle (varus < 180°).

2.1.2 HA Treatments

Patients received injections every seven days for a total of three injections. Physicians received specific instructions to standardize injection technique. All injections were performed using an anteromedial approach with a 21-gauge 1½" needle. Physicians aspirated the knee joint prior to injection of the HA product to ensure needle placement. Patients were asked to flex and extend their knee a few times following injection to maximize dispersal into the joint. Patients were provided with written post injection and standardized physical therapy instructions. Patients were allowed full weight bearing and full range of motion (active and passive) after injections but were advised to avoid strenuous activity (such as jogging, tennis, etc.) or prolonged weight bearing for the first 48 hours after injection. Patients were also instructed to use ice 30 minutes on and 30 minutes off for 48 hours and take up to 4 gram of acetaminophen per day as need for knee pain, but not to take any 24 hours prior to each visit.

Patients were not offered a second course of HA treatment within the first six months following the final injection. Following the standard clinical practice, those who received a second series of injections after the first six months were not considered treatment failures. Patients who had surgery on the target knee to relieve arthritis symptoms within the first six months following the last HA injection were considered treatment failures.

The protocol was approved by the Institutional Review Board at the data collection site and was registered with ClinicalTrials.gov (identifier: NCT01557868). A physician at the site served as the medical monitor and an independent data and safety board monitored the study.

2.1.3 Primary and Secondary Outcomes

The primary outcome was treatment responder status defined a priori by improvement in the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) Pain Scale (Hochberg et al., 1997; Riddle & Perera, 2020) between baseline and 3-month assessments. The WOMAC Pain Scale is comprised of 5 items and the response format used in this study was the 5-point rating scale. Scores were calculated to range from 0 (worst) to 100 (best). The reliability, validity and responsiveness of the WOMAC Pain Scale have been supported in numerous studies (Bellamy, et al., 2011; Burgers, et al. 2015) and the WOMAC is one of the most widely used outcome instruments in arthritis research. Patients whose pain scores decreased by 20% or more compared with their baseline scores were classified as treatment responders and those whose scores did not meet this criterion were classified as non-responders.

2.2 Machine Learning Predictive Analysis

We apply a multi-stage machine learning approach to analyze how different types of patients may respond differently to HA treatment. The system will uncover discriminatory features in the HA data that will reveal patient and treatment characteristics that predict optimal response to intra-articular injections of hyaluronic acid for knee osteoarthritis. The model determines which patient variables lead to the best outcomes of HA.

Detail of the multi-stage multi-group discriminant analysis via mixed-integer program (DAMIP) model and computational framework is reported in Lee et al. (Lee, 2017; Lee & Egan, 2022; Lee, Wang, et al., 2016; Lee et al., 2021, 2023a, 2023b). Briefly we include the DAMIP formulation below.

Let u_{hgi} represent the binary variable that indicates whether observation i in group g is classified to group h, $h \in \{0\} \cup \mathcal{G}$. Thus, $u_{ggi} = 1$ denotes a correct classification for observation i in group g. The multi-group model with a reserved judgement region is formulated as:

$$\max \sum_{g \in \mathcal{G}} \sum_{j \in \mathcal{O}_g} u_{ggj} \qquad \textbf{(DAMIP)}$$

subject to

$$L_{hgj} = \pi_g f_g(\boldsymbol{x}_j) - \sum_{h \in \mathcal{G}, h \neq g} \lambda_{hg} f_h(\boldsymbol{x}_j), \forall\, h, g \in \mathcal{G}, j \in \mathcal{O}_g \quad (1)$$

$$y_{gj} - L_{hgj} \leq M(1 - u_{hgj}), \qquad \forall\, h, g \in \mathcal{G},\ j \in \mathcal{O}_g \quad (2)$$

$$y_{gj} \le M(1 - u_{0gj}), \qquad \forall\, g \in \mathcal{G},\ j \in \mathcal{O}_g \quad (3)$$

$$y_{gj} - L_{hgj} \ge \varepsilon(1 - u_{hgj}), \qquad \forall\, h, g \in \mathcal{G},\ j \in \mathcal{O}_g \quad (4)$$

$$y_{gj} \ge \varepsilon\, u_{hgj}, \qquad \forall\, h, g \in \mathcal{G},\ j \in \mathcal{O}_g \quad (5)$$

$$\sum_{h \in \{0\} \cup \mathcal{G}} u_{hgj} = 1, \qquad \forall\, g \in \mathcal{G},\ j \in \mathcal{O}_g \quad (6)$$

$$\sum_{j \in \mathcal{O}_g} u_{hgj} \le \lfloor \alpha_{hg} n_g \rfloor, \qquad \forall\, h, g \in \mathcal{G}, g \ne h \quad (7)$$

$$u_{hgj} \in \{0,1\} \qquad \forall\, h \in \{0\} \cup \mathcal{G}, g \in \mathcal{G}, j \in \mathcal{O}_g \quad (8)$$

$$y_{gj} \ge 0, \qquad \forall\, h, g \in \mathcal{G},\ j \in \mathcal{O}_g \quad (9)$$

$$\lambda_{hg} \ge 0 \qquad \forall\, h, g \in \mathcal{G}, g \ne h \quad (10)$$

Here, π_g is the prior probability of group g and $f_g(\boldsymbol{x})$ is the conditional probability density function of group g, $g \in \mathcal{G}$ for the data point $\boldsymbol{x} \in \mathbb{R}^m$. $\mathcal{O}_g$ denote the set of observations in group g, and n_g denote the number of observations in group $g \in \mathcal{G}$. $\alpha_{hg} \in (0,\ 1)$, $h,\ g \in \mathcal{G}$, $h \ne g$ represents the predetermined limit on the inter-group misclassification rate where the observations of group g are misclassified to group h. The group assignment decisions of observations that are classified into a reserved judgment region are denoted by group $g = 0$.

Constraints (1) define the loss functions; constraints (2)-(6) guarantee an observation is uniquely assigned to the group with the maximum value of $L_g(\boldsymbol{x})$ among all group, and constraints (7) set the misclassification limits. With the reserved judgment region in place, the mathematical system ensures that a solution that satisfies the pre-set misclassification rate always exists.

Theorem 1. Given prior probabilities π_g and conditional group density functions $f_g(\boldsymbol{x})$, allocation according to modified posterior probabilities defined by the solution to (DAMIP) is a *universally strongly consistent* method for classification.

Theorem 2. The DAMIP optimization problem is $\mathcal{NP} - Complete$ when the number of groups is greater than 2. The theoretical result holds for DAMIP variants: (a) maximize the minimum value of correct classification rates among all groups; (b) maximize the minimum difference between correct classification and misclassification; and (c) maximize correct classification while constraining the percentage of reserved judgment for each group.

The multi-stage classification approach utilizes the reserved judgment region in DAMIP to improve the classification performance, especially among highly inseparable data. At each stage, DAMIP partitions the observations into an '*easy–to-classify*' subset that is classified to specific groups, and a '*difficult-to-classify*' subset that is classified to a *reserved judgment region*. The group assignment of the difficult-to-classify observations are delayed, thus allowing the DAMIP classifier to maintain a low misclassification error. The observations in the reserved judgment region are moved to the next stage where a new feature set is selected and a new DAMIP classifier is developed. In this way, the multi-stage framework constructs a chain of successive classifiers using different subsets of features. The classifier at the *i*th stage, denoted by f_i, can be represented by a discriminant function $f(\boldsymbol{x_i}, \boldsymbol{\lambda_i})$, which is determined by the feature subset $\boldsymbol{x_i}$, and the decision variables $\boldsymbol{\lambda_i}$ in DAMIP.

At each stage, two models are performed: a single-stage model that solves a DAMIP model without a reserved judgment region and a multi-stage model that solves a DAMIP model with a reserved judgment region. The computational framework selects the better of the two results. The algorithm naturally terminates when there are no observations in the reserved judgment region. To avoid overfitting using too few observations for training, two additional stopping criteria are used to terminate the process: (a) the number of observations is less than a preset minimum value, *n*, and (b) the maximum allowed depth, *d,* is reached. The parameters *n* and *d* are predetermined according to the number of observations and the number of input features in the given data.

Computationally, DAMIP classifier has some distinct characteristics: (a) it is applicable for classification of any number of groups; (b) there is always a feasible solution to the model; (c) the reserved judgement region facilitates successive stage of classification to be performed; (d) DAMIP is able to establish classification rules with good predictive accuracy even when the training set is relatively small; (e) DAMIP classifier can handle imbalanced data; and (f) DAMIP classifier is totally universally consistent.

Figure 1 shows the machine learning framework where features are first selected via an exact branch-and-bound algorithm (BB) and a fast heuristic particle swarm optimization (PSO) (Lee et al., 2023a). The resulting classification rule is subsequently established via the DAMIP classifier. To quantify the accuracy, ten-fold cross validation evaluation is performed. If the results satisfy some pre-set accuracy level, the classification rule is reported. Blind prediction using this rule is then performed. We contrast the BB-PSO/DAMIP results with eight commonly used classifiers: Bernoulli Naïve Bayes, Decision Tree, Gradient Boosting, K-nearest neighbors, Logistic Regression, Neural Network, Random Forest, and Support Vector Machine (SVM).

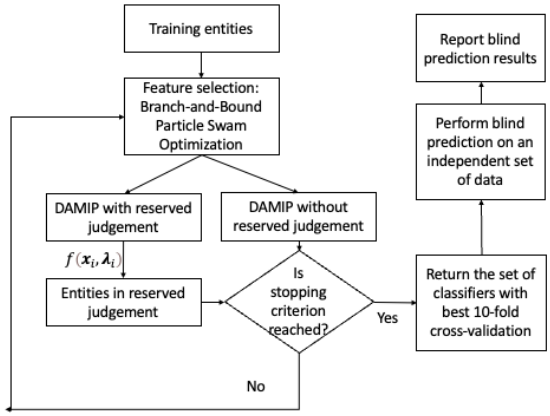


Figure 1: Multi-stage machine learning framework for HA predictive analytics.

In 10-fold cross validation, the training set is partitioned into 10 roughly equal subsets. In each run, 9-fold are selected to train and establish the rule, and the remaining 1-fold is then tested, counting how many of them are classified into which group. Through 10 folds procedure (where each fold is being validated exactly once), we obtain an unbiased estimate of the classification accuracy.

Blind prediction is performed on patients that are independent of the training set to gauge the predictive power of the established rule. These patients have never been used in the feature selection and the machine learning analysis. We run each patient in the blind set through the rule, which returns a group status of the patient. The status is then checked against the clinical status to confirm the accuracy.

The classifier response and outcome prediction rules will culminate in a clinical decision algorithm for the use of viscosupplementation in the treatment of knee OA. For example, a physician determines that HA is indicated for a particular patient. The physician would then enter specific variables (those discriminatory features identified by the classifier) into a clinical computer program and a response set would be generated for the potential outcome after using hyaluronic acid injections. The optimal HA agent(s) would be ranked. The physician would then take this information into account as part of the clinical decision process to select the HA agent for the individual patient.

3 RESULTS

3.1 Patient Characteristics

Of the 273 patients assessed for eligibility, 45 did not meet study criteria, 13 eligible patients declined to participate, and 12 eligible patients could not complete study participation due to anticipated deployment or relocation. The other 203 eligible patients were randomized to treatment: 107 assigned to the Synvisc group and 96 to the Euflexxa group. After randomization, 6 patients were non-compliant with the study protocol, 9 received an excluded intervention, 6 were reassigned, 10 were lost to follow-up and 6 missed the follow-up appointment. Consequently, these patients were not included in the analyses, leaving a total of 166 (87 in the Synvisc group and 79 in the Euflexxa group).

Table 1 summarizes the baseline characteristics of the study participants. The Synvisc and the Euflexxa groups did not differ on demographic or anthropometric variables. The groups also did not differ on co-morbid conditions with the exception that a greater proportion of patients in the Euflexxa group reported depression (21% vs. 10%, p = 0.02). The baseline scores from the patient report measures did

Table 1: Baseline Characteristics of the Study Participants.

Characteristic	Synvisc (N = 107)	Euflexxa (N = 96)	Combined Sample (N = 203)
Age – year	46±10	43±10	45±10
Female sex – no. (%)	44 (41)	36 (38)	80 (39)
Body mass index	30±5	29±5	30±5
Race			
Asian	1 (1)	5 (5)	6 (3)
Black/African-American	36 (34)	21 (22)	57 (28)
Hispanic	5 (5)	6 (6)	11 (5)
White	63 (59)	63 (66)	126 (62)
Other	2 (2)	1 (1)	3 (2)
Married – no. (%)	86 (80)	79 (82)	165 (81)
Current smoker – no. (%)	16 (15)	11(12)	27 (13)
Kellgren-Lawrence Score			
Grade I – no. (%)	28 (26)	37 (39)	65 (32)
Grade II – no. (%)	44 (41)	33 (34)	77 (38)
Grade III – no. (%)	29 (27)	18 (19)	47 (23)
Grade IV–no. (%)	6 (6)	8 (8)	14 (7)
WOMAC Pain Scale	59±17	61±19	60±18
SF-36			
Physical functioning	51±23	54±24	53±23
Mental health	79±15	74±18	77±17
Marx Activity Scale	5±5	5±5	5±5
EuroQOL EQ-5D Health Rating	71±16	70±20	70±19
Arthritis Self-Efficacy Scale	6±2	6±2	6±2
Treatment response expectation	5±1	5±1	5±1
Bilateral HA injections – no. (%)	54 (51)	45 (47)	99 (49)

not significantly differ between the two treatment groups either.

3.2 Primary End Points

Of the 166 patients who completed the 3-month assessment, 84 (50.6%) were classified as treatment responders. Within the Synvisc group, 57.5% were responders compared to 43% of the Euflexxa group ($p = 0.04$). This outcome, as well as those at the 2-week and 6-month follow-ups, is shown in Table 2. Table 3 displays the percentage of patients who were classified as "recovered" based on both statistically reliable improvement in WOMAC Pain Scale scores and a follow-up score that fell within the range of age- and sex-matched patients who reported having no knee problems or any history of knee surgery (see Mann, et al., 2012).

Table 2: Treatment Responders (20% Reduction in WOMAC Pain) by Treatment Group.

Follow-Up	Synvisc	Euflexxa	P Value
2 weeks	56.3%	56.3%	0.55
3 months	57.5%	43.0%	**0.04**
6 months	51.3%	41.5%	0.31

Table 3: Return to Normal on WOMAC Pain Scale by Treatment Group.

Follow-Up	Synvisc	Euflexxa	P Value
2 weeks	36.5%	25.4%	0.20
3 months	38.0%	22.6%	0.06
6 months	33.9%	30.0%	0.31

3.3 Response and Outcome Prediction

We analyze the HA data to uncover patient and treatment factors that predict optimal response to intra-articular injections of hyaluronic acid for knee osteoarthritis. The treatment responder status six months after final injection is measured by 'WOMACP20," Treatment Responder Status Using 20% Reduction in WOMAC Pain Scale. Recovery status is assessed via the KOOS Scale. The machine learning model determines which patient variables lead to the best outcomes of HA. We also perform the prediction for each of the two HA products to gauge their similarities and differences in treatment outcome characteristics.

Table 4 shows the number of patients in the training set and the blind prediction set for predicting reinjection status. In this analysis, for every attribute in which there is missing data, an associated binary attribute is created to capture whether data is missing or not for this field. The number of attributes at three time-points: (a) baseline screening before first injection; (b) prior to second injection (prefix: T0); and (c) six months after final injection (prefix: T5) are 27, 483, and 1215 respectively. Table 5 shows the training set and blind prediction statistics used for predicting treatment responder status and recovery status.

Table 4: Training set and blind prediction set characteristics for predicting reinjection status.

Training set			Blind Prediction Set		
Total	No reinjection	Reinjection	Total	No reinjection	Reinjection
150	111	39	53	40	13

Table 5: Training set and blind prediction set characteristics for predicting treatment responder status and recovery status.

Training set			Blind Prediction Set		
Total	Non-Responder	Responder	Total	Non-Responder	Responder
71	34	37	70	41	29
Synvisc					
40	18	22	36	19	17
Euflexxa					
35	21	14	30	17	13

We summarize herein the best predictive rules for each of the analyses. Table 6 shows the prediction accuracy for no-reinjection versus re-injection using attributes collected up to the three stated time-points.

For the baseline results, factors that appear to be critical includes "Weight," "Currently Smoke Cigarettes," and "Smoking: Number per day." Baseline prediction results are comparable to Pap Smear test accuracy (~70%).

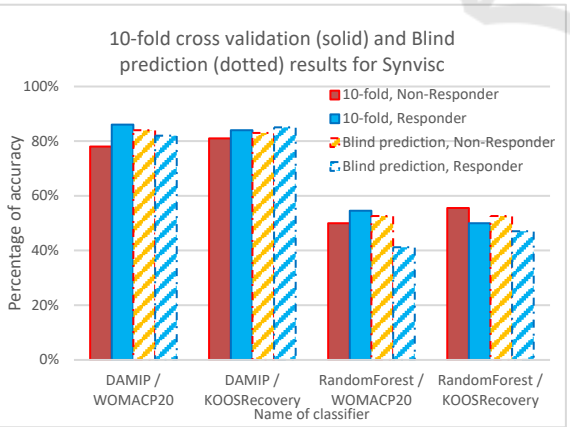
We can observe high accuracy in predicting success for patients using screening and T0 attributes alone (86% blind predictive accuracy). This is very promising for identifying patients early (just after the first injection) who should be targeted for HA intervention (with an expected success outcome). The discriminatory features selected includes the Marx Activity Scale "T0MarxCuttingSymptomFree", "T0MarxCutting", effectiveness of exercise "T0ExerciseEffective", confidence in the injector "T0ConfidenceInjector", and other medications "T0MedicationXEffective."

Including attributes until T5 significantly increases the accuracy for predicting the reinjection group (from 71% to 89%). Early attributes include "T0PhysicalTherapyEffective", "T0MedicationXEffective," and overall health "T0EQRateHealth" continue to appear among the selected features.

Table 6: Best predictive rule for re-injection status when using attributes (a) baseline screening before first injection, (b) prior to second injection, and (c) 6 months after final injection.

Input attributes	**10-fold cross validation**		**blind prediction**	
	No-reinjection	Re-injection	No-reinjection	Re-injection
Baseline screening	**71%**	**71%**	**72%**	71%
Prior to 2nd injection	**89%**	**74%**	**86%**	71%
Input attributes	84%	83%	81%	**89%**

Figure 2 show the 10-fold cross validation and blind prediction accuracies for predicting treatment responder status and recovery status for patients injected with Synvisc and Euflexxa, respectively. For each HA injection, four measurement frameworks are graphed: PSO/DAMIP results for predicting treatment responder and recovery respectively versus the best results from the eight commonly used classifiers, Random Forest. Our PSO/DAMIP framework selected 3-8 discriminatory features whereas Random Forest uses over 40 features with poor results. Although the size of the two groups is rather balanced, the challenge here is due to the highly inseparable data that makes it difficult to classify using traditional approaches. A multi-stage approach allows the partitioning of patients from the same group via different rules (associated with different features).



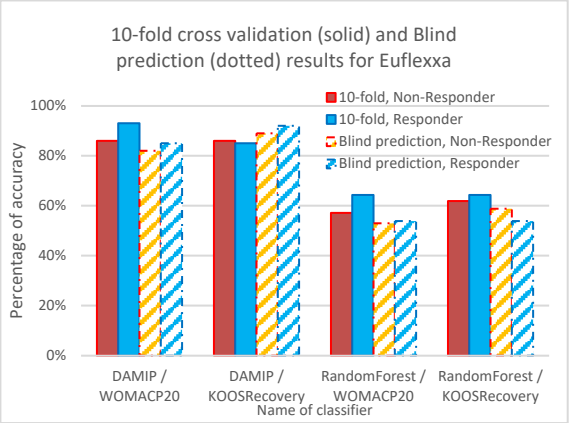


Figure 2: Comparison of the best DAMIP classification rules for predicting treatment responder status and recovery status using Synvisc(top) and Euflexxa (bottom) against the Random Forest approach.

Our study shows that early predictors can be used to determine the group of patients who benefit the most from HA injection. It also allows evidence-based correction to be made during the course of treatment. For example, after T0, the physician can quit treatment based on results from the predictive rule.

4 CONCLUSIONS

In 2019, about 528 million people worldwide were living with osteoarthritis, an increase of 113% since 1990. For 365 million, the knee was the most frequently affected joint. On average, the total cost of knee replacement surgery ranges from $30,000 to $50,000. This includes the cost of the surgery itself, the hospital stays, anesthesia and other associated medical expenses. HA treatment, on the other hand, costs about $900 to $3,000 for a full course (three to five injections administered over several weeks). The range reflects the variations due to the type of HA product and the physician's fees. Although 50% of knee osteoarthritis patients eventually receive surgical procedures, almost one third of these surgeries are unnecessary. Hence intra-articular injections of hyaluronic acid can serve as a non-invasive cost-effective alternative to surgery for knee osteoarthritis.

Unlike surgical options, HA injections do not require incisions or extensive recovery periods. HA is a substance that naturally occurs in the synovial fluid of the joints, which helps lubricate and cushion them. In osteoarthritis, this fluid becomes less effective, leading to pain and reduced mobility. Thus, HA

injected directly into the knee joint helps restore the lubricating properties of the synovial fluid and reduce inflammation. By restoring lubrication, HA injections can help improve joint mobility and reduce stiffness. The procedure is relatively low risk, with mild potential side effects, such as temporary swelling or discomfort.

However, the benefits of HA injections are not permanent; they typically last for several months. Repeated injections may be needed for ongoing relief. More importantly, controversies exist regarding its safety and efficacy, the number of injections and courses, type of preparation, duration of its effects, and combining it with other drugs or molecules. Other factors include patient characteristics such as age, weight, gender, and severity of the OA. The study uses a prospective, double-blinded clinical trial. A multi-stage, multi-group DAMIP-based machine learning model is utilized to uncover discriminatory features that can predict the response status of knee OA patients to different types of HA treatment. The algorithm can identify certain subgroups of knee OA patients who respond well (or those who don't) to HA therapy. The study's baseline result, including factors such as patients' weight, smoking status and smoking frequency, gives physicians insight for patient treatment recommendations by identifying those most suitable for HA injection.

To the best of our knowledge, this work presents the first machine learning approach that predicts patient responses to HA injections for knee osteoarthritis. Another uniqueness of this study is that this is the first prospective clinical trial designed such that in addition to clinical data, patient self-reporting data is also carefully collected. The latter is challenging since patients often refuse or bypass questionnaires or miss filling in forms. Self-reported answers may be exaggerated; respondents may be too embarrassed to reveal private details; various biases may affect the results, like social desirability bias. However, knee pains, whether patients can move or do certain activities are standard questions used by physicians and are rather routine evaluation for active-duty personnel and athletes, and hence their self-reporting are rather reliable. Further, there has been no study indicating that patients would exaggerate their pain to receive treatment to their knee pain.

Traditional data collection methods, primarily focusing on clinical settings, limits our understanding of drug efficacy and *patient* wellbeing. Patient self-reporting data is crucial for machine learning in healthcare because it provides a unique, subjective perspective on a patient's health experience, including their symptoms, quality of life, and perception of treatment effectiveness, which can be vital for accurate diagnosis, treatment planning, and overall patient care, often not captured by solely objective medical data like lab results or imaging scans. There is growing interest and support for the utility and importance of patient-reported outcome measures (PROMs) (Kingsley & Patel, 2017; Verma, et al., 2021). This is one of the strengths of our study since it includes a broad spectrum of patient wellbeing data.

DAMIP classifier was chosen partly due to earlier DAMIP models have produced good predictive accuracy on blind data for numerous clinical studies where the training patient size is relatively small (e.g., in early cancer detection to uncover genomic signatures that predict CpG islands methylation (Feltus, et al., 2003), vaccine immunogenicity prediction that accelerates vaccine design and target delivery (Lee, Nakaya, et al., 2016a; Nakaya, et al, 2011, 2015; Querec, et al., 2009;) in which DAMIP results were instrumental in the eventual world-wide clinical trial of the Malaria vaccines (Kazmin, et al., 2017; Lee, Nakaya et al., 2016a)). DAMIP has also been used for studies involving very large number patient sets with equally consistent predictive accuracy (Lee, Wang, et al., 2016b). Multi-stage is performed herein to manage the highly inseparable data.

With the established predictive rule, prior to treatment physicians can predict a patient's response to HA injections based on patient and treatment characteristics. Physicians can then make empirically informed decisions about whether to treat a particular patient with HA and perhaps which type of HA preparation is most likely to produce the best treatment response for that individual patient.

Predicting treatment response based on clinically measured variables and patient-centered well-being data will empower physicians with an evidence-based decision-making tool to administer the most cost-effective intervention for the patients.

The study's follow-up period is focused on six months after the final injection. Since knee osteoarthritis is incurable, treatment for patients includes rehabilitation, activity modification, weight loss when indicated, shoe orthoses, local modalities, and medication. For more severe cases, either HA injections or knee surgery is selected. And HA injections are typically given as a series of 3-5 injections, spaced one week apart, with *repeat courses* usually needed every six months, depending on the individual's pain relief duration and the severity of their arthritis; most people experience pain relief for several months after a full course of injections.

The data and model derived from this study allows physicians to administer HA products more selectively and effectively, which will increase the percentage of patients who experience a successful HA therapy. Information about predicted responses could easily be shared with patients to incorporate their values and preferences into treatment selection. Specifically, the classification rule can be implemented within the electronic health record system as an Application Programming Interface (API). In addition, this decision support tool would allow physicians to quickly determine whether a patient is exhibiting at least an expected treatment response and if not, to potentially take corrective action. Of note, this model can also be used to predict patient responses to other forms of treatment and conditions.

There is a clear demand for evidence-based medical decision-making in addition to expert opinion, clinical experience and case reports. Additionally, there is an increased demand for clinical studies of prospective, rather than retrospective, treatment assessment options. While each of these study types has a role, the value of evidence-based, single studies or meta-analyses of published reports is that clinical criterion or criteria are analyzed globally with respect to outcome. Quantified variables that are uncovered by predictive models are evaluated and analyzed and can serve as important decision variables to help physicians select the best course of treatment for patients. Evidence-based decision-making increases outcome success. Trends, impressions and opinions are minimized and objective, evidence-based, outcome-driven targeted delivery is maximized.

ACKNOWLEDGEMENTS

A portion of the results from this project (the machine learning advances) received the first runner-up prize at the Caterpillar and INFORMS Innovative Applications in Analytics Award. This work is partially supported by grants from the National Science Foundation (IIP-1361532), and the American Orthopedic Society for Sports Medicine. Findings and conclusions in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation and or the American Orthopedic Society for Sports Medicine. The authors extend their deepest respect and gratitude to the late Dr. Barton J. Mann PhD, with whom we collaborated on the knee osteoarthritis research, and to Dr. Captain Marlene DeMaio (retired) from Medical Corps, United States Navy, for her clinical guidance and collaboration on the project. We thank the anonymous reviewers for their insightful comments.

REFERENCES

Ayis S, Dieppe P. (2009). The natural history of disability and its determinants in adults with lower limb musculoskeletal pain. J Rheumatol. 36:583–91. doi: 10.3899/jrheum.080455

Barbour, K. E., Helmick, C. G., Boring, M., & Brady, T. J. (2017). Vital Signs: Prevalence of Doctor-Diagnosed Arthritis and Arthritis-Attributable Activity Limitation — United States, 2013–2015. MMWR. Morbidity and Mortality Weekly Report, 66(9).

Bellamy, N. (2002). WOMAC osteoarthritis index user guide. Version V. Brisbane, Australia.

Bellamy, N., Wilson, C., Hendrikz, J., Whitehouse, S. L., Patel, B., Dennison, S., & Davis, T. (2011). Osteoarthritis Index delivered by mobile phone (m-WOMAC) is valid, reliable, and responsive. Journal of Clinical Epidemiology, 64(2).

Brooks, R., & De Charro, F. (1996). EuroQol: The current state of play. Health Policy, 37(1). https://doi.org/10.1016/0168-8510(96)00822-6

Burgers, P. T. P. W., Poolman, R. W., Van Bakel, T. M. J., Tuinebreijer, W. E., Zielinski, S. M., Bhandari, M., Patka, P., & Van Lieshout, E. M. M. (2015). Reliability, validity, and responsiveness of the Western Ontario and McMaster Universities osteoarthritis index for elderly patients with a femoral neck fracture. Journal of Bone and Joint Surgery - American Volume, 97(9).

Chan, L. L. Y., Wong, A. Y. L., & Wang, M. H. (2020). Associations between sport participation and knee symptoms: A cross-sectional study involving 3053 undergraduate students. BMC Sports Science, Medicine and Rehabilitation, 12(1).

Chavda, S., Rabbani, S. A., & Wadhwa, T. (2022). Role and Effectiveness of Intra-articular Injection of Hyaluronic Acid in the Treatment of Knee Osteoarthritis: A Systematic Review. Cureus. https://doi.org/10.7759/cureus.24503

Cui, A., Li, H., Wang, D., Zhong, J., Chen, Y., & Lu, H. (2020). Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies. EClinicalMedicine, 29–30.

Driban, J. B., Hootman, J. M., Sitler, M. R., Harris, K. P., & Cattano, N. M. (2017). Is participation in certain sports associated with knee osteoarthritis? A systematic review. In Journal of Athletic Training (Vol. 52, Issue 6). https://doi.org/10.4085/1062-6050-50.2.08

Elgaddal N, Kramarow EA, Weeks JD, Reuben C. Arthritis in adults age 18 and older: United States, 2022. NCHS Data Brief, no 497. Hyattsville, MD: National Center for Health Statistics. 2024.

Fallon, E. A., Boring, M. A., Foster, A. L., Stowe, E. W., Lites, T. D., Odom, E. L., & Seth, P. (2023). Prevalence of Diagnosed Arthritis — United States, 2019–2021.

MMWR. Morbidity and Mortality Weekly Report, 72(41). https://doi.org/10.15585/mmwr.mm7241a1

Feltus, F. A., Lee, E. K., Costello, J. F., Plass, C., & Vertino, P. M. (2003). Predicting aberrant CpG island methylation. Proceedings of the National Academy of Sciences of the United States of America, 100(21).

GBD 2019: Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019.

Hays, R. D., Sherbourne, C. D., & Mazel, R. M. (1993). The rand 36-item health survey 1.0. Health Economics, 2(3). https://doi.org/10.1002/hec.4730020305

Hochberg, M. C., Altman, R. D., Brandt, K. D., & Moskowitz, R. W. (1997). Design and conduct of clinical trials in osteoarthritis: Preliminary recommendations from a Task Force of the Osteoarthritis Research Society. Journal of Rheumatology, 24(4).

Kazmin, D., Nakaya, H. I., Lee, E. K., Johnson, M. J., Van Der Most, R., Van Den Berg, R. A., Ballou, W. R., Jongert, E., Wille-Reece, U., Ockenhouse, C., Aderem, A., Zak, D. E., Sadoff, J., Hendriks, J., Wrammert, J., Ahmed, R., & Pulendran, B. (2017). Systems analysis of protective immune responses to RTS,S malaria vaccination in humans. Proceedings of the National Academy of Sciences of the United States of America, 114(9). https://doi.org/10.1073/pnas.1621489114

Kingsley, C., & Patel, S. (2017). Patient-reported outcome measures and patient-reported experience measures. BJA Education, 17(4).

Lee, E. K. (2017). Innovation in big data analytics: Applications of mathematical programming in medicine and healthcare. Proceedings - 2017 IEEE International Conference on Big Data, Big Data 2017, 2018-January.

Lee, E. K., & Egan, B. (2022). A Multi-stage Multi-group Classification Model: Applications to Knowledge Discovery for Evidence-based Patient-centered Care. International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K - Proceedings, 1.

Lee, E. K., Li, Z., Wang, Y., Hagen, M. S., Davis, R., & Egan, B. M. (2021). Multi-Site Best Practice Discovery: From Free Text to Standardized Concepts to Clinical Decisions. 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2766–2773. https://doi.org/10.1109/BIBM52615.2021.9669414

Lee, E. K., Nakaya, H. I., Yuan, F., Querec, T. D., Burel, G., Pietz, F. H., Benecke, B. A., & Pulendran, B. (2016a). Machine learning for predicting vaccine immunogenicity. Interfaces, 46(5).

Lee, E. K., Wang, Y., Hagen, M. S., Wei, X., Davis, R. A., & Egan, B. M. (2016b). Machine learning: Multi-site evidence-based best practice discovery. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10122 LNCS.

Lee, E. K., Wang, Y., He, Y., & Egan, B. M. (2019). An efficient, robust, and customizable information extraction and pre-processing pipeline for electronic health records. IC3K 2019 - Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, 1. https://doi.org/10.5220/0008071303100321

Lee, E. K., Yuan, F., Man, B. J., & Egan, B. (2023a). A General-Purpose Multi-stage Multi-group Machine Learning Framework for Knowledge Discovery and Decision Support. Communications in Computer and Information Science, 1842 CCIS. https://doi.org/10.1007/978-3-031-43471-6_4

Lee, E. K., Yuan, F., Mann, B. J., & Egan, B. (2023b). Handling Imbalanced and Poorly Separated Data: a Multi-Stage Multi-Group Machine Learning Approach. Proceedings - 2023 2023 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2023. https://doi.org/10.1109/BIBM58861.2023.10386028

Leifer, V. P., Katz, J. N., & Losina, E. (2022). The burden of OA-health services and economics. Osteoarthritis and Cartilage, 30(1). https://doi.org/10.1016/j.joca.2021.05.007

Lo, J., Chan, L., & Flynn, S. (2021). A Systematic Review of the Incidence, Prevalence, Costs, and Activity and Work Limitations of Amputation, Osteoarthritis, Rheumatoid Arthritis, Back Pain, Multiple Sclerosis, Spinal Cord Injury, Stroke, and Traumatic Brain Injury in the United States: A 2019 Update. In Archives of Physical Medicine and Rehabilitation (Vol. 102, Issue 1). https://doi.org/10.1016/j.apmr.2020.04.001

Lohmander, L. S., Englund, P. M., Dahl, L. L., & Roos, E. M. (2007). The long-term consequence of anterior cruciate ligament and meniscus injuries: Osteoarthritis. In American Journal of Sports Medicine (Vol. 35, Issue 10). https://doi.org/10.1177/0363546507307396

Long, H., Liu, Q., Yin, H., Wang, K., Diao, N., Zhang, Y., Lin, J., & Guo, A. (2022). Prevalence Trends of Site-Specific Osteoarthritis From 1990 to 2019: Findings From the Global Burden of Disease Study 2019. Arthritis and Rheumatology, 74(7). https://doi.org/10.1002/art.42089

Lorig, K., Chastain, R. L., Ung, E., Shoor, S., & Holman, H. R. (1989). Development and evaluation of a scale to measure perceived self-efficacy in people with arthritis. Arthritis & Rheumatism, 32(1). https://doi.org/10.1002/anr.1780320107

Mancuso, C. A., Sculco, T. P., Wickiewicz, T. L., Jones, E. C., Robbins, L., Warren, R. F., & Williams-Russo, P. (2001). Patients' expectations of knee surgery. Journal of Bone and Joint Surgery, 83(7). https://doi.org/10.2106/00004623-200107000-00005

Mann, B. J., Gosens, T., & Lyman, S. (2012). Quantifying clinically significant change: A brief review of methods and presentation of a hybrid approach. In American Journal of Sports Medicine (Vol. 40, Issue 10). https://doi.org/10.1177/0363546512457346

Marx, R. G., Stump, T. J., Jones, E. C., Wickiewicz, T. L., & Warren, R. F. (2001). Development and evaluation of an activity rating scale for disorders of the knee.

American Journal of Sports Medicine, 29(2). https://doi.org/10.1177/03635465010290021601

McAlindon, T. E., Wilson, P. W. F., Aliabadi, P., Weissman, B., & Felson, D. T. (1999). Level of physical activity and the risk of radiographic and symptomatic knee osteoarthritis in the elderly: The Framingham study. American Journal of Medicine, 106(2). https://doi.org/10.1016/S0002-9343(98)00413-6

Mora, J. C., Przkora, R., & Cruz-Almeida, Y. (2018). Knee osteoarthritis: Pathophysiology and current treatment modalities. In Journal of Pain Research (Vol. 11). https://doi.org/10.2147/JPR.S154002

Nakaya, H. I., Hagan, T., Duraisingham, S. S., Lee, E. K., Kwissa, M., Rouphael, N., Frasca, D., Gersten, M., Mehta, A. K., Gaujoux, R., Li, G. M., Gupta, S., Ahmed, R., Mulligan, M. J., Shen-Orr, S., Blomberg, B. B., Subramaniam, S., & Pulendran, B. (2015). Systems Analysis of Immunity to Influenza Vaccination across Multiple Years and in Diverse Populations Reveals Shared Molecular Signatures. Immunity, 43(6). https://doi.org/10.1016/j.immuni.2015.11.012

Nakaya, H. I., Wrammert, J., Lee, E. K., Racioppi, L., Marie-Kunze, S., Haining, W. N., Means, A. R., Kasturi, S. P., Khan, N., Li, G. M., McCausland, M., Kanchan, V., Kokko, K. E., Li, S., Elbein, R., Mehta, A. K., Aderem, A., Subbarao, K., Ahmed, R., & Pulendran, B. (2011). Systems biology of vaccination for seasonal influenza in humans. Nature Immunology, 12(8). https://doi.org/10.1038/ni.2067

Querec, T. D., Akondy, R. S., Lee, E. K., Cao, W., Nakaya, H. I., Teuwen, D., Pirani, A., Gernert, K., Deng, J., Marzolf, B., Kennedy, K., Wu, H., Bennouna, S., Oluoch, H., Miller, J., Vencio, R. Z., Mulligan, M., Aderem, A., Ahmed, R., & Pulendran, B. (2009). Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. Nature Immunology, 10(1). https://doi.org/10.1038/ni.1688

Riddle, D. L., Jiranek, W. A., & Hayes, C. W. (2014). Use of a validated algorithm to judge the appropriateness of total knee arthroplasty in the United States: A multicenter longitudinal cohort study. Arthritis and Rheumatology, 66(8). https://doi.org/10.1002/art.38685

Riddle, D. L., & Perera, R. A. (2020). The WOMAC pain scale and crosstalk from co-occurring pain sites in people with knee pain: A causal modeling study. Physical Therapy, 100(10). https://doi.org/10.1093/ptj/pzaa098

Salis, Z., & Sainsbury, A. (2024). Association of long-term use of non-steroidal anti-inflammatory drugs with knee osteoarthritis: a prospective multi-cohort study over 4-to-5 years. Scientific Reports, 14(1). https://doi.org/10.1038/s41598-024-56665-3

Sangha, O., Stucki, G., Liang, M. H., Fossel, A. H., & Katz, J. N. (2003). The Self-Administered Comorbidity Questionnaire: A new method to assess comorbidity for clinical and health services research. Arthritis Care and Research, 49(2). https://doi.org/10.1002/art.10993

Saxon, L., Finch, C., & Bass, S. (1999). Sports participation, sports injuries and osteoarthritis implications for prevention. In Sports Medicine (Vol. 28, Issue 2). https://doi.org/10.2165/00007256-199928020-00005

Sharma, L., Song, J., Felson, D. T., Cahue, S., Shamiyeh, E., & Dunlop, D. D. (2001). The role of knee alignment in disease progression and functional decline in knee osteoarthritis. JAMA, 286(2). https://doi.org/10.1001/jama.286.2.188

Spector, T. D., Harris, P. A., Hart, D. J., Cicuttini, F. M., Nandra, D., Etherington, J., Wolman, R. L., & Doyle, D. V. (1996). Risk of osteoarthritis associated with long-term weight-bearing sports: A radiologic survey of the hips and knees in female ex-athletes and population controls. Arthritis and Rheumatism, 39(6). https://doi.org/10.1002/art.1780390616

Theis, K. A., Steinweg, A., Helmick, C. G., Courtney-Long, E., Bolen, J. A., & Lee, R. (2019). Which one? What kind? How many? Types, causes, and prevalence of disability among U.S. adults. Disability and Health Journal, 12(3). https://doi.org/10.1016/j.dhjo.2019.03.001

Turner, A. P., Barlow, J. H., & Heathcote-Elliot, C. (2000). Long term health impact of playing professional football in the United Kingdom. British Journal of Sports Medicine, 34(5). https://doi.org/10.1136/bjsm.34.5.332

United States Bone and Joint Initiative: The Burden of Musculoskeletal Diseases in the United States (BMUS) Fourth Edition. Forthcoming; 4th:http://www.boneandjointburden.org. Accessed December 20, 2023.

United States Bone and Joint Initiative. The Burden of Musculoskeletal Diseases in the United States (BMUS). In: In. Fourth ed. Rosemont, IL. 2018: Available at https://www.boneandjointburden.org/fourth-edition. Accessed June 12, 2023

Verma, D., Bach, K., & Mork, P. J. (2021). Application of machine learning methods on patient reported outcome measurements for predicting outcomes: A literature review. In Informatics (Vol. 8, Issue 3). https://doi.org/10.3390/informatics8030056

World Health Organization. Osteoarthritis. 14 July 2023. https://www.who.int/news-room/fact-sheets/detail/osteoarthritis

Zhang, W., Nuki, G., Moskowitz, R. W., Abramson, S., Altman, R. D., Arden, N. K., Bierma-Zeinstra, S., Brandt, K. D., Croft, P., Doherty, M., Dougados, M., Hochberg, M., Hunter, D. J., Kwoh, K., Lohmander, L. S., & Tugwell, P. (2010). OARSI recommendations for the management of hip and knee osteoarthritis. Part III: Changes in evidence following systematic cumulative update of research published through January 2009. Osteoarthritis and Cartilage, 18(4).