ASSOCIATION OF SOCIOECONOMIC STATUS WITH OSTEOARTHRITIS-
INDUCED DISABILITY PROGRESSION

By
Divya Narayanan

A thesis submitted to Johns Hopkins University in conformity with the requirements for the degree of Master of Science

Baltimore, Maryland
May 2017
ABSTRACT

Introduction

Osteoarthritis (OA), commonly referred to as “wear and tear” arthritis, is the number one chronic condition of the joints. Social determinants associated with disability progression due to OA are largely unknown. Efforts to determine if socioeconomic status is associated with OA disability progression may aid in prevention efforts over time.

Objective

To describe the relationship between socioeconomic status (SES) measures with disability progression among individuals with radiographic knee and/or hip osteoarthritis from the Johnston County OA (JoCo OA) Project in rural North Carolina.

Methods

We examined the association between disability progression and each of four individual SES variables (education, occupation, income, and home ownership) and one community level SES measure (household poverty) using data from 886 participants of the JoCo OA cohort. Disability progression was defined as a change of $\geq 0.22$ points in the Health Assessment Questionnaire (HAQ) disability score over time. Risk of disability progression was evaluated using a parametric time to event Weibull model that accounted for interval censoring. Change in HAQ score was assessed using linear mixed models. Interaction terms were included to examine disability progression by age. All analyses accounted for age, race, and gender.
Results

Having less than a high school (HS) degree (HR = 1.32; 95% CI: 1.07, 1.63) was independently associated with increased risk of disability progression, among all individuals. Annual income ≤ $30,000 (HR = 1.45; 95% CI: 1.03, 2.04) was independently associated with an increased risk of disability progression among individuals who reported income. When evaluating HAQ score in the longitudinal data analysis, increased household poverty was consistently associated with increased disability status in all models, and lack of home ownership was associated with disability among the subset of participants who reported this information. Across all analyses, women were at higher risk for disability progression.

Conclusion

Educational attainment less than a HS degree, lack of home ownership, and increased household poverty are independently associated with disability progression, suggesting that societal interventions to address these specific SES factors could help to mitigate OA-related disability.

Primary Reader: Eric C. Seaberg, PhD MPH

Secondary Reader: Karen Bandeen-Roche, PhD MS
ACKNOWLEDGEMENTS

I would like to thank several individuals and institutions for their support and contributions towards conducting this research.

First, I’d like to thank my advisor, Dr. Eric Seaberg for his thoughtful feedback, and for his invaluable support throughout this process. His attention to detail and accuracy has instilled a new appreciation for the caliber of work I produce. He has created an environment which continually challenged me to push my intellectual bounds, for which I am immensely grateful.

Next, I’d like to thank my advisor at the University of North Carolina, Chapel Hill (UNC), Dr. Rebecca Cleveland. Her patience and guidance as I navigated the data was invaluable. She was readily available to bounce ideas off of throughout the development and analysis of this work, and shed new light on exploring my passion for epidemiology through osteoarthritis research.

I’m extremely grateful to Drs. Leigh Callahan and Joanne Jordan at UNC Chapel Hill for the opportunity to collaborate with them on the Johnston County Osteoarthritis Project (JoCo OA). I look forward to further collaborations in the future.

Many thanks to my thesis reader, Dr. Karen Bandeen-Roche. Her judicious comments provided fresh insights toward the analytical details of my work.

Many other faculty, staff, and students both at the Johns Hopkins Bloomberg School of Public Health and the Thurston Arthritis Research Center helped make this work...
possible. I would like to extend my sincerest thanks to all of them. They have been champions of my success and a great support system throughout this process.

Thank you to the Rheumatology Research Foundation for funding my contribution to the Johnston County Osteoarthritis Project and supporting my collaboration with UNC TARC.

A special thanks to my parents who have always supported me and been depots of boundless love and encouragement. Without them, none of this would have been possible.

Finally, I’d like to give thanks to the participants of the Johnston County Osteoarthritis Project, without whom this, and many other studies, would not be possible. I hope this work is meaningful and beneficial to those who participated and suffer from osteoarthritis. You are truly an inspiration.

Data in this manuscript were collected from the Johnston County Osteoarthritis Project. JoCo OA is funded by the National Institute of Arthritis and Musculoskeletal Skin Disease (R01-AR-053-989-01) and by the Centers for Disease Control.

Funding was also provided by the Rheumatology Research Foundation. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.
# TABLE OF CONTENTS

ABSTRACT .................................................................................................................. ii
ACKNOWLEDGEMENTS ........................................................................................ iv
LIST OF TABLES ........................................................................................................ vii
LIST OF FIGURES ...................................................................................................... viii
1. BACKGROUND ..................................................................................................... 1
   A. Introduction .......................................................................................................... 1
   B. Pathogenesis of Osteoarthritis ........................................................................ 2
   C. Epidemiology of Osteoarthritis ....................................................................... 3
   D. Clinical Manifestation and Diagnosis ........................................................... 5
   E. Disease Progression and Disability ................................................................. 6
   F. Existing Therapies ............................................................................................. 8
   G. Prior Literature on SES and Progressive Disability due to OA .................... 9
   H. Summary .......................................................................................................... 10
2. AIMS AND HYPOTHESIS .................................................................................... 12
   A. Hypothesis ......................................................................................................... 12
3. METHODS ............................................................................................................. 13
   A. The Johnston County Osteoarthritis Project Design .................................... 13
   B. Study Design and Population ........................................................................ 14
   C. Outcome Assessment ...................................................................................... 15
   D. Primary Exposures .......................................................................................... 16
   E. Potential Confounders ..................................................................................... 17
   F. Statistical Methods ........................................................................................ 17
4. RESULTS ................................................................................................................ 20
   G. Characteristics of the Study Population at Baseline .................................... 20
   H. Time-to-Event Analysis of Disability Progression ...................................... 21
   I. Linear Mixed Model Analysis of Disability Progression ............................... 24
   J. Exploratory Factor Analysis of SES Factors .................................................. 27
5. DISCUSSION .......................................................................................................... 29
6. REFERENCES ........................................................................................................ 41
7. TABLES .................................................................................................................. 46
8. FIGURES ............................................................................................................... 53
9. CURRICULUM VITAE ......................................................................................... 62
LIST OF TABLES

Table 1: Summary of Demographic and SES Measures
Table 2: Baseline Characteristics of the Study Population
Table 3: Univariate Analysis of Time to Disability Progression
Table 4: Multivariable Analysis of Time to Disability Progression
Table 5: Association between Disability Status and SES Measures and Demographic Factors by Age
Table 6: Association between Disability Progression and SES Measures and Demographic Factors by Age
Table 7: Exploratory Factor Analysis of SES Measures
LIST OF FIGURES

Figure 1: Timeline of Enrollment Waves in the Johnston County Osteoarthritis Project

Figure 2a: Inclusion and Exclusion Criteria for Time to Disability Progression Analysis

Figure 2b: Inclusion and Exclusion Criteria for Disability Progression by Age Analysis

Figure 3: Select Questions from the Health Assessment Questionnaire (HAQ)

Figure 4: Study Directed Acyclic Graph (DAG)

Figure 5: Weibull Distribution Probability Density Function – Null Model

Figure 6: Weibull Diagnostic Plot for Time to Disability Progression

Figure 7: Model Checking Plot – QQ Plot

Figure 8: Correlation Plots of Exploratory Factor Analysis of SES Measures
1. BACKGROUND

a. Introduction

Osteoarthritis (OA) is the leading cause of disability in the US among older individuals. Commonly referred to as “wear and tear” arthritis, it is the number one chronic condition of the joints, affecting nearly 27 million US adults aged 18 and older in 2005. Among older adults aged 60 and older, the prevalence is higher, occurring in approximately 13% of women and 10% of men.

Popular belief suggests that osteoarthritis is simply due to extended use and aging of the joints, as it is more prevalent in older individuals. As one in four individuals face a lifetime risk of developing symptomatic osteoarthritis by age 85, and an increasingly older population, efforts are being taken to understand the etiology of disease for prevention efforts.

According to data from the 2010-2012 National Health Interview Survey, an estimated 22.7 million adults aged 18 or older have arthritis-attributable activity limitation. Disability due to osteoarthritis is projected to increase over time, and a shift has been made to look at osteoarthritis over the course of the lifetime as opposed to solely in older individuals. However, as the population ages, prevalence of disability and uncertainty around other potential risk factors are increasing. Examining the risk factors associated with the chronic disease could prove useful in prevention mechanisms and programs.
b. Pathogenesis of Osteoarthritis

Osteoarthritis, commonly referred to as “wear and tear” arthritis, affects the cartilage between joints, as this cartilage breaks down over time and causes pain and inflammation during movement of the joint\textsuperscript{2,3}. Cartilage is a connective tissue which protects the intersection of bones at a joint\textsuperscript{9}. Its elastic and compressive properties are predominantly composed of liquid, as well as type II collagen and proteoglycans, embedded in an extracellular matrix\textsuperscript{5}. The collagen serves as a source of strength, while the proteoglycans allow compression resistance by drawing water into the matrix\textsuperscript{5,10}.

Degradation of the cartilage is a hallmark of the disease, and is largely believed to occur due to an overexpression of degrading enzymes\textsuperscript{5,10}. In addition to the degradation of cartilage, inflammation of the synovium and changes to the subchondral bone also characterize the disease\textsuperscript{10,11}. The synovium is a lining of all joints that are not lined by cartilage, and contains a layer of cells known as synoviocytes which give rise to the lubricating properties of the synovial fluid\textsuperscript{12}. This fluid-filled sac shares a border with both the synovium and cartilage, and serves to reduce the friction between cartilage of synovial joints during movement\textsuperscript{12}. It is the inflammation of this oiling fluid that creates pain during movement of the affected joint\textsuperscript{10,12}.

There has been some evidence to indicate that proteases from the synovium and/or cartilage play a critical role in the degradation of extracellular matrix components of the cartilage and joint-space narrowing\textsuperscript{10,13}. This ultimately results in loss of cartilage and joint function.
c. Epidemiology of Osteoarthritis

The prevalence and incidence of OA are characteristics of the population under study and the specific joints that are being analyzed. Among three well known studies and surveys, these numbers have roughly varied. The age-standardized prevalence of radiographic knee OA in adults aged 45 or older was 19.2% among participants in the Framingham Study, and 27.8% in the Johnston County Osteoarthritis Project. Investigators using data from the third edition of the National Health and Nutrition Examination Survey (NHANES III) estimated that roughly 37% of adults aged 60 or older nationwide had radiographic knee OA. In both the Framingham Study and Johnston County Osteoarthritis Project, the prevalence of radiographic hip OA among participants aged 45 or older was approximately 27%.

OA is known to have a multifactorial etiology, and it is likely an interplay between genetic and environmental components that play a role in diagnosis and progression. The importance of risk factors may vary across joints, stages of the disease, and type of disease (radiographic vs. symptomatic).

One of the most established risk factors for osteoarthritis is age. This can be attributed to various reasons, including prolonged and cumulative exposure to various risk factors, and biologic changes that preclude proper functioning of the joints. As individuals age, their joints experience overuse and natural wear and tear.

Other established risk factors for OA include gender, race/ethnicity, and diet. Men have a significantly lower risk of OA in the knee, but not in the hip, compared to women. Men also have significantly reduced rates for incident OA in both the knee and
Women aged 55 or over tend to have higher risks associated with prevalent OA at all sites, and also tend to have more severe OA\textsuperscript{2,14–16}.

Differences in race/ethnicity have also been shown to yield different OA outcomes. Several studies have shown that African-American individuals have an increased frequency of features that are indicative of future joint replacement, indicating potential anatomic differences or developmental variation\textsuperscript{17,18}. There is also evidence that Hispanics in the US have a lower prevalence of OA compared to Caucasian women in the US\textsuperscript{19}. In a study done on Chinese individuals through the Beijing Osteoarthritis Study, women were shown to have a higher prevalence of both radiographic and symptomatic knee OA than Caucasian women in the Framingham Study\textsuperscript{15,20}.

Several epidemiological studies have shown that heredity plays a role in common forms of osteoarthritis\textsuperscript{21–23}. Twin and family studies have shown that the heritable component of osteoarthritis in women is between 39\% and 65\% for radiographic OA, and about 60\% in the hip, independent of environmental confounders\textsuperscript{15,22,24}.

Body mass index (BMI) is often cited as strongly associated with development of OA. Increased BMI places a greater load on joints, especially the knee and hip, which could possibly lead to the breakdown of cartilage\textsuperscript{25}. Many studies have concluded that individuals with higher BMI are at increased risk for knee OA. A meta-analysis to establish a dose-response relationship between BMI and risk of knee OA showed that risk increased exponentially with an increase of BMI, while others show a 4 times increased risk for women and a 5-times increased risk for men\textsuperscript{21,26}. 
Clinical Manifestation and Diagnosis

Osteoarthritis can manifest in any joint, and symptoms can vary across joints. Of all joints, knees and hips are among the most commonly affected. Early identification of OA is crucial to improving clinical decision-making and advancing the understanding of disease progression and treatment options. Symptoms of knee osteoarthritis often include stiffness, swelling, limited range of motion, and difficulty completing daily living tasks.

Limitations in activity due to pain include sitting and standing up, walking, and many daily living activities. As a result of increased knee and hip use in these activities, development of disability is increased for these joints. Pain usually increases during movement, and subsides when inert, resulting in increased pain at the end of the day due to increased movement. Symptoms for hip osteoarthritis are similar to those of knee osteoarthritis, with additional pain often reaching towards the buttocks, and limited range of motion, specific to internal rotation.

Symptomatic OA is characteristically defined as the co-occurrence of pain, achiness, and stiffness in a joint site, while radiographic OA is often diagnosed through physical exam based on plain film radiographs. However, establishing appreciable progression by either of these methods can often be challenging. Given the difficulty with diagnosing osteoarthritis, physicians and rheumatologists typically explore a variety of methods. These methods include clinical history, physical examination, and radiographic imaging.

Clinical history gives physicians an idea about whether or not an individual’s specific condition may be predominantly genetically or environmentally motivated. Physical
function tests allow physicians to understand to what extent symptoms may affect daily living activities, and how the potentially affected joints appear and feel. Joints affected by osteoarthritis may appear swollen or feel tender\textsuperscript{27,28,34}. A distinguishable characteristic between osteoarthritis and other inflammatory types of arthritis, such as rheumatoid arthritis, is the increased temperature of the joint due to inflammation\textsuperscript{27}.

Radiographic testing can be used to confirm a potential diagnosis and rule out other conditions. The most common imaging modality used to identify osteoarthritis is X-Ray, as it readily identifies joint-space narrowing, bone damage, and cartilage loss\textsuperscript{27,34,35}. Other modalities that have become more common in the diagnostic process include magnetic resonance imaging (MRI) and ultrasound\textsuperscript{35}. MRI highlights the surrounding tissue and can occasionally provide information on cartilage degeneration at an earlier stage by visualizing morphologic changes\textsuperscript{35,36}. In addition to visual assessment, images can be graded using the Kellgren-Lawrence scale, used to quantify the severity of OA\textsuperscript{37}.

e. Disease Progression and Disability

The currently irreversible pathogenesis of osteoarthritis has centered talk of treatment around disability and disease progression. Disability is defined in many ways, depending on the condition of interest or the metric criteria being used. With regards to osteoarthritis, the CDC generally characterizes disability with regards to osteoarthritis by asking a standard case-finding question: “Because of osteoarthritis, do you have difficulty completing this task by yourself?”\textsuperscript{8}

One measure of disability that is commonly used for arthritis-related studies is the Health Assessment Questionnaire – Disability Index (HAQ – DI). The HAQ-DI, more
commonly referred to as the HAQ, is one of the most widely used measures of function when looking at chronic illnesses\textsuperscript{38}. Originating in the field of rheumatology, the HAQ was developed to longitudinally assess outcome measures for rheumatic diseases including but not limited to rheumatoid arthritis (RA) therapy, risk factor analysis in RA, and mortality risk among those with RA\textsuperscript{39}. Since its inception, the questionnaire has been shown to be a valid, reliable, and comprehensive measure across disciplines, making it a robust measure of disability over time\textsuperscript{38,40}. Test-retest correlations of the HAQ have demonstrated reproducibility, with correlations ranging from 0.87 to 0.99, and “correlations between interview and questionnaire formats” ranging from 0.85 to 0.95\textsuperscript{41}. The HAQ has also been shown to possess content and criterion validity, with correlations between questionnaire and task performance ranging from 0.71 to 0.95\textsuperscript{41}. HAQ has been used in a variety of different studies, including those involving HIV/AIDS patients, disabled workers, adults and children with rheumatic diseases, and NHANES\textsuperscript{39,42–45}. Reliance on observational data for a long-term study allows use of the HAQ to obtain valid and comprehensive data across markers.

Disease progression, or the course of disease over time, often occurs over long periods of time in osteoarthritic individuals\textsuperscript{11}. The rate of progression is unique to each individual, and can be assessed symptomatically or radiographically\textsuperscript{17}. As there is no specific lab marker or single established metric to quantify disease progression, defining disease progression is idiosyncratic. Subjective methods, such as questionnaires, can target certain aspects of progression, such as number of joints affected or experiential pain level. These can be used in conjunction with radiographic methods and monitoring of lab
results which can be markers for inflammation and other symptoms associated with osteoarthritis

f. Existing Therapies

With no currently pharmacological treatments that can reverse joint damage caused by osteoarthritis, treatments and existing therapies for osteoarthritis generally serve to meet four goals: 1) control pain 2) improve joint function 3) maintain normal body weight 4) achieve a healthy lifestyle. Some possible therapies include non-pharmacologic, pharmacologic, and surgical procedures and replacements.

One of the most beneficial ways to manage osteoarthritis is through physical activity. As an inexpensive treatment option, exercising has few side effects when performed correctly and can also help build muscle around OA-affected joints to reduce burden and pain. Aerobic exercises can assist with heart health to keep the circulatory system in shape, while range of motion exercises can help keep joints limber and reduce stiffness due to inactivity.

Pain relief medications are often used as short-term relief for individuals suffering from their OA. Oral pain medications, such as acetaminophen, are often part of the first-line approach to tackle pain. However, acetaminophen has been shown to have adverse effects on the liver, and can lead to liver damage or disease. Nonsteroidal anti-inflammatory drugs (NSAIDs) have also proven useful against pain and are commonly used despite their side effects of stomach problems and liver problems.

For individuals for whom non-pharmacologic and pharmacologic treatments do not suffice, surgical procedures may need to be considered, usually after taking factors such
as age, occupation, level of disability, and pain into consideration\textsuperscript{28,34}. Knees and hips are load-bearing joints, and often times require replacement due to extreme usage. As a minimally invasive options, cortisone or lubrication injections can often help with pain relief and cushioning of the knee\textsuperscript{47}. However, some research show that these injections may be minimally effective and can potentially cause adverse effects long term\textsuperscript{47}. Other invasive surgical options include partial or total joint replacement. Partial replacement can involve removal of loose pieces of bone or cartilage that cause buckling or locking, while total replacement requires substitution of damaged parts with plastic or metal prostheses\textsuperscript{28,34,47}.

\textbf{g. Prior Literature on SES and Progressive Disability due to OA}

Few longitudinal studies have examined the association of socioeconomic status (SES) with progressive disability due to OA. One potential influential factor in the development of osteoarthritis is socioeconomic status. A prior study by Hannan et. al. showed that low educational attainment was associated with pain and disability due to osteoarthritis at any reported joint\textsuperscript{50}. However, the study was limited to a cross-sectional analysis, and only studied individuals with knee osteoarthritis, without any consideration for other types of osteoarthritis.

Further studies from the Johnston County Osteoarthritis Project (JoCo OA) have also shown that occupation and community-based measures, such as block-group poverty, are associated with both knee and hip osteoarthritis outcomes\textsuperscript{51,52}. Each of these studies is limited to focusing on either knee or hip osteoarthritis. Furthermore, these analyses only utilized baseline data, as opposed to longitudinal data, which prevented any discussion of disability and progression. When looking at either form of osteoarthritis, both Callahan
et. al. and Cleveland et. al. found that both educational attainment and community based SES were associated with OA\textsuperscript{51,52}. Additional studies conducted across North Carolina in family practice clinics have shown that there is in fact a correlation between lower SES, specifically educational attainment, community poverty, and health assessment measures in patients with self-reported arthritis\textsuperscript{53–55}.

Only one prior study has included home ownership as a measure of SES, and showed no association with health outcomes. However, it was concluded that it remains to be an important measure of SES for future studies in search of an association with osteoarthritis\textsuperscript{53}. Few studies have looked at the association between income and chronic conditions. One population based cohort of individuals aged 45 and older in Australia concluded that there is an additional burden placed on individuals who are financially disadvantaged, especially in an aging population\textsuperscript{56}. This study solely focused on income and its association with a myriad of conditions\textsuperscript{56}. However, it gives premise to look at income as a potential risk factor in a developed country with a similar age distribution as the Johnston County Osteoarthritis Project cohort.

h. Summary

Osteoarthritis is a chronic disease that can lead to lack of mobility with increased use of the joints, specifically the hip and knee joints. Disability due to osteoarthritis is often manifested in difficulty completing daily living tasks, and is more common in older individuals. Several known risk factors, such as age, race, gender, and BMI, have that are associated with increased OA-induced disability. The association between SES-related measures and OA have also been explored, but with no conclusive results.
Collectively, current evidence suggests that socioeconomic status may be associated with osteoarthritis-related disability and consequently, progression. This study was designed to contribute to the existing body of literature about the relationship between SES and OA disability by exploring this association using longitudinal data collected from participants with radiographic knee and/or hip osteoarthritis in the Johnston County Osteoarthritis Project in rural North Carolina.
2. AIMS AND HYPOTHESIS

The aim of this study was to describe the independent relationship between SES measures of education, occupation, household poverty, home ownership, and income, with increased disability progression among individuals with radiographic knee and/or hip OA.

a. Hypothesis

i. Disability progression due to osteoarthritis, differs across individual-level SES measures of education, occupation, home-ownership, and income
3. METHODS

a. The Johnston County Osteoarthritis Project Design

This study was conducted among participants enrolled in the Johnston County Osteoarthritis Project, an ongoing population-based cohort study based in rural North Carolina. The aims of the Johnston County Osteoarthritis Project were to determine the risk factors associated with knee and hip osteoarthritis, in addition to prevalence and incidence estimates. Further aims and descriptions have been previously discussed\(^\text{57,58}\).

JoCo OA was established in 1991 in Johnston County, North Carolina as a partnership between the University of North Carolina, Chapel Hill and the Centers for Disease Control (CDC) to engage in community-based research in rural areas. Between 1991 and 2011, the study completed two enrollment waves and four waves of data collection, enrolling nearly 5000 participants.

The first enrollment wave occurred between May 1991 and December 1997. During this baseline time period (T0), 3187 individuals were enrolled and had full participation in the study. The first 5-year follow-up window (T1) occurred between 1999 and 2004 and showed a substantial loss to follow-up, with only 1733 (54.4%) individuals returning for the T1 visit. As a result, 1150 individuals were added in an enrichment cohort known as T1*. For individuals who were first enrolled at T1*, T1* was considered the baseline for these participants and were followed with the same protocol as T0 enrollees. The second 5-year follow-up visits (T2) occurred between 2006 and 2010 and showed a study population of 1708 individuals (62.0%), composed of baseline-enrolled individuals coming in for their second follow-up, as well as enrichment-enrolled (T1*) individuals coming in for their first follow-up (Figure 1).
Study participants were civilian, non-institutionalized individuals 45 and older who were living in Johnston County at the time of enrollment. Participants had to be mentally and physically capable of completing the study protocol, which included radiographic procedures, a myriad of physical examinations, and home interviews\textsuperscript{58}.

This study was approved by the Institutional Review Boards of the University of North Carolina Schools of Medicine and Public Health and the Centers for Disease Control and Prevention. All participants gave written informed consent at the time of enrollment.

b. **Study Design and Population**

This longitudinal study was designed to explore the association between SES measures and disability due to osteoarthritis, using prospectively collected data. HAQ scores were collected during time points T0, T1, T1*, and T2, and only data for individuals with knee and/or hip osteoarthritis with at least 2 HAQ scores were included. Based on these criteria, 975 individuals were eligible for analysis.

Participants were spread across the baseline cohort and the enrichment cohort. The baseline cohort included three of four different cohort designations: B0F1 (information at T0 and T1), B0F2 (information at T0 and T2, but missing at T1), and B0F1F2 (information at T0, T1, and T2). The enrichment cohort accounted for the last cohort designation, N1N2 (information at T1* and T2).

Due to varying methods of analysis in this study, data were formatted in two ways, per-person, as well as per-visit, to assess disability progression over time. Figures 2a and 2b outline the exclusion criteria for participants and visits, with Figure 2a outlining these criteria per participant, and Figure 2b per visit. In brief, of the 975 eligible individuals for
analysis, 2 observations were excluded due to missing baseline education, 50 excluded for missing baseline occupation, and 37 excluded due to missing baseline household poverty. The final analytic samples included 886 participants with knee and/or hip osteoarthritis and 1,096 complete visits contributed by 541 individuals. Overall, 42.6% of individuals showed up for all three visits, baseline and two follow-ups.

c. Outcome Assessment

The primary outcome in this study was disability among individuals with knee and/or hip osteoarthritis. A proxy measure of disability was assessed using the HAQ-DI, an 8 section questionnaire containing information on dressing, grooming, arising, walking, reaching, eating, hygiene, and grip. This information was ascertained during home interviews conducted by JoCo OA project staff members in an effort to increase retention by catering to patient convenience.

Each of the 8 sections contains 2-3 questions, each with a range of 4 answer choices on a Likert-like scale: 0 – without any difficulty, 1 – with some difficulty, 2– with much difficulty, 3 – unable to do. To account for underlying disability of an individual, the use of an aid or device increases the lowest score in the section to at least 2. For example, if an individual scored questions as 0 and 2 within a given section, the 0 in that section will increase to a score of 2. Each section receives its own score, with the highest score in that section determining the score for the category. An example of part of the questionnaire can be seen in Figure 3.

For this study, disability progression was defined in two ways. In the time-to-disability progression analysis, disability progression was defined as experiencing an increase in
the HAQ score of greater than 0.22 between baseline and last visit, which has been shown to be the minimal clinical important difference for the HAQ \(^{38}\). For the longitudinal analysis, disability progression was defined as an increase in HAQ score over time.

d. Primary Exposures

The primary exposures for this study included four measures of individual socioeconomic status (educational attainment, occupation, home-ownership, and income) and one measure of community-level socioeconomic status (household poverty level).

Education was categorized as completion of less than twelve years of formal primary education or twelve or more years. Occupation was categorized as low or high SES jobs, based on US census classifications. Home ownership was classified as ownership of one’s own home or not. Income was classified to represent an income of less than or equal to $30,000 or greater than $30,000. The cutoff of $30,000 was based on the median income at baseline.

Household poverty rate was measured as a continuous variable, determined from the percentage of households with income below the poverty level of a given US Census block group. The baseline value for household poverty level was based on the 1990 census, with the threshold values changing at follow-up visits based on the most recent census\(^{58}\). For ease of interpretation, poverty was scaled by a factor of 10 to indicate the effect of a 10% change.
e. Potential Confounders

Demographic variables were also included in our analysis as potential confounders, as they have been shown to be associated with all five SES measures, as well as disability due to osteoarthritis. The directed acyclic graph (DAG) shown in Figure 4 demonstrates this pictorially for age, race, and gender. Age was included as a continuous variable, and was centered at the mean age in the sample, and scaled by 5. Race was classified as either Caucasian or African-American, as JoCo OA explicitly studies osteoarthritis in these two racial groups. Sex was classified as male or female. A summary of how all variables were coded and their reference groups can be seen in Table 1.

f. Statistical Methods

Our exploratory analyses looked at the distributions of all covariates and potential confounders graphically and numerically through statistical summaries.

Our initial method of analyzing the JoCo OA data focused on statistical summaries of each of the five SES covariates in a time to event setting to examine the progression beyond a certain threshold. The decision to analyze these data in a survival context was unconventional, but efforts were taken to verify certain distributional assumptions that were made with this approach.

Data were reformatted from HAQ scores at each visit to time to event by dichotomizing the change in HAQ score between visits, where an increase of +0.22 from baseline indicated the event of disease progression. Due to long durations between time points per individuals, interval censoring was used to account for the uncertainty in knowing when an event occurred in a given time-interval, if at all. This was implemented within R using
the WeibullReg function in the SurvRegCensCov package, and specifying interval as the form of censoring. Under the assumption of a monotonically increasing hazard, a Weibull distribution was chosen to describe the data. This parametric modeling approach, which incorporates the proportional hazards assumption, provides a means of accounting for interval censoring. Log hazards were modeled as a linear function of demographic variables and SES covariates. In order to assess the Weibull distributional assumption, the shape parameter was assessed for significance.

Our second analysis examined the independent association between each of the SES measures and disability by assessing the HAQ score on a continuous scale as a proxy for disability. We modeled disability using linear mixed models (LMM) to account for the longitudinal nature of the data, as well as the within-person correlation over time. For univariate analyses, a baseline LMM was specified to include random effect terms in the form of subject (intercept) and age (slope), as well as fixed effects in the form of demographic variables (age, race, and gender). Age was our chosen metric of time. A first set of models considered the SES measures as main effect terms, allowing for characterization of disability status in terms of these variables. This model was then used as a comparison to independently evaluate each of the 5 SES variables in the subsequent models. A second set of models additionally included interactions between age and the SES variables, allowing for characterization of disability progression. Residual diagnostics were evaluated for all models in the form of residual vs. fitted and QQ plots to assess assumptions.

In addition to monitoring the association between disability progression and each of the SES covariates, we performed a factor analysis to determine whether latent factors might
be used as a construct to better categorize the five SES variables for future use. Two sets of exploratory polytomous factor analyses using maximum likelihood estimates were conducted to determine an appropriate combination of the 5 measures of socioeconomic status. The primary analysis looked only at education, occupation, and poverty, as these variables were present in both the baseline and enrichment cohorts. A subsequent analysis looked at these 3 covariates, in addition to home ownership and income. These analyses utilized a varimax rotation to maximize the variance of the squared loadings of a factor. Results were explored through numerical summaries in the form of factor loadings, with communalities reported as a measure of fit, as well as graphically through correlation plots.

Statistical significance was defined as a p-value less than the critical threshold of $\alpha = 0.05$ for a 2-sided test. All analyses were conducted in R version 3.3.1 (R Core Team, 2013).
4. RESULTS

a. Characteristics of the Study Population at Baseline

Baseline characteristics of the study population are shown in Table 2. Overall, 886 participants were included in the study, 678 (76.5%) of whom were enrolled during the T0 recruitment period. The average age of the study participants was 62.6 years, with those included in the baseline cohort being slightly older than those included in the enrichment cohort (62.9 and 61.8, respectively, p=0.15). Women comprised 65.7% of the participants in the entire study population, and this did not differ significantly between the two recruitment cohorts (p=0.89).

Our sample included nearly twice as many Caucasian individuals (67%) as African-Americans. There was a greater percentage of Caucasians in the initial cohort, but this did not differ significantly between the two recruitment cohorts (68.9% and 62.0% respectively, p = 0.08). African-American individuals tended to be younger than their Caucasian counterparts. However, the two race groups did not differ by age (p=0.21).

Study participants were educated, with 64% having at least a high school degree. Interestingly, educational attainment was significantly higher among the enrichment cohort than among the baseline cohort (75.5% ≥ HS vs. 60.5% ≥ HS, respectively; p < 0.001).

Home ownership and income were ascertained only during the T0 recruitment period. Among the 678 individuals in the baseline cohort, 538 (79.4%) owned a home at baseline. Income information was only available for 541 participants, resulting in an almost 20% missing rate for the covariate, as 137 individuals neglected to provide this
information during their baseline visit. Of those who reported income, 78.6% reported having an income of ≤ $30,000.

Occupation was categorized into high and low SES jobs, approximately 58% of individuals reported having a low SES job. There was no significant difference in occupation level between the baseline and enrichment cohorts (p = 0.19).

The average household poverty across all individuals was approximately 20% for both the baseline and enrichment cohorts. Overall, there was no difference in average household poverty level between the baseline and enrichment cohorts (p = 0.96). While significant differences were not seen when analyzing household poverty as a continuous variable, there were significant differences between the baseline and enrollment groups when household poverty was analyzed in tertiles based on the 1990 Census (p = 0.03). Two-thirds of individuals lived in areas with a household poverty level greater than 15.1%. 18.0% of Caucasians lived in an area where the poverty level was greater than 23.4%, compared to the 59.7% of African-Americans who lived in the highest poverty area.

b. Time-to-Event Analysis of Disability Progression

We began by fitting a parametric Weibull model with interval censoring to the disability progression data where the scale parameter indicates the variability of the distribution and the shape parameter determines whether the hazard function is constant, increasing, or decreasing. To assess the need for a Weibull distribution, as opposed to an exponential distribution, a null model with no covariates was fit to assess the significance of the shape parameter. This was found to be significant (p < 0.001) and justified use of the Weibull
distribution with over the exponential distribution for this analysis. From this model, the estimate for the scale parameter was 0.04 (SE = 0.01) and for the shape parameter was 1.43 (SE = 0.08). Across all risk factors and confounders, the scale parameter was quite close to 0, indicating that the distribution was not stretched (Figure 5).

To determine if there was an underlying difference in the risk of progression between the baseline and enrichment cohorts, we fit a Weibull model with interval censoring, including a single indicator variable for the recruitment cohort. From this model, the estimate for the scale parameter was 0.03 (SE = 0.01) and for the shape parameter was 1.56 (SE = 0.09). The estimated progression hazard ratio comparing the enrichment cohort to the baseline cohort was 1.87 (95% CI: 1.48, 2.35), demonstrating that progression differed between the two recruitment cohorts.

The univariate associations among the baseline cohort using the Weibull assumption showed that age (HR = 1.02; 95% CI: 1.01, 1.03), gender (HR = 1.58; 95% CI: 1.27, 1.97), lower educational attainment (HR = 1.48; 95% CI: 1.20, 1.81), and an income ≤ $30,000 (HR = 1.85; 95% CI: 1.38, 2.48) were each significantly associated with faster disability progression.

When the analysis included all individuals looking at education, occupation, and household poverty, similar associations between each of the covariates and disability progression were seen. Age had a significant effect on disability (HR = 1.02; 95% CI: 1.01, 1.03). Gender (HR = 1.51; 95% CI: 1.24, 1.83) and education (HR = 1.34; 95% CI: 1.12, 1.61) were also shown to be significant across all individuals.
A summary of the multivariable results is provided in Table 4. We performed two multivariable analyses using the baseline cohort, one excluding income and home ownership which were not ascertained in the enrichment cohort, and the other including these two additional SES measures. In the analysis excluding income and home ownership, the likelihood of disability progression was significantly higher among women compared to men (aHR = 1.51; 95% CI: 1.20, 1.89). Education was also independently associated with disability progression with those who did not complete high school having an increased likelihood of progression compared to those who had received a high school diploma (aHR = 1.34; 95% CI: 1.06, 1.70).

The inclusion of the income and home ownership variables in the baseline cohort analysis reduced the sample size of complete cases from 678 to 541 participants due to missing income data. After accounting for these two additional SES measures, men were again found to have a significantly higher likelihood of experiencing disability progression compared to women (aHR = 1.59; 95% CI: 1.22, 2.07). Interestingly, income replaced education attainment as the only SES measure that remained independently associated with progression. Specifically, individuals who reported making $30,000 or less were more likely (aHR = 1.45; 95% CI: 1.03, 2.04) to experience disability progression compared to those with an income higher than $30,000.

The third time to event analysis we performed included all 886 participants enrolled in the baseline and enrichment cohorts. This analysis included the same covariates that we included in the first model since data for home ownership and income were not available for the enrichment cohort. The results for this combined analysis that included all study participants yielded the same results as those obtained from the first model with older age.
(aHR = 1.01; 95% CI: 1.00, 1.02), female gender (aHR = 1.48; 95% CI: 1.21, 1.80), and lower educational attainment (aHR = 1.32; 95% CI: 1.07, 1.63) being independently associated with disability progression. We also found that progression was significantly more likely among participants enrolled in the baseline cohort (aHR = 1.98; 95% CI: 1.57, 2.50), but the estimated hazard ratios for the other variables that were independently associated with progression were very similar, regardless of whether or not the enrichment cohort was included in the analysis.

In summary, utilizing this time to disability analysis among the baseline cohort, we found that older individuals and women had a higher likelihood of disability progression compared to younger individuals and men, respectively. The same demographic and SES measures were significant in the comparable models that did and did not include data from the enrichment cohort, with income replacing education as the significant SES factor in the expanded model fit to data from the baseline cohort.

c. Linear Mixed Model Analysis of Disability Progression

The second approach we used to examine the association between SES measures and disability progression consisted of a longitudinal data analysis (LDA) using linear mixed models. For this LDA, we first fit three separate models to examine the associations between the SES variables and disability status. A summary of the models run using linear mixed models are shown in Table 5. For Model 1, we regressed HAQ score on the selected demographic and SES measures, except for income and home ownership among the baseline cohort. Age is used as a metric of time, and an increase in age by 5 years ($\beta = 0.05; 95\% \text{ CI: } 0.03, 0.07$) was independently associated with an increased disability status.
Female gender ($\beta = 0.23; \text{95\% CI: 0.15, 0.31}$) was also significant in Model 1, with women having an average HAQ score that was higher than that of men. This model suggests that a 62-year old white woman with at least a high school education has a HAQ score of 0.23 greater than a 62-year old white male with similar characteristics, when not accounting for income and home ownership. Unlike the other demographic variables, race was not significantly associated with the HAQ score ($\beta = 0.05; \text{95\% CI: -0.05, 0.15}$). Additionally, an increase in household poverty by 10% was independently associated with an increase in the HAQ scores ($\beta = 0.06; \text{95\% CI: -0.02, 0.05}$) and disability status.

For Model 2, we regressed HAQ score on all selected demographic and SES measures, including income and home ownership. Despite the addition of these 2 SES measures, all the results were consistent with those seen in Model 1, including the magnitude of the beta coefficients. Additionally, lack of home ownership was also significant ($\beta = 0.13; \text{95\% CI: 0.02, 0.24}$), with increases in the HAQ score for participants who reported home ownership information.

Model 3 included all the same covariates as in Model 1, but also introduced a baseline cohort indicator. Similar to the time to event analyses, Model 3 included the entire cohort of participants enrolled in both the baseline and enrichment cohorts. In doing so, the results for age and gender were found to be consistent with those from both Models 1 and 2. Household poverty was the only SES covariate that was independently associated with the disability score ($\beta = 0.05; \text{95\% CI: 0.01, 0.09}$). However, having a low SES occupation was found to be borderline significantly associated with disability ($\beta = 0.06; \text{95\% CI: -0.001, 0.13}$). Finally, there remained a statistically significant disability difference between the two cohorts, with the adjusted HAQ score being lower among
those enrolled at baseline compared to those included in the enrichment cohort (β = -0.16; 95% CI: -0.25, -0.08).

Our next step was to examine whether the effects of these SES measures increased with age. To accomplish this, we augmented our models by adding interaction terms between each of the SES measures and age to each of the three models described above, the results of which are summarized in Table 6.

In Model 1, the interactions of age with education (β = -0.04; 95% CI: -0.08, -0.01) and with household poverty (β = 0.02; 95% CI: 0.00, 0.04) were both statistically significant. However, the interaction of age with occupational category was not (β = 0.00; 95% CI: -0.03, 0.04). The interaction of age with education indicates that HAQ scores among those with lower educational attainment increase more as their age increases compared to those with higher education. In contrast, the significant interaction between age and poverty demonstrates a synergistic effect, where the effect of poverty is higher at older ages.

With regards to the main effects in Model 1, gender was found to be significant (β = 0.23; 95% CI: 0.15, 0.31) with women again estimated to have a higher average HAQ score than males with reference levels of other covariates. Household poverty was the only SES covariate that had an independent association with disability after accounting for the interaction terms (β = 0.05; 95% CI: 0.01, 0.10). Lack of further educational attainment and low SES occupation were not found to be significant.

In Model 2, the only interaction that was significant was that of age and household poverty (β = 0.03; 95% CI: 0.01, 0.05). For all other SES measures, none of the
interaction terms with age were independently associated with the average HAQ score, adjusting for other covariates in the model. While the interaction between home ownership and age was not significant, the main effect of lack of home ownership was significantly associated with an average increase in HAQ score by 0.14 ($\beta = 0.14$; 95% CI: 0.02, 0.25). Additionally, the effect of being a female was consistent with that seen in Model 1 ($\beta = 0.04$; 95% CI: -0.07, 0.14).

Model 3 mirrors the synergistic effect seen in the interaction between age and household poverty ($\beta = 0.02$; 95% CI: 0.00, 0.03). This was the only interaction term between age and an SES factor which was significant in this model. With regards to main effects, there was an association between female status and increased HAQ scores, as well as with increases in household poverty by 10% ($\beta = 0.05$; 95% CI: 0.01, 0.08).

d. **Exploratory Factor Analysis of Socioeconomic Status Factors**

Our next step was to examine the possibility that SES measures included in this analysis could provide evidence that one or more underlying factors explain the association between these SES measures and disability. In the first factor analysis, we focused on education, occupation, and household poverty. The factor loadings obtained using the data from the 886 participants are shown in Table 7. We then added in home ownership and income to a second factor analysis. Since this information was only available for the baseline cohort, these results are based on the subset of 541 participants with complete data.

In looking only at education, occupation, and household poverty, we observed that education and occupation loaded onto a factor with education having a correlation of 0.52
and occupation a correlation of 0.67 with this factor. Household poverty was not strongly correlated with a common underlying factor which was related to occupation and education. The amount of variance in each variable that is accounted for, or the communality, is also shown in Table 7. Occupation has the most variance explained among the three variables, with the variance in household poverty not adequately being explained.

In the second factor analysis where all 5 SES variables were used, a single factor analysis resulted in all SES measures except household poverty loading on to the factor. The communalities for all five variables are in concordance with this, as household poverty again has the lowest value. Education and occupation had the highest factor loadings of 0.60 and 0.62, respectively. We also allowed for two factors, and found that education and occupation again loaded heavily onto the first factor. Poverty was the only SES variable to load very heavily onto the second factor (0.99), and the communality of 1 reflects this finding that poverty is a largely independent variable. The amount of variance explained by each of the variables is comparable to the single factor analysis using all 5 variables, with the exception of household poverty.
5. DISCUSSION

This study of the relationship between measures of SES and disability due to knee and/or hip osteoarthritis was designed to examine disability in terms of both progression and its relationship with increasing age. In the time to disability progression analysis, females and lower educational attainment were independently associated with increased risk of progression, while income ≤ $30,000 was independently associated in the subset of the study population who provided income information. In the analysis that examined the association between disability status and SES, we demonstrated that HAQ score increased significantly with age, and was higher among females, people living in a higher household poverty area, and those who did not own their own homes. After adjusting for the effect of age, females gender remained associated with increased disability status. Interestingly, the observed disability increase with age was greater among those living in higher poverty areas compared to those living in lower poverty areas. Additionally, lack of home ownership was independently associated with increased disability. Overall, the SES measure of lower educational attainment and increased household poverty were associated with increased disability status or progression in the study population, and could be strong markers for predicting disability progression.

Few studies have examined the independent association of socioeconomic status with osteoarthritis-induced disability progression, as many rely on cross-sectional analyses. With cross-sectional analyses, no conclusions can be made about the temporal sequence between exposure and outcome. Our study focused on disability progression with aging, which allowed for assessing how the SES measures of interest were temporally related to disability progression as age increased. For this study, we defined
socioeconomic status with five different variables and disability progression based on the HAQ score. Accordingly, comparisons to prior literature can be made on a per-variable basis.

With regards to the association between education and disability status and progression, it has been shown that lower educational attainment was independently associated with both radiographic and symptomatic hip OA compared to those with higher educational attainment\textsuperscript{52}. Similar results were seen in previous studies looking at knee osteoarthritis, but these studies were based on cross-sectional analyses\textsuperscript{50,51}. Despite a different outcome by using the HAQ score, and use of age as a proxy for time in our study, as opposed to time since enrollment, our results were consistent in showing that lower levels of education were associated with increased disability.

A small number of studies have looked at poverty and its association with osteoarthritis. One such study looked at the health-related quality of life in Caucasians and African-American individuals in North Carolina, and found that better quality of life measures were associated with lower levels of household poverty\textsuperscript{54}. However, the measures used to define quality of life did not include the HAQ score as a measure of osteoarthritis disability. Using the same source population in Johnston County, disadvantaged Caucasians and African-Americans who lived in areas with increased household poverty were shown to have an increased odds of self-reported osteoarthritis compared to those that were living in areas with less household poverty, with African-American individuals having increase odds compared to Caucasians\textsuperscript{55}. All of these studies share the same overall conclusion that higher community-based poverty levels are associated with decreased quality of life with regards to osteoarthritis-induced disability, in concordance
with our mixed effects analysis where poverty was shown to accelerate the HAQ score increase as individuals aged. This is interesting to note, as poverty was a community-level measure, as opposed to the individual-level nature of the remaining SES variables.

Among studies that have included home ownership as a factor of socioeconomic status, an independent association between home ownership and increased HAQ scores was observed in a cross-sectional analysis\(^5\). These investigators took other demographic covariates into account and reported that each of the five socioeconomic factors was independently associated with poorer HAQ scores\(^5\). However, when all SES measures were included in the final model, there was no association between home ownership and HAQ scores. Our results differed in that home ownership was not associated with disability progression, but lack of home ownership was significantly associated with a higher HAQ score in the longitudinal analysis.

Several investigators have examined the association between income and disability. Korda et al. observed an increase in the prevalence of osteoarthritis and a decrease in physical function among older individuals with a lower income; however, their analysis was cross-sectional and did not account for race as a demographic covariate\(^5\). Another study also found that income was significantly associated with decreased physical function through HAQ scores, and found this to remain an independent risk factor in a multivariable analysis\(^5\). Our study was consistent with findings from both of studies that income was significantly associated with disability progression; however, our study accounted for demographic factors, such as race, and allowed for analysis of this association as individuals aged.
The relationships and interactions among the 5 SES measures are of importance for understanding the latent concept of SES. The factor analysis conducted in this study indicated that there may be some factors that underlie co-occurrence of social determinants, with particularly high correlations between education and occupation. In the 1-factor analysis conducted with education, occupation, and household poverty across all participants, household poverty did not contribute any additional information in establishing a latent factor. The two-factor analysis showed a similar result that household poverty heavily loaded onto a second factor, and was the only measure to do so. One possible explanation for this finding is that household poverty is a community-level variable, while education and occupation are both individual-level, suggesting that individual-level measures play a greater role in this study. Furthermore, the latent process that the other four SES measures are related to may not reflect the same process of household poverty.

When all 5 SES variables were used in the smaller subset of individuals, all variables except household poverty loaded onto a single latent factor and household poverty loaded completely onto the second factor. Education and occupation maintain a stronger association with the underlying factor when income and home ownership were introduced, suggesting that income and home ownership reflect an underlying concept of SES similarly as education and occupation. The second factor appears nearly entirely reflected by the community-level factor of household poverty itself. As a result, the development of programs that address both individual and community-level needs may have an impact on reducing the burden and progression of osteoarthritis-induced disability.
In our study cohort, there was a strong association between gender and disability status across analyses. This is corroborated by the fact that women are in fact at higher risk for developing osteoarthritis and suffer from more severe forms in both their knees and hips.\textsuperscript{2,14–16} Compared to previous studies which included race as a demographic risk factor, our findings that OA disability did not differ by race are inconsistent with those findings that African-Americans experience an increased risk of disability progression.\textsuperscript{2,14,15} However, our study was limited to analyzing Caucasians and African-Americans without also including other ethnic groups, such as Hispanics or Asians, which may have shed more insight on how disability progression manifests in other racial subcategories.

Our study showed that low income is strongly associated with disability progression in participants, using a time to event definition. However, under a longitudinal analysis, there was no association between income and disability status or progression. While the lack of association between income and disability progression is inconsistent with the results from previous studies,\textsuperscript{53} it is possible that our results are biased, or that these studies have less power due to the availability of data. Income information was only available from participants in the baseline cohort, and even then, there were individuals who did not respond. While the difference between those who did and did not report their income might have resulted in data that were missing at random, a selection bias is plausible such that individuals were more likely to report low income if it meant that they had a better chance of receiving care, or better care, for their osteoarthritis.

Despite the incomplete data and separate analyses for those who had maximum SES information and those who did not, there was a prevailing association between gender
and HAQ score across all analyses. When interaction terms were introduced in the longitudinal model, we were able to assess associations between each of the covariates and disability progression with age. While household poverty remained significant, the interaction with age was also significant, suggesting that household poverty status can be an important indicator of long-term progression, as poverty status can change as participants age. This relationship also holds true when looking at different subsets of the study population. However, the finding that the rate of disability progression over time decreased for those with lower education compared to higher education is unexpected. Education status was not shown to be associated with disability status, and it’s possible that as people age, having a particular educational attainment in combination with knowledge about the disease affects the way a participant manages their disease.

Our mixed models analysis of progression focused on progression with respect to SES measures and not demographic variables. SES measures were our main exposures of interest, and we sought to examine their association while controlling for demographic factors such as race and gender. By focusing our analysis on the association between the SES measures and their association with disability progression over time, we were able to maximize power to detect progression for these factors. However, because no interactions were included between age and the demographic variables, our model may conflate progression with these variables and progression with respect to the demographic variables. A sensitivity analysis was done to explore the interaction of age with gender to assess for an association between gender and disability progression, which was not found to be significant.
Strengths and Limitations

Our study population only included individuals who were 45 years and older at their baseline visit and either Caucasian or African-American. In a CDC report based on prevalence estimates of osteoarthritis between 2003 and 2005, among individuals 65 and older, 50% had osteoarthritis, while the prevalence estimate was around 29.3% for individuals aged 45-64 years\textsuperscript{60}. In contrast, our study population had a greater prevalence among 45-64-year-old individuals than 65 and above individuals, with prevalence estimates of 59% and 41%, respectively. With respect to race, Caucasians composed roughly 67% of the study population, and African-Americans 33%. Based on the same CDC report, prevalence was greatest among non-Hispanic Caucasians (24.3%) followed by non-Hispanic African-Americans (19.2%) and Hispanics (11.4%)\textsuperscript{60}. Our population was primarily located in a rural setting, and had a generally high socioeconomic status. These characteristics are specific to our study population, despite having limited access to osteoarthritis prevalence estimates in the US. As a result, our results may not be generalizable to other populations of osteoarthritic individuals where population distributions differ greatly from ours.

Another limitation of our study is the manner in which our 5 SES variables and outcome were collected and analyzed. Educational attainment, occupation, home ownership, income were all treated as dichotomous variables. More granular information about each of these measures could have helped the validity of results. All of the aforementioned variables were collected on an individual-level, in contrast to household poverty, which was a community-level variable. Household poverty tertile cutoffs were based on the
most recent Census, this can bias the association between poverty and disability or progression, as these cutoffs can change over time.

With regards to the outcome, HAQ score, which can be easily ascertained via examination or by self-report, would have ideally been collected every six months to a year with extensive efforts to follow-up with participants. However, due to funding and other resource limitations, an individual had a maximum of 3 time points worth of demographic and outcome data. Furthermore, inclusion criteria for this study required that individuals had radiographic knee and/or hip OA and had at least two time points of data, which excluded many individuals from our study. As a result, the collection of HAQ data for only three data collection periods proved to be a limiting factor in our study population. For individuals who had only two time points of data, the conclusions of our analysis that certain risk factors are associated with increased disability progression may be biased.

In the larger JoCo OA study, there was a substantial loss to follow-up between T0 and T1, which required the cohort to be refreshed at T1*. It is possible that individuals who did not return for the first follow-up were healthier individuals, and may have been leading otherwise active lives and were too busy to return. Unfortunately, lack of information about participants that were lost to follow-up precludes us from assessing whether or not being lost to follow-up was related to disability progression.

Even in an older population, SES measures may change significantly over time, with the possible exception of education. Observations would need to be confirmed in a longitudinal cohort of individuals where there was complete information about all
covariates at each visit. While there were few differences between the baseline and enrichment cohorts in terms of demographic confounders and social determinants, the study could have benefited from more frequent information over a longer period of time to better assess the temporality of the effect between SES and disability progression. Our observation that income $\leq$ $30,000 has a strong impact on disability progression may in fact change over time if income is treated as time-varying with more data points.

Additional limitations exist with the methods employed in our study. The use of parametric methods to model time to event data allowed for the incorporation of interval censoring to account for the lack of ascertainment of exactly when the outcome (progression) occurred. However, because there are long intervals and relatively little variation in these intervals, our findings may be highly dependent upon the modeling assumption. Further data collection and more frequent assessments of the outcome per individual could have provided for a better evaluation of our Weibull distributional assumption.

Despite the aforementioned limitations of our analysis, our study did have some notable strengths, including a large sample of participants with longitudinal outcome measures. There were a number of individuals that had 2 follow-up visits. This provided valuable insights into disease progression over time and allowed us to look at different measures of the outcome, both as time to event and a change in HAQ score, as missing data eliminated less than 10% of initial observations.

The use of multiple statistical models provided a means of exploring the aims of this study from different perspectives which allowed for the drawing of more robust
conclusions than would generally be possible from a single analysis. In using the mixed effects models to assess the effect of SES on the association of disability with age, we were able to make the most of our available, albeit limited, data.

Our use of multiple socioeconomic measures can also be seen as an advantage in our analysis, as it allowed us to explore the associations that these variables had not only with the outcome, but other factors as well. Future analyses could benefit by incorporating the results of this factor analysis as variables in a time to event or longitudinal model, to better understand the relationship between SES and osteoarthritis in the context of disability progression.

*Public Health Significance*

Osteoarthritis-induced disability progression is of critical importance as the U.S. population ages. Currently affecting over 27 million individuals in the US aged 18 or older, OA often interferes with productivity and daily activities, which can result in joint replacements and inordinate socioeconomic costs. Disability due to osteoarthritis is often largely overlooked, as there is a general notion that symptomatic OA is an inevitable part of aging, and not a chronic disease that can be pacified with proper management. The conversation around osteoarthritis is often centered around clinical comorbidities and modifiable risk factors, such as excess body mass and chronic conditions such as diabetes or heart disease. However, osteoarthritis affects a large number of individuals from various socioeconomic backgrounds, each of whom may face different risk factors. Evidence-based interventions to prevent and manage disability due
to osteoarthritis should be available to those who need them in a variety of community-based settings.

Future studies of osteoarthritis-induced disability outcomes should assess these indirect associations among individuals with knee and/or hip osteoarthritis. Societal interventions to address specific SES measures on an individual and community-level are needed to limit the impact of OA on disability progression. Furthermore, educational attainment is likely a marker for a constellation of things that can help individuals with OA manage their condition better, essentially self-efficacy. Individuals with OA can certainly become more educated about their condition to help defer this disability and could benefit from resources targeting self-management interventions and free access to interventions to help the slowing of disease progression.
Conclusion

In summary, our study findings suggest that lower educational attainment, and lower income may be the socioeconomic factors that are most predictive of disability progression among individuals with knee and/or hip osteoarthritis. In contrast, under a longitudinal definition where disability status and progression are assessed, increased levels of household poverty and lack of home ownership are most associated with the outcome. Although prior studies have indicated similar associations between education and osteoarthritis, these studies did not assess the longitudinal effects of all five SES measures, and were based on radiographic OA\textsuperscript{51,52}. Our study showed that across all analyses, household poverty was associated with an increased risk of disability progression, and with higher disability status over time, perhaps due to the nature of work in a rural area and the underlying relationship with lower income, which has not been extensively explored. Thus, there may be an indirect link between socioeconomic status factors and disability progression through levels of physical activity and other comorbidities, and future research will be required to examine their relationship further.
6. REFERENCES


28. Osteoarthritis: Diagnosis and Treatment - American Family Physician.


34. Ray Fleming O of C and PL. Handout on Health: Osteoarthritis.


7. TABLES

TABLE 1: Summary of Demographic and SES Measures in the Study

<table>
<thead>
<tr>
<th>VARIABLE*</th>
<th>REFERENCE GROUP</th>
<th>LEVEL OF MEASUREMENT</th>
<th>MISSING DATA (per person)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Mean age</td>
<td>Continuous (centered at mean age and scaled by a factor of 5 years)</td>
<td>No missing data</td>
</tr>
<tr>
<td>Race</td>
<td>Caucasian</td>
<td>Dichotomous (African-American vs. Caucasian)</td>
<td>No missing data</td>
</tr>
<tr>
<td>Gender</td>
<td>Men</td>
<td>Dichotomous (Women vs. Men)</td>
<td>No missing data</td>
</tr>
<tr>
<td>Education</td>
<td>Greater than High School (HS)</td>
<td>Dichotomous (&lt; HS vs. ≥ HS)</td>
<td>Negligible</td>
</tr>
<tr>
<td>Occupation</td>
<td>High SES Job</td>
<td>Dichotomous (Low SES job vs. High SES job)</td>
<td>5.1% missing</td>
</tr>
<tr>
<td>Household Poverty</td>
<td>0% Poverty</td>
<td>Continuous (scaled by a factor of 10 percentage points)</td>
<td>3.8% missing</td>
</tr>
<tr>
<td>Home Ownership</td>
<td>Own a home</td>
<td>Dichotomous (Don’t own vs. own)</td>
<td>No missing data (Only ascertained at T0)</td>
</tr>
<tr>
<td>Income</td>
<td>Greater than $30,000</td>
<td>Dichotomous (≤ $30,000 vs. &gt; $30,000)</td>
<td>20.2% missing (Only ascertained at T0)</td>
</tr>
</tbody>
</table>

* All exposures were analyzed as time-fixed, with the exception of age
TABLE 2: Baseline Characteristics of the Study Population

<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>Baseline Cohort</th>
<th>Enrichment Cohort</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>886</td>
<td>678 (76.5)</td>
<td>208 (23.5)</td>
<td>0.85</td>
</tr>
<tr>
<td>Age Categories</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45-54</td>
<td>241 (27.2)</td>
<td>181 (26.7)</td>
<td>60 (28.8)</td>
<td></td>
</tr>
<tr>
<td>55-64</td>
<td>282 (31.8)</td>
<td>214 (31.6)</td>
<td>68 (32.7)</td>
<td></td>
</tr>
<tr>
<td>65-74</td>
<td>255 (28.8)</td>
<td>198 (29.2)</td>
<td>57 (27.4)</td>
<td></td>
</tr>
<tr>
<td>75+</td>
<td>108 (12.2)</td>
<td>85 (12.5)</td>
<td>23 (11.1)</td>
<td></td>
</tr>
<tr>
<td>Mean(SD)</td>
<td>62.64 (9.94)</td>
<td>62.91 (10.03)</td>
<td>61.78 (9.61)</td>
<td>0.15</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td>0.08</td>
</tr>
<tr>
<td>Caucasian</td>
<td>596 (67.3)</td>
<td>467 (68.9)</td>
<td>129 (62.0)</td>
<td></td>
</tr>
<tr>
<td>African-American</td>
<td>290 (32.7)</td>
<td>211 (31.1)</td>
<td>79 (38.0)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td>0.89</td>
</tr>
<tr>
<td>Men</td>
<td>304 (34.3)</td>
<td>234 (34.5)</td>
<td>70 (33.7)</td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>582 (65.7)</td>
<td>444 (65.5)</td>
<td>138 (66.3)</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>≥ HS</td>
<td>567 (64.0)</td>
<td>410 (60.5)</td>
<td>157 (75.5)</td>
<td></td>
</tr>
<tr>
<td>&lt; HS</td>
<td>319 (36.0)</td>
<td>268 (39.5)</td>
<td>51 (24.5)</td>
<td></td>
</tr>
<tr>
<td>Household Poverty</td>
<td></td>
<td></td>
<td></td>
<td>0.03</td>
</tr>
<tr>
<td>&lt; 15.1%</td>
<td>280 (31.6)</td>
<td>228 (33.6)</td>
<td>52 (25.0)</td>
<td></td>
</tr>
<tr>
<td>15.1 – &lt; 23.4%</td>
<td>326 (36.8)</td>
<td>236 (24.8)</td>
<td>90 (43.3)</td>
<td></td>
</tr>
<tr>
<td>&gt; 23.4%</td>
<td>280 (31.6)</td>
<td>214 (31.6)</td>
<td>66 (31.7)</td>
<td></td>
</tr>
<tr>
<td>Mean(SD)</td>
<td>19.80 (10.39)</td>
<td>19.81 (9.61)</td>
<td>19.77 (12.64)</td>
<td>0.96</td>
</tr>
<tr>
<td>Occupation</td>
<td></td>
<td></td>
<td></td>
<td>0.19</td>
</tr>
<tr>
<td>High SES job</td>
<td>376 (42.4)</td>
<td>279 (41.2)</td>
<td>97 (46.6)</td>
<td></td>
</tr>
<tr>
<td>Low SES job</td>
<td>510 (57.6)</td>
<td>399 (58.8)</td>
<td>111 (53.4)</td>
<td></td>
</tr>
<tr>
<td>Home Ownership*</td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Own a Home</td>
<td>538 (79.4)</td>
<td>538 (79.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doesn’t Own a Home</td>
<td>140 (20.6)</td>
<td>140 (20.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income*</td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>&gt; $30,000</td>
<td>116 (21.4)</td>
<td>116 (21.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ $30,000</td>
<td>425 (78.6)</td>
<td>425 (78.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>137</td>
<td>137</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Income and Home Ownership were only ascertained for the baseline cohort
### TABLE 3: Univariate Analysis of Time to Disability Progression

<table>
<thead>
<tr>
<th></th>
<th>Baseline Cohort (N = 678)</th>
<th>All Individuals (N = 886)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>HR (95% CI)</td>
</tr>
<tr>
<td>Age (per 5 years)</td>
<td>1.02 (1.01, 1.03)</td>
<td>1.02 (1.01, 1.03)</td>
</tr>
<tr>
<td>Race (Black vs. White)</td>
<td>1.01 (0.81, 1.25)</td>
<td>1.07 (0.89, 1.29)</td>
</tr>
<tr>
<td>Gender (Female vs. Male)</td>
<td>1.58 (1.27, 1.97)</td>
<td>1.51 (1.24, 1.83)</td>
</tr>
<tr>
<td>Education (&lt;HS vs. &gt; HS)</td>
<td>1.48 (1.20, 1.81)</td>
<td>1.34 (1.12, 1.61)</td>
</tr>
<tr>
<td>Household Poverty (per 10% change)</td>
<td>1.08 (0.97, 1.19)</td>
<td>1.01 (0.97, 1.05)</td>
</tr>
<tr>
<td>Occupation (Low SES vs. High SES job)</td>
<td>1.15 (0.94, 1.41)</td>
<td>1.14 (0.95, 1.36)</td>
</tr>
<tr>
<td>Income*† (≤ $30K vs. &gt; $30K)</td>
<td>1.85 (1.38, 2.48)</td>
<td></td>
</tr>
<tr>
<td>Home Ownership* (Don’t own vs. own)</td>
<td>1.27 (0.99, 1.63)</td>
<td></td>
</tr>
</tbody>
</table>

*Income and Home Ownership were only ascertained for the baseline cohort.
†Income had missing values (N = 137), resulting in a sample size of N = 541
TABLE 4: Multivariable Analysis of Time to Disability Progression

<table>
<thead>
<tr>
<th></th>
<th>Model 1: Excluding Home Ownership and Income (N = 678)</th>
<th>Model 2: Including Home Ownership and Income (N = 541)</th>
<th>Model 3: Excluding Home Ownership and Income (N = 886)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>HR (95% CI)</td>
<td>HR (95% CI)</td>
</tr>
<tr>
<td>Age (per 5 years)</td>
<td>1.02 (1.01, 1.03)</td>
<td>1.05 (0.98, 1.12)</td>
<td>1.01 (1.00, 1.02)</td>
</tr>
<tr>
<td>Race (Black vs. White)</td>
<td>0.87 (0.68, 1.13)</td>
<td>0.80 (0.60, 1.07)</td>
<td>1.01 (0.81, 1.25)</td>
</tr>
<tr>
<td>Gender (Women vs. Men)</td>
<td>1.51 (1.20, 1.89)</td>
<td>1.59 (1.22, 2.07)</td>
<td>1.48 (1.21, 1.80)</td>
</tr>
<tr>
<td>Education (&lt; HS vs ≥ HS)</td>
<td>1.34 (1.06, 1.70)</td>
<td>1.06 (0.81, 1.39)</td>
<td>1.32 (1.07, 1.63)</td>
</tr>
<tr>
<td>Household Poverty (per 10% change)</td>
<td>1.08 (0.95, 1.22)</td>
<td>1.06 (0.93, 1.22)</td>
<td>0.99 (0.94, 1.04)</td>
</tr>
<tr>
<td>Occupation Category (Low SES vs. High SES)</td>
<td>1.00 (0.80, 1.26)</td>
<td>1.00 (0.77, 1.30)</td>
<td>1.05 (0.86, 1.28)</td>
</tr>
<tr>
<td>Income (≤ $30K vs &gt; $30K)</td>
<td></td>
<td>1.45 (1.03, 2.04)</td>
<td></td>
</tr>
<tr>
<td>Home Ownership (Don't Own vs. Own)</td>
<td>1.10 (0.81, 1.48)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline Cohort Indicator (Baseline vs. Enrichment)</td>
<td>1.98 (1.57, 2.50)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 5. Association between Disability Status and SES Measures and Demographic Factors by Age

<table>
<thead>
<tr>
<th></th>
<th>Model 1: Baseline - Excluding Home Ownership and Income</th>
<th>Model 2: Baseline - Including Home Ownership and Income</th>
<th>Model 3: All Participants - Excluding Home Ownership and Income</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β</td>
<td>95% CI</td>
<td>p</td>
</tr>
<tr>
<td>Age (per 5 years)</td>
<td>0.05</td>
<td>(0.03, 0.07)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Race (Black)</td>
<td>0.05</td>
<td>(-0.05, 0.15)</td>
<td>.33</td>
</tr>
<tr>
<td>Gender (Female)</td>
<td>0.23</td>
<td>(0.15, 0.31)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Education (&lt; HS)</td>
<td>0.01</td>
<td>(-0.08, 0.09)</td>
<td>.84</td>
</tr>
<tr>
<td>Household Poverty (per 10% change)</td>
<td>0.06</td>
<td>(-0.02, 0.05)</td>
<td>.01</td>
</tr>
<tr>
<td>Occupation Category (Low SES)</td>
<td>0.06</td>
<td>(-0.01, 0.14)</td>
<td>.08</td>
</tr>
<tr>
<td>Income (≤ $30K)</td>
<td>-0.02</td>
<td>(-0.14, 0.09)</td>
<td>.69</td>
</tr>
<tr>
<td>Home Ownership (Don't Own)</td>
<td>0.13</td>
<td>(0.02, 0.24)</td>
<td>.02</td>
</tr>
<tr>
<td>Baseline Cohort Indicator (Baseline)</td>
<td>-0.16</td>
<td>(-0.25, -0.08)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>
### TABLE 6. Association between Disability Progression and SES Measures and Demographic Factors by Age

<table>
<thead>
<tr>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline - Excluding Home Ownership and Income</strong></td>
<td><strong>Baseline - Including Home Ownership and Income</strong></td>
<td><strong>All Participants - Without Home Ownership and Income</strong></td>
</tr>
<tr>
<td><strong>β</strong></td>
<td><strong>95% CI</strong></td>
<td><strong>p</strong></td>
</tr>
<tr>
<td>Age (per 5 years)</td>
<td>0.02</td>
<td>(-0.02, 0.06)</td>
</tr>
<tr>
<td>Race (Black)</td>
<td>0.04</td>
<td>(-0.05, 0.14)</td>
</tr>
<tr>
<td>Gender (Female)</td>
<td><strong>0.23</strong></td>
<td><strong>(0.15, 0.31)</strong></td>
</tr>
<tr>
<td>Education (&lt; HS)</td>
<td>0.05</td>
<td>(-0.04, 0.14)</td>
</tr>
<tr>
<td>Household Poverty (per 10% change)</td>
<td><strong>0.05</strong></td>
<td><strong>(0.01, 0.10)</strong></td>
</tr>
<tr>
<td>Occupation Category (Low SES)</td>
<td>0.06</td>
<td>(-0.01, 0.14)</td>
</tr>
<tr>
<td>Age * Education</td>
<td><strong>-0.04</strong></td>
<td><strong>(-0.08, -0.01)</strong></td>
</tr>
<tr>
<td>Age * Household Poverty</td>
<td><strong>0.02</strong></td>
<td><strong>(0.00, 0.04)</strong></td>
</tr>
<tr>
<td>Age * Occupation</td>
<td>0.00</td>
<td>(-0.03, 0.04)</td>
</tr>
<tr>
<td>Income (≤ $30K)</td>
<td><strong>0.00</strong></td>
<td></td>
</tr>
<tr>
<td>Home Ownership (Don't Own)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age * Income</td>
<td>0.04</td>
<td>(-0.01, 0.10)</td>
</tr>
<tr>
<td>Age * Home Ownership</td>
<td>-0.00</td>
<td>(-0.05, 0.04)</td>
</tr>
<tr>
<td>Baseline Cohort Indicator</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 7: Exploratory Factor Analysis of SES Measures

<table>
<thead>
<tr>
<th>Variable</th>
<th>3 Covariates (N = 886)</th>
<th>5 Covariates (N = 541)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single Factor</td>
<td>Single Factor</td>
</tr>
<tr>
<td></td>
<td>Factor 1</td>
<td>Communalities*</td>
</tr>
<tr>
<td>Education</td>
<td>0.52</td>
<td>0.27</td>
</tr>
<tr>
<td>Occupation</td>
<td>0.67</td>
<td>0.44</td>
</tr>
<tr>
<td>Household Poverty</td>
<td>0.18</td>
<td>0.03</td>
</tr>
<tr>
<td>Home Ownership</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>0.51</td>
<td>0.26</td>
</tr>
</tbody>
</table>

*Communalities indicate the amount of variance in each variable that is accounted for by the latent factor*
8. FIGURES

Figure 1. Timeline of Enrollment Waves in the Johnston County Osteoarthritis Project. The original enrollment included 3187 participants between 1991 and 1997. However due to a high loss to follow-up rate between T0 and T1, there were only 1733 participants at the first follow-up between 1999 and 2004, known as T1. To account for this loss, an enrichment cohort, T1*, supplied 1015 new participants between 2003 and 2004. The last follow-up enrollment wave where the HAQ was used, known as T2, occurred between 2006 and 2010.
Figure 2a. Inclusion and Exclusion Criteria for Time to Disability Progression Analysis. STARD Chart of participant inclusion and exclusion criteria based on wide dataset (per patient). For exclusion into the original dataset, participants were required to have either knee and/or hip OA, and have completed at least 2 HAQs. Participants were then removed due to missing information for relevant covariates.
Figure 2b. Inclusion and Exclusion Criteria for Disability Progression by Age Analysis. STARD Chart of participant inclusion and exclusion criteria based on long dataset (per visit). Visits were removed if there was no HAQ score present, and if there was not complete data for relevant SES variables.
HEALTH ASSESSMENT QUESTIONNAIRE (HAQ-DI)®

Name: ___________________________________________ Date: __________________

Please place an "x" in the box which best describes your abilities OVER THE PAST WEEK:

<table>
<thead>
<tr>
<th>DRESSING &amp; GROOMING</th>
<th>WITHOUT ANY DIFFICULTY</th>
<th>WITH SOME DIFFICULTY</th>
<th>WITH MUCH DIFFICULTY</th>
<th>UNABLE TO DO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are you able to:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dress yourself, including shoelaces and buttons?</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>Shampoo your hair?</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ARISING</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Are you able to:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stand up from a straight chair?</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>Get in and out of bed?</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

Figure 3. Select Questions from the Health Assessment Questionnaire (HAQ). Example of two sections of the HAQ. Answering “Without any difficulty” yields a score of 0, and “Unable to do” yields a score of 3.
Figure 4. Study Directed Acyclic Graph (DAG). This DAG shows the relationship between the demographic covariates (age, race, and gender), seen here as confounders. Additionally, the association in question between SES and disability progression is the primary question at hand.
Figure 5. Weibull Distribution Probability Density Function (PDF) – Null Model. The probability density function (PDF) of the Weibull Distribution based on the determined scale and shape parameters.
Figure 6. Weibull Diagnostic Plot for Time to Disability Progression. The Weibull Regression Diagnostic plot allows for verification of the Weibull distributional assumption in comparison to an exponential distribution.
Figure 7. Model Checking Plot – QQ Plot. All the six models presented similar QQ plots, indicating that the data was in fact right skewed. This is to be expected, as there are more changes that are closer to 0 and in the lower range of change, compared to the upper extremes.
Figure 8. Correlation Plots of Exploratory Factor Analysis of SES Measures. In the panel on the left, correlations can be seen for SES variables which were present for both the baseline and enrichment cohorts. Aside from the diagonals which are expected to be highly correlated, education and occupation have the highest correlation, which is reflected in the factor loading seen in Table 7. In the panel on the right, correlations can be seen when all five SES variables are used. This visualization supports the higher correlations between education, occupation, home ownership, and income compared to almost no correlations with household poverty.
Divya Narayanan
929 N. Wolfe Street, Apt 804B, Baltimore, MD 21205 • dnaraya5@jhu.edu

EDUCATION

Master of Science (ScM) in Epidemiology
Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 05/2017

Bachelor of Arts (BA) in Statistics
University of California, Berkeley, Berkeley, CA 05/2014

PROGRAMMING SKILLS

R, STATA, LaTeX, MATLAB, SAS, SQL, 3D Slicer Visualization and Image Analysis Software

PUBLIC HEALTH EXPERIENCE

Project RISHI Member 06/ 2011 – present

Project RISHI (Rural India Social and Health Improvement) is a non-profit student-run organization whose mission is to promote the sustainable development and growth of rural Indian communities. University of California, Berkeley, CA

- **Director of Data Analytics – Strategy, National Board** (06/2016 – present)
  - Work with fundraising team to effectively obtain grants and funding for projects
  - Serve as contact between Project Initiatives and Operations teams

- **Vice President, UC Berkeley Chapter** (03/2013 – 05/2014)
  - Directed transition from rural village in South India to new village in North India
  - Organized pilot summer India trip to new village
  - Communicated with other campus-wide and external organizations
  - Administered weekly meetings with club chapter
  - Served as official campus representative of club

- **Co-Director of Project Development, UC Berkeley Chapter** (05/2012 – 03/2013)
  - Created Project RISHI India Base
  - Assessed project feasibility by planning and implementing surveys in village
  - Coordinated project teams and delegated project tasks

- **Committee Member of Health Education, UC Berkeley Chapter** (08/2011 – 05/2012)
  - Assisted in development and translation of health pamphlets for clinic in rural South India village
  - Worked with local organizations to improve clinic and health services

- **Committee Member of Community Development, UC Berkeley Chapter** (08/2011 – 05/2012)
  - Developed project centered around community togetherness and interaction
  - Organized trash pick-up day and painting of community area

- **Translator** (05/2011 – 06/2011)
  - Translated daily conversations between team members and local villagers during summer trip
  - Provided translation services for educational media posted in the village clinic

Public Health Applications for Student Experience (PHASE) Intern, IT Department
Baltimore City Health Department, Baltimore, MD 10/2015 – 05/ 2016

- Develop framework of existing database workflow and data evaluation
- Build databases, information systems, and policy recommendations
• Assist with data policies and corresponding documentation

**RESEARCH AND CLINICAL EXPERIENCE**

**Rheumatology Research Foundation Graduate Preceptorship Awardee**  
**06/2016 – 08/2016**

Thurston Arthritis Research Center, UNC Chapel Hill, Chapel Hill, NC

• Analyzed relationship between SES and progression of disability in longitudinal cohort study among individuals with osteoarthritis (OA)

• Conducted causal inference and mediation analysis of the relationship between socioeconomic status (SES) and self-reported health among those with osteoarthritis (OA)

**Intramural Research Training Awardee**  
**08/2014 – 07/2015**

NIH Clinical Center, Bethesda, MD

• Designed computer-aided detection program of various organs for targeted radiation treatment using machine learning techniques, such as random forests, for image processing

• Analyzed 1000+ CT images of patients using reliability and validity analyses (e.g. Bland-Altman) and cross-validation

• Set a new standard in the literature for similarity coefficient with an increase of 3% using random forests

• NIH Academy Certificate in Global Health Disparities  
  **09/2014 – 05/2015**

  • Participated in roundtable discussions about global health
  • Contributed to discussions about chronic disease and domestic health

**Undergraduate Researcher**  
**01/2012 – 07/2014**

UCSF Department of Radiology, San Francisco, CA

• Acquired patient data and interacted with clinicians, radiologists, and orthopedic surgeons

• Conducted study and statistical analysis for hip cartilage measurements in osteoarthritic patients and controls

**Student Research Assistant – Summer Internship Program**  
**06/2012 – 08/2012**

National Institute of Arthritis & Musculoskeletal and Skin Diseases, Bethesda, MD

• Analyzed batch effect across control and CANDLE patients by employing ELISA assays

• Sequenced genes of patient biopsy samples and cultured cells for genetic cloning

**Arthritis Foundation Research Intern**  
**06/2009 – 08/2012**

Arthritis Foundation, Rosalind Russell Medical Research Center for Arthritis, San Francisco, CA


• Acquired patient data, conducted surveys, and interacted with clinicians and residents in clinical setting

• Determined correlation between various medications/regimens and rheumatoid arthritis

**PUBLICATIONS AND PRESENTATIONS**


- Graf J, **Narayanan D**. The Prevalence of Alcohol and Over the Counter Medication Use in Treated Patients with Arthritis. Stanford Institute of Medicine Summer Research Program Poster Session, Stanford, CA, August 2009.

**CONFERENCE ABSTRACTS**


**NON-PROFIT EXPERIENCE**

**Speaker for Arthritis Foundation** 11/2010 - 12/2012

*Arthritis Foundation, San Francisco, CA*

- Spoke at various Arthritis Foundation events (4) about personal experiences and research

**HONORS AND AWARDS**

<table>
<thead>
<tr>
<th>Award</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rheumatology Research Foundation Graduate Student Research Award</td>
<td>06/2016</td>
</tr>
<tr>
<td>NIH Postbaccalaureate IRTA Outstanding Poster Award</td>
<td>05/2015</td>
</tr>
<tr>
<td>NIH Intramural Research Training Award (IRTA)</td>
<td>08/2014 – 07/2015</td>
</tr>
<tr>
<td>NIAMS Summer Internship Awardee</td>
<td>06/2012 – 08/2012</td>
</tr>
<tr>
<td>Arthritis Foundation Internship Awardee</td>
<td>06/2009 – 08/2009</td>
</tr>
<tr>
<td>Arthritis Foundation - Certificate of Achievement for Research in Arthritis</td>
<td>06/2009</td>
</tr>
</tbody>
</table>

**PROFESSIONAL DEVELOPMENT**

**Leadership:** Epidemiology Master’s Student Representative, Epidemiology Students Organization Sports Chair
Languages: Tamil (Fluent), Spanish (Conversational)

Volunteer: Johns Hopkins Child Safety Center, Baltimore, MD (2016-2017); Emergency Department, Alta Bates Summit Medical Center, Berkeley, CA (2012)