

# Evaluation of a Dynamic Bayesian Belief Network to Predict Osteoarthritic Knee Pain Using Data from the Osteoarthritis Initiative

Emily W. Watt, MLIS, Alex A.T. Bui, PhD

University of California Los Angeles, Los Angeles, CA

## Abstract

*The most common cause of disability in older adults in the United States is osteoarthritis. To address the problem of early disease prediction, we have constructed a Bayesian belief network (BBN) composed of knee OA-related symptoms to support prognostic queries. The purpose of this study is to evaluate a static and dynamic BBN-based on the NIH Osteoarthritis Initiative (OAI) data—in predicting the likelihood of a patient being diagnosed with knee OA. Initial validation results are promising: our model outperforms a logistic regression model in several designed studies. We can conclude that our model can effectively predict the symptoms that are commonly associated with the presence of knee OA.*

## Introduction

Osteoarthritis (OA) is the most common cause of disability in the United States [7]. Given the aging population, the number of individuals experiencing disabilities due to OA is projected to exceed 11 million in the US by 2020 [5]. The knee is the most common anatomy affected by this disease, with >30% of adults over age 60 experiencing functional limitations attributable to knee OA [5]. The impact and end outcome of knee OA varies greatly between subjects: in some, the disease stabilizes and symptoms may improve with time; in others, the joint damage progresses and results in severe pain and disability requiring surgical treatment. Osteoarthritis is typically diagnosed via imaging, looking for (morphological) tissue changes; but once detected, disease progression is already irreversible. Thus, the significance of being able to predict such changes early is clear. Traditionally, early detection of knee OA has relied upon various methods (*e.g.*, detecting subtle structural joint changes on imaging, patient-reported metrics such as pain [2, 11]). Unfortunately, our ability to predict the disease's onset and progression is still limited as the variables contributing to this disease are both numerous and multifaceted.

Here, we investigate the use of Bayesian belief networks (BBNs) to help predict early presence and to assess the development of knee OA. The network models were built from a dataset maintained by the Osteoarthritis Initiative (OAI), a NIH-funded (National Institute of Health) multisite endeavor aimed at

uncovering biomarkers identifying the development and progression of knee OA [8]. Subsequent follow-up studies were also incorporated into a dynamic belief network (DBN) to evaluate disease progression over time, providing a more complete understanding of end outcomes. Our goal in extending the model to a DBN is to determine how the data influences the structure and efficacy of the model in predicting the future knee health of a patient. This paper presents our initial work and evaluation of the model.

## Background

Osteoarthritis is a slowly evolving disorder of synovial joints in which a complex combination of degradative and reparative processes alters the anatomy and matrix composition of the articular cartilage and subchondral bone. Osteoarthritis studies have yet to produce conclusive results about the disease's relationship to certain symptoms (*i.e.*, disability, decreased bone mass). Although individuals may develop OA in one or more joints, it frequently occurs within the knee; moreover, the physical disability resultant from general OA in other areas may contribute specifically to disuse and knee OA. Relatively little is known about which aetiological factors and disease processes control progression or determine end outcome of knee OA, though it has been suggested that subchondral bone activity may be an indicator of the disease [12]. As such, determining which variables predict the outcome of knee OA is desirable for three reasons: first, it would be clinically valuable to be able to distinguish between patients with high and low risk of developing severe disease; second, it would help characterize the symptoms/risk factors most important in assessing disease progression, which can lead to strategies for tailored treatment [7]; and lastly it would create a unified disease paradigm to aid in prognostic and decision-making tasks.

We chose to apply the Bayesian belief network as a probabilistic approach for predicting onset and outcome of knee OA. Other approaches exist for prediction, but the BBN's utility is rooted in its capacity to integrate expert knowledge with empirical data to model a disease [3, 13]. Importantly, BBNs can accommodate for uncertainty in estimating disease outcomes and for incomplete evidence, two problems

inherent to the clinical setting. BBNs can provide a concrete understanding of how the relationships influence a particular disease outcome, thus elucidating the impact of various parameters and metrics in disease assessment. A review of BBNs in biomedicine and usage is given in [1, 3, 10].

**NIH Osteoarthritis Initiative.** The data used in this exploration consists of fields extracted from patient records in a publicly accessible (<http://www.-oai.ucsf.edu>) database for osteoarthritis subjects regularly updated and maintained by the Osteoarthritis Initiative (OAI). Started in 2004, the OAI involves five national centers participating in a program sponsored by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) and pharmaceutical partners. The primary objective of this consortium is to identify imaging and biochemical markers for the development and progression of both incident and progressive knee OA [6]. The OAI dataset consists of clinical evaluation data (medical history, physical exam, joint-specific observations), imaging (x-ray; magnetic resonance, MR), and a bio-specimen repository (biochemical and genetic) from 4,796 volunteers aged 45-79.

### Methodology

As a first step towards model construction, evidence variables were selected from tables of the OAI dataset based on a review of the literature for knee OA symptoms (Enrollees [v.4], Biomarkers00 [v.0.2.1], JointSx00 [v.0.2.1], MedHist00 [v.0.2.1], PhysExam00 [v.0.2.1], SubjectChar00 [v.0.2.1]). A total of 44 variables were identified (Table 1) from the demographic/history, imaging, quality of life scores and clinical assessment data. Although the majority of the OAI dataset is codified, continuous variables (*e.g.*, age, body mass index (BMI)) were discretized into quartiles (*i.e.*, four equally-sized bins) via a *k*-means algorithm.

**Model construction.** A network topology was defined for the BBN (Fig. 1) organizing the evidence variables according to: 1) expert clinical input based on experience with (knee) OA and bone disease, used to derive dependency relationships; and 2) a systematic review of the literature to uncover further variable associations. Notably, the OAI dataset contains information on both general OA and knee OA for each subject; we opted to follow this same paradigm as part of the BBNs, creating a relationship between general OA and onset of knee OA. The *Bayesia* software package (<http://www.bayesia.com>) was used to compute the conditional probability tables (CPTs) for the BBNs.

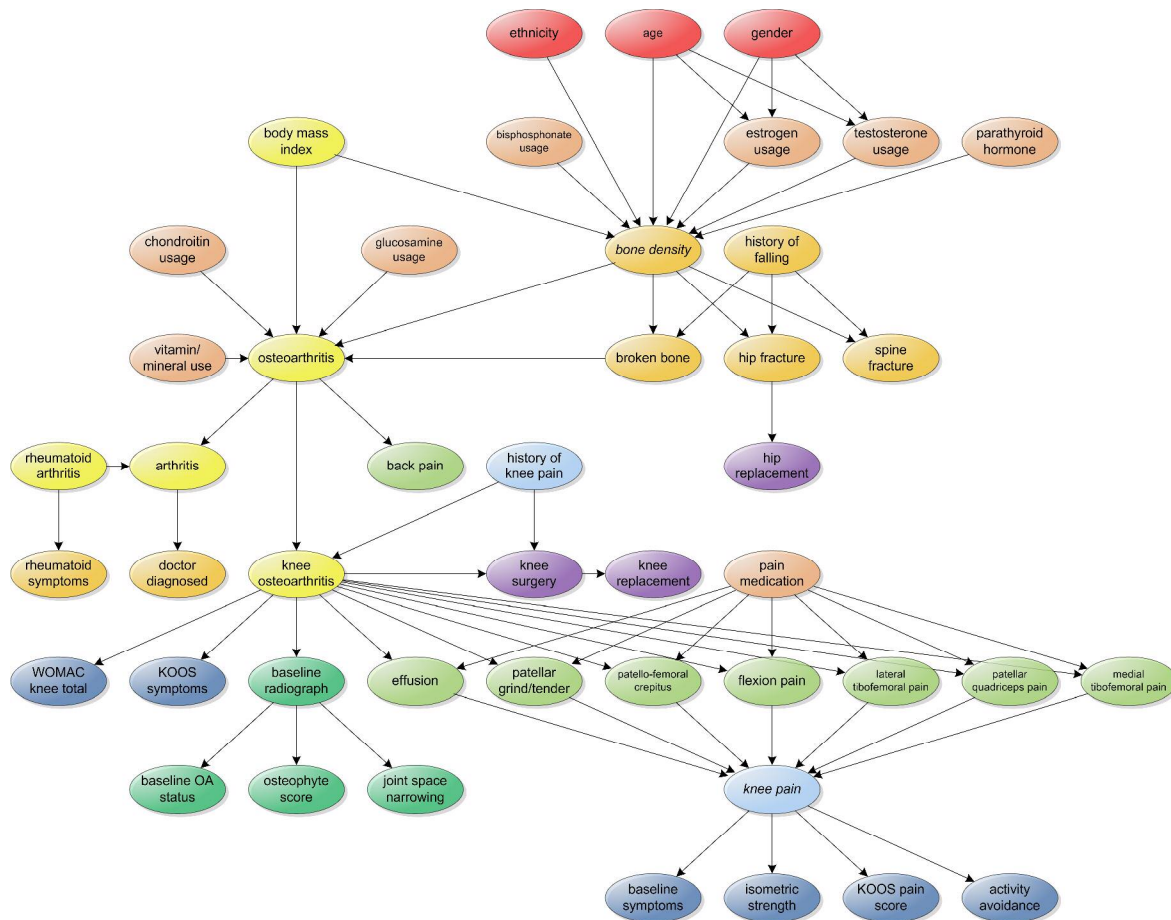
Static and dynamic belief networks (DBN) were constructed for model evaluation. Rudimentarily, a DBN

OAI Table	Selected Variables	
Subjects	Age	Ethnicity
	Gender	Body mass index (BMI)
Medical history	Estrogen usage	Vitamin usage
	Parathyroid usage	Pain medication usage
	Bisphosphonate usage	Knee surgery
	Hip replacement	Doctor diagnosed OA
	Osteoarthritis	History of falling
	Hip fracture	Spine fracture
	Rheumatoid symptoms	Testosterone usage
	Chondroitin usage	Glucosamine usage
	Arthritis	Knee osteoarthritis
	Broken bone	Rheumatoid arthritis
Physical exam	Effusion	Lateral tibiofemoral pain
	Flexion pain	Medial tibiofemoral pain
	Baseline OA symptoms	Patello-femoral crepitus
	Isometric strength	Patellar grind/tenderness
	Patellar quadriceps pain	
Biomarkers	Knee replacement	Joint space narrowing
	Baseline symptoms	Osteophyte score
Joints	Back pain	Activity avoidance
	KOOS symptoms score	KOOS pain score
	History of knee pain	Knee pain
	WOMAC total score	

**Table 1.** OAI variables predicting knee osteoarthritis.

is a temporal model imposed on a BBN to characterize changes in probabilities brought about by new information. The added temporal dimension is managed by time-indexed evidence variables at each observed time point [10]. A link between two nodes represented in different time points represents a change in belief (and thus associated probabilities) about the state of nodes (*i.e.*, a transition state). Given that the OAI study is longitudinal in nature, we fitted the DBN to model follow-up data at the 12-month interval: links between the baseline and all corresponding 12-month time point were formed as part of the DBN topology.

**Evaluation.** Two studies were designed to evaluate the static BBN: 1) a standard tenfold cross validation for predictive power; and 2) a 10-90 test (computing the CPTs from 10% of the data and testing on the remaining 90%) to identify potential over-training of the network. Each study tested the BBN's ability to predict seven different variables: the presence of knee OA; the WOMAC (Western Ontario and McMaster Universities Arthritis Index) total score; the presence of knee pain; the appearance of joint space narrowing (JSN) on imaging; the osteophyte score based on x-rays; and the presence of osteoarthritis and other baseline symptoms. The variables were selected based on current literature documenting these factors as relevant or correlated (predictive) factors of knee OA. The studies were conducted on the full baseline dataset ( $N = 4,796$ ). Accuracy of tested target vari-



**Figure 1.** Static Bayesian belief network for knee osteoarthritis

ables was additionally measured on 10 randomly selected subsets ( $n = 200$ ) for the baseline and 12-month follow-up data using *Bayesia*. Only the tenfold cross validation was conducted on the DBN using 10 random subset of 200 records from the follow-up data of 2,686 patients.

To provide context for comparison against the evaluation studies, logistic regression models were also built using the full baseline dataset. For each target variable, multiple stepwise logistic regression was performed with a selection method for variable entry based on the score statistic and removal based on the probability of a likelihood ratio statistic using conditional parameter estimates. To avoid learning bias, the CTABLE option in SAS' PROC LOGISTIC was used. This classification table uses a method approximating unbiased jackknifing for cross validation.

### Results

The validation study results are compiled in Tables 2 & 3, and are summarized as follows.

Logistic regression. The logistic regression model was able to predict the presence of knee pain with

100% accuracy. Notably, the developed models were able to predict the presence of both general osteoarthritis and knee OA with accuracy of 80% and 79%, respectively. Model accuracy ranged between 60-70% for the remaining variables.

BBN tenfold cross validation. The presence of knee pain was also predicted consistently at 100% ( $N = 4,796$ ;  $n = 200$ ). Baseline symptoms was predicted with 95% precision ( $n = 200$ ) and 80% for knee osteoarthritis ( $N = 4796$ ;  $n = 200$ ). The static BBN predicts all target variables with accuracy better than that achieved by regression, with the exception of JSN and WOMAC total score (55% and 50%, respectively). Between the two different sample sizes, we note a high variance (e.g., 87% versus 80% accuracy in predicting OA between  $n = 200$  and  $N = 4,796$ ), suggesting that the random sub-sampling effectively stabilized the results in the variables.

10-90 validation. Results from the 10-90 test reflect the accuracy seen in the tenfold cross validation. This evaluation similarly demonstrates 100% accuracy in predicting knee pain; and 85% ( $n = 200$ ) and 79% accuracy ( $n = 200$ ) in predicting the presence of base-

BBN Variable	Logistic	Tenfold		90-10	
	n=4796	200	4796	200	4796
Knee OA	79%	80%	80%	79%	78%
WOMAC score	77%	45%	50%	48%	52%
JSN	69%	55%	53%	52%	50%
Knee pain	100%	100%	100%	100%	100%
Osteophyte	68%	75%	76%	73%	75%
Osteoarthritis	80%	80%	87%	81%	75%
Baseline Sx	62%	95%	85%	85%	84%

**Table 2.** Baseline knee OA power prediction results. The static BBN outperformed logistic regression in 4 of 7 selected variables, and was equal to the logistic model in assessing the presence of knee pain.

BBN Variable	Tenfold Cross Validation
	n=200
Knee OA	100%
WOMAC score	56%
JSN	100%
Knee pain	100%
Osteophyte	100%
Osteoarthritis	89%

**Table 3.** Dynamic belief network evaluation results. The DBN improved upon the results seen with the static Bayesian network.

line symptoms and knee OA, respectively. Again, the JSN and WOMAC scores performance were subpar compared to the logistic regression. Arguably, these results may be more representative of the knee OA population, as testing is performed on 90% of the data, in contrast to the 10% of the tenfold validation.

**DBN performance.** Results from the DBN demonstrate performance exceeding the predictive power seen with logistic regression. Four of the six variables in the follow-up study were predicted at 100% accuracy with the DBN (knee OA, JSN, presence of knee pain, osteophyte score). For general osteoarthritis, the accuracy was 89%; the worst performing prediction was for the WOMAC total score, with accuracy only at 56%.

## Discussion

The static BBN consistently predicts the presence of knee pain, osteoarthritis, baseline symptoms and knee OA at baseline. The high degree of dependency between three of the evidence variables (knee OA, knee pain, osteoarthritis) to other nodes in the network may explain their comparatively better performance. Further sensitivity analysis of the static network and DBN finds that knee OA and knee pain as strong predictors for general OA; the other factors have yet to be more strongly validated. The potential for the dynamic belief network in modeling knee OA is also encouraging in this initial study: many of the metrics

used to assess and/or predict knee OA were correctly predicted with high accuracy.

There are several factors that could potentially explain the observed results:

1. **Dataset skew/bias.** Skewed data clearly impacted our ability to appropriately learn an underlying probability distribution, and thus influenced the model's performance. Further data analysis revealed, for instance, that over 80% of the patients reported presence of knee pain, suggesting a biased distribution. Similarly, the variable representing osteophyte presence was dominated by the definite or grade 1-3 state. Although such skew may be representative of the population, it indicates that different modeling should be applied to handle such variables within the BBN.
2. **Composite values.** Target variables characterizing cumulative values, such as health assessment or the quality of life metrics represented by WOMAC, are dependent on a number of underlying variables that are aggregated together. It can be argued that the selection of variables involving the collapse of various metrics (*e.g.*, knee pain) is subjective, and thus not a reliable metric in the network. Thus as survey instruments, questionnaires need to be re-examined for their validity in our model.
3. **Quantitative metrics.** Some of the imaging features in the OAI dataset are based on subjective radiographic interpretation as opposed to actual measurements (*e.g.*, JSN and WOMAC). More objective quantitative metrics (*i.e.*, direct measurement of imaging features) may provide improved assessment and an opportunity to provide more granular discretization.
4. **Discretization.** Using only *k*-means hierarchical classification to codify continuous variables also affected prediction outcomes. Our classification was (naïvely) applied to create equal-sized bins.
5. **Network construction.** The subjectivity of expert knowledge can introduce bias into the modeled BBN relations. Although a literature review to introduce/remove dependencies was performed, alternative methods that automatically suggest network topology from the data may be of use.

Little research has been done specifically in the area of predicting knee OA using probabilistic networks, however work does extend to other anatomical areas that are affected by arthritis, namely the hip [5]. The literature supports the combination of 44 variables that were selected in this investigation as strong predictors of assessing joint damage associated with osteoarthritis. However, predicting early disease presence is limited to indirect metrics: it has been suggested that JSN measures are an early predictor of

knee OA, but location and joint space narrowing severity were poorly predicted [5]. Other studies describe data and results that are used to determine other potential predictors for change or how the variable could affect outcome of the patient's knee health. Significant correlations do demonstrate that certain biomarkers (e.g., crepitus, joint swelling, knee surgery and knee pain) and demographic data (e.g., patient age, sex, body mass index) had little or no effect on predicting OA outcomes [2].

Continuing work in this project will involve a more thorough construction and re-evaluation of expert-input supplied to the underlying network. Given the encouraging results seen from the DBN, further work will also refine the DBN given subsequent follow-up studies and testing on diverse populations outside of the OAI dataset. The rationale for selecting the target and qualitative variables defining knee pain, as well as the WOMAC composite score, should be examined further and modeled accordingly. Alternative discretizations are being explored including the use of accepted categorical scales for mapping the JSN, osteophyte and WOMAC values to better suit a variable's underlying distribution.

#### Acknowledgements

This research was supported in part by the NLM Training Grant LM007356 and NIH R01 EB000362. We would also like to thank Dr. James Sayre for his assistance with the analysis.

The OAI is a public-private partnership comprised of five contracts (N01-AR-2-2258; N01-AR-2-2259; N01-AR-2-2260; N01-AR-2-2261; N01-AR-2-2262) funded by the National Institutes of Health, a branch of the Department of Health and Human Services, and conducted by the OAI Study Investigators. Private funding partners include Merck Research Laboratories; Novartis Pharmaceuticals Corporation, GlaxoSmithKline; and Pfizer, Inc. Private sector funding for the OAI is managed by the Foundation for the National Institutes of Health. This manuscript was prepared using an OAI public use data set and does not necessarily reflect the opinions or views of the OAI investigators, the NIH, or the private funding partners.

#### References

1. Acid S, Campos LM, Luna JM, et al. A comparison of learning algorithms for Bayesian networks: A case study based on data from an emergency medical service. *Artificial Intel Medicine*, 2004; 30: 215-232.
2. Boniatis I, Cavouras D, Costaridou L, et al. Computer-aided grading and quantification of hip osteoarthritis severity employing shape descriptors of radiographic hip joint space. *Comp Biology and Medicine*, 2007; 37: 1786-1795.
3. Boniatis I, Costaridou L, et al. Assessing hip osteoarthritis severity utilizing a probabilistic neural network based classification scheme. *Med Engineering & Phys*, 2007; 29: 227-237.
4. Bruyere O, Collette J, Kothari M, et al. Osteoarthritis, magnetic resonance imaging, and biochemical markers: a one year prospective study. *Ann Rheum Dis*, 2006; 65: 1050-1054.
5. Currier LL, Forehlich PJ, Carow SD, et al. Development of a clinical prediction rule to identify patients with knee pain and clinical evidence of knee osteoarthritis who demonstrate a favorable short-term response to hip mobilization. *Physical Therapy*, 2007; 87: 1106-1119.
6. Dieppe PA. Recommended methodology for assessing the progression of osteoarthritis of the hip and knee joints. *Osteoarthritis and Cartilage*, 1995; 3: 73-77.
7. Dieppe P, Cushnaghan J, et al. Prediction of the progression of joint space narrowing in osteoarthritis of the knee by bone scintigraphy. *Ann Rheumatic Diseases*, 1993; 52: 557-563.
8. Eckstein F, Kunz M, Schutzer M, et al. Two year longitudinal change and test-pretest-precision of knee cartilage morphology in a pilot study for the osteoarthritis initiative. *Osteoarthritis Cartilage*, Jun 2007.
9. Lai W, Chan C, Tang Y, Bronson R, Tung C. Early diagnosis of osteoarthritis using cathepsin B sensitive near-infrared fluorescent probes. *Osteoarthritis and Cartilage*, 2004; 12: 239-244.
10. Lucas PJF, van der Gaag LC, Abu-Hanna A. Bayesian networks in biomedicine and healthcare. *Artificial Intelligence in Medicine*, 30(3):201-214, 2004
11. Mansell JP, Tarlton JF, Bailey AJ. Biochemical evidence for altered subchondral bone collagen metabolism in osteoarthritis of the hip. *British Journal of Rheumatology*, 1997; 36: 16-19.
12. Vicari R, Flores C, Silvestre A, Seixas L, et al. A multi-agent intelligent environment for medical knowledge. *Artificial Intel Med*, 2003; 27: 335-366.
13. Vincent F, Vignon E, Brandt K, et al. Superiority of the Lyon schuss view over the standing anteroposterior view for detecting joint space narrowing, especially in the lateral tibiofemoral compartment, in early knee osteoarthritis. *Ann Rheumatic Diseases*, 2007; 747-753.
14. Weber P, Jouffe L. Complex system reliability modeling with dynamic object oriented Bayesian networks (DOOBN). *Reliability Engineering & System Safety*, 2006; 91: 149-162.