Published in final edited form as:

Ann Biomed Eng. 2015 September ; 43(9): 2056–2068. doi:10.1007/s10439-015-1269-8.

Coefficient of Friction Patterns Can Identify Damage in Native and Engineered Cartilage Subjected to Frictional-Shear Stress

G. A. Whitney^{1,3}, J. M. Mansour², and J. E. Dennis^{1,4}

¹Department of Biomedical Engineering, Case Western Reserve University, Wickenden, Room 319, 2071 Martin Luther King Jr. Drive, Cleveland, OH 44106, USA

²Department of Mechanical and Aerospace Engineering, Case Western Reserve University, Glennan Building, 2123 Martin Luther King Jr. Drive, Cleveland, OH, USA

³Benaroya Research Institute, Matrix Biology Program, 1201 9th Ave., Seattle, WA 98101, USA

⁴Department of Orthopedic Surgery, Baylor College of Medicine, 7200 Cambridge, Ste. 10A (281), Houston, TX 77030, USA

Abstract

The mechanical loading environment encountered by articular cartilage *in situ* makes frictional-shear testing an invaluable technique for assessing engineered cartilage. Despite the important information that is gained from this testing, it remains under-utilized, especially for determining damage behavior. Currently, extensive visual inspection is required to assess damage; this is cumbersome and subjective. Tools to simplify, automate, and remove subjectivity from the analysis may increase the accessibility and usefulness of frictional-shear testing as an evaluation method. The objective of this study was to determine if the friction signal could be used to detect damage that occurred during the testing. This study proceeded in two phases: first, a simplified model of biphasic lubrication that does not require knowledge of interstitial fluid pressure was developed. In the second phase, frictional-shear tests were performed on 74 cartilage samples, and the simplified model was used to extract characteristic features from the friction signals. Using support vector machine classifiers, the extracted features were able to detect damage with a median accuracy of approximately 90%. The accuracy remained high even in samples with minimal damage. In conclusion, the friction signal acquired during frictional-shear testing can be used to detect resultant damage to a high level of accuracy.

Keywords

Cartilage; Frictional-shear; Damage; Biphasic lubrication model; Damage identification; Quality; Assessment

Address correspondence to J. E. Dennis, Department of Orthopedic Surgery, Baylor College of Medicine, 7200 Cambridge, Ste. 10A (281), Houston, TX 77030, USA. James. Dennis@bcm.edu.

Electronic Supplementary Material: The online version of this article (doi: 10.1007/s10439-015-1269-8) contains supplementary material, which is available to authorized users.

Conflict of Interest: None.

Introduction

Engineered cartilage (EC) is being explored by many groups as a potential treatment for osteoarthritis.^{17,26} Various methodologies exist for producing EC, yet the mechanical properties of EC generated *in vitro* are consistently inferior to those of native cartilage (NC).^{7,12,28} Thus, mechanical testing plays an important role in cartilage tissue engineering, both in characterizing EC constructs, and in directing work for the improvement of EC constructs. While uniaxial properties, such as Young's modulus or aggregate modulus, are most commonly reported for EC, the forces experienced *in vivo* are time-varying and multi-axial: both compressive and frictional shear forces are present.^{19,25} One of the functions of cartilage is to limit the amount of frictional shear force transmitted across the joint by maintaining a low coefficient of friction (CoF). Tribological evaluation, where tissue is subjected to both compression and frictional shear forces, has been used extensively to characterize the friction, lubrication, and wear properties, of cartilage *in vitro*.^{1,4,9,10,13,24} Much opportunity still exists for applying this type of multi-axial testing to characterize the damage properties of EC.

We recently reported a study in which we characterized the damage of scaffold-free EC resulting from exposure to frictional shear.²⁷ Lizhang *et al.* have reported similar testing in NC.¹⁶ In both studies, damage was identified visually, at the conclusion of the test. In some cases, detecting damage required staining of the sample surface with dye, and in other cases, internal cracking occurred and could only be detected by cross-sectioning of the sample for histological observation. Visual damage detection is not optimal since it is limited to surface analysis, and cross-sectioning, which destroys the sample, is required to verify that no internal damage has occurred. Even if only surface damage is characterized macroscopically, the addition of dye modifies the bearing surface. In addition to requiring cross-sectioning of the sample, visual and histological methods only allow the experimenter to subjectively determine *if* damage has occurred during the test, not *when* it occurred during the test. Lastly, histological preparation of samples can be costly and time consuming. An alternate damage detection method that eliminates the need for cross-sectioning or surface staining would be valuable for preserving samples for additional studies, such as *in vivo* studies, or biochemical characterization.

Signal processing techniques have already been applied to detect and predict damage in mechanical devices, such as wind-turbines.⁶ Similar techniques based on vibration signals have been applied for the early detection of osteoarthritis *in vivo*.^{20–22} Based on the friction signal patterns we observed in our previous study,²⁷ we hypothesized that temporal properties of the friction signal could be used to identify cartilage samples that sustain damage during friction testing. To address this hypothesis, we chose to use the receiver operator characteristic (ROC) curve,^{11,23} and support vector machine (SVM) analysis,¹⁵ both of which are common approaches taken from the field of machine learning for analyzing patterns.

Use of ROC and SVM analysis for the purposes outlined above first requires extracting features from the individual friction signal of each sample tested. In this study, we used features intrinsic to the measured friction signal, such as the magnitude of the change in the

CoF between consecutive oscillation cycles, as well as features determined by comparison of the measured CoF to an ideal model of cartilage biphasic lubrication.^{8,14} This study is therefore presented in two parts: first, the development of a modified biphasic lubrication model which does not require knowledge of interstitial fluid pressure, and second, the use of features of the CoF signal to detect damage incurred during friction testing.

The purpose of the technique described here is as a non-destructive method for identifying samples which are damaged during frictional-shear testing, i.e., a method which does not require cross-sectioning of samples after frictional-shear testing. Successful development of such a technique could potentially be used as a quality assurance test for EC prior to implantation.

Materials and Methods

Cartilage Sample Preparation

Scaffold-free EC constructs were generated as previously described.²⁸ Briefly, chondrocytes were isolated from the joints or ears of skeletally mature New Zealand White rabbits in accordance with the University of Washington IRB. Subsequently, they were culture expanded, then re-differentiated at high density on porous membranes suspended in chondrogenic medium. This cartilage fabrication method generated constructs as large as 4×4 cm, from which multiple samples were harvested for friction testing. Samples were either 3 or 5 mm biopsy punches (discs) or 5×14 mm rectangular strips. Since experiments were dependent on the response of each sample individually, and not the response of the larger construct as a whole, each sample was counted as an individual observation.

NC for friction testing was harvested from the knees or shoulders of steers near 1 year of age obtained from an abattoir. Bovine cartilage was used because flat samples of rabbit cartilage of equal size to those of EC were not attainable. Samples were frozen at -80°C until tested.

Friction Signal Acquisition and Processing

Friction testing was performed on a custom device with rectangular constructs being tested in a curved configuration, and disc constructs being tested in a flat configuration as previously described.²⁷ In both configurations, constructs were adhered to metal platens (curved or flat) with cyanoacrylate, and then oscillated in PBS against a glass counterface. In both configurations, the normal load was applied to each sample through a linear-slide bearing, and the frictional-shear force was applied by the counterface which was attached to an oscillating table (see Whitney *et al.*,²⁷ for detailed setup). Sliding speed was 1 mm s^{-1} with a 56 mm path length in each direction. These parameters were chosen to conform to our previous study on damage resulting from frictional shear.²⁷ Only a portion of the rectangular construct made contact with the counterface in the curved configuration, whilst in the flat configuration, a ball joint between the device and the platen aligned the entire surface of the disc construct to the counterface. In both cases, the region of cartilage that was in contact with the counterface was constant.

During testing, the friction force and normal force were measured by separate single point load cells that were insensitive to off-axis loading. Due to the reciprocation of the

counterface against the construct, an undulating friction pattern was produced (Figs. 1a and 1b). Prior to friction model fitting, the raw friction force signal was converted to CoF by dividing by the normal load, and the raw CoF signal was rectified and filtered with a finite impulse response lowpass filter to smooth transients resulting from reciprocation of the counterface (Fig. 1c). The rectified and filtered version of the CoF will be referred to as μ .

Modeling the Friction Signal with Modified Biphasic Lubrication Models

In order to determine features of μ which might indicate construct damage, a model of μ in undamaged constructs was developed. Samples tested in the flat and curved configurations were pooled together for model analysis for a total of 19 TE samples and 14 native samples. Three samples of each tissue type were tested in the curved configuration. Since the friction response is dependent on the normal stress, a range of normal loads from 2.2 N to 11.8 N was used to ensure that the model accuracy was not dependent on the normal load. Each sample was only tested once. This resulted in a total of 20 damaged samples, and 13 undamaged samples for model evaluation.

Two models for describing time varying friction, μ , were evaluated. These were created by modifying the biphasic lubrication model (Eq. 1) of Krishnan *et al.*,¹⁴

$$\mu(t) = \mu(t)_{Eq} [1 - (1 - \phi)(W_p/W)], \quad (1)$$

where $\mu(t)$ is the predicted CoF, $\mu(t)_{Eq}$ is the predicted $\mu(t)$ at equilibrium (long time scales), ϕ is the product of the solid area fractions of cartilage and counterface at the interface, and (W_p/W) is the ratio of interstitial fluid pressurization to normal load.^{8,14} We modeled the fluid pressurization ratio with either one exponential (1 exp model), or a sum of two exponentials (2 exp model). The modified model is shown in Eq. (2),

$$\mu(t) = \mu(t)_{Eq} \left[1 - (1 - \phi) \sum_{n=1}^N a_n e^{-t/\tau_n} \right] \quad (2)$$

where $\mu(t)$ is the predicted value of μ , $N = [1, 2]$, and $\mu(t)_{Eq}$, a_n , τ_n , ϕ were fit as parameters of the model. All parameter estimation for model fitting was done using global optimization (to decrease chances of finding local minima) in Matlab (Mathworks, Natick, MA) by a least squares error model-fitting routine subject to constraints which ensured that $\mu(t = 0)$, $\mu(t)_{Eq}$, and $\tau > 0$, while $0\% \leq \phi \leq 100\%$. Using these models, knowledge of the construct interstitial fluid pressure during testing was unnecessary.

Models were assessed by multiple measures. Overall agreement of models to μ was assessed by R^2 , calculated as $1 - \text{SSE} \div \text{SST}$ where SSE is sum of the squared residual error and SST is the sum of the squared difference between μ_i and the mean of μ . Since the two models had varying numbers of parameters, the Akaike Information Criterion (AIC) was calculated as an indication of benefit of increased model complexity on overall agreement to measured data.³ AIC was calculated as $n[\log(\text{SSE} \div n)] + 2k$ where n is number of samples, and k is the number of parameters in the model. Since initial friction and final friction (equilibrium

friction for undamaged samples) values are often used to characterize the friction response, the magnitude of the model error at both the initial friction and final friction values was also estimated. For model error calculations, measured initial friction ($\mu_i=0$) and measured final friction (μ_{Eq}) values were calculated by taking the average of μ during the first and last oscillation cycle respectively. The errors between the calculated friction values and the model determined values were then calculated as (modeled-calculated) $\div$ calculated. Summary statistics are presented as means with standard deviations when applicable.

Visual Characterization of Damage

The counterface-contacting surface of the samples was imaged under a stereomicroscope before and after friction testing. Post-testing, India ink (Figs. 2b–2e) was used to increase contrast, whereas ink was not applied before testing to avoid modification of the bearing surface, however, one sample was stained with India ink and included for comparison purposes (Fig. 2a). ImageJ (NIH, Bethesda, MD) was used to quantify the percentage of the surface displaying damage (surface peeling or missing portions of the construct—both can be observed in Fig. 2d). Post-testing, samples were also processed for histology by fixation in 10% formalin and sequential dehydration in ethanol. Two 5- μ m cross-sections from each of two different locations (along the diameter for discs, or length for strips) of the specimen were then stained with safranin O and fast green; examples are seen in Figs. 2g–2j. These were then imaged under high magnification to visualize the internal integrity of the specimens. Specimens with greater than 5% surface damage, or with internal cracking (Figs. 2d and 2i), or with surface cracking, were labeled as “damaged.” Those with internal cracking and less than 5% surface damage were labeled as “internally damaged” (Figs. 2c and 2h).

Friction Signal Pattern Classification

To assess whether the friction signal could be used to determine if constructs were damaged during friction testing, 70 new EC constructs were generated and 3 new NC samples were collected, resulting in a total of 73 samples for pattern classification. These samples were divided between curved and flat testing configurations as indicated in Table 1, which also includes normal load ranges. Test duration was 2500 s since in the modeling study damage was seen to occur prior to that time.

Feature Extraction—Friction signal features that were intrinsic to the raw CoF signal, as well as some generated by comparison of μ to the 1 exp model, were assessed for their ability to classify damaged and undamaged samples. Intrinsic features included the number of consecutive undulations in the raw CoF signal that exhibited decreasing peak CoF (*consec. drops*), the maximum decrease in peak CoF between two adjacent undulations in the raw CoF signal (*max drop*), and $(\mu_{max} - \mu_{Eq})/\mu_{Eq}$ where μ_{max} is the maximum measured μ and μ_{Eq} is as above. Using the 1 exp model, the following features were calculated: R^2 and endpoint error (identical to final friction error) were calculated as before, and *max error* was calculated as the maximum error at any point between $\mu(t)$ and μ_i . An additional feature, *max drop %*, was included only in SVM analysis, and was calculated by normalizing *max drop* by the pre-drop CoF. See Fig. 3 for further clarification on features.

ROC Curve Analysis—To determine the ability of individual features in pattern classification, ROC curves with their following characteristics were calculated in GraphPad Prism 6: true positive rate (TPR), false positive rate (FPR), area under the ROC curve (AUROC), standard error, 95% confidence interval, and p value of AUROC. The p -value tests the hypothesis that the AUROC = 0.5, meaning that the classifier is no better than random chance. For detailed explanation on ROC curve construction, see the supplemental material. In this analysis, significance was set to $p < 0.008$ (rounded down from $p < 0.05 \div 6$) to account for 6 multiple comparisons, i.e., each curve compared to random chance. Confidence areas of selected ROC curves were calculated in R (R foundation for statistical computing, Vienna, Austria) using the open-source pROC package.²³ Selected curves were also tested to determine if their AUROCs were significantly different by using the bootstrapping method for correlated ROC curves since the same dataset was used for all features.²³ Significance was set to $p < 0.05$.

SVM Analysis—To determine the ability of multiple features to detect damage, SVM analysis was performed in Matlab R2014a using the built-in `FITCSVM` and `PREDICT` functions. Optimization was performed to determine the SVM model with the highest accuracy. Here, an SVM model refers to the combination of the SVM parameters, along with a specific feature set. The SVM parameters optimized in this study were: kernel transformation (linear, 2nd through 5th order polynomials, and Gaussian); box constraint (C), which is a penalty on misclassification when a perfect separation between the two classes does not exist; and kernel scale (KS), a scaling factor by which all inputs are divided. Further development of the SVM equations are beyond the scope of this manuscript, however, they have been described extensively in the literature, and Burges gives a thorough description.² To account for the disparity between the number of samples in the undamaged vs. the damaged group, for all SVM training described below, including during feature selection, and parameter optimization, the equal prior condition was enforced. With this condition enforced, undamaged and damaged groups carried equal weight in the training procedure by penalizing misclassification of undamaged samples more than misclassification of damaged samples. Also, the values of each feature were standardized by dividing by the maximum value for that feature, and shrinkage period was set to 1000 as suggested by the Matlab documentation.

Feature Selection—Forward sequential feature selection was performed in order to select features for further testing. For this selection, one-half of the damaged and one-half of the undamaged samples were randomly partitioned for training the SVM, and then the remaining one-half were used as the test set to determine classification accuracy (stratified cross-validation). This was performed iteratively 50 times in each step of the forward selection procedure to allow a classification distribution to be determined. Since classification distributions were found to be non-normal distributions, notched box-plots were used for feature selection. The location of the notch is calculated as

$$\text{median} \pm 1.7 \times (1.25 \times IQR \div (1.35 \sqrt{N})), \quad (3)$$

where *IQR* is interquartile range, and *N* is the number of iterations. Distributions with non-overlapping notch ranges are significantly different at approximately the 95% confidence range.¹⁸ In the forward selection procedure, all possible combinations of features were tested for $C = [0.001, 0.01, 0.1, 1, 5, 10]$ with $KS = 1$. At each parameter value, the feature which caused the median of the upper notch for the misclassification rate to decrease the most was added to the model. This was performed iteratively until no further decrease in the upper notch of the misclassification rate could be achieved. The best feature set for each kernel was selected for parameter optimization. During feature selection for the Gaussian kernel, two feature sets produced a tie, so parameter optimization was performed for both.

Parameter Optimization—Parameter optimization of C and KS was performed using the built-in `PATTERNSEARCH` Matlab routine. Lower and upper search bounds were set to $e^{-6.5}$ and e^5 respectively, since preliminary trials converged to values within that range. Defaults settings were used except in the case of tolerances and “Search Method” since there appeared to be many local minima of the loss function, and adjustment of these led to lower estimates of the misclassification rate upper-notch values. Function, mesh, and X tolerances were set to 1×10^{-7} , 1×10^{-8} , and 1×10^{-8} , respectively. Search Method was set to “Genetic Algorithm.”

Model Selection by Cross-Validation—Optimized models for each kernel were compared by stratified cross-validation as before, but 2000 repetitions were used for each model. This large number was chosen to make the cross-validation more robust, i.e., to limit variability of distributions based on the subsets of train/test partitions randomly selected during cross-validation.

Total classification accuracy of each model was calculated as (number of observations correctly classified ÷ total number of observations in test set). Normality of classification accuracy distributions was tested by Chi squared hypothesis testing, and all were found to be non-normal. Since distributions were non-normal, classification accuracy distributions were shown graphically by estimation using a kernel distribution with a bandwidth of 0.013 generated from the built-in Matlab `KSDENSITY` function. The kernel distribution generates an estimated probability density function (pdf) by summing continuous estimates of each data point, resulting in a continuous function, rather than a discrete histogram. The bandwidth was chosen as the lowest bandwidth which produced smooth distributions. Total classification accuracies of models were assessed for significant differences by notched boxplots as described above. Classification accuracy distributions of damaged samples, internally damaged samples, and undamaged samples, were all calculated simultaneously. All were assessed for significant differences by notched boxplots. Supplemental Figure S1 gives a diagram of the SVM model selection process.

Results

Modified Biphasic Lubrication Model

Both 1 exp and 2 exp modified biphasic models (Eq. 2) were able to characterize the friction response of undamaged samples (Figs. 2b, 2e, 2g, and 2j) as shown by R^2 values (Table 2), which were not significantly different between the two models. The increase in AIC between 1 exp and 2 exp models (Table 2) indicates that increasing model complexity may

not be significantly beneficial to the overall agreement of the model to the measured data. Despite the high level of overall agreement between models and data, $\mu_{t=0}$ error magnitude was high for both models (Table 2). μ_{Eq} error magnitude did not benefit from increased model complexity (Table 2).

Model agreement to the friction response of samples that were damaged during the testing (Figs. 2c, 2d, 2h, and 2i) was not as good as that of undamaged samples as measured by R^2 values (Table 2). Also, μ_{Eq} error magnitude plus standard deviation exceeded 50% in damaged samples for both models (Table 2). Since there appeared to be no benefit of the added complexity of the 2 exp model, the 1 exp model was chosen for subsequent experiments.

ROC Curve Analysis

To assess the performance of individual features in classifying samples as damaged or undamaged, ROC curves were generated and the respective AUROCs were compared. The R^2 and *max drop* features appeared to be the best indicators of damage out of the six features examined (Fig. 4a). Only AUROC for R^2 and *max drop* were significant at $p < 0.008$ (Table 3). The confidence intervals overlapped, suggesting that the classification accuracy of these two features was not significantly different. Confidence areas for these curves (Figs. 4b and 4c) confirmed overlap between the estimates of these two curves. This was further confirmed by the hypothesis test using the bootstrapping method for correlated ROC curves which resulted in $p = 0.2268$.

Damage Detection by SVM

Feature Selection and Parameter Optimization—Features selected for each kernel by forward selection are shown in Table 4. The R^2 feature was the 1st feature selected for each kernel, indicating that it was the feature with the greatest influence on misclassification rate. After feature selection at various C values and $KS = 1$, C and KS were optimized for the selected kernel-feature combinations. Kernel, features, and C and KS values, of the complete optimized models are shown in Table 4. Models will be referenced as model 1 through model 7 as indicated in Table 4.

Model Selection by Cross-Validation—Classification accuracy of the seven parameter-optimized models was determined by cross-validation. Median classification accuracy for damaged samples was greater than 80% for all models, and greater than 90% for models 3, 5, 6, and 7 (Fig. 5). Median total classification accuracy (Figs. 6a and 6b) and median classification accuracy of undamaged samples (Fig. 6d) were also quantified to determine if classification accuracy of undamaged samples was sacrificed in favor of correct identification of damaged samples. While total classification accuracy appeared to be similar for models 3, 6, and 7 (Figs. 6a and 6b), classification accuracy of undamaged samples was significantly better for model 3, and model 7, vs. model 6 (Fig. 6d). The median classification accuracy of damaged samples using model 3 was significantly greater than that of model 7 (100 vs. 80%), although there were many outliers in the lower range of the model 3 distribution (Fig. 6c). Note that in Fig. 7 offset has been applied so the number of outliers can be better assessed.

Based on the results of the various classification accuracies, model 3 was selected for additional characterization. To validate the findings from the forward feature selection process, we compared the classification accuracies of the full model, to the model when only R^2 was included as a feature. This was performed by first optimizing the C and KS parameters for the 3rd order polynomial kernel with R^2 as the only feature. Table 5 shows parameters of the original model 3 (full model) alongside the parameters for the model with a reduced feature set (partial model). Median classification accuracy of the full model was significantly better for total classification (Figs. 7a and 7c), damaged sample classification (Fig. 7b), and internally damaged sample classification (Fig. 7d), but not significantly different than the partial model for undamaged sample classification (Fig. 7e).

Figure 8 provides a visual of classification by the full model. For this example, 20% of samples from each class (damaged and undamaged) were held out during SVM training then classified with the trained SVM. Training data are shown as grey dots. Those selected by the algorithm as support vectors (those nearest the boundary) are circled in grey. Test samples are red dots or green squares for damaged and undamaged samples respectively. The classification by the SVM is indicated by red and green circles. Test samples circled in their own color were classified correctly, while those circled in the opposite color were classified incorrectly. In this particular train/test set, all samples were classified correctly except one undamaged sample.

From this visual it is seen that undamaged samples are typically found in the high R^2 -low $(\mu_{max} - \mu_{Eq})/\mu_{Eq}$ region. Also, that this model results in an overtrained SVM, i.e., the 2nd decision boundary in the high R^2 -high $(\mu_{max} - \mu_{Eq})/\mu_{Eq}$ region is superfluous. However, in this example, it has no effect on sample classification.

Discussion

The potential of friction testing to provide critical information other than that typically investigated during tribology is significant. Specifically, friction testing can provide understanding into the damage characteristics of cartilage under multi-axial loading, such as those recently reported by Lizhang *et al.*, for NC.¹⁶ We also recently published a study on this topic for EC,²⁷ in which histological cross-sectioning was required to verify the presence or absence of damage.

In the current study, we investigated a method to objectively detect damage that occurred during friction testing based on the observed friction response. We found that the SVM classifier tested here was able to detect damaged samples with a median accuracy of 90%. Median classification accuracy of undamaged samples was slightly lower, at approximately 82%. However, the number of undamaged samples was only a fraction of the number of damaged samples, and additional studies with an increased number of undamaged samples are warranted. The low ratio of undamaged to damaged samples, even at sub-physiologically relevant normal loads, is likely a result of inferior biochemical and biomechanical properties typical of constructs generated using these methods, as reported in a previous study.²⁶ While it would be useful to increase the number of undamaged samples, a large number of damaged samples was important for this study in order to observe the effect of damage on

the friction pattern. The friction testing and damage detection method described here could potentially be used as a quality assurance test for EC prior to implantation; specifically, constructs that fail during a pre-implantation “stress-test” could be eliminated with 90% accuracy.

To facilitate feature extraction from the friction signal, we also developed a modified model of biphasic lubrication for comparing the measured friction response to an ideal response. Features of the friction signal derived from this comparison, as well as intrinsic features of the friction signal were tested for their ability to detect damage in EC caused by frictional shear loads. To model the ideal response, we developed simplified lubrication models (Eq. 2) that were able to characterize the overall response of undamaged cartilage (Figs. 1c and 1d), as demonstrated by R^2 values of 0.87 or greater for pooled TE and native samples (Table 1). Of the two models investigated, we favor use of the less complex single exponential model, since there were no statistical differences between R^2 , initial friction error, and final friction error values, between the single and double exponential models. AIC analysis also supports selection of the single exponential model since the AIC value is lower for the single exponential model.

The final friction error magnitude of the single exponential model was acceptable for undamaged samples (<5%) at final (equilibrium) friction, but not for initial friction. However, we have not seen the accuracy of initial friction determined from the biphasic lubrication model reported in literature. Given that the initial friction values are very small compared to equilibrium friction, relatively small deviations result in large initial friction errors, so the initial friction error may be large in the original model (Eq. 1) as well. The modified models (Eq. 2) were unable to characterize the response of damaged samples. One limitation of our modified biphasic lubrication model (Eq. 2) is that ϕ loses physical significance due to the interdependence of this parameter and the coefficients of the exponentials. Large errors between measured solid fraction and ϕ when using the full model have been reported,¹⁴ so the usefulness of the model would not typically be compromised by the modification employed here.

Other limitations of this study include that other machine learning classifiers such as K-means, decision trees, and neural networks, were not tested. Additionally, the results of this study are dependent on the assumption that samples which were only internally damaged incurred that damage during friction testing, since internal damage could not be assayed prior to friction testing. However, the internal cracking counted as damage in this study (Fig. 2h) was observed to be a result of exposure to frictional shear in our previous study.²⁷

Another limitation of this study is the use of notched boxplots to compare models, since this method cannot strictly determine if distributions are significantly different. However, given the non-normal distributions, notched boxplots are an accepted method to determine significant difference at an approximate p -value of $p < 0.05$. Furthermore, the limitations of the notched boxplot only affect model comparisons, and do not affect the reported accuracies. Additional limitations of this study are that the visual classification of sample damage was subjective, that this friction-based method doesn't quantify the amount of damage, that only scaffold-free EC was used so other types should be tested as well, and that

μ is a systems variable so classifiers would need to be trained for each study in which different friction devices and test parameters are used.

If a damage detection method could determine not only *if*, but also *when* damage occurs during testing, it would enable other valuable measures of EC quality, such as work-to-failure. While this study was not designed specifically to determine *when* damage occurred during friction testing, two of the six features tested in this study were independent of the length of the friction test, and were not dependent on comparison to the modified biphasic model: *consec. drops*, and *max drop*. If these two features had been able to classify damaged vs. undamaged samples in this study, it would be reasonable to hypothesize that they might be useful in determining when damage occurred. Forward feature selection for the SVM models resulted in models which did not contain either of these features except in the case of model 1, where *max drop* % was chosen as the 5th and final feature. Therefore, while these models are useful for determining *if* damage occurred, they are likely inadequate for determining *when* damage occurred. However, ROC analysis showed that 42% of damaged samples had a *max drop* value greater than 0.021, while all undamaged samples had *max drop* values lower than this threshold (Figs. 4a and 4c). Given this relationship, it is not surprising that AUROC of *max drop* was significant at the specified $p < 0.008$ level. Clearly a relationship between *max drop* and damage does exist. Therefore, further studies into the usefulness of *max drop* to detect when damage occurs may be warranted. That this feature was not selected in most SVM models during forward feature selection is likely due to the dependent nature of the features, i.e., R^2 and possibly other features are influenced by *max drop*.

In this study, mechanisms explaining the concurrence of construct damage and aberrations in the friction signal (reduction of R^2 , etc.) have not been investigated. However, simple physical interpretations may exist. One such explanation is that as the surface of the construct adjacent to the counterface is peeled away from the surface of the construct which remains attached to the metal platen (Fig. 2d), the peeled portion becomes attached to the counterface and moves with the counterface. Thus, the bearing surface is shifted such that it is internal to the construct, creating a cartilage on cartilage bearing. A rapid change in the bearing surface, from cartilage on glass to cartilage on cartilage, could, theoretically, produce an abrupt change in friction force, since the CoF for cartilage on cartilage bearings are reported to differ from CoF of cartilage on glass.⁵ In addition, for constructs exhibiting surface peeling or internal damage, it is likely that damage results in at least partial loss of the pressurized interstitial fluid which is known to affect the cartilage friction response.¹³ These physical changes to the bearing surface or fluid pressurization are likely to be, at least partially, responsible for the observed aberrations in the friction pattern. Another possible physical phenomenon which could account for abrupt changes in the CoF at the initiation of damage is a transient pooling of fluid in cracks or holes that are at the counterface-contacting surface of the construct; fluid shearing would likely reduce the CoF.

In conclusion, this study demonstrates that the friction signal can be used to identify, with a high level of accuracy, which samples are damaged during friction testing. Of the features and methods tested, R^2 and $(\mu_{max} - \mu_{Eq})/\mu_{Eq}$ used in a SVM model yielded the highest classification accuracy. Even samples that were damaged only internally were classified

correctly with a high level of accuracy. The ability of this technique to distinguish damaged from undamaged constructs, without the need for construct cross-sectioning, enables frictional-shear testing as a candidate for construct quality assurance testing prior to construct implantation. With further development, this approach to damage detection during friction testing may be able to yield further measures of cartilage function, including those requiring knowledge of when damage occurs during the friction test.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NIH NIDCR grant number R01 DE015322 (J.E.D.), and NIH NIAMS Grant Number P01 AR053622 (J.M.M), and under the Ruth L. Kirschstein National Research Service Award T32 AR007505 from the NIH NIAMS (G.A.W.).

References

1. Basalo IM, Chahine NO, Kaplun M, Chen FH, Hung CT, Ateshian GA. Chondroitin sulfate reduces the friction coefficient of articular cartilage. *J Biomech.* 2007; 40:1847–1854. [PubMed: 17084404]
2. Burges CJC. A tutorial on support vector machines for pattern recognition. *Data Min Knowl Disc.* 1998; 2:121–167.
3. Burnham KP, Anderson DR. Multimodel inference understanding AIC and BIC in model selection. *Sociol Methods Res.* 2004; 33:261–304.
4. Caligaris M, Ateshian GA. Effects of sustained interstitial fluid pressurization under migrating contact area, and boundary lubrication by synovial fluid, on cartilage friction. *Osteoarthr Cartil.* 2008; 16:1220–1227. [PubMed: 18395475]
5. Caligaris M, Canal CE, Ahmad CS, Gardner TR, Ateshian GA. Investigation of the frictional response of osteoarthritic human tibiofemoral joints and the potential beneficial tribological effect of healthy synovial fluid. *Osteoarthr Cartil.* 2009; 17:1327–1332. [PubMed: 19410031]
6. Changzheng C, Changcheng S, Yu Z, Nan W. Fault diagnosis for large-scale wind turbine rolling bearing using stress wave and wavelet analysis. *ICEMS.* 2005; 3:2239–2244.
7. Chung C, Erickson IE, Mauck RL, Burdick JA. Differential behavior of auricular and articular chondrocytes in hyaluronic acid hydrogels. *Tissue Eng A.* 2008; 14:1121–1131.
8. Forster H, Fisher J. The influence of loading time and lubricant on the friction of articular cartilage. *Proc Inst Mech Eng H.* 1996; 210:109–119. [PubMed: 8688115]
9. Forster H, Fisher J. The influence of continuous sliding and subsequent surface wear on the friction of articular cartilage. *Proc Inst Mech Eng H.* 1999; 213:329–345. [PubMed: 10466364]
10. Graindorge S, Stachowiak G. Changes occurring in the surface morphology of articular cartilage during wear. *Wear.* 2000; 241:143–150.
11. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982; 143:29–36. [PubMed: 7063747]
12. Henderson JH, Welter JF, Mansour JM, Niyibizi C, Caplan AI, Dennis JE. Cartilage tissue engineering for laryngotracheal reconstruction: comparison of chondrocytes from three anatomic locations in the rabbit. *Tissue Eng.* 2007; 13:843–853. [PubMed: 17394383]
13. Jones AR, Gleghorn JP, Hughes CE, Fitz LJ, Zollner R, Wainwright SD, Catterson B, Morris EA, Bonassar LJ, Flannery CR. Binding and localization of recombinant lubricin to articular cartilage surfaces. *J Orthop Res.* 2007; 25:283–292. [PubMed: 17106881]
14. Krishnan R, Kopacz M, Ateshian GA. Experimental verification of the role of interstitial fluid pressurization in cartilage lubrication. *J Orthop Res.* 2004; 22:565–570. [PubMed: 15099636]
15. Liu, B. *Web Data Mining.* Berlin: Springer; 2011.

16. Lizhang J, Fisher J, Jin Z, Burton A, Williams S. The effect of contact stress on cartilage friction, deformation and wear. *Proc Inst Mech Eng H*. 2011; 225:461–475. [PubMed: 21755776]
17. Martin I, Miot S, Barbero A, Jakob M, Wendt D. Osteochondral tissue engineering. *J Biomech*. 2007; 40:750–765. [PubMed: 16730354]
18. McGill R, Tukey JW, Larsen WA. Variations of box plots. *Am Stat*. 1978; 32:12–16.
19. Morrell KC. Corroboration of in vivo cartilage pressures with implications for synovial joint tribology and osteoarthritis causation. *Proc Natl Acad Sci USA*. 2005; 102:14819–14824. [PubMed: 16203974]
20. Mu T, Nandi AK, Rangayyan RM. Screening of knee-joint vibroarthrographic signals using the strict 2-surface proximal classifier and genetic algorithm. *Comput Biol Med*. 2008; 38:1103–1111. [PubMed: 18823882]
21. Rangayyan RM, Oloumi F, Wu Y, Cai S. Fractal analysis of knee-joint vibroarthrographic signals via power spectral analysis. *Biomed Signal Process Control*. 2013; 8:23–29.
22. Reddy NP, Rothschild BM, Mandal M, Gupta V, Suryanarayanan S. Noninvasive acceleration measurements to characterize knee arthritis and chondromalacia. *Ann Biomed Eng*. 1995; 23:78–84. [PubMed: 7762885]
23. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform*. 2011; 12:77.
24. Schmidt TA, Gastelum NS, Nguyen QT, Schumacher BL, Sah RL. Boundary lubrication of articular cartilage: role of synovial fluid constituents. *Arthritis Rheum*. 2007; 56:882–891. [PubMed: 17328061]
25. Unsworth A, Dowson D, Wright V. The frictional behavior of human synovial joints—part i: natural joints. *J Lubr Technol*. 1975; 97:369–376.
26. Wada Y, Enjo M, Isogai N, Jacquet R, Lowder E, Landis WJ. Development of bone and cartilage in tissue-engineered human middle phalanx models. *Tissue Eng A*. 2009; 15:3765–3778.
27. Whitney GA, Jayaraman K, Dennis JE, Mansour JM. Scaffold-free cartilage subjected to frictional shear stress demonstrates damage by cracking and surface peeling. *J Tissue Eng Regen Med*. 2014; 10:1002/term.1925.
28. Whitney GA, Mera H, Weidenbecher M, Awadallah A, Mansour JM, Dennis JE. Methods for producing scaffold-free engineered cartilage sheets from auricular and articular chondrocyte cell sources and attachment to porous tantalum. *Biores Open Access*. 2012; 1:157–165. [PubMed: 23514898]

Figure 1.

Friction signal processing and modeling. The raw CoF signal was measured as a sinusoidal wave, a result of the oscillation of the glass counterface (a and b, (b) is a zoom of (a) from the region highlighted in blue). The individual oscillations are not seen in (a) due to the short length of the x-axis. The raw CoF (a) was then rectified and smoothed with a low pass FIR filter resulting μ , as shown in gray in (c). (a–c) represent 1 native sample. The response from engineered cartilage often showed large drops in the magnitude of μ , depicted by the arrow in (d), which represents 1 engineered cartilage construct. The fit of the single exponential model (1 exp) which gives $\mu(t)$ is shown as a black dotted line in (c and d).

Figure 2.

Damage during friction testing. Representative India ink stained stereo micrographs (a–e) and safranin O stained histological cross-sections (f–j) of engineered (a–d, f–i) and native (e, j) cartilage prior to friction testing (a, f) and post friction testing (b–e, g–j). The red dotted line in (b) marks the boundary of the cartilage construct; outside the red dotted line is the metal platen. An example of an internally damaged sample is shown in (h), and the corresponding surface image is in (c). The arrow in (c) shows a possible crack that covers a very small portion of the surface. An example of complete surface peeling is shown in (d) and (i). Histological samples are oriented with the articular surface (or surface contacting the counterface) towards the bottom of the image. The apparent damage in (j) (arrow) was caused by removal from the subchondral bone. The scale bar in (i) applies to (g–h) as well. The cartilage punch in (a) is 3 mm in diameter, and the punches in (b–e) are 5 mm in diameter.

Figure 3.

Friction signal feature extraction. (a) The features *consec. drops* and *max drop* were calculated from the CoF signal, i.e. the friction signal before processing. Undulations with dots over them indicate undulations whose maximum value was less than the maximum value of preceding undulation. When multiple sets of undulations with consecutive decreases were present in the CoF signal, as in (a), the *consec. drops* feature was set to the longest number of consecutively decreasing undulations. (b) The features *max error*, *endpoint error*, $\mu_{max} - \mu_{Eq}$, and R^2 (not shown), were calculated from μ and the 1-exp fit. (a and b) See the text for further explanation features labeled on these plots.

Figure 4.

ROC curve analysis. (a) ROC curves of all features. Text associated with arrows indicates the threshold value for the point indicated by the arrow. For ease of comparison, the R^2 curve has been transposed to orient it such that the area is greater than 0.5 since it is the only feature for which smaller values are associated with damage. (b and c) Curves from (a) with AUROCs significantly greater than 0.5 at $p < 0.008$. The curves in these two panels were further analyzed for significant differences between them. The black line is the ROC curve, while the grey line is the identity line, or the line represented by random chance. Blue shaded areas are the confidence areas of the ROC curves. Note that the reverse ordering of the x-axis in (b and c), and that R^2 has not been transposed in (b). This second format is that used by pROC in the R statistical program. TPR = True Positive Ratio; FPR = False Positive Ratio.

Figure 5. Damaged sample classification accuracy by optimized models. Notched box plots represent the distributions of the various optimized models. Distributions with non-overlapping notches are significantly different at approximately the 95% confidence level. The median and notch for model 6 are at the upper limit of the interquartile range.

Figure 6.

Additional classification accuracies of optimized models. (a) Total classification accuracy distributions. (b) Total classification accuracy notched box plots. (c) Classification accuracy of internally damaged samples. (d) Classification accuracy of undamaged samples.

Figure 7.

Influence of individual features in model 3. (a) Total classification accuracy distributions. (b) Classification accuracy of damaged samples. (c) Total classification accuracy. (d) Classification accuracy of internally damaged samples. (e) Classification accuracy of undamaged samples.

Figure 8.

SVM classification of damaged and undamaged samples. The full model (model 3) was used to classify 20% of the samples in each class (damaged and undamaged). Training data are in grey, and test data are red or green for damaged and undamaged samples respectively. Circled training data were selected by the SVM as the support vectors used to generate the decision boundary. Test data (colored markers) circled in their own color are correctly classified. Thus, the SVM labels samples falling in the region between the two boundaries as damaged, and those outside the boundaries as undamaged. (b) is a zoom of (a) from the region bounded in blue.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Number of samples in various groups

	Total	Damaged	Internal damage	Curved	Flat	Min load (kg)	Max load (kg)
Native	3	0	0	2	1	0.39	1.1
Engineered	70	62	6	10	60	0.33	3.09
Total	73	62	6	12	61	0.33	3.09

Table 2

Friction signal model results

1 Exp		2 Exp	
Mean	σ^2	Mean	σ^2
Initial error magnitude			
Undamaged	94.5% ±	89.1% ±	62.5% ± 77.9%
Damaged	52.4% ±	92.0% ±	38.5% ± 58.4%
Final error magnitude			
Undamaged	3.4% ±	2.5% ±	3.7% ± 2.7%
Damaged	22.9% ±	36.3% ±	22.9% ± 36.1%
Average R^2			
Undamaged	0.87 ±	0.13	0.88 ± 0.14
Damaged	0.42 ±	0.34	0.42 ± 0.34
AIC			
Undamaged	31.2		32.0
Damaged	41.9		46.4

AIC Akaike Information Criterion.

Table 3

AUROC statistics

Identifier	Area	Std. error	95% CI		p-value
			Lower	Upper	
R^2	0.870	0.058	0.755	0.984	0.0001
Endpoint error	0.528	0.072	0.386	0.670	0.7696
Max error	0.727	0.080	0.571	0.883	0.0169
Consec. drops	0.670	0.110	0.454	0.886	0.0737
$(\mu_{max} - \mu_{Eq})/\mu_{Eq}$	0.543	0.082	0.382	0.703	0.6548
Max drop	0.792	0.076	0.643	0.941	0.0022

CI confidence interval.

Table 4

Optimized models

Model	Kernel	Features							Loss UN
		1	2	3	4	5	KS	BC	
1	Linear	R^2	Max error	Consec. drops	μ_{max}/μ_{Eq}	Max drop %	0.08037	4.2640	0.13501
2	2nd order	R^2	Max error	μ_{max}/μ_{Eq}			0.64059	39.374	0.17543
3	3rd order	R^2	μ_{max}/μ_{Eq}				0.27589	0.00646	0.06751
4	4th order	R^2	μ_{max}/μ_{Eq}				1.84968	24.312	0.09453
5	5th order	R^2	Max error	μ_{max}/μ_{Eq}			1.22561	0.27437	0.09453
6	Gaussian	R^2	Max error				0.15597	0.02223	0.08108
7	Gaussian	R^2					0.08951	0.05580	0.09453

KS kernel space, BC box constraint, Loss UN misclassification of undamaged samples.

Table 5

Feature significance of selected model

Model	Original model	Kernel	Features			KS	BC
			1	2	μ_{max}/μ_{Eq}		
Full model	3	3	R^2	R^2	μ_{max}/μ_{Eq}	0.27589	0.00646
Partial model		3	R^2			0.34212	0.02403