2017-03-27

Impact of Intra-Articular Injection Use on Patient-Reported Outcomes Among Patients with Knee Osteoarthritis

Shao-Hsien Liu

University of Massachusetts Medical School

IMPACT OF INTRA-ARTICULAR INJECTION USE ON PATIENT-REPORTED OUTCOMES AMONG PATIENTS WITH KNEE OSTEOARTHRITIS

A Dissertation Presented

By

SHAO-HSIEN LIU

Submitted to the Faculty of the
University of Massachusetts Graduate School of Biomedical Sciences, Worcester
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

March 27, 2017

CLINICAL AND POPULATION HEALTH RESEARCH

IMPACT OF INTRA-ARTICULAR INJECTION USE ON PATIENT-REPORTED OUTCOMES AMONG PATIENTS WITH KNEE OSTEOARTHRITIS

A Dissertation Presented
By

SHAO-HSIEN LIU

The signatures of the Dissertation Defense Committee signify completion and approval as to style and content of the Dissertation

_____

Kate L. Lapane, Ph.D., Thesis Advisor

_____

Leslie Harrold, M.D., MPH, Member of Committee

_____

Sharina D. Person, Ph.D., Member of Committee

_____

Catherine Dube, Ed.D., Member of Committee

_____

Tuhina Neogi, M.D., Ph.D., FRCPC, Member of Committee

The signature of the Chair of the Committee signifies that the written dissertation meets the requirements of the Dissertation Committee

_____

Jennifer Tjia, MD, MSCE, Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies that the student has met all graduation requirements of the school.

_____

Anthony Carruthers, Ph.D.
Dean of the Graduate School of Biomedical Sciences

Clinical and Population Health Research

March 27, 2017

DEDICATION

To my family, especially my parents and wife (Dr. Qiuzhi Chang), for their unconditional support, encouragement, and love over the years.

ACKNOWLEDGEMENTS

This dissertation could not have been accomplished without the support and guidance from a great number of individuals.

I first would like to thank Dr. Kate L. Lapane, for her invaluable guidance, wisdom, and moral support over the years as my mentor in the completion of my dissertation. Her support and guidance has made this a thoughtful and rewarding journey.

I am also grateful to my dissertation committee of Jennifer Tjia, MD, MSCE; Leslie Harrold, M.D., MPH; Sharina Person, PhD; Catherine Dube, Ed.D.; and Tuhina Neogi, M.D., Ph.D., FRCPC for their comments, patience, and support as I moved from an idea to a completed study through my research.

I am also indebted to the participants of the OAI study and staff for their tremendous contribution to help advance knowledge regarding treatment use in patients with osteoarthritis. This dissertation would not have been possible without them. In addition, I would like to thank Dr. Charles Eaton, Dr. Jeffrey Driban, and Dr. Timothy McAlindon for their sage advice, feedback, and time commitments toward studies using data from OAI as well as this dissertation.

I have also benefitted enormously from the support of the faculty and staff of the Clinical and Population Health Research Program/Department of Quantitative Health Sciences, especially: Dr. William Jesdale, Dr. Christine Ulbricht, Dr. Stavroula Chrysanthopoulou, Dr. Stephenie Lemon, Kelley Baron, and Sandy Stankus. I would also like to thank my fellow CPHR students and members from PHARE study group for their feedback, collaboration and of course friendship.

Finally, I must express my very profound gratitude to my family: my parents, my wife (Dr. Qiuzhi Chang), and my brother (Dr. Shao-Ting Liu) for continuous encouragement throughout the study and supporting my life in general. I am also beyond grateful to my friends, especially: Michael Kung, Nancy Kung, Willy Yen, and Ralph Chiang. Thank you so much for your “spiritual” and “bottomless” support (i.e., beer, wine, and scotch) over the years.

ABSTRACT

Background: Knee osteoarthritis (OA) is the most common type of OA and is a major cause of pain and thus results in disability for daily activities among persons living in the community. OA currently has no cure. In addition to the conflicting recommendations from clinical guidelines, evidence about the extent to which long-term use of intra-articular injections improves patient outcomes is also lacking.

Methods: Using data from the Osteoarthritis Initiative (OAI), marginal structural models (MSMs) applying inverse probability treatment weights (IPTW) were used to examine the effectiveness of intra-articular injections and changes in symptoms over time. The specific aims of this dissertation were to: 1) evaluate longitudinal use of intra-articular injections after treatment initiation among persons with radiographic knee OA; 2) quantify the extent to which intra-articular injection relieves symptoms among persons with radiographic knee OA; and 3) evaluate the performance of missing data techniques under the setting of MSMs.

Results: Of those initiating injections, ~19% switched, ~21% continued injection type, and ~60% did not report any additional injections. For participants initiating corticosteroid (CO) injections, greater symptoms post-initial injection rather than changes in symptoms over time were associated with continued use compared to one-time use. Among participants with radiographic evidence of knee OA, initiating treatments with either CO or hyaluronic acid (HA) injections was not associated with reduced symptoms compared to non-users over two years. Compared to inverse probability weighting (IPW), missing data techniques such as multiple imputation (MI) produced less biased marginal

causal effects (IPW: -2.33% to 15.74%; -1.88% to 4.24%). For most scenarios, estimates using MI had smaller mean square error (range: 0.013 to 0.024) than IPW (range: 0.027 to 0.22).

Conclusions: Among participants with radiographic evidence of knee OA living in the community, the proportion of those switching injection use and one-time users was substantial after treatment initiation. In addition, initiating injection use was not associated with reduced symptoms over time. With respect to issues of missing data, using MI may confer an advantage over IPW in MSMs applications. The results of this work highlight the importance of using comparative effectiveness research with non-experimental data to study these commonly used injections and may help to understand the usefulness of these treatments for patients with knee OA.

TABLE OF CONTENTS

TITLE PAGE……..... i

SIGNATURE PAGE...……....…… ii

DEDICATION ……...……..…….. iii

ACKNOWLEDGEMENTS……..…….. iv

ABSTRACT……..... vi

TABLE OF CONTENTS…….……... viii

LIST OF TABLES…….……. ix

LIST OF FIGURES…… xi

LIST OF ABBREVIATIONS……. xii

PREFACE……... xiv

CHAPTER I: INTRODUCTION…… 1

CHAPTER II: PATTERNS OF INTRA-ARTICULAR INJECTION USE AFTER INITIATION OF TREATMENT IN PATIENTS WITH KNEE OSTEOARTHRITIS: DATA FROM THE OSTEOARTHRITIS INITIATIVE……... 17

CHAPTER III: EFFECT OF INTRA-ARTICULAR INJECTIONS ON PATIENT-REPORTED SYMPTOMS IN PERSONS WITH OSTEOARTHRITIS: ANALYSIS WITH MARGINAL STRUCTURAL MODELS...……. 42

CHAPTER IV: MISSING DATA IN MARGINAL STRUCTURAL MODELS: A PLASMODE SIMULATION STUDY COMPARING MULTIPLE IMPUTATION AND INVERSE PROBABILITY WEIGHTING……..…… 76

CHAPTER V: DISCUSSION AND CONCLUSIONS…….. 100

REFERENCES…… 110

LIST OF TABLES

Table 2.1: Clinical characteristics by patterns of hyaluronic acid injection use (N=96).……..…….... 36

Table 2.2: Clinical characteristics by patterns of corticosteroid injection use (N=316)…….... 38

Table 2.3: Association between symptoms post-initial injection and first switching/continuation of corticosteroid injections among participants with radiographic confirmed knee osteoarthritis……...…….... 40

Table 2.4: Association between average change in symptoms and first switching/continuation of corticosteroid injections among participants with radiographic confirmed knee osteoarthritis……. 41

Table 3.1: Sociodemographic and clinical characteristics among participants with radiographically confirmed knee OA by use of injections.……. 63

Table 3.2: Concomitant use of medications and supplements among persons with radiographically confirmed knee OA by the use of injections….……... 67

Table 3.3: Estimated effects of injection use compared with non-users on symptoms and disease progression among persons with radiographically confirmed knee OA……...…….. 69

Supplementary Table 3.1: Model specifications for the construction of inverse probability weights at the index visit….……. 71

Supplementary Table 3.2: Sociodemographic and clinical factors among people with radiographically confirmed knee OA by censoring mechanisms……... 73

Table 4.1: Values of parameters for data generation using plasmode simulation framework and Osteoarthritis Initiative data……..……. 94

Table 4.2: Comparison of percent bias, mean square error, and empirical power for methods to handle missing data (CC, MI, IPW) under various mechanisms for missing data and extent of missing data: the case of a missing data in a time-independent confounder……...……... 96

Table 4.3: Comparison of percent bias, mean square error, and empirical power for methods to handle missing data (CC, MI, IPW) under various mechanisms for missing data and extent of missing data: the case of missing data in a time-varying confounder.……...….. 97

Table 4.4: Comparison of percent bias, mean square error, and empirical power for methods to handle missing data (CC, MI, IPW) under various mechanisms for missing data and extent of missing data: the case of missing data in either a time-independent confounder, time-varying confounder, or both.....……..... 98

Supplementary Table 4.1: A brief summary comparing approaches of IPW and MI…… 99

LIST OF FIGURES

Figure 1.1: Directed acyclic graph (causal diagram) between treatments use, study outcomes, censoring, and potential time-varying confounders……..…… 15

Figure 1.2: Causal diagrams depicting relationship in simulated datasets…….…. 16

Figure 2.1: Flowchart of study participants.....….. 34

Figure 2.2: Scheme of the study design to define changes between visits among injection users……. 35

Figure 3.1: Flowchart of study participants……..…….. 61

Figure 3.2: Directed acyclic graph (causal diagram) between the initiation of injection use, study outcomes, censoring, and potential time-varying confounder… 62

Supplementary Figure 3.1: Graphical representation of evaluation the experimental treatment assumption…….……. 75

Figure 4.1: Causal diagrams depicting relationship in simulated datasets.……. 93

LIST OF ABBREVIATIONS

AIC – Akaike information criterion

BMI – Body mass index

CC – Complete case analysis

CES-D – Centers for Epidemiologic Studies

CO – Corticosteroid

GEE – Generalized estimating equations

HA – Hyaluronic acid

IPCW – Inverse probability of censoring weights

IPTW – Inverse probability of treatment weights

IPW – Inverse probability weighting

JSW – Joint space width

K-L – Kellgren – Lawrence

KOOS-QoL – Knee injury and Osteoarthritis Outcome Score Quality of Life Subscale

MAR – Missing at random

MCAR – Missing complete at random

MCS – SF-12 Mental Component Summary Scores

MI – -Multiple imputation

MNAR – Missing not complete at random

MRI – Magnetic resonance imaging

MSMs – Marginal structural models

NSAID – Non-steroidal anti-inflammatory drugs

OA – Osteoarthritis

OAI – Osteoarthritis Initiative

PACE – Physical Activity Scale for the Elderly

PCS – SF-12 Physical Component Summary Scores

RCTs – Randomized clinical trials

SF-12 – 12-item Short-Form Health Survey

WOMAC – Western Ontario and McMaster Universities Arthritis Index

PREFACE

Chapter II of this dissertation is under preparation as:

> Liu SH, Dube C, Driban JB, McAlindon TE, Eaton CB, Lapane KL. Patterns of intra-articular injection use after initiation of treatment in patients with knee osteoarthritis: data from the Osteoarthritis Initiative.

Chapter III of this dissertation is under preparation as:

> Liu SH, Dube C, Eaton CB, Driban JB, McAlindon TE, Lapane KL. Effect of intra-articular injections on patient reported symptoms in persons with osteoarthritis: analysis with marginal structural models.

Chapter IV of this dissertation is under preparation as:

> Liu SH, Chrysanthopoulou SA, Chang Q, Hunnicutt JN, Lapane KL. Missing data in a point-exposure study using marginal structural models: A plasmode simulation study to compare multiple imputation to inverse probability weighting.

CHAPTER I

INTRODUCTION

Knee Osteoarthritis among the U.S. population

Osteoarthritis (OA) is a chronically degenerative condition involving changes in muscular and cartilage tissues around the joint. OA may develop in any joints, but the most commonly affected joints are the knee, hip, hand, spine, and foot. In 2005, nearly 27 million of U.S. adults were affected by OA.[1] Knee OA is the most common type of OA with nearly 6.1% of U.S. adults who are aged 30 or older having this disease.[2–4] The prevalence and incidence of knee OA increases with age. Among participants in the Framingham Study, the prevalence of knee OA was 19.2% in those aged over 45 years and 43.7% in those over 80 years.[5] In adults aged over 20 years, the incidence rate of knee OA has been estimated to be 240/100,000 person-years and approximately 2% per year developed incident radiographic disease among women.[6,7]

The risk factors of developing OA, including both systematic and local elements, have been identified. Systemic factors are those that may act by increasing the susceptibility of joints to injury or by impairing the process of repair such as age, gender and hormones, race/ethnicity, and genetics.[8,9] Local factors are those biomechanical elements that adversely affect the forces applied to the joints such as obesity, history of injury or surgery, occupation, physical activity and sports.[8,9] Clinical symptoms of OA include joint pain, stiffness, and limitations of movements. Symptomatic knee OA is defined by the presence of pain, aching, or stiffness in a joint with clinical radiographic evidence. Approximately 10% of women and 13.6% of men over 60 years of age have symptomatic knee OA.[10] Although disease progression of OA is usually slow, it can ultimately lead to joint failure with pain and disability. Disease in weight bearing joints

such as the knee has a greater clinical impact. Knee OA is 1 of the 5 leading causes of disability among non-institutionalized adults.[11]

Use of Intra-articular Injections and Outcomes in Knee Osteoarthritis

Although knee OA is a major cause of pain and disability among persons living in the community, OA currently has no cure, as the mechanism by which it arises and progresses remains incompletely understood. Goals of OA treatment using non-pharmacologic or pharmacologic modalities include pain relief, improved mobility, delayed disease progression, and improved quality of life. Patients with a clinical diagnosis of knee OA are often treated primarily with conservative treatment plans including a combination of non-pharmacologic and pharmacologic agents. If non-pharmacologic intervention such as exercise or weight management and/or the use of orally administered medications such as acetaminophen are ineffective, intra-articular injections may offer symptom relief. [12–14]

Currently, there are two primary types of intra-articular injections used in OA: corticosteroid (CO) injections and viscosupplementation (hyaluronic acid (HA) injections). For CO injections, there are 5 formulations that have been approved by the US Food and Drug Administration.[14] CO injections reduce inflammation which indirectly may relieve pain for up to 4 weeks. Typically, no more than three corticosteroid injections a year are recommended. Although there are several formulations available for corticosteroid injections, they are thought to be equally efficacious[15–18] with most clinical guidelines advising consideration of such injections for acute exacerbations or short-term

relief.[19–21] On the other hand, HA is a naturally occurring constituent of cartilage and synovial fluid in the joints. HA injections have a different mechanism of action including enhancing and maintaining inter-articular lubrication while potentially providing additional protection such as anti-inflammatory action, analgesic and chondroprotective effects.[22] Treatment cycles for these injections can consist of up to five weekly injections and may offer relief for up to six months thus providing longer term benefits compared to corticosteroids.[23]

Patient-focused evidence-based recommendations for the management of knee OA have been issued in several treatment guidelines.[12,24,25] For CO injections, the Osteoarthritis Research Society International 2014 treatment guidelines are consistent with the American College of Rheumatology's 2012 guidelines suggesting that intra-articular CO injections for patients in the absence of relevant comorbidities are appropriate because these treatments demonstrate clinically significant short-term decreases in pain.[12,24] Nevertheless, this treatment is not suggested by the 2013 American Academy of Orthopaedic Surgeons guidelines, which cites inconclusive evidence to recommend the treatment use.[25] Despite a large number of studies, the safety and efficacy of HA injections also remains inconclusive leading to a lack of consensus across clinical guidelines.[12,24,25]

In addition to the conflicting recommendations from clinical guidelines, evidence about the extent to which long-term use of intra-articular injections improves patient outcomes is lacking. Recent systematic reviews and meta-analyses suggest that the effect of US-approved viscosupplement injections can last through 26 weeks but there is no

similar evidence for corticosteroid injections for persons with knee OA.[17,26,27] Among patients with milder disease, receiving intra-articular sodium hyaluronate appears to slow joint space narrowing compared to placebo.[28] For corticosteroid injections, there is no difference between treatment and placebo groups in joint space changes over two years of follow-up.[29] Studies documenting the relationship between long-term use of intra-articular injections and changes in symptoms using patient-reported outcomes on knee OA symptoms among general populations are scarce. Furthermore, no previous studies have evaluated treatment patterns for intra-articular injection use among patients with knee OA living in the community.

Specific Aims

This dissertation examined the effectiveness of commonly used intra-articular injections in relieving symptoms among persons with knee OA over 2 years using marginal structural models (MSMs). The use of advanced statistical techniques such as MSMs using inverse probability of treatment weights (IPTW) allowed us to quantify the effect of injections use over time in a more heterogeneous population than those typically recruited in clinical trials. In addition, it can also allow for improved adjustment of confounding in the situations where a special type of confounding (e.g., confounding by indication) that can occur when studying the effects of drugs using observational (non-experimental) studies.[30] However, when applying such a technique to estimate the effect of injection use on treating knee OA, we were required to fully understand the

complexity of treatment use over time as well as issues such as missing data approaches under the settings of MSMs. The specific aims of this dissertation were as follows.

Aim 1. To evaluate longitudinal use of intra-articular injections after treatment initiation among persons with radiographic knee OA:

In this aim, patterns of injection use among patients with newly initiated injection were described. Whether more severe OA symptoms after treatment initiation and/or changes in symptoms over time were associated with patterns of injection use was examined.

Aim 2. To quantify the extent to which intra-articular injection relieves symptoms among persons with radiographic knee OA:

In this aim, the effect of injection use patterns over a two-year period on changes in patient-reported symptoms was estimated in a real-world setting.

Aim 3. To evaluate the performance of missing data techniques under the setting of MSMs:

In this aim, simulated datasets under the plasmode simulation framework were generated. The bias and precision of estimates obtained from three missing data approaches in the setting of MSMs including complete case analysis (CC), multiple imputation (MI), and inverse probability weighting (IPW) were compared.

Data Source and Study Population

This dissertation analyzed publicly available data from the Osteoarthritis Initiative (OAI) (http://oai.epi-ucsf.org/), a multi-center (i.e., Baltimore, MD; Columbus, OH;

Pittsburgh, PA; and Pawtucket, RI), longitudinal, prospective observational study examining not only the development and progression of knee OA but also the effectiveness of disease-modifying therapies. The study cohort included 4,796 men and women ages 45-79 enrolled between February 2004 and May 2006 and followed for 9 years. Participants who have either symptomatic OA in at least one knee or have at least one from a set of established risk factors for developing knee OA such as having pain, aching, or stiffness in or around the knee, radiographic evidence of tibiofemoral osteophytes, the use of medications to treat for symptoms were eligible for the study. Detailed inclusion and exclusion criteria for the OAI are described elsewhere.[31]

For this dissertation, 2,550 participants who had radiographic knee OA (Kellgren-Lawrence (K-L) grade ≥ 2) at baseline served as the basis of the eligible sample for analyses. From this group, participants who reported already receiving injections at baseline, did not have any follow-up assessments, or were missing > half of follow-up assessments were excluded. From the remaining 2,150, we further excluded those reporting use of both injections, not reporting any injection use over the 9 years of follow-up, reporting first initiation at year 9, and reporting injections after total knee replacement. The final analytic sample for Aim 1 consisted of 412 participants initiating injection use. In addition to the new users selected for Aim 1, a group of participants not reporting any injection use over the 9 years of follow-up (non-user) was used in Aim 2. To mimic the study design from clinical trials,[29,32] participants whose age < 45 years and did not have symptomatic knee OA at baseline were considered ineligible for injection use and thus further excluded for non-users. The final analytic sample included 412

participants initiating injection use and 576 non-users. Using the sample derived in Aim 2, we used participants initiating CO injections and non-users to construct the cohort. Only complete cases that provided all of the information were used for generating simulated datasets. This resulted in a cohort consisting of 646 participants in Aim 3.

Analytic Methods

A causal relationship between a treatment and its associated outcome becomes ambiguous in the presence of a confounder; the treatment effect is confounded when one or more risk factors for the outcome are also associated with the treatment. Observational studies, in which randomization cannot be performed like clinical trials, typically address confounding by applying statistical techniques such as stratification and multivariable regression analysis. For a point treatment study in which the treatment is administered once, multivariable regression models may be sufficient to control for confounding.

However, in longitudinal studies with repeated treatments over time, the estimates from regression models may still be biased if (1) there exists a time-dependent covariate that is a risk factor for, or predictor of, the event of interest and also predicts subsequent exposure, and (2) past exposure history predicts subsequent level of the covariate.[30,33] In addition, condition (1) and (2) will always hold in many in pharmacoepidemiologic studies, particularly those in which there is confounding by indication and there are time-dependent covariates that are simultaneously confounders and intermediate variables. For instance, it is often difficult to assess medical indications and underlying disease severity

and prognosis over time and thus confounding by indication may arise when a drug treatment appears to cause outcomes they are meant to prevent.

Traditionally, multivariable regression is used to account for differences in measured covariates between subjects. However, this method may not fully adjust for confounding by indication occurring if the health status of patients affects treatment allocation. Alternative analytic approaches such as propensity scores have been proposed that may provide more precise estimates of the treatment effect in observational studies in which confounding by indication may occur.[34] Despite that propensity score-based analytic methods have been widely used in pharmacoepidemiologic studies, conditioning or stratification on time-dependent propensity score could induce substantial collider-stratification or confounding bias when treatment and confounders vary during follow-up.[35]

Another approach is instrumental variable analysis when substantial uncontrolled confounding is likely due to confounding by indication or confounding by disease severity.[36] However, the primary barrier to the use of instrumental variable analysis is the need to have a plausible instrumental variable. Unfortunately, such variables have been difficult to find in epidemiology and medicine. Furthermore, inclusion of variables that are strongly related to the exposure, but unrelated to the outcome (i.e., instrumental variables), can increase the variance and bias of an estimated exposure effect when added to a statistical model. As such, any plausible instrumental variable could potentially introduce Z-bias in the presence of uncontrolled confounding.[37,38]

Marginal Structural Models

In this dissertation work, since we sought to examine the complex relationship between the use of intra-articular injections (exposure) and changes in symptoms (outcome) over time using observational data, using traditional analytic methods to assess time-varying covariates and predictors will yield biased estimates especially when the time-varying confounders lie on the causal pathway between prior and subsequent exposures, and also affect the outcome measures.[30] Marginal structural models (MSMs) are a class of causal models for the estimation, from observational data, of the causal effect of a time-dependent exposure in the presence of time-dependent covariates that may be simultaneously confounders and intermediate variables.[30] As such, through applying IPTW, MSMs can allow us to account for changing values of the confounders over time and thus minimize the effect of confounding by indication that is compounded by the fact that treatment decisions may be affected not only by difference in baseline disease severity but by the natural course of the disease and by treatment response.

MSMs has advantages of eliminating bias from two sources when estimating the effect from a time-varying treatment. First, through applying IPTW, it can control for the time-varying confounding while avoiding two types of bias that may arise in analyses with standard regression models.[39] The first type of bias occurs when the time-varying confounder is simultaneously a confounder and intermediate variable. Conditioning analysis on such a variable (as performed in standard regression models) will block the indirect effect from previous treatment on study outcome that is mediated by this variable. Another type of bias (called collider-stratification bias[33] or selection bias[30])

occurs in standard regression models when the time-varying confounder is a common effect (i.e., a collider) of previous treatment and an unmeasured risk factor for the study outcome. Conditioning analysis on this time-varying confounder induces a non-causal relationship between previous treatment and the unmeasured risk factor, which introduces bias in the effect estimate of previous treatment use.[30]

In addition the potential to minimize bias, using IPTW can provide several analytical advantages compared to propensity score methods. First, it requires fewer distributional assumptions about the underlying data. Second, it can avoid potential residual confounding arising from setting an arbitrary fixed number of strata. Last but not least, it uses entire sample and avoid losing people in the matching process. However, despite that weights created by IPTW can take on extreme values and thus affect the stability and precision of exposure effect estimates when the number of intervals increases, stabilization and truncation of weights can improve efficiency.[40]

Before modeling the effect of injection use on knee OA, we used a causal diagram (Directed Acyclic Graph) to help identify potential confounders as well as methods to control for confounding over time (Figure 1.1). In this graph, Y(t) denotes outcomes (e.g., symptoms of knee) at visit t (e.g., 0, 1, 2,…t years) and L(t) corresponds to the measured time-varying confounders at visit t. L(t-1) includes the potential confounders available at baseline and covariates measured before injection initiation.

As shown in Figure 1.1, previously measured study outcomes and time-varying confounders may be simultaneously confounders and intermediate variables. For instance, when studying knee OA symptoms as the outcomes (i.e., Y(t)), the severity of

symptoms measured at the previous visit (i.e., Y(t-1)) can be a potential confounder because 1) it correlates with symptoms measured at current visit, and 2) patients with more severe pain are more likely to use treatments (e.g., injection). Furthermore, if treatments are effective in relieving symptoms, previously measured symptoms (i.e., Y(t-1)) lie on the causal path from prior treatment use and currently measured symptoms. Under such assumed causal relationships, using standard regression models adjusting for previous severity of symptoms will generate a biased estimate of the overall treatment effect of injection use on the outcome.[41] In addition to the relationship between outcomes and potential confounders displayed in Figure 1.1, C(t+1) indicates censoring status during the follow-up periods which represents that analysis are only restricted to participants who have not been censored.

Since patients may change their treatment use and result in the loss to follow-up, we first evaluated whether factors such as OA symptoms after treatment initiation and/or changes in symptoms over time were associated with patterns of injection use among patients with knee OA in Aim 1. We first conducted descriptive statistics for continuous variables and percentages for categorical variables to describe socio-demographic and clinical characteristics according to patterns of injections use. Before starting the modeling process, we checked whether there were strong linear dependencies among the potential correlates. Multicollinearity was evaluated and ruled out before the model building process by evaluating the correlations between the covariates of interest. Multinomial logistic models were then built to evaluate the association between two operational expressions of exposure (symptoms post-initial injection; and change in

symptoms) and three symptoms (pain, stiffness, and physical function) for patterns of injection use.

In Aim 2, we sought to quantify the extent to which intra-articular injection relieves symptoms among persons with radiographic knee OA. To properly control for the bias by time-varying confounders that may be affected by previous treatment, we used MSMs through applying IPTW estimation.[42] IPTW reduces confounding through assigning a weight to each participant, which is proportional to the inverse of the conditional probability of receiving his/her observed treatment given those time-varying confounders.[42] In the resulting weighted pseudo-population, treated participants and untreated participants are balanced over those time-varying confounders.[42] Since the analysis is not conditioned on the confounders, IPTW can properly estimate overall treatment effect.

The analytical approach for MSMs was conducted in three steps. First, each individual was weighted by the inverse of the conditional probability of receiving the treatment that was actually received to construct IPTW. This created a pseudo-population in which the differences in the distribution of confounders between those receiving the injection of interest and those not receiving were minimized. Stabilized weights were used to increase precision.[30] The stabilized IPTW for each individual i at each visit t is: $SW_i^t(t) = \prod \frac{\Pr[INJECTIONS_{it}=injections_{it}|\overline{INJECTIONS}_{it}=\overline{injections}_{it}CON=con_{it}]}{\Pr[INJECTIONS_{it}=injections_{it}|\overline{INJECTIONS}_{i(t-1)}=\overline{injections}_{i(t-1)}CON=con_{it}]}$, where capital letters indicated a random variable, lowercase letters indicated the observed value for that random variable, and an over bar indicates the history of that variable until time

t.[30] The IPTW was created using logistic models to estimate the probability of treatment for each individual with the covariates.[30]

Second, the stabilized inverse probability of censoring weights (IPCW) was created in a similar fashion. Since bias can also arise from loss to follow-up when individuals discontinue study participation for reasons associated with predictors or confounders of interest, using stabilized IPCW can reduce bias due to censoring.[30] The stabilized IPCW for each individual i at each visit t is: $SW_i^t(t) = \prod \frac{\Pr[C_t=0|\bar{C}_{t-1}=0,\ \overline{INJECTIONS}_{t-1}CON,\ T>t]}{\Pr[C_t=0|\bar{C}_{t-1}=0,\ \overline{INJECTIONS}_{t-1}\overline{TVC}_{(t-1)},\ T>t]}$, where C is a dichotomous variable (yes/no) indicating whether or not the participant has been censored at time t. Combined weights were then derived by multiplying the stabilized IPTW and IPCW.

Lastly, weighted generalized estimating equation (GEE) linear models were fit. To determine the mean difference in changes in symptoms (Y), caused by the use of injections over time, weighted GEE models were used to fit linear models to assess the average causal effect of use of injections use on the difference in changes in symptoms. Robust variance estimators were obtained to address within-subject correlations induced by weighting. Model fit was assessed using Akaike's Information Criterion (AIC) as well as regression diagnostics to identify potentially influential observations.

Missing Data Approaches and Plasmode Simulation

Despite the advantages of MSMs using IPTW to estimate the unbiased causal effect when time-varying confounding is a concern,[30,41,43] missing data of the covariate is another issue in longitudinal settings. The objective of Aim 3 was to compare validity

and precision obtained from commonly used missing data approaches (i.e., MI and IPW) through simulation studies in the presence of either missing time-independent or time-varying confounders in cohort studies using MSMs. We simulated datasets under the framework of plasmode simulation using diagrams depicting causal relationship among the exposure (A), outcome (Y), covariates, and missing data mechanisms (Figure 1.2). While L_0 indicates a set of all pre-specified confounders at baseline, L_t represents time-dependent confounders measured at time t. C' indicates a set of potential confounders in addition to the pre-specified confounders that are selected in the data generating mechanism. Some arrows are omitted due to the simplicity for presentation. Missing data mechanisms including missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR) are represented by M.

Overall, the analytic approach for Aim 3 was carried out in three steps. First, we used the plasmode simulation frame work to create simulated data sets.[44] We began by constructing the cohort using the study sample in Aim 2 which included complete cases of participants initiating CO injections and non-users. We then estimated a linear regression model for the observed study outcome as a function of the exposure status, baseline covariates, and a subset of the potential confounders. This estimated model served as the basis of outcome-generating model. Next, to create a simulated data set, we sampled with replacement among exposed and unexposed participants from the constructed cohort to achieve the desired sample size. Since we preserved the information on covariates and exposure data for each participant without modification, associations among these variables remained intact in the sampled population. We then used the

outcome-generating model described above to generate outcome values through replacing a pre-identified treatment effect on the estimated coefficient exposure. The value of all other model coefficients remained unchanged and was used to generate values of outcome status for each patient in the simulated data. This process was repeated 1000 times to yield 1000 simulated data sets in each simulation scenario.

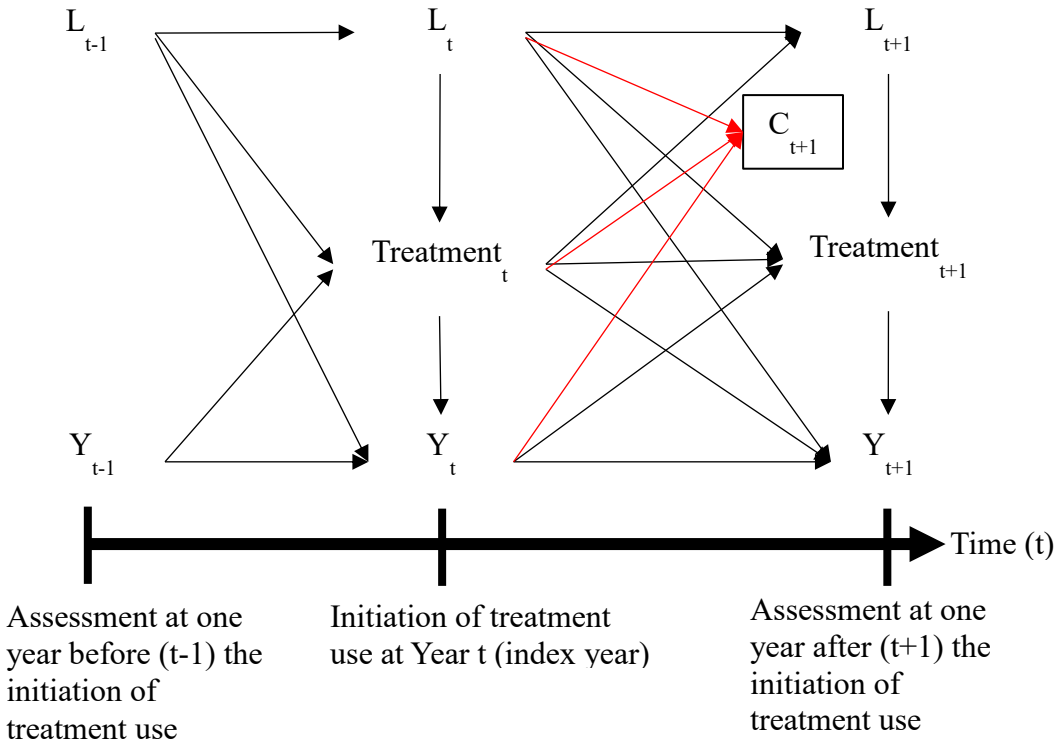
Second, we introduced missing data mechanisms and scenarios to the simulated data sets. The missing data mechanisms missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR) were used.[45,46] Missing data is considered as MCAR when the probability of missing does not depend on the values of observed covariates. If the probability of missing depends on values of observed covariates, it is considered as MAR. If it is informative, MNAR is considered when the probability of missing depends on values of unobserved covariates. In this study, we assumed that participants with more severe knee OA status (for the time-independent confounder) and higher knee pain (for the time-dependent confounder) were more likely to be missing using the empirical data.

In addition to complete case analysis, we then applied IPW and MI to take into account the missing data in the analysis. In the setting of MSMs, IPW is not a new approach for handling missing data and is particularly straightforward to use.[47,48] It is similar to the weight building process for the inverse probability of observed treatment or censoring weights that are performed under MSMs.[30] IPW proceeds by calculating the probability of having complete data for each individual in the study. Using logistic regression models, each individual is weighted by the inverse probability of having

complete data conditional on other relevant covariates. In MI, on the other hand, missing values in the incomplete observed data are actually imputed through generating *m* complete datasets. Each of the imputed datasets is then analyzed using the same outcome model or method of estimation. Estimates from each of the imputed datasets are then combined to produce a single estimate that incorporates the sampling variability as well as the variability of the missing data.[49]

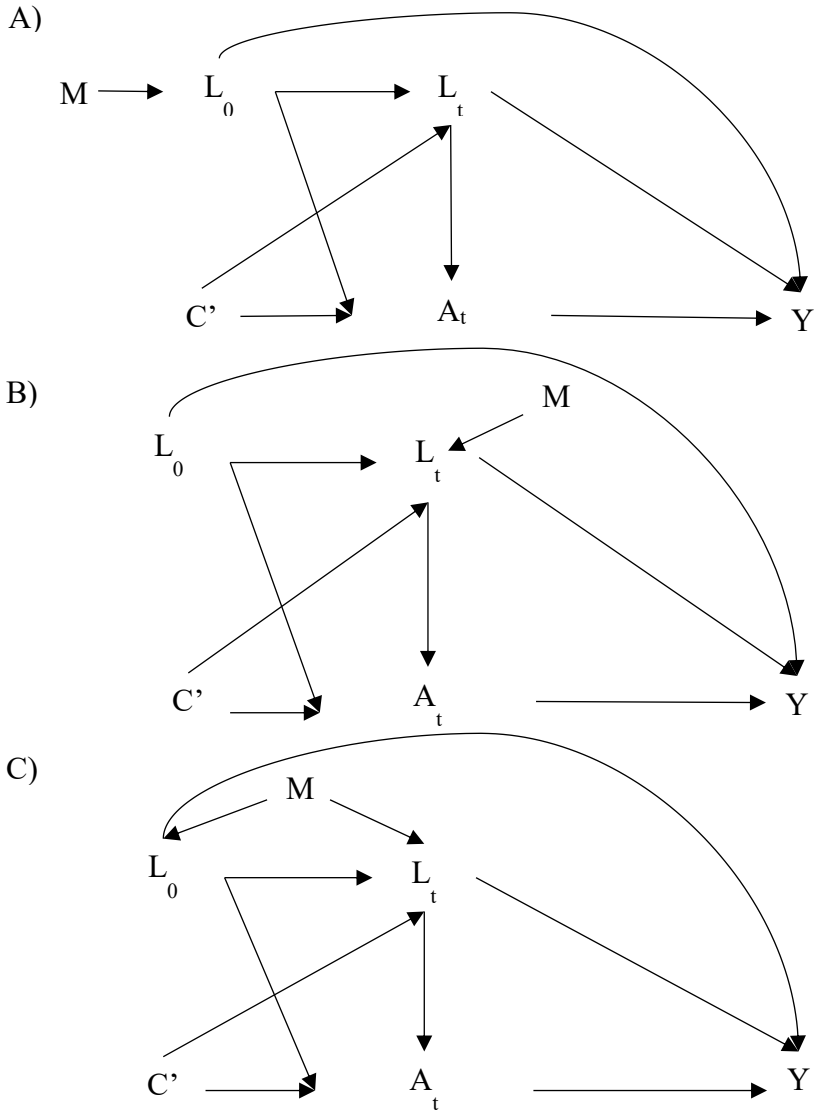
Lastly, weighted GEE linear models were fit using CC, IPW and MI to estimate the average causal effect in MSMs over increasingly problematic scenarios of missingness. Performance of methods was compared using relative bias, mean squared error, and empirical power of the estimates of interest.

Figure 1.1: Directed acyclic graph (causal diagram) between the initiation of treatment use, study outcomes, censoring, and potential time-varying confounders.



LEGEND: Hypothesized causal relationships between treatment use, study outcomes and potential time-varying confounders. Y(t) denotes outcomes at visit t (e.g., 0, 1, 2,…t years) and L(t) corresponds to the measured time-varying confounders at visit t. L(t-1) includes the potential confounders available at baseline and covariates measured before injection initiation Y(t)). C(t+1) indicates censoring status during the follow-up periods which represents that analysis are only restricted to participants who have not been censored.

Figure 1.2: Causal diagrams depicting relationship in simulated datasets.



LEGEND: Data generation and missing mechanisms for simulation studies. L_0 indicates a set of all pre-specified confounders measured at baseline, L_t represents time-varying confounders of the exposure (A) and outcome (Y) association measured at time t. C' indicates a set of potential confounders in addition to the pre-specified confounders that are selected in the data generating mechanism. M indicates missing mechanisms. A) This causal diagram shows that missingness (M) is present in the baseline confounder L_0; B) This causal diagram shows missingness (M) in the time-varying confounder L_t; and C) This causal diagram indicates that missingess (M) is present in either or both baseline or time-varying confounders.

CHAPTER II

PATTERNS OF INTRA-ARTICULAR INJECTION USE AFTER INITIATION OF TREATMENT IN PATIENTS WITH KNEE OSTEOARTHRITIS: DATA FROM THE OSTEOARTHRITIS INITIATIVE

Abstract

Objective: We sought to describe and evaluate longitudinal use of intra-articular injections after treatment initiation among adults with radiographically confirmed knee osteoarthritis (OA).

Methods: Using data from the Osteoarthritis Initiative, we included participants with radiographically confirmed OA (Kellgren-Lawrence grade (K-L) ≥ 2) in ≥ 1 knee at baseline. With 9 years of follow-up data, 412 participants newly initiating hyaluronic acid or corticosteroid injections at their index visit were identified. For each type of injection initiated, socio-demographic and clinical characteristics were described by patterns of treatments (one-time use, switched, or continued injections). Multinomial logistic models estimated the extent to which patient-reported symptoms (post-initial injection and changes over time) were associated with patterns of injection use.

Results: Of those initiating injections, ~19% switched, ~21% continued injection type, and ~60% did not report any additional injections. For participants initiating corticosteroid injections, greater symptoms post-initial injection were associated with lower odds of continued use compared to one-time users (adjusted odds ratio (aOR) for WOMAC pain: 0.91; 95%, confidence interval (CI): 0.83 to 0.99; $aOR_{stiffness}$: 0.77; CI: 0.63 to 0.94; $aOR_{physical\ function}$: 0.97; CI: 0.94 to 1.00). Symptom changes over time (e.g., worsened or improved) were not associated with patterns of injections use.

Conclusions: After treatment initiation, the proportion of patients switching injection use and one-time users was substantial. Symptoms post-initial injection rather than changes in symptoms over time appear to be associated with patterns of injection use. The extent

to which these patterns are an indication of lack of impact on patient-reported symptoms should be explored.

Introduction

Osteoarthritis (OA) is the most commonly seen arthritis among U.S. adults.[11,50,51] Among OA-affected joints, knee OA is one of the leading causes of disability among adults living in the community.[11] OA is a slowly progressive joint disease and currently has no cure. Generally, the treatment goals for non-pharmacologic or pharmacologic treatment of OA include pain relief, improved mobility, delayed disease progression, and improved quality of life. For those whose non-pharmacologic interventions or symptom-relieving medications are ineffective, intra-articular injections may be recommended to attempt to more directly target underlying pathophysiological processes.[12–14,52]

Although several guidelines for the treatment of knee OA have been issued, the recommendations for use of intra-articular injections such as corticosteroid or hyaluronic acid injections are inconsistent.[12,24,25] Costs contributing to the long-term use of injections due to the chronic symptoms of knee OA could be substantial.[53] Intra-articular injections are increasingly common, particularly among patients newly diagnosed with knee OA.[54] Given the concern regarding increase in use and the associated economic burden for patients with knee OA, evaluating how patients use these modalities over time is important. Examining patterns of injections use can help understand the switching and/or augmentation of treatment related to clinical outcomes.[55,56] However, no previous studies have evaluated treatment patterns for intra-articular injection use among patients with knee OA living in the community.

This study sought 1) to describe and evaluate longitudinal use of intra-articular injections after treatment initiation; and 2) to identify factors associated with patterns of

treatment use among adults with radiographically confirmed knee OA. With the data derived from yearly visits from the Osteoarthritis Initiative (OAI), we were able to identify newly initiated injection users and examine factors associated with treatment patterns. We hypothesized that the patterns of injection use among participants initiating injections would be associated with more severe OA symptoms and/or changes in symptoms over time compared to those received one-time injections.

Methods

Study sample

Publicly available data from the OAI were used.[31] The OAI study was originally a prospective cohort which enrolled participants from Baltimore, MD; Columbus, OH; Pittsburgh, PA; and Pawtucket, RI from 2004 through 2006. Using these four study sites, 4,796 patients with established knee OA or at high risk for developing knee OA were enrolled. For this present retrospective cohort study, we used information from annual assessments from baseline through year 9. Figure 2.1 shows the study sample for the current study. We first included patients with radiographically confirmed knee OA (defined as having a Kellgren-Lawrence grade (K-L) >2) at baseline (n=2,550). From this group, participants who reported already receiving injections (corticosteroid or hyaluronic) at baseline (n=97) were excluded. In addition, participants with no follow-up assessments or missing > half of these assessments were also excluded (n=303). From the remaining 2,150, we further excluded those reporting use of both injections (concurrent users, n=52), not reporting any injection use over the 9 years of follow-up (n=1,636),

reporting first initiation at year 9 (n=47), and reporting injections after total knee replacement (n=3). The final analytic sample consisted 412 participants initiating injection use.

Index Knee for Analysis

In OAI, symptoms and x-rays for both knees were collected separately. We selected an index knee for use in the analysis based on the presence of symptoms (knee pain) and radiographic evidence of OA. For participants with only one radiographically confirmed OA knee at baseline, that knee was used as the index knee. If both knees were radiographically confirmed with OA, then the knee with greater pain at baseline measured by Western Ontario and McMaster Universities Arthritis Index (WOMAC) pain scale was used. If both pain scores were equal, then the knee with worse K-L grade was used as the index knee.

Injections use, index visit, and patterns defined

Injection use was assessed separately for both knees in OAI. Participants were first asked "During the past 6 months, have you had any injections in either of your knees for treatment of arthritis?" For those who answered yes, participants were then asked questions "During the past 6 months, have you had an injection of hyaluronic acid (Synvisc or Hyalgan) in either of your knees for treatment of your arthritis? These injections are given as a series of 3 to 5 weekly injections." To assess corticosteroid injections, participants were asked: "During the past 6 months, have you had an injection

of steroids (cortisone, corticosteroids) in either of your knees for treatment of your arthritis?" The visit that participants reported their initial injection was used as the index visit.

After identifying the initial injection and index visit, switching injection users were then defined as reporting at least one injection other than their initial injection during follow-up. For example, for participants who initiated corticosteroid injection use, reporting a hyaluronic acid injection constituted a switch. Continued users were defined as reporting more than one injection of the same type during follow-up. Those who did not report either a switching or continuing use during follow-up were considered as one-time injection users. To determine the first switching/continued injection use among one-time users, we matched by the distribution of follow-up time intervals between injection initiations and first reported switching/continuation among those switched and continued users.

Symptoms

In the OAI, knee-related symptoms such as pain, stiffness, and physical function were examined annually using WOMAC, with a 5-point Likert scale. The range of scores was 0 to 20 for pain, 0 to 8 for stiffness, and 0 to 68 for physical function.[57] For each subscale, higher numbers indicated worse symptoms. We were interested in evaluating patient-reported symptoms in two ways: 1) symptoms post-initial injection; and 2) longitudinal change in symptoms. We believed that these two metrics offered complementary (but distinct) information. While we considered that higher symptoms

post-initial injection represent a surrogate for chronic and persistent symptoms of the knee,[58–60] we also evaluated change in symptoms over the disease course which could also be associated with treatment use.

Symptoms post-initial injection were measured at the index visit in which the initial injection was reported. Change in symptoms were assessed from the index visit to the visit before the reported switching/continued injection use. To evaluate average changes of symptoms between visits, we had to account for the varied number of years between visits. To do so, the difference of symptoms between index visit of injections use and one visit before the visit reporting switched/continued injection use was first calculated and then divided by the appropriate time intervals (Figure 2.2.A). For example, if the participants reported an initial injection at visit 2 and reported switching/continued injection use at visit 5, the average change of symptoms (e.g., WOMAC pain) between visits was calculated as: (pain score (visit 4) – pain score (visit 2)) / 2. If participants initiated injections use at visit 2 and reported switching/continued injection use at one year after initiation at visit 3, the difference between the index visit and one year before was calculated (Figure 2.2.B). Using these change scores adjusted for time intervals, we created 3 categories to define minimal clinically important changes for each symptom: 1) improved, 2) no change, and 3) worsened. A negative change in WOMAC scores indicated improvement ranging from -4.6 to -1.2 for pain, -1.5 to -0.5 for stiffness, and -9.9 to -4.1 for physical function.[61–64] The minimal threshold was used for creating the categories (e.g., improved pain was defined as < -1.2; worsened pain was defined as > 1.2; and no change was defined as -1.2 to 1.2).

Covariates

Sociodemographic characteristics including age, sex, race/ethnicity, income, and education were evaluated. Clinical characteristics such disease severity, knee-related symptoms, insurance coverage, general health status, and physical activity were also assessed. Sociodemographic variables (e.g., sex, race/ethnicity, and household income) were considered time-invariant variables and thus using information from the enrollment. Information from the index visit were used for other variables.

Sociodemographic variables were self-reported. Race/ethnicity was categorized as non-Hispanic White, non-Hispanic Black, and other. Educational levels were collapsed and categorized as high school or less, some college, college graduate, and at least some graduate school. Annual household income was categorized into three levels: < $25,000, $25,000-50,000, and > $50,000.

Clinical characteristics such as disease severity was measured based on K-L grade and joint space width (JSW). A detailed protocol regarding the measurement of JSW has been documented elsewhere.[31,65] If JSW measures were implausible (e.g., the distance between plateau and rim was > 6.5mm), we treated these as missing as these measures could be due to poorly positioned knees.[31] Multi-joint symptoms were present if participants had frequent pain, aching, or stiffness in ≥ two joints other than knee. Cumulative measures were used to assess the history of knee injury and surgery. Information was collected on prior knee injuries if participants had limited ability to walk

for ≥ two days reported at any previous visit. A history of having knee surgery included arthroscopy, ligament repair or meniscectomy at any previous visit.

General health status was evaluated using the 12-item Short-Form Health Survey (SF-12),[66] with summary scores for physical and mental health ranging from 0 to 100 and higher scores indicating better health status. Body mass index (BMI) was calculated using measured height and weight [weight (kg)/height $(m)^2$]. Participants were categorized into: < 25, normal weight; 25 to < 30, overweight; and ≥ 30, obese.[67] Information on depressive symptoms were collected using Centers for Epidemiologic Studies Depression Scale (CES-D). A CES-D score >16 indicated elevated depressive symptoms.[68] Comorbidity status was evaluated using the Charlson index and then categorized into 0, 1, and ≥ 2.[69] Health coverage status and physical activity were self-reported. The Physical Activity Scale for the Elderly (PASE) consisting of 26 items was used to assess activities including occupational, household, and leisure work over the past week, with higher scores indicating greater activities.[70]

Statistical analyses

Descriptive statistics including means and standard deviations (SD) for continuous variables and percentages for categorical variables were first calculated to describe socio-demographic and clinical characteristics by patterns of injections use for types of injections initiated at the index visit. Average yearly changes in symptoms between treatment initiation and switching/continued treatment use were also examined. Multicollinearity was evaluated and ruled out before the model building process by

evaluating the correlations between the covariates of interest. Six multinomial logistic models were then built to evaluate the association between two operational expressions (symptoms post-initial injection; and change in symptoms) for three symptoms (pain, stiffness, and physical function) with patterns of injection use. Adjusted odds ratios (aOR) and 95% confidence intervals (CI) for each group compared to a common reference group (one-time users) were estimated after adjusting for socio-demographics and clinical/functional factors. Due to the small sample size, we were not able to examine such relationships among participants initiating hyaluronic acid use. To address the possibility of misclassification bias, and specifically the case that those classified as "not continued" users may include patients who went on to total knee replacement (TKR) during the follow-up, we also conducted a sensitivity analysis. We first examined the overall proportion of participants who received TKR in follow-up among the group who were classified as not continued users. We then repeated the analysis and compared the findings with the main analysis after removing those participants who were initially classified as "not continued" but had TKR during follow-up.

Results

Overall, 96 initiated of hyaluronic acid injections and 316 initiated corticosteroid injections. Regardless of the type of initial injection, nearly 1 in 5 participants switched or continued the initial injection use and approximately 60% of participants did not receive any additional injections. Socio-demographic and clinical characteristics of participants initiating hyaluronic acid injections are presented in Table 2.1. The average

age of those who switched from hyaluronic injections to corticosteroids was 62.0 years whereas the average age for those continuing with hyaluronic injections or who had only one injection was 65.6 years and 67.0 years, respectively. Fifty-two percent of those who switched injections were women, while 40.0% of those who continued and 46.6% of one-time users were women. Nearly two thirds of continued users and one-time users had at least some graduate school education whereas half of those switching injections use had graduate level education. Half of participants who switched injections had K-L grade 4 and reports of multi-joint symptoms and history of knee injuries were common. Mean WOMAC pain at the index visit was 5.8 (SD: 3.7) for switching users whereas those who continued use and one-time users were 4.0 (SD: 2.9) and 6.3 (SD: 3.7), respectively.

For participants initiating corticosteroid injections (Table 2.2), the average age of one-time users was 68.7 years (SD: 9.5 years) whereas those switched or continued were 67.2 years and 67 years of age, respectively. The majority of participants initiating corticosteroid injections had annual household income >$50,000. For clinical characteristics, nearly half of participants who switched injection use had K-L grade 2 relative to the other groups (e.g., continued: 38.6%, one-time users: 30.2%). Overall, the majority of participants initiating corticosteroid injections had multi-joint symptoms. Knee-specific symptoms such as mean WOMAC pain at the index visit was 7.0 (SD: 4.4) for switching users whereas those who continued and one-time users were 5.2 (SD: 4.0) and 6.3 (SD: 4.1), respectively.

Table 2.3 shows the association between symptoms post-initial injection and switching or continuing corticosteroid injections compared to one-time users. In relation

to one-time users, greater knee-specific symptoms post-initial injection were associated with lower odds of continued injections use (adjusted odds ratio (aOR) for WOMAC pain: 0.91; 95%, confidence interval (CI): 0.83 to 0.99; $aOR_{stiffness}$: 0.77; CI: 0.63 to 0.94; $aOR_{physical\ function}$: 0.97; CI: 0.94 to 1.00). Compared with one-time users, symptoms post-initial injection and changes in symptoms over time did not appear to be associated with switching injections use among initiators of corticosteroid injections.

Table 2.4 shows the association between the clinically relevant change in symptoms with first switching/continuation compared to one-time users among participants initiating corticosteroid injections use. Reaching a priori defined minimally clinical important differences in worsened or improved pain was not associated with patterns of switching or continued injections use compared to one-time users. No association was observed for stiffness and physical function.

To address the possibility of misclassification bias in specification of "not continued" users, we found that approximately 8.5% of participants in this group had TKR after the initiation of injection use. After excluding those participants, results were comparable to the main findings.

Discussion

To our knowledge, this present study is the first study to describe the patterns of intra-articular injection use in patients with knee OA. Our data suggest that a substantial proportion of knee OA patients initiating intra-articular injections switched their treatment or used it for one time, regardless of the initial therapy or actual symptoms change. We found that approximately 1 in 5 participants initiating injections had

switched injection type, but that channeling into the hyaluronic injections was not apparent. While 24.0% of hyaluronic acid initiators switched to corticosteroid injections, 17.8% of corticosteroid initiators switched to hyaluronic acid injections. We also observed that it was reported symptoms post-initial injection rather than changes in symptoms (e.g., improved or worsened) over time that appeared to be associated with patterns of use among corticosteroid users.

We hypothesized that the patterns of injection use among persons with knee OA would be associated with more severe symptoms and/or changes in symptoms over time compared to those who only received one-time injections. Although we were not able to examine such a relationship among participants initiating hyaluronic acid injections owing to a limited sample size available for analysis, our findings did support the hypothesis that greater symptoms post-initial injection were associated with lower odds of continued treatment use among participants initiating corticosteroid injections. Further, we observed that changes in knee symptoms over time were not associated with either switching or continued injections compared to one-time users. Several explanations could be responsible for this observation. It might be the channeling effects of injection use rather than the actual symptom changes since current clinical guidelines are still conflicting.[12,24,25,71,72] A similar phenomenon was found in a study of celecoxib compared to other NSAIDs in OA patients.[73] Another potential explanation is that the relation between change in symptoms and actual treatments received may not be a linear. As such, it may affect the treatment decisions, particularly in persons with OA.[58,59] Since the sensitization of symptoms could be both influenced by both physical and psychological

factors, it could be that sustained chronic symptoms rather than the success or failure of treatment affect treatment decisions.

In addition to symptoms of the knee, the decision to continue or switch intra-articular injections among patients might be driven by several factors including patient choice, physician specialty, insurance reimbursement, and cost .[74,75] Indeed, the cost to manage OA could be substantial.[76–78] Typically, patients switch from cheaper treatments (e.g., corticosteroid injections) to newer and more expensive therapies that may target more directly the underlying pathophysiological effects (e.g., hyaluronic acid injections). However, we observed that approximately 1 in 4 participants initiating hyaluronic acid injection use switched to corticosteroid injections. Currently, treatment guidelines for patients with knee OA do not suggest a "step-up" approach for intra-articular injections use compared to other pain treatments, such as NSAIDs.[12,24,25] Another factor could be that "clinical equipoise" still exists and physicians and patients frequently switch between available options.[79,80] The end result may be additional economic and human burden of OA. Although general practitioners can administer corticosteroid injections, hyaluronic acid injections are typically provided by specialists such as an orthopedic surgeons or rheumatologists. The extent of switching from hyaluronic injections to corticosteroid injections observed in this study may be a reflection of ease of convenience, rather than preference of injection type.[75]

Treatment switching, continuation, and discontinuation in populations with chronic disease might be due to suboptimal efficacy, safety, and tolerability of treatment modalities.[55,56] We did observe that there were high percentages of switching in both

groups and this could be due to patients' attempts to manage chronic pain with potentially suboptimal treatments. Indeed, recent evidence shows that viscosupplements can last through 26 weeks but no similar evidence exists for corticosteroid injections among persons with knee OA.[26,27] More studies regarding the long-term efficacy using multiple treatments might be needed. We observed a high proportion of one-time injection use in this study. While in some clinical scenarios this may indicate potentially intolerable side effects of the modalities, adverse events and side effects of intra-articular injections are rare and more often due to localized reactions (e.g., swelling or redness on the injection site) rather than systematic effects from the agent.[81] In this present study, we were not able to evaluate this relationship since information related to adverse effects is limited in the OAI. The high proportion of one-time use may also be reflective of lack of perceived efficacy of the treatment.[82]

Strengths and limitations of this study are acknowledged. This is the first study examining patterns of intra-articular injections use for patients with knee OA. Although participants were not newly diagnosed with OA and thus may have had injection use before entering into the study, a new user design was used to minimize the bias.[83] Using data from the OAI, we were able to evaluate the associations between clinical outcomes and patterns of injections use longitudinally. The low number of participants initiating hyaluronic acid injections limited our ability to develop models of patterns of injection use in this group. Despite the comprehensive assessment in OAI, physicians' prescribing notes or chart information are lacking. Therefore, the extent to which treatment switching and continued use observed in this study related to safety or tolerability of injections use

is unknown. Although the indication of injections was self-reported and assessed for the 6 months before annual visits in the study, the proportion of participants receiving either injections is comparable to previous study.[84] In addition, there is a potential for mismatch between the timing of actual injection use and assessments of symptoms. We may not have the optimal window to examine the associations of injection use and symptoms for all participants. However, we examined symptoms using two operational definitions which provided complementary but distinct information. Last, since those "not continued" users may include patients who went on to TKR, there is a potential for misclassification bias. However, results from sensitivity analysis were comparable to the main findings.

In conclusion, we found that approximately 19% of patients with radiographically confirmed knee OA who initiated injections switched injection type and ~60% did not receive any injections after their initial injection. Furthermore, among those initiating corticosteroid injection use, we also observed that it was the symptoms post-initial injection rather than changes in symptoms (e.g., improved or worsened) over time that might be associated with patterns of injections use. Despite that the proportion of patients switching injection use and one-time users was substantial in both types of commonly used treatment agents, there is currently no "step-up" approach for intra-articular injections use in treatment guidelines for patients with knee OA. Further, these phenomena may suggest that longer-term efficacy regarding symptom relief and/or slowing disease progression of these agents may be suboptimal among patients with OA in the real-world setting.

Figure 2.1: Flowchart of study participants.

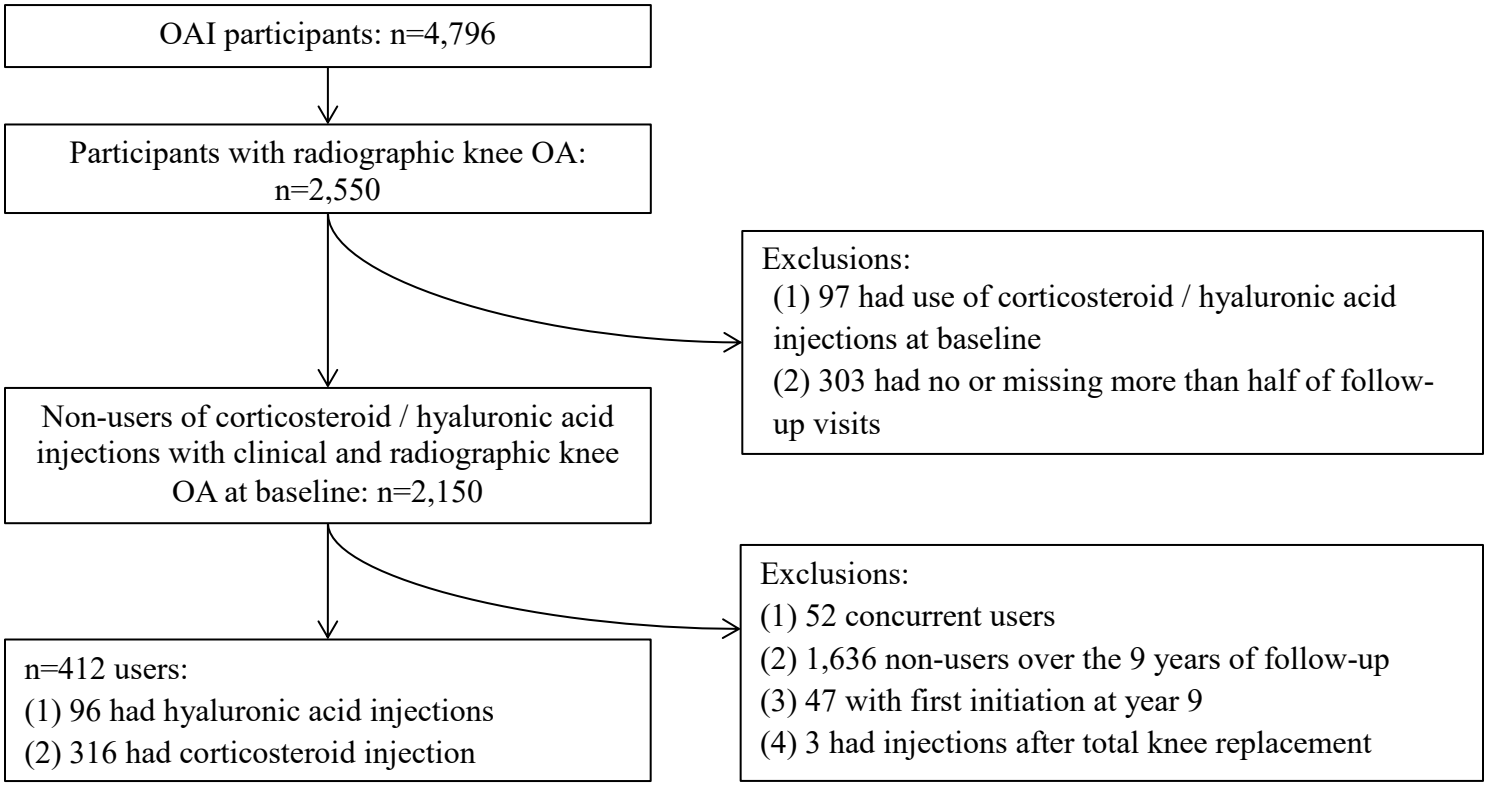


Figure 2.2: Scheme of the study design to define changes between visits among injection users.

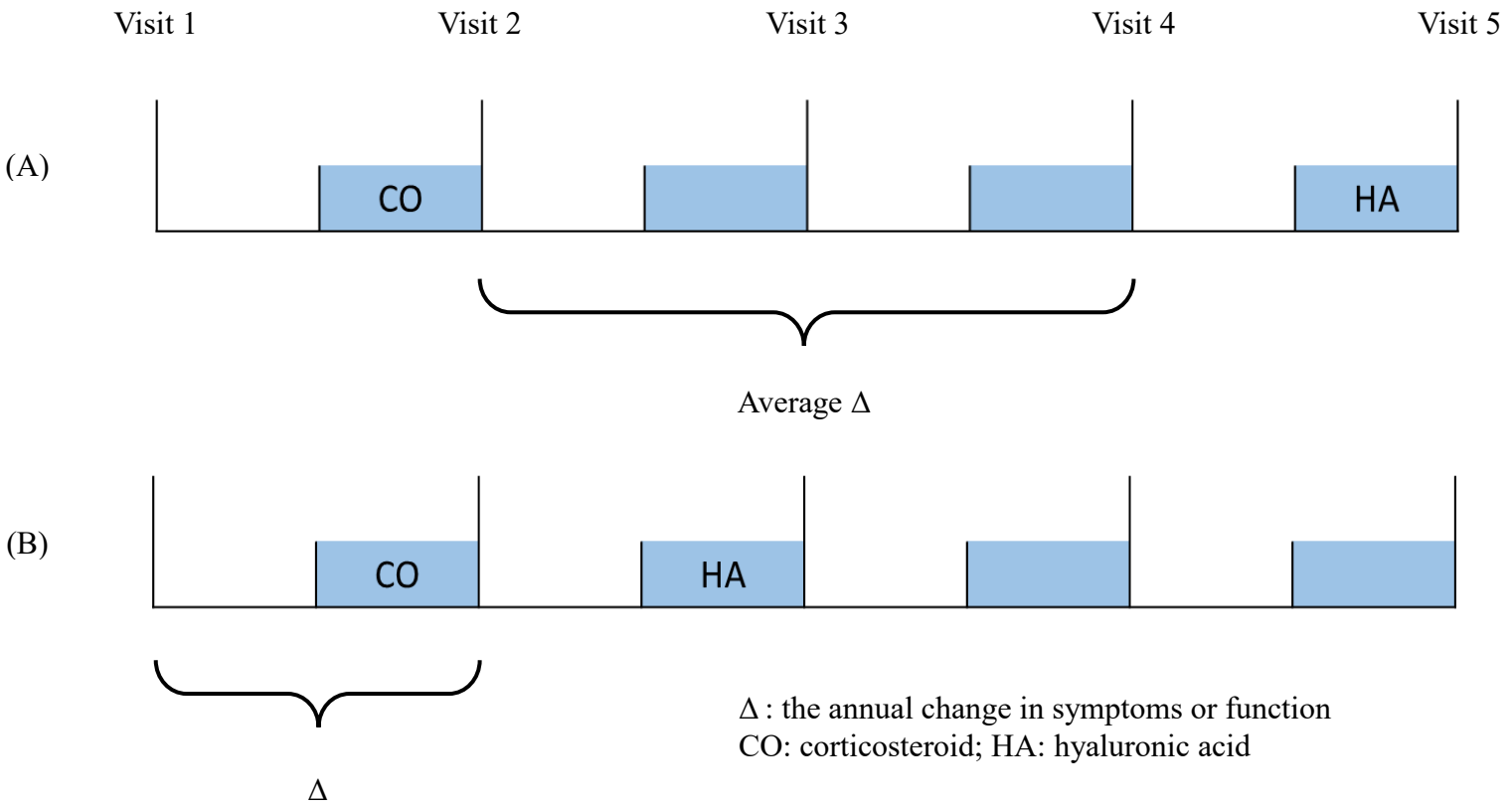


Table 2.1: Clinical characteristics by patterns of hyaluronic acid injection use (N=96*).

Characteristics	Switching users (n=23)	Continued users (n=15)	One-time users (n=58)
Mean (SD) age in years	62.0 (8.0)	65.6 (7.4)	67.0 (8.6)
Mean (SD) intervals of visit	2.4 (1.9)	1.5 (0.9)	NA
Women (%)[a]	52.2	40.0	46.6
Race/ethnicity (%)[a]			
Non-Hispanic White	82.6	93.3	89.7
Non-Hispanic Black	8.7	0	6.9
Other	8.7	6.7	3.5
Education (%)[a]			
High school or less	8.7	20.0	12.3
Some college	39.1	20.0	22.8
College graduate	8.7	33.3	19.3
Some graduate school or above	43.5	26.7	45.6
Income (%)[a]			
<$25,000	0	13.3	3.5
$25,000 - $50,000	34.8	6.7	22.8
>$50,000	65.2	80.0	73.7
Body mass index (%)			
Normal	13.0	20.0	15.5
Overweight	34.8	20.0	31.0
Obese	52.2	60.0	53.5
Health care coverage (%)	100.0	100.0	100.0
Kellgren-Lawrence grade (%)			
2	15.0	15.4	18.4
3	35.0	46.2	46.9
4	50.0	38.5	34.7
Multi-joint symptoms (%)	78.3	46.7	60.3
History of knee injury (%)	69.6	46.7	56.9

History of knee surgery (%)	56.5	46.7	53.5
Depressive symptoms (CES-D >16) (%)	13.0	20.0	9.3
Comorbidity status (%)			
0	73.9	64.3	72.4
1	17.4	28.6	19.0
≥2	8.7	7.1	8.6
WOMAC scores, mean (SD)			
Pain	5.8 (3.7)	4.0 (2.9)	6.3 (3.7)
Stiffness	2.7 (1.7)	2.3 (1.5)	3.0 (1.6)
Physical function	17.1 (11.5)	14.2 (9.9)	18.7 (12.1)
Joint space width, mean (SD)	3.4 (1.9)	4.6 (1.7)	3.9 (2.2)
SF-12 PCS, mean (SD)	40.2 (9.9)	41.5 (9.1)	41.2 (10.4)
SF-12 MCS, mean (SD)	54.0 (7.6)	53.5 (11.5)	56.2 (7.6)
PACE, mean (SD)	154.2 (70.4)	177.4 (64.2)	145.1 (82.3)

Abbreviation: CES-D, Centers for Epidemiologic Studies Depression Scale; MCS, SF-12 Mental Component Summary scores; PACE, Physical Activity Scale for the Elderly; PCS, SF-12 Physical Component Summary scores; SD, standard deviation; WOMAC, The Western Ontario and McMaster Universities Arthritis Index.

[a] Information at enrollment was used.

* Number of participants with missing information: education (1), annual household income (1), health care coverage (5), Kellgren-Lawrence grade (14), CES-D (4), comorbidity status (1), WOMAC Physical function (2), joint space width (24), SF-12 PCS (5), SF-12 MCS (5), PACE (2).

Table 2.2: Clinical characteristics by patterns of corticosteroid injection use (N=316*).

Characteristics	Switching users (n=56)	Continued users (n=75)	One-time users (n=185)
Mean (SD) age in years	67.2 (9.5)	67.0 (7.9)	68.7 (9.5)
Mean intervals	2.3 (1.7)	1.9 (1.5)	NA
Women (%)[a]	66.1	60.0	64.3
Race/ethnicity (%)[a]			
Non-Hispanic White	76.8	74.7	78.4
Non-Hispanic Black	14.3	20.0	18.4
Other	8.9	5.3	3.2
Education (%)[a]			
High school or less	14.3	14.7	21.7
Some college	21.4	26.7	27.7
College graduate	28.6	24.0	20.1
Some graduate school or above	35.7	34.7	30.4
Income (%)[a]			
<$25,000	14.3	12.0	15.8
$25,000 - $50,000	28.6	28.0	33.7
>$50,000	57.1	60.0	50.5
Body mass index (%)			
Normal	14.3	12.0	13.0
Overweight	33.9	45.3	41.1
Obese	51.8	42.7	46.0
Health care coverage (%)	100.0	100.0	98.9
Kellgren-Lawrence grade (%)			
2	48.9	38.6	30.2
3	28.9	38.6	39.0
4	22.2	22.9	30.8
Multi-joint symptoms (%)	66.1	58.7	51.9
History of knee injury (%)	42.9	49.3	51.4

History of knee surgery (%)	23.2	32.0	37.3
Depressive symptoms (CES-D >16) (%)	18.2	13.3	14.0
Comorbidity status			
0	52.7	58.7	65.8
1	32.7	24.0	15.2
≥2	14.6	17.3	19.0
WOMAC scores, mean (SD)			
Pain	7.0 (4.4)	5.2 (4.0)	6.3 (4.1)
Stiffness	3.0 (1.8)	2.5 (1.8)	2.9 (1.7)
Physical function	20.2 (12.7)	17.2 (13.6)	19.6 (13.2)
Joint space width, mean (SD)	4.7 (1.9)	4.4 (2.0)	4.2 (1.9)
SF-12 PCS, mean (SD)	39.4 (8.7)	41.1 (10.3)	41.3 (9.9)
SF-12 MCS, mean (SD)	54.8 (10.4)	55.1 (9.5)	53.8 (9.3)
PACE, mean (SD)	141.0 (81.0)	137.9 (71.8)	137.3 (73.8)

Abbreviation: CES-D, Centers for Epidemiologic Studies Depression Scale; MCS, SF-12 Mental Component Summary scores; PACE, Physical activity scale for the elderly; PCS, SF-12 Physical Component Summary scores; SD, standard deviation; WOMAC, The Western Ontario and McMaster Universities Arthritis Index.

[a] Information at enrollment was used.

* Number of participants with missing information: education (1), annual household income (1), health care coverage (10), Kellgren-Lawrence grade (42), CES-D (7), comorbidity status (2), WOMAC Physical function (7), joint space width (94), SF-12 PCS (16), SF-12 MCS (16), PACE (20).

Table 2.3: Association* between symptoms post-initial injection and first switching/continuation of corticosteroid injections among participants with radiographic confirmed knee osteoarthritis.

	Switching users	Continued users
Pain	Odds Ratio (95% Confidence Interval)	
Crude	1.04 (0.97 to 1.12)	0.93 (0.87 to 1.00)
Adjusted[a]	0.97 (0.88 to 1.07)	0.91 (0.83 to 0.99)
Stiffness		
Crude	1.02 (0.86 to 1.21)	0.88 (0.75 to 1.03)
Adjusted[a]	0.83 (0.66 to 1.05)	0.77 (0.63 to 0.94)
Physical function		
Crude	1.00 (0.98 to 1.03)	0.99 (0.97 to 1.01)
Adjusted[a]	0.97 (0.94 to 1.00)	0.97 (0.94 to 1.00)

* Odds ratios (95% Confidence Intervals) were estimated using participants with one-time use of injections as the reference group.

[a] Adjusted for age at the index visit, sex, K-L grade, comorbidity status, and SF-12 physical component scores.

Table 2.4: Association* between average change in symptoms and first switching/continuation of corticosteroid injections among participants with radiographic confirmed knee osteoarthritis.

	Switching			Continued		
Pain	Improved Pain¶	No changes	Worsened Pain	Improved Pain¶	No changes	Worsened Pain
Crude	0.69 (0.32 to 1.49)	Reference	1.11 (0.54 to 2.31)	0.48 (0.24 to 0.99)	Reference	1.00 (0.52 to 1.89)
Adjusted[a]	0.61 (0.25 to 1.47)	Reference	0.66 (0.27 to 1.61)	0.56 (0.26 to 1.22)	Reference	0.99 (0.49 to 2.01)
Stiffness	Improved Stiffness¶	No changes	Worsened Stiffness	Improved Stiffness¶	No changes	Worsened Stiffness
Crude	0.96 (0.46 to 2.00)	Reference	0.66 (0.31 to 1.43)	1.39 (0.70 to 2.78)	Reference	1.19 (0.59 to 2.38)
Adjusted[a]	1.02 (0.44 to 2.33)	Reference	0.47 (0.18 to 1.19)	1.53 (0.72 to 3.24)	Reference	1.34 (0.64 to 2.82)
Physical function	Improved Function¶	No changes	Worsened Function	Improved Function¶	No changes	Worsened Function
Crude	0.57 (0.26 to 1.23)	Reference	0.63 (0.30 to 1.31)	0.58 (0.29 to 1.15)	Reference	0.58 (0.30 to 1.12)
Adjusted[a]	0.55 (0.23 to 1.33)	Reference	0.42 (0.17 to 1.05)	0.71 (0.34 to 1.49)	Reference	0.68 (0.33 to 1.41)

* Odds ratios (95% Confidence Intervals) were estimated using participants with one-time use of injections as the reference group.

[a] Adjusted for age at the index visit, sex, K-L grade, comorbidity status, and SF-12 physical component scores

¶ A negative change in WOMAC scores indicated improvement ranging from -4.6 to -1.2 for pain, -1.5 to -0.5 for stiffness, and -9.9 to -4.1 for physical function. The minimal threshold was used for creating the categories (e.g., improved pain was defined as < -1.2; worsened pain was defined as > 1.2; and no change was defined as -1.2 to 1.2).

CHAPTER III

EFFECT OF INTRA-ARTICULAR INJECTIONS ON PATIENT-REPORTED SYMPTOMS IN PERSONS WITH OSTEOARTHRITIS: ANALYSIS WITH MARGINAL STRUCTURAL MODELS

Abstract

Objective: Parameters for use of intra-articular injections lacks consensus across clinical guidelines. This study sought to examine the effectiveness of corticosteroid or hyaluronic acid injections in relieving symptoms among persons with knee osteoarthritis (OA).

Methods: Using Osteoarthritis Initiative (OAI) data, we applied a new-user design to identify participants who initiated corticosteroid or hyaluronic acid injections during the study. We identified 988 participants with follow-up information for at least one year who met our eligibility criteria. The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) was used to measure knee symptoms (pain, stiffness, function). We used marginal structural models controlling for time-varying confounders to estimate the effects of newly initiated injection use compared to non-users over two years of follow-up.

Results: Among 412 participants initiating injections, 77.2% used corticosteroid injections and 22.8% used hyaluronic acid use. Compared to non-users, on average, participants reporting corticosteroid injection initiation experienced a worsening of pain (yearly worsening: 1.24 points; 95% confidence interval [95% CI]: 0.82 to 1.66), stiffness (yearly worsening: 0.30 points; 95% CI: 0.10 to 0.49), and physical functioning (yearly worsening: 2.62 points; 95% CI: 0.94 to 4.29) after adjusting for potential confounders with marginal structural models. The hyaluronic acid injections did not show improvements of WOMAC subscales (pain: 0.50; 95% CI: -0.11 to 1.11, stiffness: -0.07; 95% CI: -0.38 to 0.24, and functioning: 0.49; 95% CI: -1.34 to 2.32).

Conclusions: The initiation of corticosteroid or hyaluronic injection use did not appear to reduce symptoms over the two years of follow-up.

Introduction

The prevalence of osteoarthritis (OA) is increasing and in the U.S. with ~ 27 million people are afflicted with the disease.[1] Knee OA is 1 of the 5 leading causes of disability among non-institutionalized adults in the U.S.[11] OA currently has no cure. The goals of OA treatment are to improve pain relief, mobility, quality of life, and delay disease progression.[12,24,25] Treatments include non-pharmacologic or pharmacologic modalities. If non-pharmacologic intervention such as exercise or weight management, and/or the use of orally administered medications such as acetaminophen are ineffective, intra-articular injections of corticosteroids or hyaluronic acid may offer symptom relief for patients with knee OA.[12–14]

Despite a large number of studies, the safety and efficacy of intra-articular injections remains inconclusive leading to a lack of consensus across clinical guidelines.[12,24,25] Evidence about the extent to which long-term use of intra-articular injections improves patient outcomes is lacking. Recent systematic reviews and meta-analyses suggest that the effect of US-approved viscosupplement injections can last through 26 weeks but there is no similar evidence for corticosteroid injections for persons with knee OA.[17,26,27] Among patients with milder disease, receiving intra-articular sodium hyaluronate appears to slow joint space narrowing compared to placebo.[28] For corticosteroid injections, there is no difference between treatment and placebo groups in joint space changes over two years of follow-up.[29] Studies documenting the longitudinal impact of patient-reported outcomes on knee OA symptoms are scarce.

Despite the lack of evidence and conflicting practice recommendations from guidelines, the use of injections is increasing among Medicare beneficiaries newly diagnosed with knee OA.[54] The cost of long-term injection could be substantial (i.e., $1700 to $3700 for viscosupplementation treatments).[53] Given the widespread use of injections and the rising costs of these treatments,[53,85,86] understanding the long-term effectiveness of intra-articular injections among persons with knee OA is warranted.

The aim of this present study was to estimate the effect of intra-articular injections use on changes in patient reported symptoms. This study builds on previous research in several areas. First, we used data from the Osteoarthritis Initiative (OAI), a multi-center study that enrolled participants with radiographically confirmed knee OA and conducted annual assessments with validated patient-reported outcomes and measures of disease progression. Second, compared to clinical trials,[26,27] this longitudinal and non-experimental study enabled us to examine injection use over a longer period of time and to evaluate treatment benefits in a real-world setting. Last, advanced statistical techniques were used that allow us to quantify the effect of injections use over time in a more heterogeneous population compared to clinical trials.

Methods

Data source

We used publicly available data from the OAI (http://oai.epi-ucsf.org/). The OAI was a longitudinal and prospective cohort study enrolling 4,796 adults aged 45 to 79 years at baseline using four study sites (i.e., Baltimore, MD; Columbus,

OH; Pittsburgh, PA; and Pawtucket, RI). The aims of the OAI study were to examine the development and progression of knee OA among adults with symptomatic OA in at least 1 knee or at least 1 established risk factor. Participants had annual follow-up assessments for up to 9 years. Detailed information about the OAI protocol has been described elsewhere.[31]

Study sample and design

Figure 3.1 shows the inclusion/exclusion criteria for our study sample. Only participants with radiographically confirmed knee OA in at least 1 knee at baseline (Kellgren - Lawrence grade ≥ 2) were included (n=2,550). To improve validity of the study, we restricted our analysis to "new users" of knee injections.[83] As such, participants who had reported injection use at baseline were not eligible (n=97). In addition, participants indicating no injection use but having missing values for more than half of the follow-up visits over the 9 years were also excluded (n=303). From the remaining group, we identified participants with and without initiation of injection use during the follow-up period. Among initiators, we excluded those reporting use of both injection types (concurrent hyaluronic acid and corticosteroid injection users, n=52), those reporting the first initiation at year 9 because we had no follow-up data after the injections (n=47), and those who reported injection in the affected knee after total knee replacement (n=3). To mimic the study design from clinical trials,[29,32] participants whose age < 45 years and did not have symptomatic knee OA at baseline were considered ineligible for injection use and thus further excluded for non-users. The final analytic

sample included 412 participants initiating injection use and 576 non-users. Among those initiating injection use with available follow-up information for at least one year, 94 initiated hyaluronic acid injections and 318 initiated corticosteroid injections.

Use of index knee

We used an index knee for the analysis based on: 1) radiographic evidence of OA and 2) the presence of symptoms (e.g., pain) in the same knee. If only one knee had radiographically confirmed OA at baseline, then that knee was used as index knee. If participants had radiographically confirmed OA for both knees, then the knee with higher pain scores at baseline measured by Western Ontario and McMaster Universities Arthritis Index (WOMAC) pain subscale was used as the index knee. If pain scores for both knees were equal, then index knee was the one with worse K-L grade.

Assessment of injection use

In OAI, injection use was assessed separately for both knees. Participants were first asked "During the past 6 months, have you had any injections in either of your knees for treatment of arthritis?" For those answering "yes", two separate questions were posed regarding hyaluronic acid or corticosteroid injections use. For hyaluronic acid injections, participants were asked: "During the past 6 months, have you had an injection of hyaluronic acid (Synvisc or Hyalgan) in either of your knees for treatment of your arthritis?" These injections are given as a series of 3 to 5 weekly injections. To assess corticosteroid injection use, participants were asked: "During the past 6 months, have you

had an injection of steroids (cortisone, corticosteroids) in either of your knees for treatment of your arthritis?" For participants whose index knees were censored during the follow-up (e.g., due to death, switching injection, and/or having total knee replacement), we used available information from the other knee to recapture the sample (16 out of 412).

Assessment of OA symptoms

Knee symptoms were evaluated annually using the WOMAC scales (Likert version 3.1) including three subscales: pain (5 items), stiffness (2 items), and physical function (17 items).[57] Each item of the subscale ranged from 0 to 4 (0=none and 4=extreme). Responses to items in each subscale were summed to produce the individual summary score ranging from 0-20 for pain, 0-8 for stiffness, and 0-68 for physical function. Higher WOMAC scores indicate worse symptoms/function. The primary outcome was change in each subscale between baseline visit (one year before the injection), index visit, and one year after the index visit.

Covariates

We considered covariates in two groups: time-invariant and time-dependent. Sociodemographic factors including sex, race/ethnicity, and income were considered as time-fixed covariates measured only at the time of enrollment. Age of participants at time of injection treatment initiation was also considered as a time-fixed covariate. Income

(from all sources) was measured using self-reported personal family income for the year before the enrollment.

Covariates that were collected annually included clinical characteristics of OA, indices of general health status, body mass index (BMI), and use of medications and biologically related supplements. These were treated as time-varying covariates. The OAI collected comprehensive measurements of OA related clinical characteristics including K-L grade, multi-joint symptoms, history of knee injury or surgery, and knee alignment.[31] In the OAI, K-L grade was measured from enrollment to year 4 for every participant. Thereafter, K-L grade was measured on a subset of the participants. Among participants initiating injections after year 5, we carried the last observation forward for the measures of K-L grade from year 4[87] if no information was available. Multi-joint symptoms were present if participants had frequent pain, aching, or stiffness in at least 2 joints other than the knee.[88] Knee malalignment including varus or valgus deformity was measured and recorded using a goniometer. History of knee injuries was present if a prior injury limited the participant's ability to walk for at least 2 days indicating on any previous visit. A history of having knee surgery was present if participants indicated that they had arthroscopy, ligament repair or meniscectomy on any previous visit.

The 12-item Short-Form (SF-12) health survey was used to assess general health status.[66] Physical and mental component summary scores were calculated and range from 0 to 100. Higher scores indicate better health status. The Centers for Epidemiologic Studies Depression Scale (CES-D) was used to evaluate depressive symptoms. Elevated depressive symptoms were considered present if participants had a CES-D score >16.[68]

A validated self-administered questionnaire modeled after the Charlson index[69] was used to develop a comorbidity score which sums weights assigned to comorbid conditions (range: 0 to 32, higher scores indicating greater severity in comorbid conditions). The comorbidity score was categorized into 0, 1, and ≥2.[69] BMI was calculated from measured height and weight [weight (kg)/height $(m)^2$] and categorized as less than 25, normal weight; 25 to less than 30, overweight; and 30 and over, obese.[89]

We considered use of pharmacological treatments such as analgesics and biologically related supplements as potential confounders. At each visit, analgesic use including acetaminophen, non-steroidal anti-inflammatory drugs (NSAIDs), COX-2 selective inhibitors, opioids, and doxycycline was assessed for the previous 30 days. Biologically based supplements including glucosamine, chondroitin sulfate, Methylsulfonylmethane and S-adenosylmethionine were also assessed for the previous 30 days. Both over-the-counter and prescription medications captured in the Medications Inventory File or reported by patients in the medication history survey were used to define use.

Statistical analysis

To understand the potential selection bias that may arise due to "lost to follow-up", we first compared the characteristics of sociodemographic and clinical factors and concurrent pharmacological treatment use at baseline (one year before initiation), index year, and one year after initiation. We also examined the distribution of the outcome variables and ruled out departures from normality. We then developed a series of models

to derive crude estimates, an estimate adjusted for baseline covariates, and adjusted for time-varying confounders using generalized estimating equations (GEEs) for continuous outcomes adjusted for within-participant correlation with an unstructured correlation matrix.[90]

Given the OAI data structure, we considered that previously measured study outcomes and time-varying confounders may be simultaneously confounders and intermediate variables (Figure 3.2). As a result, the estimated overall treatment effects would likely be biased using standard regression models.[41] To account for time-varying confounders that may lie on the causal pathway from previous treatments to the study outcomes, we used marginal structural models to estimate the overall treatment effects of injection use through inverse probability of treatment weights.[30,33]

The weights were calculated in three steps. First, we estimated time-varying stabilized inverse probability treatment weights separately for hyaluronic acid injection and corticosteroid injection use using non-users as the comparator at the index and follow-up visit. While the numerator was estimated using the conditional probability of observed injection use given the baseline characteristics, the denominator was the predicted probability of observed injection use at the index and follow-up visit conditional on baseline covariates and time-varying confounders (e.g., WOMAC subscale scores measured at the prior visit and the same visit as use of injections). To construct appropriate weights, we also explored the sensitivity of weights to different model specifications at the index visit (Supplementary Table 3.1).[40] For three different outcomes, we adjusted for the previously measured WOMAC subscales as a potential

confounder. Second, since participants were censored due to death, switching injection, and/or having total knee replacement during the follow-up, we estimated and incorporated the inverse probability of censoring weight to account for the potential selection bias due to differential censoring by injections use.[30,33] Sociodemographic and clinical factors among participants who were censored by censoring mechanisms were also examined (Supplementary Table 3.2). Censoring weights were calculated using similar approach as treatment weights, except that past treatment use was also added into models to estimate the probability of having observed censoring status. Lastly, final weights were then calculated as the products of treatment (including index and follow-up visits) and censoring weights. In addition to checking the distributions of the final weights, we also plotted the log odds of injection use conditional on the covariates to examine if there was an adequate degree of variation given observed values against the predicted injection use (Supplementary Figure 3.1).[91] To minimize the impact of potential violations of the positivity assumption, we also truncated the weights at the first and 99th percentile.[40]

With the final estimated weights, we used weighted linear models adjusted for baseline covariates to estimate effects of injection use on changes in symptoms and disease progression with 95% confidence intervals (95% CIs). Under the assumptions of no unmeasured confounding with correct specifications of treatment and outcome models, the beta coefficients from marginal structural models indicated the effects of hyaluronic acid or corticosteroid injection use compared to non-users on yearly changes in WOMAC scores. Minimal clinically important changes for improvements were

defined using previous validation studies ranging from -4.6 to -1.2 for WOMAC pain, -1.5 to -0.5 for WOMAC stiffness, and -9.9 to -4.1 for WOMAC physical function.[61–64]

Sensitivity analysis

To examine the robustness of findings, we conducted sensitivity analyses to account for missing values of covariates. Multiple imputation was performed to handle missing data in the context of marginal structural model analyses.[92] We applied the Fully Conditional Specification method for imputation of missing data using SAS PROC MI FCS.[93] We first used all available information from the covariates (including the outcome variable) as variables in the imputation model to impute the missing values.[94] Twenty imputed datasets were created. We then incorporated the imputed values to rebuild the inverse probability treatment weights and fit the outcome models for each imputed dataset. Finally, we combined estimates and generated valid inferences using SAS PROC MIANALYZE to compare results.

Results

Sociodemographic and clinical characteristics of study participants

Table 3.1 shows sociodemographic and clinical characteristics of study participants at baseline (one year before the injection initiation), index year, and one year after the injection initiation among those remaining uncensored during follow-up. This provides insight into potential bias due to differential censoring mechanisms over time.

Overall, the majority of study participants were women, non-Hispanic white, college graduate or above, and had a household income >$50,000. Among participants initiating injection use, most of them had K-L ≥ 3. Both CO injection initiators and non-users had similar distributions of sociodemographic and clinical characteristics at baseline. Men and those with higher household income (e.g., > $50,000) comprised the majority of HA injection initiators relative to non-users. Among HA injection initiators, 33.7% of had K-L grade 4, while 17.4% of non-users had K-L grade 4. During follow-up, the proportion of those censored at one year after initiation was 29.8% for HA injection initiators relative to the other groups (e.g., CO: 19.9%; non-users: 3.0%). Among those who remained uncensored, the distribution of characteristics was similar over time compared to the distribution at baseline.

Concurrent pharmacological treatment use

NSAIDs were the most commonly reported concurrent-use pharmacological treatments among the study groups at baseline (Table 3.2). The majority of injection initiators reported analgesic use. Similar to sociodemographic and clinical characteristics, both CO injection initiators and non-users had similar distributions of concurrent pharmacological treatments use at baseline. Among HA injection initiators, 48.9% reported concurrent use of supplements such as glucosamine whereas CO injection initiators and non-users reported 30.8% and 27.1%, respectively. During follow-up, the distribution of concurrent pharmacological treatment use remained similar over time between CO injection initiators and non-users group. However, among HA injection

initiators, the use of glucosamine or chondroitin sulfate decreased from baseline to one-year after initiation (e.g., glucosamine: 48.9% to 36.4%; chondroitin sulfate: 44.7% to 30.3%).

Effects of injections use on knee OA

Table 3.3 shows average effects of initiating corticosteroid or hyaluronic acid injection use compared to non-users on patient-reported outcomes. After adjusting for potential confounders with marginal structural models, the use of corticosteroid injections did not improve WOMAC subscales compared with non-users. On average, the yearly changes were 1.24 (95% CI: 0.82 to 1.66) for WOMAC pain, 0.30 (95% CI: 0.10 to 0.49) for WOMAC stiffness, and 2.62 (95% CI: 0.94 to 4.29) for WOMAC physical function. While results from sensitivity analyses were qualitatively similar to our main findings, the effect of estimates for WOMAC pain did not meet a priori definitions of minimal clinically important differences.

While the use of hyaluronic acid injections did not show improvements of WOMAC subscales compared with non-users, the magnitude of effects were relatively smaller compared to corticosteroid injections use. On average, the yearly changes were 0.50 (95% CI: -0.11 to 1.11) for WOMAC pain, -0.07 (95% CI: -0.38 to 0.24) for WOMAC stiffness, and 0.49 (95% CI: -1.34 to 2.32) for WOMAC physical function. The findings from sensitivity analyses remained similar.

Discussion

Using data from the OAI, a longitudinal, multi-center, and prospective cohort study, we identified 412 participants with radiographically confirmed knee OA initiated injection use. Among those, 77.2% initiated corticosteroid injections and 22.8% initiated hyaluronic acid injection use. In the 2 years of follow-up, we did not observe reduced symptoms associated with the initiation of corticosteroid or hyaluronic injections compared with non-users after carefully controlling for potential time-varying and time-independent confounders with marginal structural models.

Among participants initiating corticosteroid injection use compared to non-users, our study findings are consistent with a newly updated review.[26] It suggests that the use of intra-articular corticosteroids does not support benefits of improved symptoms in the long-term after stratifying results by length of follow-up. However, the study duration included in the review ranged from two weeks to one year and mixed single injection and multiple injection use. To our knowledge, there are only 2 published trials that are comparable to our study design that assessed the effect of continuous intra-articular corticosteroid use over two years.[29,95] Our results are consistent with both studies and demonstrate that the use of corticosteroid injections over two years do not appear to reduce symptoms.[29,95] Despite the fact that, more studies with adequate power and proper design may still be needed, our findings do contribute to the growing body of evidence produced using non-experimental study design with advanced analytical techniques.

With respect to changes in symptoms, our results did not appear to support the use of hyaluronic acid injections. Although our findings are not consistent with evidence from reviews and meta-analysis,[27,96] there are some issues that may hamper the

comparison. First, the follow-up periods in trials included in the reviews are mostly short-term with only one treatment cycle. As such, the beneficial effects of long-term use remain unclear. Studies with longer follow-up periods using multiple treatment cycles may be needed. In addition, the potential efficacy of hyaluronic injections to patients with more severe disease remains unknown since some trials excluded patients with severe knee OA. In our study, we observed a substantial percentage of participants initiating these injections had K-L 4.

Currently, the evidence from systematic reviews and meta-analysis is not conclusive about the effects of hyaluronic acid injections use.[27,96–98] One of the explanations is the potential for publication bias since some reviews selected small and/or poor quality trials with positive results.[97,98] In addition, evidence from clinical trials is also mostly generated from small studies within a shorter period.[99] When using only results from larger trials with better quality, later updated reviews suggested that the use of hyaluronic acid injections compared to non-users is associated with small but not clinically important improvement in knee symptoms.[97,98] Indeed, for changes in symptoms, our results are consistent with the study with a larger sample size over one year of follow-up.[100]

The efficacy of both corticosteroid and hyaluronic injections for knee OA patients remains in question. Nevertheless, the use of both types of injections is increasing.[54] More information about the comparative efficacy of these common treatments is needed for patients, clinicians, and decision makers. Regarding the efficacy of corticosteroid and hyaluronic injections, most trials only compare with placebo and the differential effect

between the two injections is less clear. While a systematic review suggested corticosteroid injections were more effective for pain in the first 4 weeks, a recently published trial suggested that the use of corticosteroid injection had similar effects on symptom relief for the first two weeks relative to hyaluronic acid injection.[17,101] We recognize that an active-comparator design offers several advantages.[102] However, we were not able to use this design with the OAI dataset. Because the proportion of participants continuing injection use in the subsequent year was small in our study, we were not able to construct inverse probability treatment weights needed for active comparison.[103]

Our study has several strengths. To avoid overestimating the treatment benefits, we used a new-user design by studying treatment initiators.[83] To minimize confounding by indication, we included comparable participants who did not receive injections by using detailed information regarding disease severity. To address threats to the validity of the study such as time-varying confounders and "lost to follow-up", we used advanced statistical techniques of inverse probability treatment weights with marginal structural models.[30] We further performed sensitivity analyses using multiple imputation to evaluate the robustness of the main results. Under the same study design, analytic techniques, and outcome definition, the results from sensitivity analyses showed consistent findings.

Several limitations are also acknowledged. First, no information was available on the formulation of injections as well as dosages used. Currently, there are several molecular weights available for hyaluronic acid injections[98] and several different

corticosteroid formulations. Second, there is a potential for mismatch between the time of injection use and outcome assessments. For example, at annual assessment visits when participants were asked about injections used, they could be in the middle of treatment cycle. Therefore, the treatment effects could be underestimated. Third, residual confounding is still a possibility despite the comprehensive assessments of disease severity and concurrent treatment use in the OAI. Lastly, in the practice of constructing appropriate weights, there could be a model misspecification and/or violation of positivity assumption.[40] However, we carefully constructed weights using an iterative process and graphically examined if there was an adequate degree of variation given observed and predicted injection use. We also truncated weights to minimize the potential impact of violating the positivity assumption.[40]

In conclusion, in patients with knee OA, initiating treatments with either corticosteroid or hyaluronic acid injections was not associated with reduced symptoms compared to non-users over two years. Future research targeting comparative effectiveness of these commonly used injections may be helpful to understand the usefulness of these treatments for patients with knee OA.

Figure 3.1: Flowchart of study participants.

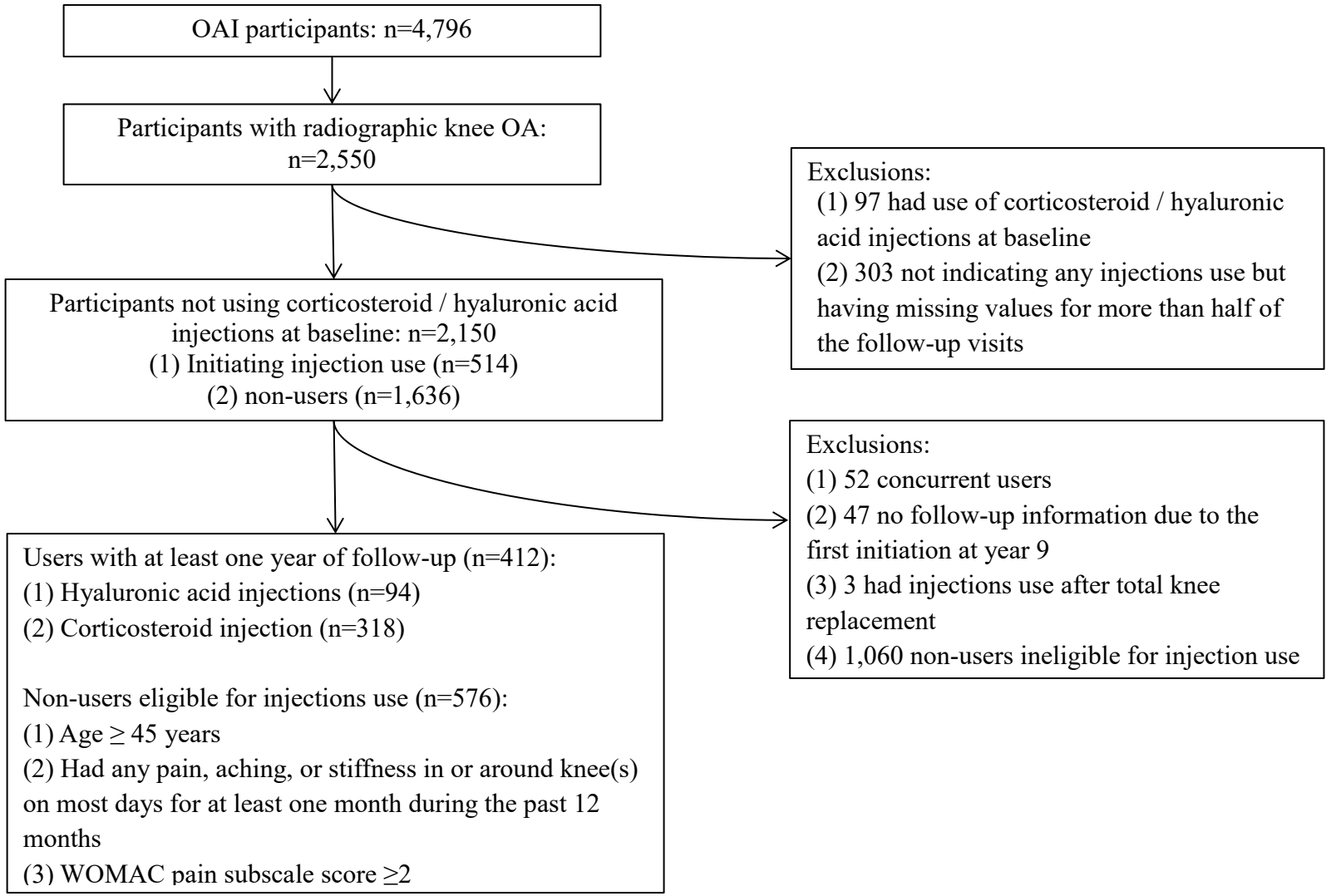


Figure 3.2: Directed acyclic graph (causal diagram) between the initiation of injection use, study outcomes, censoring, and potential time-varying confounders.

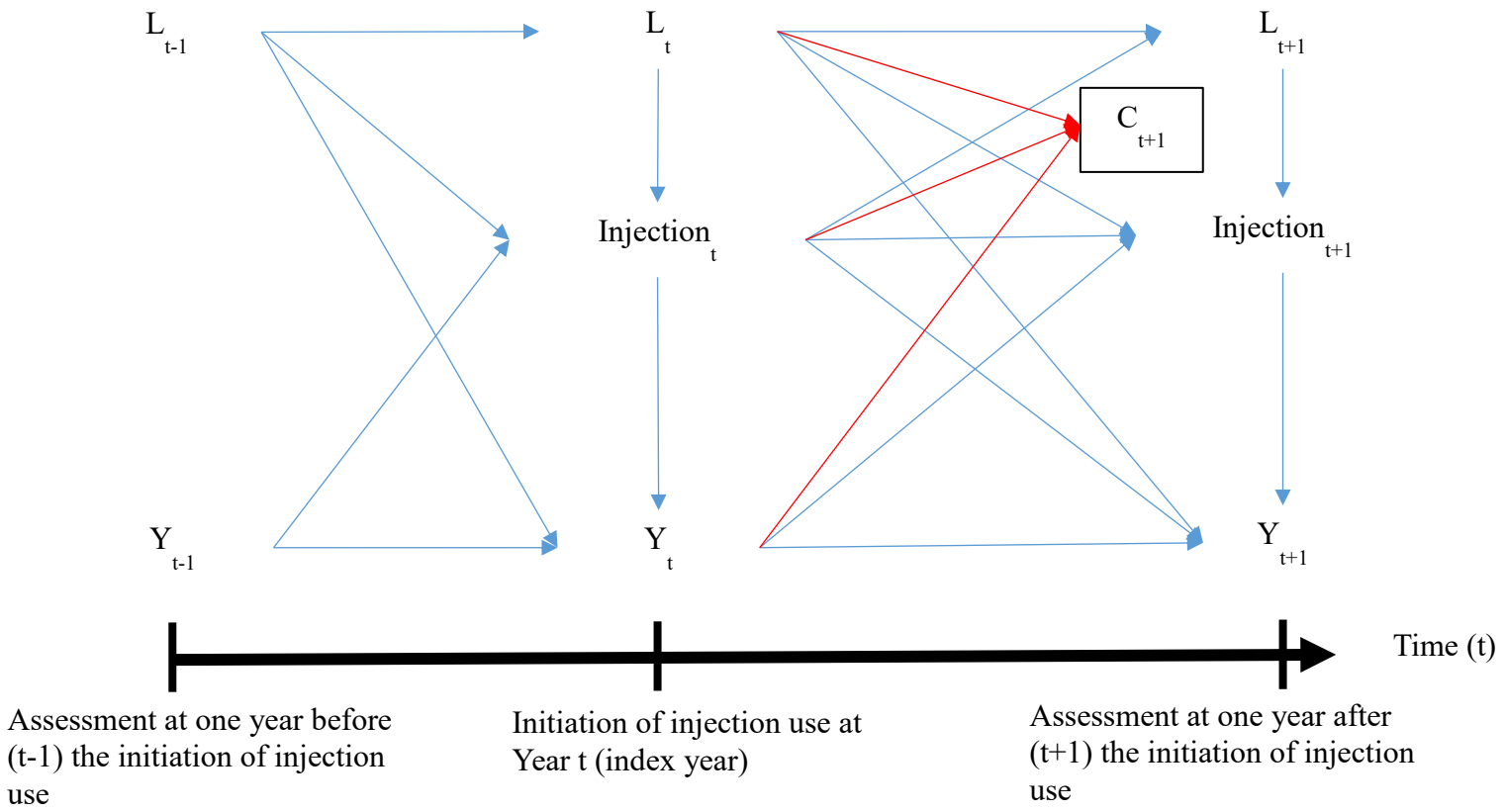


Table 3.1: Sociodemographic and clinical characteristics among participants with radiographically confirmed knee OA by use of injections.

Characteristics	Baseline[a]			Index year			One year after injection initiation		
	CO	HA	Non-user	CO	HA	Non-user	CO	HA	Non-user
Total n*	318	94	576	318	94	576	257	66	559
Proportions relative to baseline	100	100	100	100	100	100	80.1	70.2	97.0
Injection use (n)	0	0	0	318	94	0	63	15	0
	Mean and Percentage								
Mean age (years, (SD))	66.9 (9.2)	65.0 (8.6)	64.0 (9.3)	67.9 (9.2)	66.0 (8.6)	65.0 (9.3)	68.5 (9.1)	66.0 (8.8)	65.8 (9.4)
Women	64.2	44.7	55.6	64.2	44.7	55.6	65.4	43.9	55.5
Ethnicity/Race									
Non-Hispanic white	77.4	88.3	67.9	77.4	88.3	67.9	75.5	87.9	67.4
Non-Hispanic black	17.9	6.4	30.0	17.9	6.4	30.0	21.0	7.6	30.4
Other	4.7	5.3	2.1	4.7	5.3	2.1	3.5	4.6	2.2
Education									
High school or less	18.6	12.9	20.9	18.6	12.9	20.9	20.2	10.8	20.
Some college	27.1	23.7	26.7	27.1	23.7	26.7	28.8	24.6	27.1
College graduate	22.1	20.4	19.2	22.1	20.4	19.2	19.5	18.5	19.8
Graduate school	32.2	43.0	33.3	32.2	43.0	33.3	31.5	46.2	32.9
Income ($)									

<25,000	15.5	4.3	17.9	15.5	4.3	17.9	17.1	4.6	18.1
25,000 - 50,000	30.9	22.6	27.8	30.9	22.6	27.8	32.7	16.9	27.6
>50,000	53.6	73.1	54.3	53.6	73.1	54.3	50.2	78.5	54.3
K-L grade									
2	40.1	22.1	52.3	35.7	16.3	50.0	38.9	22.2	50.7
3	39.1	44.2	30.3	37.9	45.0	30.8	41.2	46.3	31.2
4	20.8	33.7	17.4	26.4	38.8	19.3	19.9	31.5	18.2
Symptom-related multi-joint OA	55.4	57.5	55.2	55.7	60.6	50.4	56.0	62.1	49.0
History of knee injury	40.6	57.5	49.3	47.2	58.5	50.2	49.4	57.6	51.2
History of knee surgery	28.6	48.9	34.6	32.1	54.3	35.8	35.8	50.0	35.2
Body Mass Index (kg/m^2)									
<25	12.9	14.9	13.2	12.9	16.0	13.0	12.1	15.2	14.5
25 - <30	40.6	31.9	39.7	40.6	30.9	37.7	40.5	31.8	36.3
≥30	46.5	53.2	47.1	46.5	53.2	49.3	47.5	53.0	49.2
Knee alignment									
Normal	20.5	18.9	18.4	17.4	15.9	15.9	18.4	11.9	16.0
Varus	39.3	40.0	41.9	42.9	44.3	45.9	38.2	44.1	46.5
Valgus	40.3	41.1	39.7	39.7	39.8	38.2	43.4	44.1	37.5
CES-D (>16)	11.0	8.9	11.0	14.3	12.2	13.3	13.2	12.9	12.1
Charlson Comorbidity Index									
0	65.4	70.2	65.7	62.0	71.0	63.4	56.7	78.5	63.0

1	18.2	22.3	19.3	19.9	20.4	19.9	21.7	15.4	19.9
≥2	16.4	7.5	15.1	18.0	8.6	16.7	21.7	6.2	17.1
	Mean (standard deviation)								
WOMAC Pain	5.0 (3.9)	5.1 (3.9)	4.9 (4.1)	6.2 (4.2)	5.7 (3.6)	4.4 (3.9)	5.6 (4.0)	5.2 (3.7)	4.2 (3.8)
WOMAC Stiffness	2.7 (1.7)	2.5(1.7)	2.5 (1.8)	2.8 (1.7)	2.8 (1.6)	2.4 (1.8)	2.7 (1.7)	2.5 (1.6)	2.2 (1.8)
WOMAC Physical Function	16.8 (13.0)	16.7 (11.8)	15.3 (13.0)	19.2 (13.2)	17.1 (11.3)	14.4 (12.8)	17.9 (13.2)	16.9 (11.1)	14.1 (12.8)
KOOS-QoL	54.6 (19.9)	52.2 (18.8)	56.0 (21.7)	48.3 (20.2)	49.0 (20.1)	57.4 (22.4)	50.6 (20.3)	50.5 (19.5)	58.4 (22.0)
SF-12 Physical Component Score	42.8 (9.6)	42.7 (9.2)	44.8 (9.7)	40.8 (9.4)	40.8 (9.9)	44.4 (10.1)	41.1 (9.5)	40.6 (9.3)	44.4 (10.0)
SF-12 Mental Component Score	54.1 (8.4)	55.4 (7.1)	53.8 (8.3)	54.1 (9.5)	55.4 (8.2)	53.4 (9.3)	53.6 (9.1)	55.1 (8.5)	53.2 (9.2)
Joint space width (mm)	4.6 (1.9)	4.2 (2.2)	5.1 (1.6)	4.4 (1.9)	4.0 (2.1)	4.9 (1.7)	4.4 (1.9)	4.3 (2.1)	4.9 (1.6)

Abbreviations: SD, standard deviation; K-L grade, The Kellgren–Lawrence grade; CES-D, Centers for Epidemiologic Studies Depression Scale; WOMAC, The Western Ontario and McMaster Universities Osteoarthritis Index; KOOS-QoL, Knee injury and Osteoarthritis Outcome Score Quality of life subscale.

[a] One year before the index year.

* Number of participants with missing information:
At baseline: education (4), income (3), KL grade (114), body mass index (1) knee alignment (99), CES-D (63), Charlson Comorbidity Index (5), WOMAC Pain (29), WOMAC Stiffness (27), WOMAC Physical function (3), KOOS-QoL (27), SF-12 Physical Component Score (70), SF-12 Mental Component Score (70), joint space width (234).

At index year: education (4), income (3), KL grade (153), knee alignment (123), CES-D (72), Charlson Comorbidity Index (5),WOMAC Pain (25), WOMAC Stiffness (27), WOMAC Physical function (41), KOOS-QoL (27), SF-12 Physical Component Score (93), SF-12 Mental Component Score (93), joint space width (300).

One year after injection initiation: education (3), income (2), KL grade (145), knee alignment (120), CES-D (80), Charlson Comorbidity Index (6), WOMAC Pain (56), WOMAC Stiffness (56),WOMAC Physical function (71), KOOS-QoL (56), SF-12 Physical Component Score (93), SF-12 Mental Component Score (93), joint space width (292).

Table 3.2: Concomitant use of medications and supplements among persons with radiographically confirmed knee OA by the use of injections.

	Baseline			Index year			One year after injection initiation		
	CO	HA	Non-user	CO	HA	Non-user	CO	HA	Non-user
	Percentage								
Use of analgesics									
Acetaminophen	23.6	21.3	16.2	22.0	19.2	12.3	21.4	18.2	14.5
NSAIDs[a]	35.2	39.4	31.5	37.3	42.6	28.3	38.4	40.9	27.7
COX-2 inhibitors	10.4	11.7	4.2	7.2	12.8	3.1	7.8	7.6	3.4
Opioids	9.4	8.5	8.7	12.9	11.7	6.9	14.8	12.1	8.1
Any use of analgesics	55.0	56.4	44.4	58.5	63.8	39.6	57.6	57.6	38.6
3+	4.7	8.5	1.9	3.5	5.3	1.7	5.8	6.1	2.0
2+	17.9	14.9	12.7	17.3	17.0	8.0	16.3	13.6	10.9
Use of supplements									
Glucosamine	30.8	48.9	27.1	34.6	44.7	25.9	26.9	36.4	21.8
Chondroitin sulfate	27.4	44.7	24.8	30.5	41.5	22.9	24.5	30.3	19.5
Methylsulfonylmethane	8.2	14.9	6.9	10.1	20.2	8.3	9.0	19.7	9.3
S-adenosylmethionine	0.3	1.1	0.4	0.9	0	0.5	0.8	0	0.5
Other current prescribed medications									
Doxycycline	0	0	0.5	0.3	0	0	0	0	0.2
Vitamin D	0.3	3.2	0.5	0.6	5.3	2.2	2.8	3.0	2.5
Medications for osteoporosis	11.0	7.5	8.0	8.8	7.5	7.1	7.4	4.6	6.1

Abbreviations: NSAIDs, nonsteroidal anti-inflammatory drugs.

[a] Including self-reported over-the-counter use and current prescriptions such as Aspirin, Ibuprofen, and Salicylate.

* Number of participants with missing information:
At baseline: NSAIDs (29), Doxycycline (27), Vitamin D (27).
At index year: NSAIDs (31), Doxycycline (26), Vitamin D (31).

One year after injection initiation: NSAIDs (58), Doxycycline (40), Vitamin D (43).

Table 3.3: Estimated effects of injection use compared with non-users on symptoms among persons with radiographically confirmed knee OA.

	Use of CO (β* coefficient (95% CI))	Use of HA (β* coefficient (95% CI))
WOMAC pain subscale		
Crude[a]	1.17 (0.80 to 1.54)	0.77 (0.17 to 1.38)
Baseline covariates[b]	3.49 (-2.36 to 9.33)	0.58 (0.07 to 1.09)
Baseline plus time-varying covariates[c]	2.11 (0.70 to 3.53)	0.40 (-0.30 to 1.10)
Marginal structural model[d]	1.24 (0.82 to 1.66)	0.50 (-0.11 to 1.11)
Sensitivity analysis	0.51 (-0.31 to 1.32)	0.42 (-1.44 to 2.29)
WOMAC stiffness subscale		
Crude[a]	0.23 (0.04 to 0.43)	0.28 (-0.03 to 0.59)
Baseline covariates[b]	-0.55 (-1.27 to 0.17)	-2.10 (-2.35 to -1.85)
Baseline plus time-varying covariates[c]	0.05 (-1.08 to 1.18)	0.17 (-0.17 to 0.51)
Marginal structural model[d]	0.30 (0.10 to 0.49)	-0.07 (-0.38 to 0.24)
Sensitivity analysis	0.14 (-0.21 to 0.50)	-0.43 (-1.54 to 0.67)
WOMAC physical function subscale		
Crude[a]	2.60 (1.36 to 3.84)	1.21 (-0.70 to 3.12)
Baseline covariates[b]	-1.73 (-9.50 to 6.05)	1.16 (-0.50 to 2.83)
Baseline plus time-varying covariates[c]	0.06 (-0.51 to 0.62)	-0.37 (-2.45 to 1.71)
Marginal structural model[d]	2.62 (0.94 to 4.29)	0.49 (-1.34 to 2.32)
Sensitivity analysis	1.05 (-1.23 to 3.33)	3.84 (-3.72 to 11.40)

Abbreviations: CO, corticosteroid; CI, confidence interval; GEE, generalized estimating equations; HA, hyaluronic acid; WOMAC, The Western Ontario and McMaster Universities Arthritis Index.

* A negative β coefficient indicates improvement for the Western WOMAC subscales and worsening for the joint space width.

[a] Values derived from GEE models used an unstructured correlation matrix.

[b] Models were adjusted for baseline characteristics including age, sex, race/ethnicity, education, income, Kellgren-Lawrence grade, body mass index, history of knee injury, history of knee surgery, Short Form 12 physical and mental component summary scores, WOMAC subscales, and use of analgesics and glucosamine.

[c] In addition to baseline covariates, time-varying confounders including Kellgren-Lawrence grade, history of knee injury and surgery, WOMAC subscales, and Short Form 12 physical and mental component summary scores measured at the same visit as injection use were also adjusted.

[d] Inverse probability-weighted analyses with truncated weights were used.

Supplementary Table 3.1: Model specifications for the construction of inverse probability weights at the index visit.

Specification	Description[a]	Estimated weights					
		Mean	SD	Median	IQR	Minimum	Maximum
1	Numerator includes linear terms for baseline WOMAC subscales, SF12 PCS &MCS, and KOOS QoL in addition to baseline categorical variables such as KL grade, injury & surgery Hx, and BMI Denominator also includes WOMAC subscales, KL grade, and KOOS QoL at the index visit	1.02	0.64	0.89	0.53	0.15	6.69
2	Numerator and denominator are as in specification 1, but without KL grade and	1.03	0.52	0.93	0.48	0.10	4.02

	KOOS QoL at the index visit in denominator						
3	Numerator and denominator are as in specification 2, plus adding WOMAC subscale*WOMAC subscale at the index visit in denominator	1.00	0.54	0.88	0.46	0.18	3.96
4	Numerator and denominator are as in specification 3, but replace linear terms for time-varying WOMAC subscales with categories	0.96	0.24	0.93	0.19	0.50	2.37

Abbreviations: SD, standard deviation; IQR, interquartile range

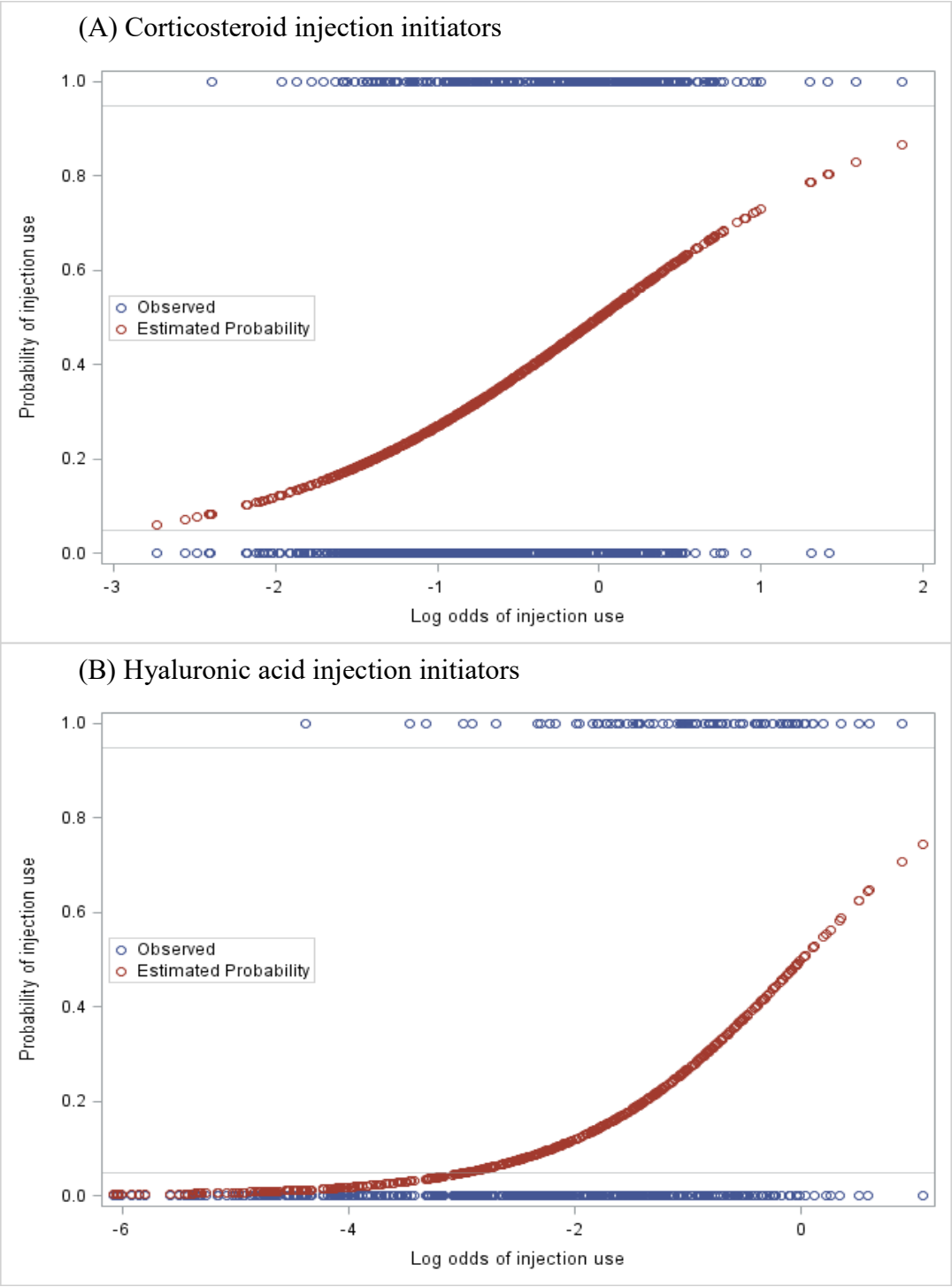
[a] All models include age, sex, race/ethnicity, education, and household income at enrollment.

Supplementary Table 3.2: Sociodemographic and clinical factors among people with radiographically confirmed knee OA by censoring mechanisms.

Characteristics	Death (n=3)	Total knee replacement (n=52)	Switching (n=34)	Non-censored (n=323)
	Mean and Percentage			
HA initiator	33.3	32.7	29.4	20.4
Mean age (years, (SD))	65.7 (6.7)	68.8 (8.9)	65.4 (9.9)	66.2 (9.0)
Women	33.3	48.1	67.7	61.0
Ethnicity/Race				
Non-Hispanic White	100.0	90.4	79.4	78.0
Non-Hispanic Black	0	3.9	5.9	18.3
Other	0	5.8	14.7	3.7
Education				
High school or less	0	17.6	8.8	18.3
Some college	0	19.6	23.5	28.0
College graduate	0	31.4	32.4	19.3
Graduate school	100.0	31.4	35.3	34.5
Income ($)				
<25,000	0	7.8	5.9	14.6
25,000 - 50,000	0	27.5	29.4	29.8
>50,000	100.0	64.7	64.7	55.6
KL grade				
2	0	4.1	41.9	41.1
3	33.3	42.9	25.8	41.4
4	66.7	53.1	32.3	17.5
Symptom-related multi-joint OA	33.3	59.6	50.0	56.0
History of knee injury	33.3	57.7	44.1	42.4
History of knee surgery	33.3	38.5	32.4	32.5
Body Mass Index (kg/m^2)				
<25	0	17.3	17.7	12.4
25 - <30	33.3	42.3	44.1	37.5
≥30	66.7	40.4	38.2	50.2
Knee alignment				

Normal	0	17.7	28.1	19.9
Varus	66.7	54.9	34.4	37.0
Valgus	33.3	27.5	37.5	43.1
CES-D (>16)	100.0	6.1	6.5	11.7
Charlson Comorbidity Index				
0	33.3	63.5	70.6	66.9
1	33.3	19.2	17.7	19.2
≥2	33.3	17.3	11.8	13.9
	Mean (standard deviation)			
WOMAC Pain	2.0 (3.5)	5.5 (3.1)	5.2 (3.5)	5.0 (4.0)
WOMAC Stiffness	1.0 (1.7)	2.7 (1.7)	2.8 (1.5)	2.6 (1.7)
WOMAC Physical Function	7.8(12.6)	19.4 (11.2)	17.4 (11.4)	16.4 (13.0)
KOOS-QoL	66.7 (20.1)	48.7 (18.5)	52.7 (14.4)	54.9 (20.2)
SF-12 Physical Component Score	30.6 (12.1)	41.6 (9.0)	40.6 (7.8)	43.3 (9.6)
SF-12 Mental Component Score	59.9 (4.8)	56.6 (7.0)	56.5 (6.8)	53.8 (8.4)
Joint space width (mm)	4.5 (2.1)	2.7 (1.9)	4.1 (1.8)	4.8 (1.9)

Supplementary Figure 3.1: Graphical representation of evaluation the experimental treatment assumption.



CHAPTER IV

MISSING DATA IN MARGINAL STRUCTURAL MODELS: A PLASMODE SIMULATION STUDY COMPARING MULTIPLE IMPUTATION AND INVERSE PROBABILITY WEIGHTING

Abstract

Objective: The use of marginal structural models (MSMs) to estimate unbiased causal effects in the presence of time-varying confounding has increased in pharmacoepidemiologic studies. Longitudinal studies are prone to missing data, however, recommendations for missing data techniques used in MSMs are contradictory. We compared the validity and precision of MSM estimates using multiple imputation (MI), inverse probability weighting (IPW), and complete case analysis (CC) in the presence of missing data on a time-independent, a time-varying , or both confounders.

Methods: Datasets were generated using the plasmode simulation framework which preserved underlying associations without modifying the exposure or other covariates. We constructed the cohort sub-study using data from the Osteoarthritis Initiative which estimated the marginal causal effect of intra-articular injection use (binary treatment) on one-year symptom change (continuous variable). We simulated scenarios through introducing three missing data mechanisms: 1) missing completely at random (MCAR); 2) missing at random (MAR); and 3) missing not at random (MNAR). We also varied the proportion of missingness (10%, 30%, and 50%) and whether the confounder subject to missing data was fixed to the measurement at baseline or time-varying. Overall, 81 simulated scenarios were generated. Performance of methods was compared using relative bias, mean squared error (MSE) of the estimates of interest, and empirical power.

Results: Regardless of scenarios, estimates of relative bias using CC and IPW were similar (relative bias: CC, -0.18% to 3.70 %; IPW, -2.33% to 4.64%), with estimates exceeding 15% for scenarios with 50% MNAR on either a time-independent or a time-

varying confounder, or both. From MI procedures, relative bias estimates ranged from -1.88% to 4.24%. For most scenarios, estimates using MI had smaller MSE (range: 0.013 to 0.024) than IPW (range: 0.027 to 0.22) and CC (range: 0.027 to 0.215). While the MI procedure maintained empirical power across scenarios, the power decreased using CC and IPW given increasing proportion of missingess across different types of miss data. **Conclusions:** Compared to CC and IPW, MI produced less biased estimates with better precision over a range of type and extent of missingness. MI may confer an advantage over IPW in MSMs applications.

Introduction

Marginal structural models (MSMs) using inverse-probability-of-treatment–weighted estimation (IPTW) have been proposed to estimate unbiased causal effects when time-varying confounding is a concern.[30,41,43] Briefly, this technique creates a pseudo-population in which bias has been eliminated by simultaneously adjusting for time-varying confounding (without blocking indirect effects from former exposures[30] and avoiding collider-stratification bias[104]) and selection bias owing to informative censoring.[41] Methodologic research has provided guidance on appropriate weight construction,[40] how best to build the outcome models,[91,105] and what assumptions are needed to identify consistent causal effects.[106,107] Little guidance, however, exists regarding how to handling missing data in the longitudinal data used by MSMs.

To our knowledge, there are two studies comparing methods to handle missing information of covariates in the setting of MSMs.[92,108] Both of these studies limited their evaluation to situations with missing data in time-varying confounders. The guidance provided from these studies appears to be contradictory. While one study recommends that multiple imputation (MI) is superior to inverse probability weighting (IPW) where missing data are strongly predicted by the available data,[92] the other suggests that IPW performs better than MI.[108] Whether the differences in the study findings could be due to the effect of the confounder on the outcome[108] or some artifact of the simulations is unknown. Specific guidance regarding which missing data techniques should be used for scenarios of associations between confounders and exposure-outcome relationship is still unclear. Further, the extent of biases resulting from applying commonly used missing

data techniques for time-independent confounding in MSMs settings has yet to be explored.

Given the increase in applying MSMs,[109] the objective of this study was to compare validity and precision derived from commonly used missing data approaches (i.e., MI and IPW) through simulation studies in the presence of either missing time-independent or time-varying confounders in cohort studies using MSMs. We evaluated the performance of these missing data techniques using plasmode simulation which generates simulated data with preserved underlying association among observed covariates and exposure data from an empirical cohort study.[44]

Methods

The University of Massachusetts Institutional Review Board considered this study exempt since we used publicly available data to construct the cohort for simulation.

Empirical data

To create simulated datasets using a plasmode simulation framework, we used data from a previously published retrospective cohort study using publicly available data from Osteoarthritis Initiative (OAI).[110] OAI is a multi-center (i.e., Baltimore, MD; Columbus, OH; Pittsburgh, PA; and Pawtucket, RI), longitudinal, prospective observational study examining not only the development and progression of knee osteoarthritis (OA) but also the effectiveness of disease-modifying therapies. The original cohort includes 4,796 men and women ages 45-79 enrolled between February 2004 and

May 2006 and followed for 9 years (http://oai.epi-ucsf.org/). The OAI collects not only clinical assessments such as symptoms and function of the knee, quality of life, physical performance, health behaviors, medications, and supplements of all participants but also biologic specimens including blood and urine for up to 9 years of follow up. In addition, radiographic and magnetic resonance imaging (MRI) studies were collected. Data on clinical, joint status, and risk factors for the progression and development of knee OA were collected at baseline and the yearly follow-up clinic visits. In the published study, the use of intra-articular injection was evaluated using a "new-user design" among participants with knee OA.[110] This is an ideal setting to have when applying plasmode simulation.[44] Using the sample derived in the study, we used participants initiating corticosteroid injection and non-users to construct the cohort. Only complete cases that provided all of the information were used for generating simulated datasets. This resulted in a cohort consisting of 646 participants.

Data generation: plasmode simulation

We simulated datasets under the framework of plasmode simulation.[44] The causal diagrams in Figure 4.1 depict the causal relationship among the injection use, outcome, covariates, and missing data mechanisms. While L_0 indicates a set of all pre-specified confounders from baseline, L_t represents time-varying confounders measured at time t. C' indicates a set of potential confounders in addition to the pre-specified confounders that are selected in the data generating mechanism. Some arrows are omitted due to the simplicity for presentation. Missing data mechanisms including missing completely at

random (MCAR), missing at random (MAR), and missing not at random (MNAR) are represented by M.

The first step to generate data under the plasmode simulation framework was to estimate a linear model with the observed study outcome as a function of the exposure status, baseline covariates, and a subset of the potential confounders using data from the constructed cohort.[110] This estimated model was then used as the basis of the outcome-generating model later in the data generation process. Next, we sampled with replacement among those exposed and unexposed participants from the constructed cohort to achieve the desired sample size to compare findings with previous studies.[92,108] Since the information on covariates and exposure data for each participant was preserved without modification, associations among these variables remained complete in the sampled population. We then used the outcome-generating model described above to generate outcome values through replacing a pre-identified treatment effect on the estimated coefficient exposure.

Table 4.1 shows values of parameters for data generation. We used 1.2 as the pre-identified treatment effect since this value was considered as the threshold for achieving minimal clinically important difference on symptoms change (i.e., knee pain in the present study) in patients with knee OA.[61–64] Values of all other model coefficients remained unchanged and were used to generate the value of outcome status for each patient in the simulated data. As such, the outcome Y (i.e., one-year change in knee pain) was normally distributed with the expected values generated from the outcome-

generating model dependent on the pre-identified treatment effect for injection use and values of covariates:

$$Y = \beta_0 + \beta_1(Injection_t) + \beta_2 L_0 + \beta_3 L_t + \beta_4 C' + \varepsilon$$

This process was then repeated 1000 times to yield 1000 simulated datasets for each scenario of the simulation.

Missing data mechanisms and missing information

Based on our experience with missing data in the OAI database in previous research,[110] we selected the severity of knee OA status (Kellgren-Lawrence grade) measured at baseline as the time-independent confounder of interest. Knee pain measured at the visit of reporting injection use was used as time-varying confounder.[110] We sought to evaluate a range of missingness that previous studies used, where 10%, 30%, and 50% of the data were imposed as missing.[92,108]

Separate simulated datasets were generated for each missing data mechanism including MCAR, MAR, and MNAR.[45,46] Missing data is considered as MCAR when the probability of missing does not depend on the values of observed covariates. To impose MCAR in the simulated data sets, we randomly selected participants and forced the information on either the time-independent or time-varying confounder to missing. The proportion of participants randomly selected varied from 10%, 30%, or 50%. If the probability of missing depends on values of observed covariates, it is considered as MAR.[45,46] To achieve this, we assumed the probability of missing on the time-independent confounder was a joint function of observed covariates (i.e., age, sex, and

visits of reporting injection initiation) that were associated with the measure. For the time-varying confounder, we assumed the probability of missing was a joint function of observed covariates using information from age, sex, household income, and race/ethnicity. This process was accomplished using R package "simstudy".[111] We again selected participants to achieve the desired proportion of missingness and forced their information on the time-independent or time-varying confounder to missing in each simulated dataset.

MNAR, if it is informative, is when the probability of missing depends on values of unobserved covariates.[45,46] Using the empirical information from the constructed cohort, missing data distributions were imposed so that participants with more severe knee OA status (for time-independent confounder) and higher knee pain (for time-varying confounder) were more likely to be missing. Among all participants missing information on Kellgren-Lawrence grade, 30% were randomly selected from Kellgren-Lawrence grade 2, 50% from grade 3, and 20% from grade 4. For participants with knee pain greater than 4, 80% of them were set to be missing.

Inverse probability weighting

IPW is not a new approach to handling missing data and is particularly straightforward to use in MSMs settings.[47,48] It shares the similarity of weight building process for the inverse probability of observed treatment or censoring weights that are performed by MSMs in estimating causal treatment effects.[30] IPW proceeds by calculating the probability of having complete data for each individual in the study. Using

logistic regression models, contributions of each individual are weighted by the inverse probability of having complete data conditional on other relevant covariates.

In the missing data mechanisms of MCAR and MAR on the time-independent confounder, we modeled the weights of missingness proportional to the inverse probability of the value being observed using the logistic regression of M conditional on all other available information at baseline from L_0 (Figure 4.1). For MNAR, the missingess introduced was considered related to some unobserved covariates in the data. Since this information was not available, we estimated weights through further included other available information at baseline from C' which potentially were considered to be correlated with the missing variable. For missing information on the time-varying confounder, a similar approach was used except that now the value being missing was at time t. Therefore, we generated the weights of missingness proportional to the inverse probability of the value being observed at time t using a logistic model conditional on all relevant covariates including injection use, time of initiating injection use, and observed values from L_0, L_t, and/or C'.

Multiple imputation

In general, the approach of MI proceeds by generating *m* complete datasets where missing values in the incomplete observed data are imputed. Each of the *m* datasets is then analyzed using the same model and estimation method. The estimates from each of the *m* datasets are then combined to produce a single estimate that incorporates the usual sampling variability as well as the variability of the missing data.[49] We used the approach

that suggested including all available information (including the outcome) in the data to predict missing values for the time-independent and time-varying confounders.[112] The MI procedure was completed using R package "mice".[113] For each simulated dataset, five imputed datasets were generated for the missing information on the time-independent and time-varying confounders.[114]

Analytical approaches and evaluation of methods performance

Overall, the analytic approach for this study was carried out in four steps. First, we created simulated data sets using the plasmode simulation framework. Second, we introduced missing data mechanisms and scenarios to the simulated datasets. Third, we then applied IPW and MI to take into account the missing data in the analysis in addition to complete case analysis (CC). Lastly, weighted generalized estimating equation (GEE) linear models were fit using CC, IPW and MI to estimate the average causal effect in MSMs over increasingly problematic scenarios of missingness. In all analyses, we used stabilized weights to yield estimates with greater precision compared to the unstabilized weight.[30,40] The stabilized inverse probability treatment weight (IPTW) was:

$$SW_i^t(t) = \prod \frac{\Pr[INJECTIONS_{it}=injections_{it}|\overline{INJECTIONS}_{it}=\overline{injections}_{it}CON=con_{it}]}{\Pr[INJECTIONS_{it}=injections_{it}|\overline{INJECTIONS}_{i(t-1)}=\overline{injections}_{i(t-1)}CON=con_{it}]},$$

While the numerator was the conditional probability of receiving observed treatment given baseline confounders, the denominator was the conditional probability of receiving observed treatment given time-varying confounders in addition to baseline confounders. For the application of IPW, the final weights incorporated in the GEE linear models were calculated as the product of stabilized treatment weights and censoring (missing) weights

which used a similar approach as described above. For the application of MI, no missing weights were needed. However, the analysis was conducted for each imputed dataset. Since multiply imputed datasets were used, we used *mi.meld* function in R package "Amelia".[115] Results generated from the function reflected the average estimates with standard errors that accounted for average uncertainty and disagreement in the estimated values across the models.[116]

The performance of methods in each scenario was compared using relative bias, mean squared error (MSE), and empirical power of the estimates of interest. Relative bias was calculated as $\left(\frac{\hat{\beta}-\beta_{truth}}{\beta_{truth}}\right) * 100\%$. MSE was calculated combining bias and true variance (bias^2 + standard error$\left(\hat{\beta}\right)^2$), where standard error$\left(\hat{\beta}\right)$ was calculated as $\sqrt{\frac{1}{B-1}\sum_{i=1}^{B}(\hat{\beta}_i - \bar{\beta})^2}$. Empirical power was defined as (1 – empirical type II error) given that empirical type II error was calculated as $\frac{\#(p>0.05)}{B}$. While MSE was used to evaluate the accuracy of estimates, empirical power displayed the percentage of time that it will reject a false null hypothesis.

Results

We overall simulated 81 scenarios with parameter values varied on missing mechanisms (MCAR, MAR, and MNAR), percentages of missing (10%, 30%, and 50%), type of confounders (time-independent, time-varying, either or both), and analytical approaches (CC, IPW, and MI).

Table 4.2 shows results from missing values in a time-independent confounder. Estimates of relative bias using three missing data approach show similar results while the range of IPW is slightly larger (range: CC, -0.18% to 1.61%; MI, -1.16% to -1.88%; IPW, -2.33% to 2.01%). Regardless of missing data mechanisms, the performance measured by MSE using MI is consistent and smaller across scenarios while CC and IPW show a trend of increased estimates given the increasing proportion of missingness. Similarly, while the performance of empirical power using MI is maintained across scenarios, CC and IPW show a trend of decreased power given the increasing proportion of missingness.

Table 4.3 shows results from missing values in a time-varying confounder. Similar to the performance of missing on the time-independent confounder, estimates of relative bias using three missing data approaches show similar results regardless of scenarios (range: CC, -0.04% to 3.70%; MI, -0.41% to -4.16%; IPW, -0.54% to 4.64%). For the three approaches, the largest bias happen when there is a 50% of missingess in MNAR (CC: 3.70%; MI: 4.16%; IPW: 4.64%). While MI gives a smaller MSE and maintains empirical power across all scenarios, CC and IPW both show a trend of increased MSE and decreased power given the increasing proportion of missingness.

Results from missing values in either a time-independent, time-varying, or both confounders is displayed in Table 4.4. While MI shows a relatively smaller bias across scenarios (range: -0.55% to 4.24%), CC and IPW give larger estimates of relative bias (range: CC, -0.69% to 15.37%; IPW -0.50% to 15.74%). Regardless of missing data mechanisms, the relative bias increases to over 10% using CC or IPW when missing

information on the confounders reaches to 50%, but not with MI. While MI gives a consistently smaller MSE and maintains empirical power across all scenarios, CC and IPW both show a trend of increased MSE and decreased power given the increasing proportion of missingness. Similar to the estimates of relative bias, the worst scenario happens when missing information on the confounders reaches 50% regardless of imposed missing data mechanisms.

Discussion

Using a plasmode simulation framework with imposed missing data on time-independent and/or time-varying confounders, our simulation study demonstrated the performance of commonly used missing data approaches including CC, IPW, and MI in the context of MSMs analyses. Compared to CC and IPW, MI consistently produced less biased marginal estimates with better precision regardless of missing data mechanisms and the extent of missingness. In addition, while empirical power performed by the MI procedure was consistent across scenarios, CC and IPW both displayed a trend of decreased power given the increasing proportion of missingness.

Our findings are aligned with one of the previous studies showing that the MI procedure provided less biased marginal estimates and noticeably less variability given imposed missing data mechanisms and the extent of missingness.[92] In the implementation of the MI technique, baseline information of the confounding variables was provided to impute the missing values in both simulation settings. In addition, we also included the outcome variable in the prediction model for the time-independent and time-varying

confounders.[112] However, it is not clear if such approaches were used in the other study.[108] Since the purpose of MI is to model the missing values, our approaches may give additional advantages of MI over IPW if the information provided was predictive of the missing variables.

On the other hand, the goal of the IPW technique focuses on predicting missing data mechanisms.[117] There may well be some situations where IPW outperforms MI. One plausible example includes situations lacking strong predictors of the missing values or the missingness mechanism is well-understood. In our setting, we used only fully observed covariates to model the missingness. Using this approach, it can ensure that all the data needed to fit the missingness model and to estimate individuals' weights are observed.[118] However, when data are 'monotone missing', non-fully observed predictors can be included to model the missingness. In the previous study, missed visits was also discussed for one of the missing-data scenarios.[108] As scenarios resulted in large numbers of consecutive missing values due to missed visits (considered as monotone missing), the IPW may yield more satisfactory performance relative to MI. Yet, whether this scenario explains the differences among these studies remains unclear owing to lack of transparency in reporting. A brief summary to compare IPW and MI approach is displayed in Supplementary Table 4.1.

In addition to comparing method performance using bias and precision, our study also demonstrated using MI procedure was able to maintain empirical power consistently across scenarios whereas the power decreased using CC and IPW given increasing proportion of missingess across different types of miss data. Typically, a strength of a

MSMs analysis through applying IPTW using observational studies is that it partially mimics a sequentially randomized trial design and thus allows estimation of the marginal treatment effect.[119] However, similar to findings from CC, we noticed that the statistical power to detect a pre-identified non-null treatment effect using IPW was decreased compared to MI in our simulated scenarios. This is an important issue since statistical power supplies both investigators and readers with information to help interpret potentially null conclusions using such an analytic approach. Our findings may thus provide additional perspectives regarding the choice of analytic methods when dealing with missing data under MSMs settings.

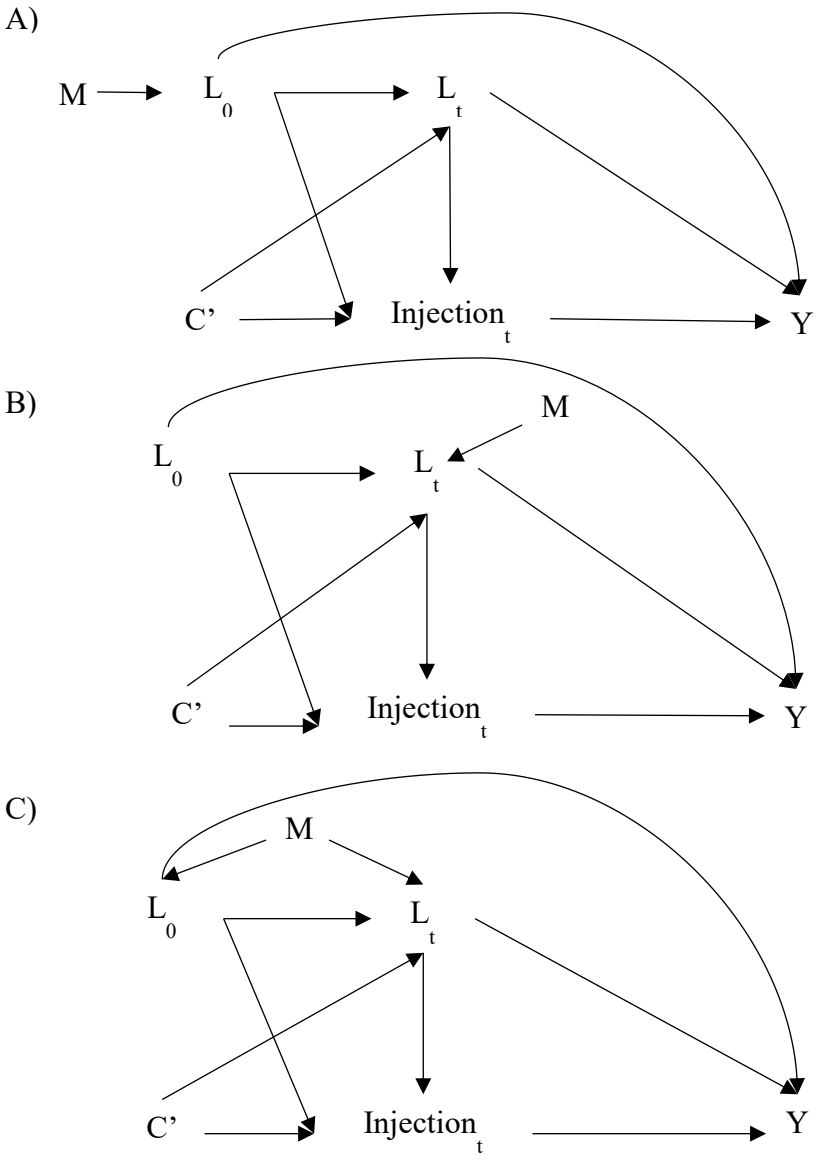
Several limitations must be acknowledged. First, we considered a simplified context which only used a single-time interval setting for data-generating scenarios related to treatment use and outcome. For situations involving more time intervals, the mechanisms regarding continued treatment use become more complicated. Information on time-varying confounders affecting treatment use is needed to correctly model the complex mechanism of treatment use.[120] Second, the model used to generate outcome values was based on a set of pre-defined covariates from a cohort sub-study using data from OAI. It is possible that outcomes generated from a much larger set of factors (e.g., using data from claims datasets) may be different due to the influence of both measured and unmeasured covariates.[44] Therefore, the performance of methods observed based on this cohort sub-study may not extend to other studies simulated from claims data. Third, while missing data can also occur in the exposure of interest,[121] our simulation study only

introduced missingness on confounders. Whether our findings extend to different types of variables, including the outcome and exposure of interest, needs to be explored.

Despite the limitations, strengths of our study include the use of plasmode simulation in which the covariate data and associations among covariates remain unchanged with the advantage to manipulate other parameters such as strength of confounders and exposure of interest.[44] In addition, using the cohort sub-study from OAI to construct the cohort for simulating datasets provided measurements not only on clinical assessments (e.g., symptoms and joint status) but also risk factors and concurrent medication use for the progression of knee OA. Given the pre-identified potential confounders and treatment effect,[110] data-generating scenarios in our study may perform better than approaches using ordinary methods or healthcare claims which may not capture important features of this population. While previous studies focused on time-varying confounders,[92,108] our study also assessed the methods performance using baseline confounders and mixed scenarios which provided a more comprehensive and realistic evaluation under MSMs analyses.

In conclusion, with a range of simulated scenarios under MSMs analyses, our simulation study demonstrated that MI generally produced less biased estimates with better precision over a range of missing data mechanisms and extent of missingness. Moreover, the MI procedure maintained empirical power across scenarios, the power decreased using CC and IPW given increasing proportion of missingness across different types of missing data. Under simple yet realistically constructed scenarios, MI may confer an advantage over IPW in MSMs applications.

Figure 4.1: Causal diagrams depicting relationship in simulated datasets.



LEGEND: Data generation and missing mechanisms for simulation studies. L_0 indicates a set of all pre-specified confounders measured at baseline, L_t represents time-varying confounders of the Injection-Y outcome association measured at time t. C' indicates a set of potential confounders in addition to the pre-specified confounders that are selected in the data generating mechanism. M indicates missing mechanisms. A) This causal diagram shows that missingness (M) is present in the baseline confounder L_0; B) This causal diagram shows missingness (M) in the time-varying confounder L_t; and C) This causal diagram indicates that missingess (M) is present in either or both baseline or time-varying confounders.

Table 4.1: Values of parameters for data generation using plasmode simulation framework and Osteoarthritis Initiative data.

Parameter	Meaning	Value
N	Sample size	500
B	Total simulations	1000
M	Number of imputed data sets	5
Missing data mechanism for a time-independent or a time-varying confounder (M)*		
MCAR	Missing complete at random	10%, 30%, 50%
MAR¶	Missing at random	10%, 30%, 50%
MNAR§	Missing not at random	10%, 30%, 50%
$Pr(Injection_t = 1)$	Probability of initiating intra-articular corticosteroid injection at time t	Empirical distribution from the constructed cohort: ~33.0%
β	True simulated effect: the difference in one year change in knee pain (Y) of initiating injection use	1.2
Y	Predicted values of one year change in knee pain depend on observed values of exposure status and confounders	$Y = \beta_0 + \beta_1(Injection_t) + \beta_2 L_0 + \beta_3 L_t + \beta_4 C' + \varepsilon$
E	Error term	$\sim N(0,1)$
L	Confounders including predefined important baseline confounders (L_0), time-varying confounders (L_1), and other pre-specified confounders (C')	Empirical distribution from the constructed cohort

* We introduced missing data within the context of a data source given the complete information from the measured covariates. The severity of knee OA status (Kellgren-Lawrence grade) measured at baseline was used as the time-independent confounder. Knee pain measured at the visit of reporting injection use was used as time-varying confounder.

¶ The probability of missing data for the time-independent confounder (Kellgren-Lawrence grade) or time-varying confounder (knee pain) was a joint function of observed

covariates (i.e., age, sex, and time of initiating injection use) associated with each variable.

§ For the time-independent confounder (Kellgren-Lawrence grade), we imposed the following missing data distributions. Among all participants missing information on Kellgren-Lawrence grade, 30% were randomly selected from Kellgren-Lawrence grade 2, 50% from grade 3, and 20% from grade 4. As such, participants with more severe knee OA status were more likely to be missing given the original distribution from the constructed cohort.

For the time-varying confounder (knee pain score), among all participants missing information on knee pain scores, 80% of them were randomly selected from pain scores greater than 4.

Table 4.2: Comparison of percent bias, mean square error, and empirical power for methods to handle missing data (CC, MI, IPW) under various mechanisms for missing data and extent of missing data: the case of a missing data in a time-independent confounder.

Missing mechanism and % missing data	Bias (%)			MSE			Empirical power		
	CC	MI	IPW	CC	MI	IPW	CC	MI	IPW
MCAR									
10%	-1.19	-1.58	-1.01	0.027	0.015	0.027	99.1	99.9	99.3
30%	-1.68	-1.43	-1.74	0.036	0.014	0.042	96.1	99.6	94.4
50%	1.61	-1.56	2.01	0.053	0.014	0.064	85.4	99.9	83.7
MAR									
10%	-1.45	-1.46	-1.32	0.028	0.015	0.029	98.9	99.8	98.9
30%	-0.18	-1.16	-0.47	0.035	0.015	0.038	96.7	99.8	95.8
50%	-1.21	-1.83	-0.67	0.054	0.015	0.063	82.1	99.5	80.7
MNAR									
10%	-1.54	-1.88	-1.84	0.028	0.014	0.03	98.9	99.9	98.6
30%	-1.92	-1.20	-3.29	0.039	0.014	0.045	94.0	99.9	92.5
50%	0.74	-1.72	-2.33	0.067	0.015	0.083	79.4	99.7	71.0

Abbreviations: CC, complete case; IPW, inverse probability weighting; MAR; missing at random; MCAR; missing completely at random; MNAR, missing not at random; MI, multiple imputation; MSE, mean squared error.

Table 4.3: Comparison of percent bias, mean square error, and empirical power for methods to handle missing data (CC, MI, IPW) under various mechanisms for missing data and extent of missing data: the case of missing data in a time-varying confounder.

Missing mechanism and % missing data	Bias (%)			MSE			Empirical power		
	CC	MI	IPW	CC	MI	IPW	CC	MI	IPW
MCAR									
10%	-1.13	-1.59	-1.11	0.028	0.014	0.028	99.0	99.8	99.0
30%	-1.51	-1.01	-1.38	0.04	0.014	0.04	95.3	99.9	95.2
50%	1.93	-0.84	1.98	0.059	0.015	0.061	86.5	99.4	86.0
MAR									
10%	-0.96	-1.82	-0.94	0.027	0.015	0.027	98.8	99.7	98.8
30%	-0.61	-1.02	-0.54	0.039	0.014	0.039	95.2	99.8	94.7
50%	1.36	-0.91	1.55	0.061	0.015	0.062	84.4	99.9	84.8
MNAR									
10%	-0.94	-1.44	-0.74	0.028	0.015	0.028	98.9	99.6	98.8
30%	-0.04	-0.41	0.50	0.037	0.014	0.064	97.5	99.8	97.2
50%	3.70	4.16	4.64	0.062	0.021	0.037	91.4	99.7	92.0

Abbreviations: CC, complete case; IPW, inverse probability weighting; MAR; missing at random; MCAR; missing completely at random; MNAR, missing not at random; MI, multiple imputation; MSE, mean squared error.

Table 4.4: Comparison of percent bias, mean square error, and empirical power for methods to handle missing data (CC, MI, IPW) under various mechanisms for missing data and extent of missing data: the case of missing data in either a time-independent confounder, time-varying confounder, or both.

Missing mechanism and % missing data	Bias (%)			MSE			Empirical power		
	CC	MI	IPW	CC	MI	IPW	CC	MI	IPW
MCAR									
10%	-1.82	-0.94	-1.82	0.029	0.014	0.029	98.4	99.7	98.2
30%	1.56	-1.12	1.50	0.066	0.013	0.067	83.6	100	82.5
50%	11.14	0.39	11.36	0.202	0.014	0.21	47.3	100	46.7
MAR									
10%	-1.14	-0.57	-1.01	0.03	0.013	0.03	98.0	99.9	97.9
30%	2.27	-0.78	2.33	0.062	0.015	0.063	84.8	99.6	84.3
50%	10.68	-0.81	10.51	0.198	0.014	0.204	47.8	99.7	47.0
MNAR									
10%	-0.69	-0.55	-0.50	0.028	0.014	0.028	98.4	99.9	98.6
30%	2.36	0.09	2.81	0.064	0.014	0.066	85.1	100	85.5
50%	15.37	4.24	15.74	0.215	0.024	0.22	59.0	99.8	59.1

Abbreviations: CC, complete case; IPW, inverse probability weighting; MAR; missing at random; MCAR; missing completely at random; MNAR, missing not at random; MI, multiple imputation; MSE, mean squared error.

Supplementary Table 4.1: A brief summary comparing approaches of IPW and MI.

IPW	MI
• Only use information on complete cases • Assume model for probability of complete cases ("missingness model") • Require data to be MCAR or MAR • Restriction to complete cases can lead to bias • If data are "monotone missing", non-fully observed predictors can be used	• Use data on all subjects • Assume model for joint distribution ("imputation model") • More efficient if some incomplete cases have some information • Generally more efficient • Can deal with MAR and MNAR

CHAPTER V

DISCUSSION AND CONCLUSIONS

The overall purpose of this dissertation was to examine the effectiveness of intra-articular injection use and changes in knee symptoms over time among patients with knee OA living in the community. CO and HA injections have different mechanisms of action and are the two primary types of intra-articular injections used in OA. Despite a large number of clinical trials addressing the safety and efficacy of intra-articular injection use, recommendations across clinical guidelines still remain inconclusive. More importantly, evidence about the effects of long-term use of these treatment modalities is still scarce. In this dissertation, we used advanced analytical methods such as MSMs and plasmode simulations to address methodological challenges including time-varying confounders and treatment initiation over time. Findings from this dissertation are summarized below.

In Aim 1, we first evaluated the complexity of treatment use of intra-articular injections after treatment initiation among persons with radiographic knee OA over time. We hypothesized that the patterns of injection use among persons with knee OA would be associated with more severe symptoms and/or changes in symptoms over time compared to those who only received one-time injections. This investigation yielded several findings. First, we found that a substantial proportion of knee OA patients who initiated intra-articular injections either switched their treatment or used it just once during the follow-up period, regardless of the initial therapy or actual change in symptoms. Second, approximately 1 in 5 participants initiating injections switched injection type. Third, while nearly 1 in 4 hyaluronic acid initiators switched to corticosteroid injections, we were surprised to find that almost 1 in 5 corticosteroid initiators switched over to hyaluronic acid injections. Given that HA is a newer and more

costly treatment for persons with more severe OA, the channeling effect into the hyaluronic injections was not clear. Last but not least, our findings did support the hypothesis that among initial users of CO injections, worse symptoms after the first injection were associated with lower odds of continued treatment use, while changes in knee symptoms had no such association.

With the information regarding pattern of injection use from Aim 1, we were able to examine the extent to which intra-articular injections relieve symptoms among persons with radiographic knee OA over time in Aim 2. For this we employed MSMs which applied weighting methods to adjust for time-varying confounding and censoring over the course of treatment. After carefully controlling for potential time-varying and time-independent confounders, we did not observe reduced symptoms associated with the initiation of CO or HA injections compared with non-users in 2 years of follow-up data. However, we did find that the power to detect the difference may be insufficient due to the missing information on some potentially important confounders. To address this issue, we used MI to conduct sensitivity analyses to account for missing values of covariates. Results from sensitivity analysis were qualitatively similar to our main findings.

For Aim 2, we recognized that the methodologic research on how to handle missing data in the longitudinal dataset used by MSMs is limited and conflicting. Thus in Aim 3, we set out to evaluate the performance of missing data techniques including CC, IPW, and MI within MSMs using a plasmode simulation framework. With imposed missing data on time-independent and/or time-varying confounders, we demonstrated MI

consistently produced less biased marginal estimates with better precision regardless of missing data mechanism and regardless of the extent of missingness compared to CC and IPW. In addition, while the MI procedure demonstrated a consistent performance measured by empirical power across scenarios, CC and IPW both displayed a trend of decreased power given the increasing proportion of missingness.

Limitations and strengths

While there are several important questions addressed in this dissertation, there are several limitations. First, there may be misclassification in reporting injection use. In OAI, treatment use was assessed annually. Given the questionnaires used (i.e., 6 months window before the annual visit to assess treatment use), we may miss the identification of new-users or continued users during the intervals of assessments. For instance, when participants were asked about their injection use, the time for their annual assessment may be out of the window to report their injection use. On the other hand, given that the knee injection is an invasive procedure and patients may be more likely to remember and report the use of injections, some participants may still report having injection use even when the time of receiving injection was out of assessment window. With these possible scenarios, it could be the non-differential misclassification given that these groups may have the same probability of being misclassified across all study subjects. If this misclassification was non-differential, the treatment effects could be toward the null.[122] Another potential for misclassification could be those "not continued" users may include patients who went on to total knee replacement (TKR). However, among those who did

not continue injection use, only a small portion of participants in this group had TKR after the initiation of injection use. After excluding those participants, results were comparable to the main findings.

Second, confounding by indication may arise when individuals have an "indication" for use of the drug even if the study population consists of subjects with the same disease (e. g., patients with OA). However, confounding by indication is not conceptually different from confounding by other factors, and the approaches to control for confounding by indication are the same such as matching, stratification, restriction, and multivariate adjustment.[123] Through applying IPTW, MSMs allowed us to account for changing values of the confounders over time and thus minimize the effect of confounding by indication. Moreover, in Aim 2, we further restricted participants who had symptomatic knee OA at baseline for non-users to minimize the potential impact of this issue.

Despite that we were able to adjust the indices to deal with the potential confounding by indication given the comprehensive measurement and information on the disease severity that might affect patients in seeking treatment in OAI,[123] we could not rule out the possibility that our findings may still be biased by unmeasured confounding (e.g., the patients' preference of physicians or specialists). In addition, the IPTW approach may be biased by a few observations with very high weights in the process of constructing appropriate weights. Violations of the positivity assumption are also possible due to model misspecification.[40] To address these concerns, we carefully constructed weights using an iterative process and graphically examined whether there

was an adequate degree of variation given observed and predicted injection use. We also truncated weights to minimize the potential impact of violating the positivity assumption.[40,124] In addition, residual confounding is still a possibility despite the comprehensive assessments of disease severity and concurrent treatment use in the OAI. Nevertheless, several contrasting models were used for illustration including results from sensitivity analyses that allowed readers to evaluate the observed associations within the context of any limitations of the data.

Third, despite an initial large sample size, the proportion of participants continuing injection use in the subsequent year was small and thus we were not able to use an active-comparator design with this dataset. Furthermore, no information was available on the formulation of injections and dosages used, so all injections were analyzed together regardless of drug. Lastly, we used a single-time interval as the simplified setting for data generation and the time-varying confounder used in the study was not affected by treatment history. Further information on time-varying confounders affecting treatment use is needed to correctly model the complex mechanism of treatment use.[120]

Despite these limitations, this dissertation has several strengths. Use of advanced analytic techniques adds to the body of literature in OA research on intra-articular injection use. Our studies were conducted within the OAI, a large, prospective cohort that collected a wealth of data on a large number of people with radiographically confirmed knee OA. With the data derived from yearly visits within the OAI, we were able to identify patients newly initiating injection use. To our knowledge, Aim 1 is the first study

examining patterns of intra-articular injections use for patients with knee OA. In addition, we applied several advanced analytic methods such as MSMs to estimate the effect of time-varying exposures and time-dependent confounders that are affected by prior treatment use. The use of MSMs allowed us to obtain less-biased effect estimates than we would have observed using a traditional approach. Lastly, compared to ordinary simulation methods, using the plasmode simulation framework provided the advantage of specifying values of parameters such as strength of confounders and exposure of interest given that covariate data and associations among covariates remain intact. While previous studies focused on time-varying confounders,[92,108] our study also assessed the performance of missing data techniques with imposed baseline confounders and mixed scenarios which provided a more comprehensive and realistic evaluation under MSMs analyses.

Implications and future research

Given the summary findings from the three specific aims of this dissertation, we suggest several clinical and research implications for future work in this area.

Clinical implications

Findings from this work have yielded several important clinical implications that could affect providers' or patients' clinical management of knee OA. First, a consistent recommendation is needed for commonly used clinical guidelines. In this work, we observed that a substantial proportion of knee OA patients initiating intra-articular

injection use switched their treatment or simply used it for one time, regardless of the initial therapy or actual symptoms change. Rather than the actual effectiveness of treatment, this phenomenon may be due to conflicting recommendations in the guidelines. Indeed, it has been suggested that the clinical guidelines for the treatment of patients with knee OA should employ a standard "appropriate methodology" to avoid confusion among physicians.[125] Second, similar to other pain treatments such as NSAIDs,[12,24,25] a "step-up" approach for intra-articular injections for patients with knee OA may be needed. Typically, clinicians may be channeling persons with more severe disease into the newest and more costly treatments (e.g., HA injection in this work). However, we observed that approximately 1 in 4 participants initiating HA injections switched to CO injections and that channeling into the hyaluronic injections was not apparent. These findings may suggest a potential lack of increased efficacy. Last but not least, given concern about cost-effectiveness, we suggest that patients be informed about the long-term efficacy of these commonly used modalities to set accurate expectations before deciding on the treatment.

Patients with more severe disease may receive HA injections as a matter of convenience since HA injections may offer longer term benefits compared to CO.[25] However, our work suggests that initiating either CO or HA injection use was not associated with reduced symptoms when compared to non-users over two years. Despite current literature showing that injections are generally safe and have positive effects for patients' satisfaction,[126–128] it is still not clear what proportion of observed measurements

are actually from the real disease modifying effect of these treatments or simply a result of a placebo effect.

Research implications

While randomized clinical trials (RCTs) are considered the "gold standard" in medical research, using observational data has several advantages. It allows investigators to assess effectiveness and safety with the availability of a longer follow-up period.[129] Using a larger sample size with the inclusion of a wide range of patients living in the community, researchers can also examine differences in non-primary outcomes and improve generalizability of study findings.[130] However, unlike RCTs that use randomization to make treatment groups balanced, observational studies are limited due to non-random comparison groups with confounding factors. Furthermore, using a traditional approach to control for confounding including time-varying confounders may result in biased estimates of treatment use in longitudinal observational studies.[30,41,43]

In this dissertation, we used MSMs with IPTW to estimate the unbiased causal effect when time-varying confounding is a concern.[30,41,43] Despite the advantages of MSMs, we identified several key methodologic areas that may inform future research efforts. First, given the cost associated with treatment modalities and inconclusive evidence from the literature regarding injection use, future studies may consider use of active-comparator design, especially for the head-to-head comparison, when conducting a comparative-effectiveness research. While our work had to employ non-users as the comparison group to examine the long-term effectiveness of injection use, we recognize

that an active-comparator design may build on this work in several areas.[102,103] Further, considering the use of multiple treatment cycles is also important when studying long-term effectiveness in the context of injection use. When using evidence from systematic reviews, we noticed that follow-up periods in those trials are mostly short-term with only one treatment cycle. Therefore, the beneficial effects of long-term use, especially for continuous users, remain unclear. In our study, despite evaluation of multiple treatment use (i.e., two time intervals), there was a potential for mismatch between the time of injection use and outcome assessments. For example, at annual assessment visits when participants were asked about injections used, they could be in the middle of a treatment cycle. Therefore, the treatment effects could be underestimated. Moreover, it is also vital to take into account the information regarding the formulation of injections as well as dosages used given that there are several molecular weights available for hyaluronic acid injections[98] and several different corticosteroid formulations. These factors may affect how we define the operational definition for the exposure of interest as well as the outcomes observed.[131] Finally, given that OA is a heterogeneous disease, it is important to identify the subgroups that may potentially benefit most from treatment in future clinical research.[132,133] This is especially important because sometimes it could be the failure to identify and examine subgroups that results in finding no long-term effects on structural modification in the knee and/or symptom relief. Therefore, future work may require a methodologically robust method to test for a treatment effect in subgroups in order to inform novel therapeutic opportunities and clinical applications.

Conclusions

Taken together, the studies in this dissertation add to the body of literature by examining patterns and effects of intra-articular injection use among person with knee OA living in the community. The findings from this dissertation suggested that the proportion of switching between injection drugs and the proportion of one-time users was substantial after treatment initiation. In addition, initiating injection use was not associated with reduced symptoms over time. With respect to issues of missing data, using MI may confer an advantage over IPW in MSMs applications. The results from this work highlight the importance of conducting comparative effectiveness research with non-experimental data to study these commonly used injections and may shed light on strategic non-surgical treatment solutions by informing patients and clinical prescribers who treat persons with knee OA. Future studies should include information on the formulation of injections and dosages used and identify subgroups that may benefit most from the treatment use.

REFERENCES

1. LawrenceRC, FelsonDT, HelmickCG, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. *Arthritis Rheum.* 2008;58:26-35. doi:10.1002/art.23176.

2. FelsonDT, ZhangY. An update on the epidemiology of knee and hip osteoarthritis with a view to prevention. *Arthritis Rheum.* 1998;41(8):1343-1355. doi:10.1002/1529-0131(199808)41:8<1343::AID-ART3>3.0.CO;2-9.

3. FelsonDT, CouropmitreeNN, ChaissonCE, et al. Evidence for a Mendelian gene in a segregation analysis of generalized radiographic osteoarthritis: The Framingham Study. *Arthritis Rheum.* 1998;41:1064-1071.

4. FelsonDT, NevittMC. Epidemiologic studies for osteoarthritis: New versus conventional study design approaches. *Rheum. Dis. Clin. North Am.* 2004;30:783-797. doi:10.1016/j.rdc.2004.07.005.

5. FelsonDT, NaimarkA, AndersonJ, KazisL, CastelliW, MeenanRF. The prevalence of knee osteoarthritis in the elderly. The Framingham Osteoarthritis Study. *Arthritis Rheum.* 1987;30:914-918. doi:10.1002/art.1780300811.

6. OliveriaSA, FelsonDT, ReedJI, CirilloPA, WalkerAM. Incidence of symptomatic hand, hip, and knee osteoarthritis among patients in a health maintenance organization. *Arthritis Rheum.* 1995;38:1134-1141. doi:10.1002/art.1780380817.

7. FelsonDT, ZhangY, HannanMT, et al. The incidence and natural history of knee osteoarthritis in the elderly: The Framingham osteoarthritis study. *Arthritis Rheum.* 1995;38:1500-1505. doi:10.1002/art.1780381017.

8. NeogiT, ZhangY. Epidemiology of Osteoarthritis. *Rheum. Dis. Clin. North Am.* 2013;39:1-19. doi:10.1016/j.rdc.2012.10.004.

9. LitwicA, EdwardsMH, DennisonEM, CooperC. Epidemiology and burden of osteoarthritis. *Br. Med. Bull.* 2013;105:185-199. doi:10.1093/bmb/lds038.

10. DillonCF, RaschEK, GuQ, HirschR. Prevalence of knee osteoarthritis in the United States: arthritis data from the Third National Health and Nutrition Examination Survey 1991-94. *J. Rheumatol.* 2006;33:2271-2279.

11. GuccioneAA, FelsonDT, AndersonJJ, et al. The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. *Am. J. Public Health* 1994;84:351-358. doi:10.2105/AJPH.84.3.351.

12. McAlindonTE, BannuruRR, SullivanMC, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. *Osteoarthr. Cartil.* 2014;22:363-388.

doi:10.1016/j.joca.2014.01.003.

13. KonE, FilardoG, DrobnicM, et al. Non-surgical management of early knee osteoarthritis. *Knee Surg. Sports Traumatol. Arthrosc.* 2012;20(3):436-49. doi:10.1007/s00167-011-1713-8.

14. AyhanE, KesmezacarH, AkgunI. Intraarticular injections (corticosteroid, hyaluronic acid, platelet rich plasma) for the knee osteoarthritis. *World J. Orthop.* 2014;5(3):351-61. doi:10.5312/wjo.v5.i3.351.

15. BellamyN, CampbellJ, RobinsonV, GeeT, BourneR, WellsG. Intraarticular corticosteroid for treatment of osteoarthritis of the knee. *Cochrane database Syst. Rev.* 2006:CD005328. doi:10.1002/14651858.CD005328.pub2.

16. DouglasRJ. Corticosteroid injection into the osteoarthritic knee: Drug selection, dose, and injection frequency. *Int. J. Clin. Pract.* 2012;66(7):699-704. doi:10.1111/j.1742-1241.2012.02963.x.

17. BannuruRR, NatovNS, ObadanIE, PriceLL, SchmidCH, McAlindonTE. Therapeutic trajectory of hyaluronic acid versus corticosteroids in the treatment of knee osteoarthritis: a systematic review and meta-analysis. *Arthritis Rheum.* 2009;61(12):1704-1711. doi:10.1002/art.24925.

18. HepperCT, HalvorsonJJ, DuncanST, GregoryAJM, DunnWR, SpindlerKP. The efficacy and duration of intra-articular corticosteroid injection for knee osteoarthritis: a systematic review of level I studies. *J. Am. Acad. Orthop. Surg.* 2009;17(10):638-46.

19. HochbergMC, AltmanRD, AprilKT, et al. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. *Arthritis Care Res. (Hoboken).* 2012;64:465-474. doi:10.1002/acr.21596.

20. PendletonA, ArdenN, DougadosM, et al. EULAR recommendations for the management of knee osteoarthritis: report of a task force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT). *Ann.Rheum.Dis.* 2000;59:936-944.

21. The Royal Australian College of General Practitioners: Osteoarthritis Working Group. Guideline for the non-surgical management of hip and knee osteoarthritis. Available at: https://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/cp117-hip-knee-osteoarthritis.pdf. Accessed February10, 2016.

22. HunterDJ. Viscosupplementation for Osteoarthritis of the Knee. *N. Engl. J. Med.* 2015;372(11):1040-1047. doi:10.1056/NEJMct1215534.

23. BellamyN, CampbellJ, RobinsonV, GeeT, BourneR, WellsG. Viscosupplementation for the treatment of osteoarthritis of the knee. *Cochrane Database Syst. Rev.* 2006;(2):CD005321. doi:10.1002/14651858.CD005321.pub2.

24. HochbergMC, AltmanRD, AprilKT, et al. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. *Arthritis Care Res. (Hoboken).* 2012;64(4):465-474. doi:10.1002/acr.21596.

25. JevsevarDS. Treatment of osteoarthritis of the knee: evidence-based guideline, 2nd edition. *J. Am. Acad. Orthop. Surg.* 2013;21:571-6. doi:10.5435/JAAOS-21-09-571.

26. JüniP, HariR, RutjesAWS, et al. Intra-articular corticosteroid for knee osteoarthritis. *Cochrane Database Syst. Rev.* 2015;(10). doi:10.1002/14651858.CD005328.pub3.

27. StrandV, McIntyreLF, BeachWR, MillerLE, BlockJE. Safety and efficacy of US-approved viscosupplements for knee osteoarthritis : a systematic review and meta-analysis of randomized , saline-controlled trials. *J. Pain Res.* 2015;8:217-228. doi:10.2147/JPR.S83076.

28. JubbRW, PivaS, BeinatL, DacreJ, GishenP. *A One-Year, Randomised, Placebo (Saline) Controlled Clinical Trial of 500-730 kDa Sodium Hyaluronate (Hyalgan) on the Radiological Change in Osteoarthritis of the Knee.*; 2003.

29. RaynauldJ-P, Buckland-WrightC, WardR, et al. Safety and Efficacy of Long-Term Intraarticular Steroid Injections in Osteoarthritis of the Knee A Randomized, Double-Blind, Placebo-Controlled Trial. *Arthritis Rheum.* 2003;48(2):370-377. doi:10.1002/art.10777.

30. RobinsJM, HernánM a, BrumbackB. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11(5):550-560. doi:10.1097/00001648-200009000-00011.

31. Nevitt MC, Felson DTLG. The Osteoarthritis Initiative Protocol for the Cohort Study. 2006. Available at: http://oai.epi-ucsf.org/datarelease/docs/StudyDesignProtocol.pdf. Accessed October6, 2016.

32. LeopoldSS, ReddBB, WarmeWJ, WehrleP a, PettisPD, ShottS. Corticosteroid compared with hyaluronic acid injections for the treatment of osteoarthritis of the knee. A prospective, randomized trial. *J. Bone Joint Surg. Am.* 2003;85-A(7):1197-203.

33. HernanMA, BrumbackB, RobinsJM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 2000;11(5):561-570.

34. RosenbaumP, RubinD. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70(1):41-55. doi:10.1093/biomet/70.1.41.

35. Sanni AliM, GroenwoldRHH, PestmanWR, et al. Time-dependent propensity score and collider-stratification bias: An example of beta2-agonist use and the risk of coronary heart disease. *Eur. J. Epidemiol.* 2013;28(4):291-299. doi:10.1007/s10654-013-9766-2.

36. AMW. Confounding by indication. *Epidemiology* 1996;7:335-336.

37. BhattacharyaJ, VogtWB. Do Instrumental Variables Belong in Propensity Scores? *Int. J. Stat. Econ.* 2012;9(A12):107-127. doi:http://www.nber.org/papers/t0343.pdf.

38. BrookhartMA, SchneeweissS, RothmanKJ, GlynnRJ, AvornJ, StürmerT. Variable selection for propensity score models. *Am. J. Epidemiol.* 2006;163(12):1149-1156. doi:10.1093/aje/kwj149.

39. HernanMA, Hernandez-DiazS, RobinsJM. A structural approach to selection bias. *Epidemiology* 2004;15:615-625. doi:00001648-200409000-00020 [pii].

40. ColeSR, HernánMA. Constructing inverse probability weights for marginal structural models. *Am. J. Epidemiol.* 2008;168(6):656-664. doi:10.1093/aje/kwn164.

41. HernánMA, Hernández-DíazS, RobinsJM. A structural approach to selection bias. *Epidemiology* 2004;15(5):615-625. doi:10.1097/01.ede.0000135174.63482.43.

42. RobinsJM, HernánMA, BrumbackB. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11(5):550-560. doi:10.1097/00001648-200009000-00011.

43. RobinsJ. A new approach to causal inference in mortality studies with a sustained exposure period-application to control of the healthy worker survivor effect. *Math. Model.* 1986;7(9-12):1393-1512. doi:10.1016/0270-0255(86)90088-6.

44. FranklinJM, SchneeweissS, PolinskiJM, RassenJA. Plasmode simulation for the evaluation of pharmacoepidemiologic methods in complex healthcare databases. *Comput. Stat. Data Anal.* 2014;72:219-226. doi:10.1016/j.csda.2013.10.018.

45. RubinDB. Inference and missing data. *Biometrika* 1976;63(3):581-592. doi:10.1093/biomet/63.3.581.

46. Little, RJA and RublinD. Statistical Analysis with Missing Data. *Wiley, New York.* 1987:381. doi:10.1002/9781119013563.

47. RobinsJM, RotnitzkyA. Recovery of information and adjustment for dependent censoring using surrogate markers. In: *AIDS Epidemiology, Methodological Issues*.; 1992:297-331. doi:10.1007/978-1-4757-1229-2_14.

48. RobinsJM, RotnitzkyA, ZhaoLP. Estimation of regression-coefficients when some regressors are not always observed. *J. Am. Stat. Assoc.* 1994;89:846-866.

49. RubinDB. *Multiple Imputation for Nonresponse in Surveys.* New York: John Wiley & Sons Inc.; 2004.

50. BraultM, HootmanJ, HelmickCG, ThesisKA, ArmourBS. Prevalence and common causes of disability among adults - United States, 2005. *Morb. Mortal. Wkly. Rep.* 2009;58:421-426.

51. FelsonDT, LawrenceRC, DieppePA, et al. Osteoarthritis: new insights. Part 1: the disease and its risk factors. In: *Annals of Internal Medicine*.Vol 133.; 2000:635-646. doi:10.7326/0003-4819-133-8-200010170-00016.

52. WehlingP, MoserC, MaixnerW. How does surgery compare with advanced intra-articular therapies in knee osteoarthritis: current thoughts. *Ther. Adv. Musculoskelet. Dis.* 2016;8(3):72-85. doi:10.1177/1759720X16642405.

53. ArnoldW, FullertonDSP, HolderS, MayCS. Viscosupplementation: managed care issues for osteoarthritis of the knee. *J. Manag. Care Pharm.* 2007;13(4 Suppl):S3-19-2.

54. KoenigKM, OngKL, LauEC, et al. The Use of Hyaluronic Acid and Corticosteroid Injections Among Medicare Patients With Knee Osteoarthritis. *J. Arthroplasty* 2016;31(2):351-355. doi:http://dx.doi.org/10.1016/j.arth.2015.08.024.

55. ZhaoSZ, WentworthC, BurkeTA, MakuchRW. Drug switching patterns among patients with rheumatoid arthritis and osteoarthritis using COX-2 specific inhibitors and non-specific NSAIDs. *Pharmacoepidemiol. Drug Saf.* 2004;13(5):277-287. doi:10.1002/pds.909.

56. M.L, K.H.K, S.X.K, et al. Drug switching patterns among patients taking non-steroidal anti-inflammatory drugs: A retrospective cohort study of a general practitioners database in the United Kingdom. *Pharmacoepidemiol. Drug Saf.* 2001;10(6):517-524. doi:10.1002/pds.653.

57. RoosEM, KlässboM, LohmanderLS. WOMAC osteoarthritis index. Reliability, validity, and responsiveness in patients with arthroscopically assessed osteoarthritis. Western Ontario and MacMaster Universities. *Scand. J. Rheumatol.* 1999;28(4):210-215. doi:10.1016/S0161-4754(00)90101-5.

58. DieppePA. Relationship between symptoms and structural change in osteoarthritis: What are the important targets for therapy? In: *Journal of Rheumatology*.Vol 32.;

2005:1147-1149.

59. SofatN, KuttapitiyaA. Future directions for the management of pain in osteoarthritis. *Int. J. Clin. Rheumtol.* 2014;9(2):197-276. doi:10.2217/ijr.14.10.

60. ModiA, SenS, AdachiJD, et al. Gastrointestinal symptoms and association with medication use patterns, adherence, treatment satisfaction, quality of life, and resource use in osteoporosis: baseline results of the MUSIC-OS study. *Osteoporos. Int.* 2016;27:1227-1238. doi:10.1007/s00198-015-3388-3.

61. VignonE, PipernoM, LeGraverandMPH, et al. Measurement of radiographic joint space width in the tibiofemoral compartment of the osteoarthritic knee: Comparison of standing anteroposterior and Lyon schuss views. *Arthritis Rheum.* 2003;48(2):378-384. doi:10.1002/art.10773.

62. DougadosM, HawkerG, LohmanderS, et al. OARSI/OMERACT criteria of being considered a candidate for total joint replacement in knee/hip osteoarthritis as an endpoint in clinical trials evaluating potential disease modifying osteoarthritic drugs. *J. Rheumatol.* 2009;36(9):2097-9. doi:10.3899/jrheum.090365.

63. AngstF, AeschlimannA, StuckiG. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower ex. *Arthritis Rheum.* 2001;45(4):384-91. doi:10.1002/1529-0131(200108)45:4<384::AID-ART352>3.0.CO;2-0.

64. GrecoNJ, AndersonAF, MannBJ, et al. Responsiveness of the International Knee Documentation Committee Subjective Knee Form in comparison to the Western Ontario and McMaster Universities Osteoarthritis Index, modified Cincinnati Knee Rating System, and Short Form 36 in patients with focal art. *Am. J. Sports Med.* 2010;38(5):891-902. doi:10.1177/0363546509354163.

65. DuryeaJ, LiJ, PeterfyCG, GordonC, GenantHK. Trainable rule-based algorithm for the measurement of joint space width in digital radiographic images of the knee. *Med. Phys.* 2000;27(3):580-591. doi:10.1118/1.598897.

66. WareJ, KosinskiM, KellerSD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med. Care* 1996;34(3):220-233. doi:10.2307/3766749.

67. JensenMD, RyanDH, ApovianCM, et al. 2013 AHA/ACC/TOS Guideline for the Management of Overweight and Obesity in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. *Circulation* 2013:1-70. doi:10.1161/01.cir.0000437739.71477.ee.

68. RadloffLS. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Appl. Psychol. Meas.* 1977;1:385-401. doi:10.1177/014662167700100306.

69. KatzJN, ChangLC, SanghaO, FosselAH, BatesDW. Can comorbidity be measured by questionnaire rather than medical record review? *Med. Care* 1996;34:73-84. doi:10.1097/00005650-199601000-00006.

70. WashburnRA, FickerJL. Physical Activity Scale for the Elderly (PASE): the relationship with activity measured by a portable accelerometer. *J. Sports Med. Phys. Fitness* 1999;39:336-340.

71. KasparJ, KasparS, OrmeC, DeBeerJDV. Intra-articular steroid hip injection for osteoarthritis: A survey of orthopedic surgeons in Ontario. *Can. J. Surg.* 2005;48(6):461-469.

72. AJS, SMB-Z, DekkerJ, et al. Agreement of general practitioners with the guideline-based stepped-care strategy for patients with osteoarthritis of the hip or knee: a cross-sectional study. *BMC Fam. Pract.* 2013;14:33. doi:http://dx.doi.org/10.1186/1471-2296-14-33.

73. MosisG, StijnenT, CastellsagueJ, et al. Channeling and prevalence of cardiovascular contraindications in users of cyclooxygenase 2 selective nonsteroidal antiinflammatory drugs. *Arthritis Care Res.* 2006;55(4):537-542. doi:10.1002/art.22096.

74. BerchiC, DegieuxP, HalholH, DanelB, BennaniM, PhilippeC. Impact of falling reimbursement rates on physician preferences regarding drug therapy for osteoarthritis using a discrete choice experiment. *Int. J. Pharm. Pract.* 2015. doi:10.1111/ijpp.12220.

75. LawA, RayMD, KnappK, BaleshJ. Unmet Needs in the Medication Use Process: Perceptions of Physicians, Pharmacists, and Patients. *J. Am. Pharm. Assoc.* 2003;43(3):394. doi:10.1331/154434503321831111.

76. RabendaV, ManetteC, LemmensR, MarianiAM, StruvayN, ReginsterJY. Direct and indirect costs attributable to osteoarthritis in active subjects. *J. Rheumatol.* 2006;33(6):1152-1158.

77. KotlarzH, GunnarssonCL, FangH, RizzoJA. Osteoarthritis and absenteeism costs: evidence from US National Survey Data. *J. Occup. Environ. Med.* 2010;52(3):263-8. doi:10.1097/JOM.0b013e3181cf00aa.

78. WhiteAG, BirnbaumHG, JanagapC, ButeauS, ScheinJ. Direct and indirect costs of pain therapy for osteoarthritis in an insured population in the United States. *J. Occup. Environ. Med.* 2008;50(9):998-1005. doi:10.1097/JOM.0b013e3181715111.

79. FreedmanB. Equipoise and the ethics of clinical research. *N. Engl. J. Med.* 1987;317(3):141-145. doi:10.1056/NEJM198707163170304.

80. PerryDC, GriffinXL, ParsonsN, CostaML. Designing clinical trials in trauma surgery: overcoming research barriers. *Bone Joint Res.* 2014;3(4):123-9. doi:10.1302/2046-3758.34.2000283.

81. EvansCH, KrausVB, SettonLA. Progress in intra-articular therapy. *Nat. Rev. Rheumatol.* 2014;10(1):11-22. doi:10.1038/nrrheum.2013.159.

82. ThomasE, CroftPR, PatersonSM, DziedzicK, HayEM. What influences participants' treatment preference and can it influence outcome? Results from a primary care-based randomised trial for shoulder pain. *Br. J. Gen. Pract.* 2004;54(499):93-96.

83. RayWA. Evaluating Medication Effects Outside of Clinical Trials: New-User Designs. *Am. J. Epidemiol.* 2003;158(9):915-920. doi:10.1093/aje/kwg231.

84. DhawanA, Mather IIIRC, KarasV, et al. An epidemiologic analysis of clinical practice guidelines for non-arthroplasty treatment of osteoarthritis of the knee. *Arthrosc. - J. Arthrosc. Relat. Surg.* 2014;30(1):65-71. doi:10.1016/j.arthro.2013.09.002.

85. WenDY. Intra-articular hyaluronic acid injections for knee osteoarthritis. *Am. Fam. Physician* 2000;62(3):565-570.

86. WattersonJR, EsdaileJM. Viscosupplementation: therapeutic mechanisms and clinical potential in osteoarthritis of the knee. *J. Am. Acad. Orthop. Surg.* 2000;8(5):277-284.

87. GreenlandS, FinkleWD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. *Am. J. Epidemiol.* 1995;142(12):1255-64. doi:10.1093/aje/kwi188.

88. Okma-KeulenP, Hopman-RockM. The onset of generalized osteoarthritis in older women: a qualitative approach. *Arthritis Rheum.* 2001;45(2):183-90. doi:10.1002/1529-0131(200104)45:2<183::AID-ANR172>3.0.CO;2-0.

89. JensenMD, RyanDH, ApovianCM, et al. 2013 AHA/ACC/TOS Guideline for the Management of Overweight and Obesity in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. *Circulation* 2014;129(25 suppl 2):S102-S138. doi:10.1161/01.cir.0000437739.71477.ee.

90. PlattRW, BrookhartMA, ColeSR, WestreichD, SchistermanEF. An information criterion for marginal structural models. *Stat. Med.* 2013;32(8):1383-1393. doi:10.1002/sim.5599.

91. MortimerKM, NeugebauerR, Van DerLaanM, TagerIB. An application of model-fitting procedures for marginal structural models. *Am. J. Epidemiol.* 2005;162(4):382-388. doi:10.1093/aje/kwi208.

92. MoodieEEM, DelaneyJAC, LefebvreG, PlattRW. Missing Confounding Data in Marginal Structural Models: A Comparison of Inverse Probability Weighting and Multiple Imputation. *Int. J. Biostat.* 2008;4(1):1-23. doi:10.2202/1557-4679.1106.

93. WhiteIR, RoystonP, WoodAM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat. Med.* 2011;30(4):377-399. doi:10.1002/sim.4067.

94. MoonsKGM, DondersRART, StijnenT, Jr.FEH. Using the outcome for imputation of missing predictor values was preferred. *J. Clin. Epidemiol.* 2006;59(10):1092-1101. doi:http://dx.doi.org/10.1016/j.jclinepi.2006.01.009.

95. Driban J, Lavalley MP, Price LL, Harvey WFMT. Intra-Articular Corticosteroids Are Safe and Have No Major Effect on Structural Progression of Synovitic Knee OA: A 2-Year Randomized Controlled Trial of 3-Monthly Triamcinolone Hexacetonide [abstract]. *Arthritis Rheumatol.* 2015;67(Suppl 10). Available at: http://acrabstracts.org/abstract/intra-articular-corticosteroids-are-safe-and-have-no-major-effect-on-structural-progression-of-synovitic-knee-oa-a-2-year-randomized-controlled-trial-of-3-monthly-triamcinolone-hexacetonide/.

96. BellamyN, CampbellJ, RobinsonV, GeeT, BourneR, WellsG. Viscosupplementation for the treatment of osteoarthritis of the knee. *Cochrane database Syst. Rev.* 2006;(2):CD005321. doi:10.1002/14651858.CD005321.pub2.

97. RutjesAW, JuniP, daCostaBR, TrelleS, NueschE, ReichenbachS. Viscosupplementation for osteoarthritis of the knee: a systematic review and meta-analysis. *Ann. Intern. Med.* 2012;157(3):180-191. doi:10.7326/0003-4819-157-3-201208070-00473.

98. JevsevarD, DonnellyP, BrownGA, CumminsDS. Viscosupplementation for Osteoarthritis of the Knee. *J. Bone Jt. Surg.* 2015;97(24):2047 LP-2060.

99. GoldbergVM, BuckwalterJA. Hyaluronans in the treatment of osteoarthritis of the knee: Evidence for disease-modifying activity. *Osteoarthr. Cartil.* 2005;13(3):216-224. doi:10.1016/j.joca.2004.11.010.

100. PhamT, LeHenanffA, RavaudP, DieppeP, PaolozziL, DougadosM. Evaluation of the symptomatic and structural efficacy of a new hyaluronic acid compound, NRD101, in comparison with diacerein and placebo in a 1 year randomised controlled study in symptomatic knee osteoarthritis. *Ann. Rheum. Dis.* 2004;63(12):1611-1617. doi:10.1136/ard.2003.019703.

101. TammachoteN, KanitnateS, YakumporT, PanichkulP. Intra-Articular, Single-Shot

Hylan G-F 20 Hyaluronic Acid Injection Compared with Corticosteroid in Knee Osteoarthritis. *J. Bone Jt. Surg.* 2016;98(11):885-92.

102. YoshidaK, SolomonDH, KimSC. Active-comparator design and new-user design in observational studies. *Nat. Rev. Rheumatol.* 2015;11:437–441. doi:10.1038/nrrheum.2015.30.

103. DesaiRJ, SolomonDH, SchneeweissS, DanaeiG, LiaoKP, KimSC. Tumor Necrosis Factor-α Inhibitor Use and the Risk of Incident Hypertension in Patients with Rheumatoid Arthritis. *Epidemiology* 2016;27(3):414-422. doi:10.1097/EDE.0000000000000446.

104. GreenlandS. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology* 2003;14(3):300-306. doi:10.1097/01.EDE.0000042804.12056.6C.

105. LefebvreG, DelaneyJAC, PlattRW. Impact of mis-specification of the treatment model on estimates from a marginal structural model. *Stat. Med.* 2008;27(18):3629-3642. doi:10.1002/sim.3200.

106. PetersenML, PorterKE, GruberS, WangY, van derLaanMJ. Diagnosing and responding to violations in the positivity assumption. *Stat. Methods Med. Res.* 2012;21(1):31-54. doi:10.1177/0962280210386207.

107. BrumbackBA, HernánMA, HaneuseSJPA, RobinsJM. Sensitivity analyses for unmeasured confounding assuming a marginal structural model for repeated measures. *Stat. Med.* 2004;23(5):749-767. doi:10.1002/sim.1657.

108. VourliG, TouloumiG. Performance of the marginal structural models under various scenarios of incomplete marker's values: A simulation study. *Biometrical J.* 2015;57(2):254-270. doi:10.1002/bimj.201300159.

109. YangS, EatonCB, LuJ, LapaneKL. Application of marginal structural models in pharmacoepidemiologic studies: A systematic review. *Pharmacoepidemiol. Drug Saf.* 2014;23(6):560-571. doi:10.1002/pds.3569.

110. LapaneKL, LiuS-H, DubéCE, DribanJB, McAlindonTE, EatonCB. Factors Associated with the Use of Hyaluronic Acid and Corticosteroid Injections among Patients with Radiographically Confirmed Knee Osteoarthritis: A Retrospective Data Analysis. *Clin. Ther.* 2017;39(2):347-358. doi:10.1016/j.clinthera.2017.01.006.

111. GoldfeldK. simstudy: Simulation of Study Data. 2016. Available at: https://cran.r-project.org/package=simstudy.

112. MoonsKGM, DondersRART, StijnenT, HarrellFE. Using the outcome for imputation of missing predictor values was preferred. *J. Clin. Epidemiol.*

2006;59(10):1092-1101. doi:10.1016/j.jclinepi.2006.01.009.

113. vanBuuren, Stef; Groothuis-OudshoornK. {mice}: Multivariate Imputation by Chained Equations in R. *J. Stat. Softw.* 2011;45(3):1-67. Available at: http://www.jstatsoft.org/v45/i03/.

114. SchaferJL. Multiple imputation: a primer. *Stat Methods Med Res* 1999;8(1):3-15. doi:10.1191/096228099671525676.

115. Honaker, James; King, Gary; BlackwellM. {Amelia II}: A Program for Missing Data. *J. Stat. Softw.* 2011;45(7):1-47. Available at: http://www.jstatsoft.org/v45/i07/.

116. RubinD. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley; 1987.

117. SchaferJL, GrahamJW. Missing data: Our view of the state of the art. *Psychol. Methods* 2002;7(2):147-177. doi:10.1037//1082-989X.7.2.147.

118. SeamanSR, WhiteIR. Review of inverse probability weighting for dealing with missing data. *Stat. Methods Med. Res.* 2013;22(3):278-295. doi:10.1177/0962280210395740.

119. JoffeMM, PistilliM, KempenJH. Marginal structural models for comparing alternative treatment strategies in ophthalmology using observational data. *Ophthalmic Epidemiol.* 2013;20(4):197-200. doi:10.3109/09286586.2013.792939.

120. VelentgasP, DreyerN, NourjahP, SmithS, TorchiaM. *Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide*.; 2013. Available at: www.effectivehealthcare.ahrq.gov/.

121. ShortreedSM, ForbesAB. Missing data in the exposure of interest and marginal structural models: A simulation study based on the Framingham Heart Study. *Stat. Med.* 2010;29(4):431-443. doi:10.1002/sim.3801.

122. WacholderS, HartgeP, LubinJH, DosemeciM. Non-differential misclassification and bias towards the null: a clarification. *Occup. Environ. Med.* 1995;52(8):557-558. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1128296/.

123. PsatyBM, KoepsellTD, LinD, et al. Assessment and control for confounding by indication in observational studies. *J. Am. Geriatr. Soc.* 1999;47(6):749-54. doi:10.1111/j.1532-5415.1999.tb01603.x.

124. PlattRW, DelaneyJAC, SuissaS. The positivity assumption and marginal structural models: The example of warfarin use and risk of bleeding. *Eur. J. Epidemiol.* 2012;27(2):77-83. doi:10.1007/s10654-011-9637-7.

125. AltmanRD, SchemitschE, BediA. Assessment of clinical practice guideline methodology for the treatment of knee osteoarthritis with intra-articular hyaluronic acid. *Semin. Arthritis Rheum.* 2015;45(2):132-139. doi:10.1016/j.semarthrit.2015.04.013.

126. DivineJG, ShafferMD. Use of viscosupplementation for knee osteoarthritis: An update. *Curr. Sports Med. Rep.* 2011;10(5):279-284. doi:10.1249/JSR.0b013e31822ed1b4.

127. BeckerLC, BergfeldWF, BelsitoDV, et al. Final report of the safety assessment of hyaluronic acid, potassium hyaluronate, and sodium hyaluronate. *Int. J. Toxicol.* 2009;28(4 Suppl):5-67. doi:10.1177/1091581809337738.

128. BhagraA, SyedH, ReedDA, et al. Efficacy of musculoskeletal injections by primary care providers in the office: a retrospective cohort study. *Int. J. Gen. Med.* 2013;6:237-43. doi:10.2147/IJGM.S41893.

129. SchneeweissS, GagneJJ, GlynnRJ, RuhlM, RassenJA. Assessing the comparative effectiveness of newly marketed medications: methodological challenges and implications for drug development. *Clin. Pharmacol. Ther.* 2011;90(6):777-90. doi:10.1038/clpt.2011.235.

130. Brian L. Strom, Stephen E. KimmelSH. *Textbook of Pharmacoepidemiology*. West Sussex, England: Wiley-Blackwell; 2013.

131. NeustadtDH. Intra-articular injections for osteoarthritis of the knee. *Cleve. Clin. J. Med.* 2006;73(10):897-911. doi:10.3949/ccjm.73.10.897.

132. Bierma-ZeinstraSMA, VerhagenAP. Osteoarthritis subpopulations and implications for clinical trial design. *Arthritis Res. Ther.* 2011;13(2):213. doi:10.1186/ar3299.

133. WaarsingJH, Bierma-ZeinstraSMA, WeinansH. Distinct subtypes of knee osteoarthritis: data from the Osteoarthritis Initiative. *Rheumatology (Oxford).* 2015;54(9):1650-1658. doi:10.1093/rheumatology/kev100.