

A machine learning approach to predict cellular mechanical stresses in response to chemical perturbation

VigneshAravind SubramanianBalachandar,^{1,2} Md. Mydul Islam,¹ and R. L. Steward, Jr.^{1,3,*}

¹Department of Mechanical and Aerospace Engineering, College of Engineering, University of Central Florida, Orlando, Florida; ²Department of Cell Biology, University of Virginia, Charlottesville, Virginia; and ³Burnett School of Biomedical Sciences, College of Medicine, University of Central Florida, Orlando, Florida

ABSTRACT Mechanical stresses generated at the cell-cell level and cell-substrate level have been suggested to be important in a host of physiological and pathological processes. However, the influence various chemical compounds have on the mechanical stresses mentioned above is poorly understood, hindering the discovery of novel therapeutics, and representing a barrier in the field. To overcome this barrier, we implemented two approaches: 1) monolayer boundary predictor and 2) discretized window predictor utilizing either stepwise linear regression or quadratic support vector machine machine learning model to predict the dose-dependent response of tractions and intercellular stresses to chemical perturbation. We used experimental traction and intercellular stress data gathered from samples subject to 0.2 or 2 $\mu\text{g/mL}$ drug concentrations along with cell morphological properties extracted from the bright-field images as predictors to train our model. To demonstrate the predictive capability of our machine learning models, we predicted tractions and intercellular stresses in response to 0 and 1 $\mu\text{g/mL}$ drug concentrations which were not utilized in the training sets. Results revealed the discretized window predictor trained just with four samples (292 images) to best predict both intercellular stresses and tractions using the quadratic support vector machine and stepwise linear regression models, respectively, for the unseen sample images.

SIGNIFICANCE The ML framework we present here can be used to predict the mechanical response of any cell type capable of adherence to a flexible substrate as a function of chemical perturbation. The proposed ML can directly predict the intercellular stresses or tractions as a function of drug dosage and/or monolayer/cell coverage area which could potentially reduce the experimental time on studying the mechanics of cells to external chemicals or mechanical constraints. We believe our findings could be helpful in accelerating drug discovery and increase our understanding in the role of cellular stresses during disease progression.

INTRODUCTION

The mechanosensing ability of cells is critical for many biological processes such as cell migration, growth, and differentiation and is therefore physiologically and pathologically relevant (1). As the adherent cell migrates through and probes its environment contractile forces must be generated and, when migrating as a collective, these same adherent cells also interact with each other by transmitting intercellular stresses through cell-cell junctions. Cell-cell junctions enable fast, long-distance mechanical force communication, which subse-

quently yields intercellular stresses (2). We measure intercellular stresses using monolayer stress microscopy (MSM) and tractions using traction force microscopy (TFM) (1).

Tractions are generated via actomyosin contractility and actin polymerization and were initially measured by observing wrinkles exerted by single cells on thin, silicone membranes (3). This method would later be extended to measure tractions generated by cells attached to extracellular matrix (ECM)-coated polymers and tractions generated by cells attached to flexible, micropost force sensor arrays (3). Three-dimensional (3D) tractions have also been measured by cells cultured on top of flexible polymers and cells embedded in 3D matrices (4). A novel Förster resonance energy transfer sensor-based approach has also been employed to measure tractions by estimating the change in excitation energy of fluorescent protein

Submitted March 16, 2023, and accepted for publication July 24, 2023.

*Correspondence: rstewardjr@ucf.edu

Editor: Baohua Ji.

<https://doi.org/10.1016/j.bpj.2023.07.016>

© 2023 Biophysical Society.

markers that are sensitive to external forces (5,6). The many variations by which tractions can be measured has led to a host of studies revealing the importance of tractions both physiologically and pathologically. In fact, tractions have been shown to be important in cell adhesion, spreading, migration, and ECM remodeling (7–10) and are linked to various pathologies including cancer metastasis, fibrosis, and inflammation (11–14).

For cells in a monolayer, in addition to tractions being generated at the cell-substrate interface, intercellular stresses are generated at the cell-cell interface. Intercellular stresses have been suggested to be important in tissue morphogenesis, epithelial-mesenchymal transition, wound healing, and tumor progression (15–17). Intercellular stresses can be calculated using MSM, which was first described by Tambe et al. (18). In brief, both normal and shear intercellular stresses are recovered from tractions by assuming a monolayer sheet of cells as elastic thin plates and imposing Newton's force balance and strain compatibility equations (19). However, other groups would also develop additional alternative methods to calculate intercellular stresses as well, as discussed below.

More recently, Bayesian inversion stress microscopy (BISM) (20) and Kalman inversion stress microscopy (KISM) (21) were presented as predictive models for internal stress fields based on corresponding traction force data. BISM and KISM use experimental traction force data as a likelihood function to make intercellular stress field predictions using Bayesian statistics (Bayes' theorem). BISM can infer internal stress fields only from a single traction field image, but a dimensionless regularization parameter must be calculated from the experimental data to make predictions. KISM, however, is capable of estimating internal stress fields from a time lapse of traction data (movie) with its accuracy depending on time resolution of the traction data (20,21). There are several mechanical factors, notably actin filament orientation, substrate stiffness, cell area, local cell curvature, and external forces, that have been demonstrated to affect tractions and intercellular stresses (22–27). Kang et al. simulated rearrangement of actin filaments in response to mechanical stretch (28). Ghosh et al. predicted endothelial cell tractions by using substrate stiffness and cell area (27). A positive correlation was observed for both stiffness and area with respect to the tractions (23). Contrastingly, Han et al. (24) reported that larger cell area lowered average traction forces in human pulmonary artery endothelial cells.

The physiological relevance of tractions and intercellular stresses are equally as important as their pathological ramifications. For example, upregulation of endothelial contractility and increase in tractions via actin stress fiber and ECM remodeling are linked to higher cellular and vascular stiffness and vascular hyperpermeability, as seen in hypertension and atherogenesis (29–32). Furthermore, tractions and intercellular stresses have been linked to vascular hyperpermeability via ROCK1/2 and thrombin-mediated pathways (33). Such hyperpermeability induces loss of blood-brain

barrier integrity and has also been linked with several neurological disorders such as multiple sclerosis, stroke, and traumatic brain injury (34).

Taking into account the physiological and pathological relevance of both tractions and intercellular stresses, we propose that TFM and MSM are powerful tools that can be utilized to clarify the biomechanical mechanisms of various diseases, some of which have been mentioned above and potentially lead to novel therapeutics. Thus far, the development of novel regenerative medicine and drugs has been heavily researched to treat numerous vascular-related diseases such as, for example, hypertension, atherosclerosis, stroke, and coronary artery disease (35–37). However, a barrier exists in the field as each of these mechanics-related diseases often require treatment with drugs, whose influence is dose dependent. We propose that studying the biomechanical mechanism by which certain drugs influence cells and the subsequent tissues they constitute could improve drug efficacy. However, the time and the financial resources required to evaluate the impact of various drugs on cell behavior can be overwhelming and expensive. Machine learning (ML) offers an alternative approach that has the potential to resolve or at the bare minimum mitigate the issues previously mentioned. It was therefore our objective to apply the ML approach to predict the dose-dependent cellular, biomechanical response to chemical stimulation.

In this paper, we utilized ML to predict both tractions and intercellular stresses as a function of drug concentration and cell morphological parameters such as monolayer perimeter and cell area. Predictive models were created using stepwise linear regression (SLR) and quadratic support vector machine (QSVM) regression learners. The SLR and QSVM models were trained using two different training sets: 1) a monolayer boundary set (MBS) that utilizes monolayer area, monolayer perimeter, and drug concentration as predictors, and 2) a discretized window set (DWS) that utilizes endothelial cell area, endothelial cell, perimeter, and drug concentration as predictors.

MATERIALS AND METHODS

Cell culture

Human umbilical vein endothelial cells (HUVECs) were purchased commercially and cultured in Medium 200 supplemented with 1% penicillin-streptomycin (Corning, Corning, NY) and large vessel endothelial supplement. HUVECs, Medium 200, and large vessel endothelial supplement were purchased from Thermo Fisher Scientific, Waltham, MA. HUVECs were cultured on 0.1% gelatin (Sigma-Aldrich, St. Louis, MO)-coated flasks at 37°C and 5% CO₂.

Preparation of polyacrylamide gel and cellular micropatterning

The protocol for preparing polyacrylamide (PA) gels can be found in Steward et al. (38). Glass-bottom petri dishes (35 mm, Cellvis, Mountain View, CA) were treated with bind silane solution for 45 min after which the dishes were rinsed with deionized water and air dried. The PA solution is made by mixing ultrapure water, 40% acrylamide (Bio-Rad), 2% bis-acrylamide (Bio-Rad,

Hercules, CA), and fluorescent beads (yellow or Texas Red, $\sim 0.5 \mu\text{m}$ diameter, Invitrogen, Waltham, MA). The stiffness of the PA gel can be fine-tuned by changing the ratio of bis-acrylamide and acrylamide solutions. The PA solution is kept in a vacuum chamber for 40 min. Ten percent of ammonia persulfate and N,N,N',N' -tetramethylethane-1,2-diamine is added to the degassed PA solution, which initiates polymerization reaction. After the addition of ammonia persulfate and N,N,N',N' -tetramethylethane-1,2-diamine, the PA solution is mixed well and plated on the petri dish wells. Hydrophobic coverslips are then placed on top of the PA solution, after which the dishes are inverted to allow more fluorescent beads to settle on top layer of the polymerizing gel. The subsequent PA gels with stiffness $\sim 1.2 \text{ kPa}$ and height $\sim 100 \mu\text{m}$ were used for the experiments as described by Stroka and Aranda-Espinoza (39).

Cellular micropattern preparation

Polydimethylsiloxane (PDMS) was used to fabricate thin micropatterns as described previously (38). A thin cross section of PDMS (Dow Corning, Midland, MI) was prepared by mixing silicone base with a curing agent (20:1) and the mixture was then poured into a 100 mm petri dish. The PDMS mixture in the petri dish with no air bubbles was then incubated at 70°C overnight. Thin, circular cross sections of cured PDMS (16 mm) were fabricated using a hole puncher. Small 2 mm holes were made on the circular PDMS section using a biopsy punch. The fabricated micropatterns were gently placed on the top layer of the PA gels.

SANPAH burning and Col I-FN treatment

The petri dish samples with PDMS micropatterns stamped on PA gels were then subject to treatment with sulfosuccinimidyl-6-(4-azido-2-nitrophenylamino) hexanoate (Sulfo-SANPAH, ProteoChem, Hurricane, UT) dissolved in 0.1 M HEPES buffer solution (Thermo Fisher Scientific) and kept under a UV lamp for 8 min. After SANPAH and UV treatments, the samples were treated with type Collage I (Col-I).

After the treatment of Col-I (Advanced BioMatrix, Carlsbad, CA) overnight at 4°C , the excess protein solution was carefully removed and HUVECs were seeded at a density of $\sim 50 \times 10^4$ cells/mL. After 60–75 min, micropatterns were cautiously removed using a tweezer. The HUVEC monolayer samples were incubated at 37°C and 5% CO_2 for at least 24 h to allow enough time for the formation of cell-cell junctions before experimentation.

Pharmacological perturbation experiments

We have previously demonstrated tractions and intercellular stresses to exhibit a dose-dependent response to 2,5-dihydroxychalcone (chalcone) (40). Chalcone is a unique drug in that it has been reported to solely disrupt the gap junction connexin 43 (40). Therefore, for this study HUVECs were seeded at a density of 50×10^4 cells/mL onto polyacrylamide gels for at least 36 h. After this time, independent experiments were conducted where chalcone was added at the following concentrations: $0.2 \mu\text{g/mL}$ (low concentration) and $2 \mu\text{g/mL}$ (high concentration). Experiments performed to validate our ML model consisted of controls and exposing HUVECs to $1 \mu\text{g/mL}$ of chalcone.

Time-lapse microscopy

Time-lapse microscopy was performed using a Zeiss inverted microscope with a $10\times$ objective and Hamamatsu camera. Fluorescent and phase contrast images were acquired at 5 min intervals for 1 h before the addition of chalcone. After this time, medium with chalcone was added at the concentrations mentioned above and imaging was done for an additional 5 h at 5 min intervals after removing the control medium. These experiments culminated with HUVEC monolayers being treated with $10\times$ trypsin to

detach the cells from the substrate and acquire a “stress free” image of our gel surface, an image essential for traction calculation.

TFM and MSM

First, in-plane displacement of fluorescent beads located on the top surface of the gel was computed using a custom-written particle image velocimetry routine (41). A window size of 32 pixels with an overlap of 0.75 was utilized for each region of interest. Cross correlation between each window in the reference image with no cells attached to the gel (stress-free configuration) was computed against a window occupying the same coordinates in the fluorescent image with cells attached (stressed configuration) sequentially across all the window blocks. The displacements were calculated in the x and y coordinate system in pixels using an iterative cross correlation function. The displacement calculated from peak cross correlation function between each reference-fluorescence window pair was assigned to the center coordinates of those windows (3). TFM (42) and MSM (7,18) were used as described by Butler et al. and Trepatt et al. to calculate the cell-substrate tractions and cell-cell intercellular stresses, respectively. In brief, the gel deformations described above were used to calculate the tractions and the intercellular stresses were subsequently recovered from traction force maps by using straightforward force balance equations imposed by Newton’s law (19). We computed the local 2D stress tensor within the monolayer by converting the maximum principal stress (σ_{max}) and minimum principal stress (σ_{min}) along the principal plane by rotating the local coordinate system along the principal orientation. The average normal stress $(\sigma_{\text{max}} + \sigma_{\text{min}})/2$ and maximum shear stress $(\sigma_{\text{max}} - \sigma_{\text{min}})/2$ were calculated at each point in the elements within the monolayer.

Building of training tables

Experimental TFM and MSM data used for building our training tables were solely obtained from the 0.2 and $2 \mu\text{g/mL}$ chalcone experiments. We used results gathered from four time-lapse experiments (two per each chalcone concentration: 0.2 and $2 \mu\text{g/mL}$ chalcone experiments, respectively) and concatenated the data sets row-wise into one large data set. The size of experimental data dictates the number of predictions required for our ML model. In general, we based our training data size on the size of the experimental data gathered. We found model prediction accuracy to be dependent on training data size, as higher training size led to higher accuracies. As far as our ML models are concerned, large data sets can help with increasing prediction accuracies. However, accuracy was found to no longer increase once a certain data size threshold was reached. The data gathered from TFM and MSM experiments were in the form of a cropped, 246×246 square matrix, which contained approximately 60,516 data points per time series once converted into column form. This data column was subsequently used to build training tables for our predictor and response variables. The three response variables were RMS (root-mean square) tractions, maximum principal stress, and minimum principal stress. Column training tables were created separately for each of these response variables with the help of predictors before training them with ML models. Predictors utilized were drug concentration and either of the two morphological predictors: monolayer boundary or discretized monolayer windows. In addition, the principal-component analysis tool in MATLAB was used to confirm that our selected predictors explained more than 95% of variability in the data.

MBS

The monolayer boundary training set was generated from the monolayer area. A contour of the monolayer shape was obtained from phase contrast images and converted to a binary image. Utilizing this binary image, monolayer area and perimeter were then calculated using the “regionprops” command in MATLAB (The MathWorks, Natick, MA). The monolayer area and

perimeter remain the same for all the time frames as our monolayer was seeded as a constrained, micropattern during experiments. Monolayer boundary data cannot capture individual HUVEC movement and individual HUVEC geometry within the monolayer since measurements are only made for the entire monolayer as a whole. Furthermore, as the single monolayer area and perimeter do not change, the monolayer boundary data set can only train ML models to predict the average values of our response variables. Ramifications of this include the lack of ability to predict the dynamic behavior of our response variables, which would be revealed through a time series. To overcome the monolayer boundary data's lack of ability to predict time series information we developed the discretized window training table.

DWS

The discretized window training set was generated by converting binarized phase images of cellular monolayers into multiple overlapping grids, which we call here “windows.” For any given frame we used a grid resolution of 32×32 to generate our windows and from these windows we calculated an area and perimeter, which were then stored as a column matrix for each image. The area and perimeter of each discretized window gives information about coverage of endothelium within that window. We choose this method as computing cell area and perimeter in an automated fashion would be difficult and time consuming because these cell properties change from frame to frame. The discretized window training method eliminates the need to track individual cell properties and is $2.6\times$ faster compared with our traditionally used fast Fourier transform-based, cross correlation method used for cell displacement tracking for a single image pair of 32 square pixels each (Fig. S4).

Selection of ML models

The generated predictor variables and response variables mentioned above were utilized in several ML models in MATLAB using the Machine Learning Toolbox. Since the data were relatively small, we were able to utilize all the available ML models in MATLAB. A fivefold variable cross validation was chosen to validate the models where the inputted training table was randomly divided into five groups (see supporting material, section 3) and one group was held as the “test data” from which the model makes predictions after training data from the remaining four sets. ML models that 1) had the highest R^2 and lowest root-mean square error (RMSE) values when predicting our response variables (tractions and intercellular stresses) and 2) were most sensitive to our predictor variables (area, perimeter, and drug concentration) served as our selection criteria. Utilizing this criterion, the support vector machine (SVM)- and SLR-based models emerged as our best candidates for predicting TFM and MSM data as a function of area, perimeter, and drug concentration. Trained models based on SVM and SLR were found to be sensitive to new data with less overfit and high R^2 values compared with the other models. SVM and SLR are described below.

SVM

SVM is widely used in classification and regression analysis. SVM makes predictions by utilizing various kernelization techniques to effectively find the hyperplane that separates the support vectors by transforming the data in higher dimensions in classification problems. In regression, the hyperplane is the best fit line or curve that effectively fits most of the points or support vectors. SVM can use different kernels such as linear, quadratic, or Gaussian to compute the transformation in higher dimensions at a reduced computational cost, where it is easier to find the hyperplane that is closer to most of the data points. SVM optimizes the hyperplane so that the distance from each support vector is minimized within the chosen decision boundary. The decision boundary encompasses the data that are closer to the hy-

perplane. A margin of tolerance (ϵ) can be inputted by the user to increase the tolerance level from the decision boundary. SVM is well equipped to handle complex, nonlinear data, robust to outliers and has better prediction accuracy compared with linear regression models. However, SVM is susceptible to noise and not preferred for very large or very small data sets (43).

SLR

SLR evaluates the independent or exploratory variables one by one through a forward selection rule (variables added at each step), backward-elimination rule (all variables included), or a bidirectional rule (combination of both forward selection and backward elimination) by computing the t statistics for the coefficients of the selected variable at each step. The SLR regression model helps choose the best independent variables efficiently based on statistical significance. Although very effective in minimizing the number of predictors, the SLR model is prone to choosing wrong variables if there is a large number of predictors and small amount of data (44).

RESULTS

A schematic showing the overview of the two proposed ML methods for the prediction of tractions and intercellular stresses is shown in Fig. 1. The defined predictor sets and corresponding features are shown in Fig. S1.

RMS tractions are best predicted by the SLR model

Both ML models (SLR and QSVM) were generated using data from cells exposed to two different chalcone concentrations (0.2 and 2 $\mu\text{g/mL}$). Particularly, the QSVM model utilizing MBS and the SLR model utilizing DWS yielded the best results for traction predictions, while MBS-SLR and DWS-QSVM yielded the best results for intercellular stress predictions based on validation (training) data. The values, coefficient of determination (R^2), and RMSE were calculated for different chalcone concentrations (0, 0.2, 1, and 2 $\mu\text{g/mL}$) based on the average from three unseen monolayers (test data) for each concentration. The results for 0 and 1 $\mu\text{g/mL}$ chalcone concentrations were broken down and shown separately in the model validation section comparing it with the corresponding actual experimental results.

Monolayer boundary predictor

A relatively low R^2 value of 0.54 was seen for RMS tractions predicted by the QSVM model for different chalcone concentrations (0, 0.2, 1, and 2 $\mu\text{g/mL}$) (from Fig. 2 *a* [MBS] and Table S1). RMSE for the RMS traction predictions from the QSVM model was 5 Pa as shown in Fig. 2 *b* MBS and Table S1. The predicted RMS tractions distributions as a function of chalcone concentrations (0.2 and 2 $\mu\text{g/mL}$) is shown in Fig. 3, *b* and *e* MBS-QSVM panel. The corresponding averages are shown in Table S2 and Fig. 2 *g*.

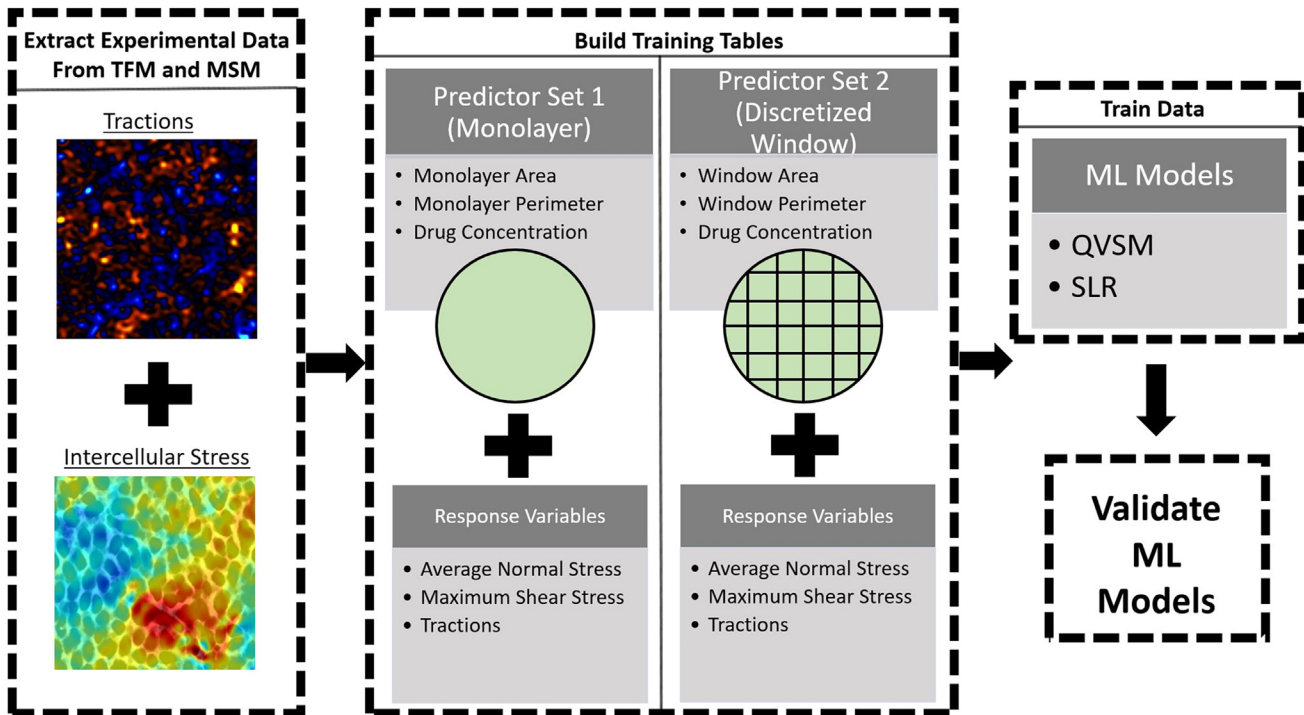


FIGURE 1 Flowchart showing implementation of machine learning approach to predict intercellular stress and tractions. The QSVSM and SLR machine learning models were used along with two predictor sets: 1) monolayer boundary set and 2) discretized window set. To see this figure in color, go online.

Discretized window predictor

Similar to results discussed above, predictions were made using the SLR model for three unseen samples for 0.2 and 2 $\mu\text{g/mL}$ chalcone concentrations using discretized window training sets and compared against the experimental results (Fig. 3, *c* and *f* in the DWS-SLR panel). The R^2 and RMSE values for RMS tractions predicted by SLR were 0.85 and 4.86 Pa for different chalcone concentrations (0, 0.2, 1, and 2 $\mu\text{g/mL}$), as shown in Fig. 2, *a* and *b* and Table S1 in the DWS panel. The average tractions for 0.2 and 2 $\mu\text{g/mL}$ are shown in Fig. 3 *g* and Table S2 in the discretized window sections, respectively.

Intercellular stresses are best predicted by the QSVSM model

Monolayer boundary predictor

The overall coefficient of determination (R^2) for the average normal stress predicted by the SLR model for different chalcone concentrations (0, 0.2, 1, and 2 $\mu\text{g/mL}$) was 0.74, while the R^2 for maximum shear stress predicted by SLR was 0.88 (Fig. 2 *a* and Table S3, monolayer boundary training column). RMSE values for the predictions from the SLR model were 24.18 and 47.26 Pa, respectively, for average normal and maximum shear stresses, respectively. The predicted intercellular stresses (average normal and maximum shear) distributions for 0.2 and 2 $\mu\text{g/mL}$ chalcone concentrations

are shown in Fig. 4 (average normal stress, Fig. 4, *b* and *h*) and (maximum shear stress, Fig. 4, *e* and *k*) in the MBS-SLR panel and the corresponding averages are shown in Fig. 4, *m* and *n* (bar plots) and Table S4 in the monolayer boundary training sections.

Discretized window predictor

The R^2 values for average normal stress predicted by QSVSM was 0.81 and the RMSE value was 22.20 Pa (Fig. 2, *a* and *b* and Table S3). A higher R^2 value of 0.93 was seen for maximum shear stress predicted by the QSVSM model with RMSE of 52.67 Pa, representing our best predicted values for different chalcone concentrations 0, 0.2, 1, and 2 $\mu\text{g/mL}$, respectively. The predicted intercellular stresses: average normal and maximum shear distributions for 0.2 and 2 $\mu\text{g/mL}$ chalcone concentrations are shown in Fig. 4 (average normal stress, Fig. 4, *c* and *i* and maximum shear stress, Fig. 4, *f* and *l*). The average values for the same are shown in Table S4 and Fig. 4, *m* and *n* (bar plots) in the discretized window section.

Validation of our ML model

We determined (based on validation score for the trained data) the prediction accuracy was to be optimally determined overall with SLR for traction predictions and the QSVSM model for

R² Accuracy and Root Mean Square Error (RMSE) Plots

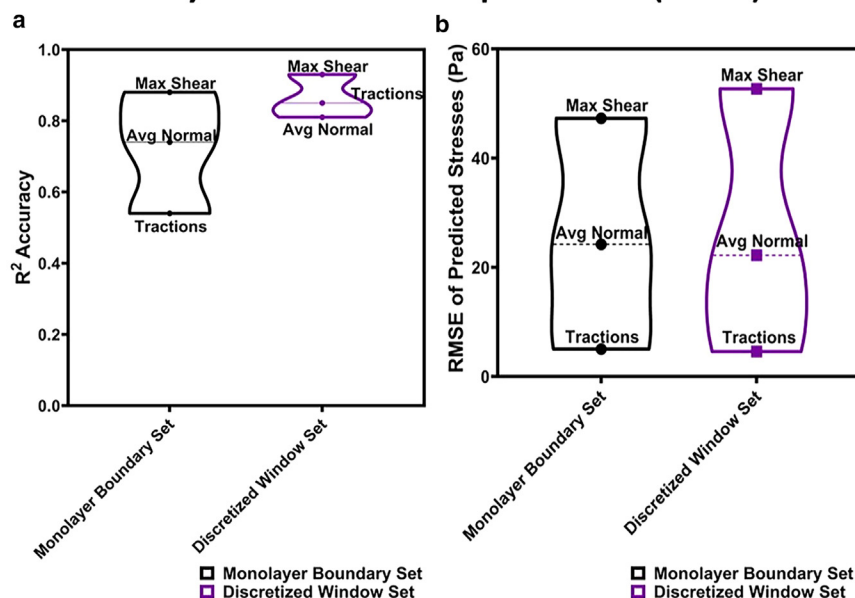


FIGURE 2 R² accuracy and root-mean square error (RMSE) plots for monolayer boundary and discretized window predictors' overall R² accuracy (a) and RMSE in Pa (b) for RMS traction, average normal stress, and maximum shear stress predictions using the monolayer boundary set (MBS) and the discretized window set (DWS). MBS utilizes QSVM for traction predictions and the SLR model for intercellular stresses predictions while DWS utilizes SLR for traction and QSVM for intercellular stresses predictions. To see this figure in color, go online.

intercellular stress predictions utilizing discretized window sets. Similarly, for monolayer boundary sets, QSVM and SLR models were determined to best predict tractions and intercellular stresses, respectively. Therefore, it was our next objective to evaluate the predictive capability of our ML model. To this end, we predicted intercellular stresses and RMS tractions in response to 0 and 1 $\mu\text{g/mL}$ of chalcone and compared our predictions to experimental values obtained from these concentrations. We highlight the fact that 0 and 1 $\mu\text{g/mL}$ chalcone data were not used in either of the training sets to build our model and the results we present here represent a true test prediction and not a fit. Predicted spatial distributions of RMS tractions for monolayers exposed to 0 and 1 $\mu\text{g/mL}$ compared with experimental results are shown in the Fig. 5, *a–n* (MBS-SLR and DWS-QSVM panels) and intercellular stresses in Fig. 6, *a–g* (MBS-QSVM and DWS-SLR panels). The raw values shown in Tables S5 and S6 are for traction and intercellular stresses predictions. Overall, higher accuracy (R²) of 0.85, 0.81, and 0.93 for traction, average normal, and maximum shear stresses predictions, respectively, were observed for discretized window sets compared with the monolayer boundary set (Fig. 2 *a*).

DISCUSSION

In this paper, we present both tractions and intercellular stresses as a function of monolayer morphology and drug concentration for the first time using regression ML models QSVM and SLR with two different training sets, utilizing one of the monolayer boundary or discretized window sets. Overall, SLR models exhibited optimal predictive capability with monolayer boundary training sets for intercellular stress predictions while the QSVM model did

perform more optimally than SLR for discretized window training sets for the same. The R² value of intercellular stress predictions was highest for the QSVM model utilizing a discretized window training set (R² = 0.81 for average normal stress and R² = 0.93 for maximum shear stress). The SLR model utilizing the discretized window training set had the highest R² value of 0.85 for traction predictions. While QSVM best predicted intercellular stresses, traction predictions were relatively poor. We believe this could be attributed to the fact that the discretized window training predictor was used and the potential boundary effects from edges that were used to train the regression models. Discretized window training requires preprocessing to extract input variables from the new data (image) for making predictions, while the monolayer boundary method requires very little preprocessing to extract the independent variables. The prediction time for a new (unseen) image was relatively same across all the ML models built on the two different training sets (~10–20 min).

Out of the two proposed training sets, ML models built on discretized window training were very much sensitive to the outliers. In the discretized window method, the predictions from all three regression models (SLR, LSVM, and QSVM) were subject to a filter that excludes the outliers < −1000 Pa and > +1000 Pa (values corresponding to the 0.15th, 99.95th percentiles) for intercellular stresses and > +600 Pa (values corresponding to the 99.95th percentile) for RMS tractions to filter noise due to overfitting. Also, the most important feature of the discretized window method is the ability to generate time series predictions of tractions and intercellular stresses based on the changes in area and perimeter observed in each small overlapping grid across the entire field of view. The R² values especially

Tractions (Pa) ML Predictions

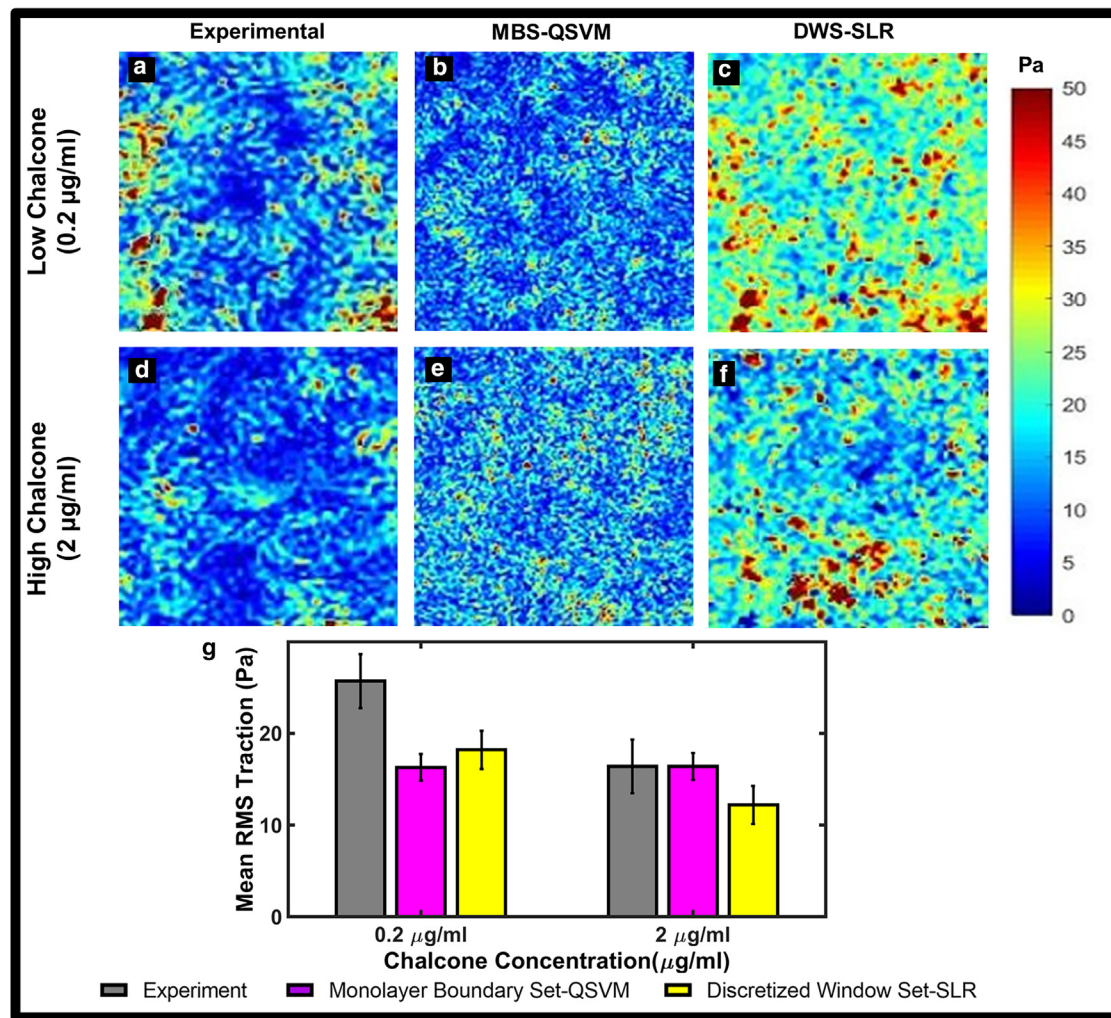


FIGURE 3 Predicted tractions using monolayer boundary and discretized window predictors. Experimental-, MBS-QSVM-, and DWS-SLR-predicted RMS traction distributions for 0.2 $\mu\text{g/ml}$ chalcone concentration (a–c) and 2 $\mu\text{g/ml}$ chalcone concentration (d–f) and the corresponding averages from three samples for each condition with standard errors shown in the bar plot (g). To see this figure in color, go online.

for tractions can be improved by using more predictors such as substrate stiffness, cell orientation, cell velocity, etc. Also, training the models using a cropped monolayer section inside the circular monolayer samples could reduce the noise due to boundary effects and improve prediction accuracy. Time series prediction is possible with discretized window training, but the R^2 value was too low because of overfitting due to noise for both tractions and intercellular stresses. One of the future works is to reduce the boundary effects of the training sets by using cropped subsets inside the monolayer and employing cell velocity as an independent variable for time series predictions. However, including cell velocity as a predictor can increase the preprocessing and computation time compared with window-based and monolayer area predictors. Other future works include making time series predictions for expanding monolayers using

monolayer boundary training sets and explore using more predictors such as cell velocity, cell orientation, curvature, cell area, ECM concentration, and substrate stiffness.

Comparison of predicted results with experimental results

Predictions made for each response variable were computed for a new image with different monolayer area and drug concentration. Although, 292 images were used to build the models, they correspond to only 4 monolayer samples with just two chalcone concentrations (0.2 and 2 $\mu\text{g/ml}$). Predictions were made for new phase images with chalcone concentrations set to 0 and 1 $\mu\text{g/ml}$, respectively. Master averages of the predicted results (RMS traction, average normal stress, maximum shear stress) were compared with the actual master

Intercellular Stresses (Pa) ML Predictions

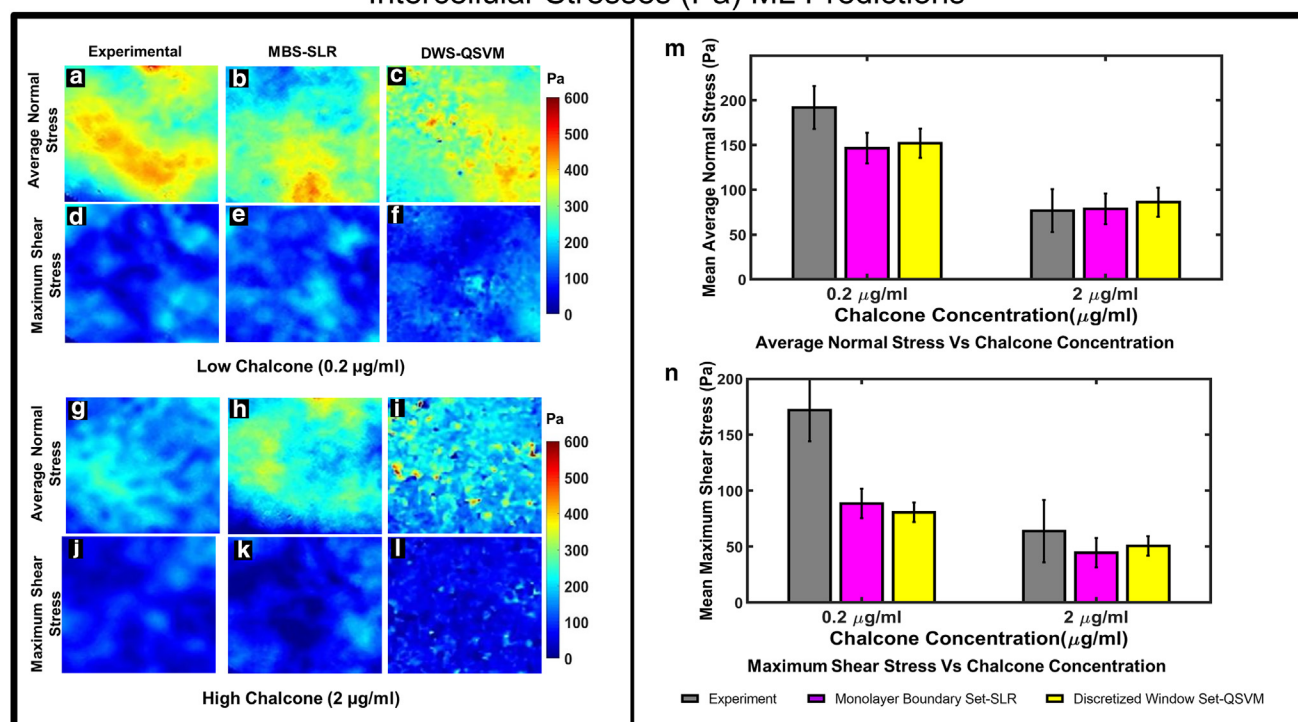


FIGURE 4 Predicted intercellular stresses using monolayer boundary and discretized window predictors. Experimental-, MBS-SLR-, and DWS-QSVM-predicted average normal stress distributions for 0.2 µg/mL chalcone concentration (a–c) and 2 µg/mL chalcone concentration (g–i). Experimental-, MBS-SLR-, and DWS-QSVM-predicted maximum shear stress distributions for 0.2 µg/mL chalcone concentration (d–f) and 2 µg/mL chalcone concentration (j–l). Corresponding averages of the distributions from three monolayers for each condition with standard errors are shown in (m) and (n) for average normal and maximum shear stresses, respectively. To see this figure in color, go online.

average of the results computed from TFM and MSM. The master average of the results is the average of the response variable in all the overlapping blocks that make up the image and the average taken across the entire image time series that follow in the MSM experiment. Predictions were made for three monolayer samples for each of the four chalcone concentrations, using the proposed ML models based on two different training sets and compared against the actual experimental result averages from three samples for each chalcone concentration obtained from MSM. Pearson correlation coefficient (R) was calculated in Excel 2016, using the formula: $R = \text{Covariance}(A, B) / (\text{SD } A * \text{SD } B)$. If the trend between MSM and predicted results were close, then we get high R values and vice versa. R^2 (coefficient of determination) is a much-preferred metric in statistics, which represents the percent of variability in data that can be explained by the ML model. RMSE was also computed for the predicted averages of the response variable against the experimental MSM data based on averages from 3 samples, across 60,516 grid points in the time series for each chalcone concentration.

CONCLUSION

The ML model proposed here was built using intercellular stresses and traction data derived from MSM and TFM,

respectively. Traction and intercellular stress measurement in part relies upon the following premises: 1) the cell type to be studied is physically adhered to a substrate, 2) the underlying substrate is soft enough such that underlying substrate deformations can be measured, and 3) if in a monolayer, cells are physically connected through cell-cell junctions. Intercellular stresses and tractions measured from a HUVEC monolayer can depend on a wide variety of factors such as monolayer area, substrate stiffness, cell area, curvature, external forces and biochemical substances, or drugs. Cellular tractions have been correlated to factors such as substrate stiffness, cell area, local curvature, and cell geometry (9). BISM and KISM were recently introduced to predict intercellular stresses from TFM. However, we believe this is the very first work to predict both tractions and intercellular stresses from the following independent variables: drug concentration, monolayer area/perimeter, and discretized window area/perimeter based on prior knowledge of tractions and intercellular stresses. We predicted both tractions and intercellular stresses using QSVM and SLR regression learners built on two different training sets. Accuracy was higher for shear stress compared with normal stress and traction predictions.

Although this study only measured tractions and intercellular stresses in endothelial cells, both tractions and

Tractions (Pa) ML Predictions

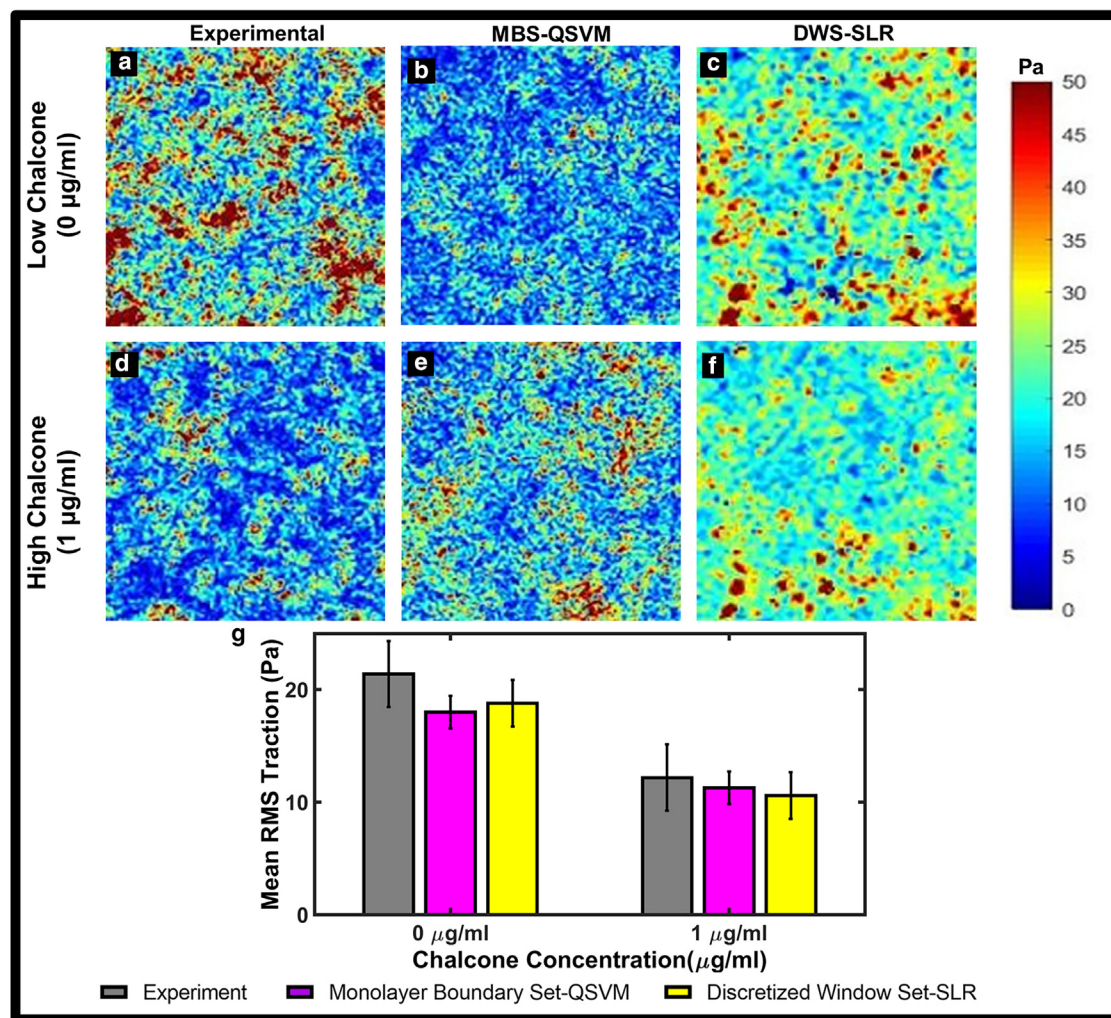


FIGURE 5 Model validation of tractions using monolayer boundary and discretized window predictors. Experimental-, MBS-QSVM-, and DWS-SLR-predicted RMS traction distributions for 0 $\mu\text{g/ml}$ chalcone concentration (a–c) and 1 $\mu\text{g/ml}$ chalcone concentration (d–f) and the corresponding averages from three samples for each condition with standard errors is shown in the bar plot (g). To see this figure in color, go online.

intercellular stresses have been measured by other groups in several cell types such as epithelial cells, fibroblasts, cardiomyocytes, osteocytes, cancer cells, and leukocytes, for example (7,40,45–48). Taken into account, we believe the model proposed here can be applied to any cell type that satisfies the conditions mentioned above. Inclusion of more predictors such as substrate stiffness should increase the prediction accuracy (R^2) of the regression learners. With these promising results, more accurate predictions can be obtained by adding additional predictors such as cell velocity, substrate stiffness, ECM concentration, for example, to improve the accuracy of the current models to make more reliable predictions of stress distributions and time series predictions instead of just the average trend. Furthermore, although not possible with our current experimental setup we believe that high-throughput prediction is possible as

several groups have already reported high-throughput traction measurement (49–51). We reason that, if high-throughput tractions can be measured, high-throughput intercellular stresses can also potentially be measured as well. For example, Yoshie et al. utilized PDMS-coated multiwell plates to measure high-throughput tractions generated by EMT, which reflects the impact our method could have on the field.

Although we tested the impact of chalcone on HUVECS and measured subsequent tractions and intercellular stresses, other groups have measured the impact of barrier-disruptive and barrier-protective compounds on the endothelium. For example, Hardin et al. demonstrated barrier-disruptive compounds (thrombin, histamine, and H2O2) to increase intercellular stresses, whereas barrier-protective compounds (Y27632, S1P, and HGF) decreased intercellular stresses

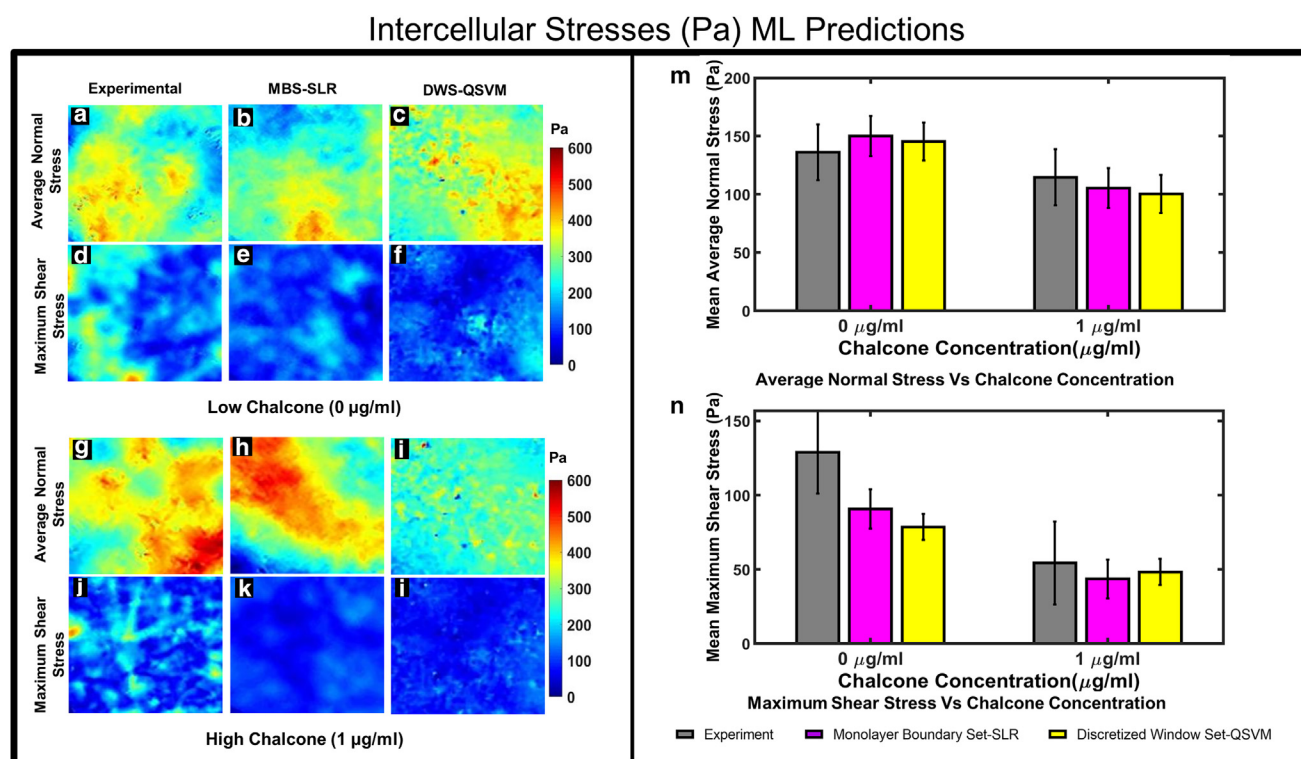


FIGURE 6 Model validation of intercellular stresses using monolayer boundary and discretized window predictors. Experimental-, MBS-SLR-, and DWS-QSVM-predicted average normal stress distributions for 0 µg/mL chalcone concentration (a–c) and 1 µg/mL chalcone concentration (g–i). Experimental-, MBS-SLR-, and DWS-QSVM-predicted maximum shear stress distributions for 0 µg/mL chalcone concentration (d–f) and 1 µg/mL chalcone concentration (j–l). Corresponding averages of the distributions from three monolayers for each condition with standard errors are shown in (m and n) for average normal and maximum shear stresses, respectively. To see this figure in color, go online.

(52). However, the experiments mentioned above were done using human lung microvascular endothelial cells and with only one drug concentration. Taken into account these drugs do have a documented impact on cell-derived mechanical forces we suggest these drugs to be an additional potential application for our model. In summary, we believe this work will be helpful for accelerating research in experimental drugs that target cell mechanical activity such as cellular contractility and tissue barrier strength and function. Proposed ML models could be applicable for testing cell mechanics of any anchorage-dependent cells as a function of pharmacological and other morphological parameters that can influence cell mechanics.

SUPPORTING MATERIAL

Supporting material can be found online at <https://doi.org/10.1016/j.bpj.2023.07.016>.

AUTHOR CONTRIBUTIONS

M.M.I. ran experiments and provided data for building the ML model. V.A.S.B. built the ML models, analyzed data, and generated figures. V.A.S.B. and R.L.S. interpreted the data, and V.A.S.B., M.M.I., and R.L.S. contributed to the writing of the manuscript.

ACKNOWLEDGMENTS

This work was funded by the National Heart, Lung, and Blood Institute of the National Institute of Health under award K25HL132098. This material is based on work supported by the National Science Foundation CAREER Award under grant no. 2045750.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

- Kurzawa, L., B. Vianay, ..., M. Théry. 2017. Dissipation of contractile forces: the missing piece in cell mechanics. *Mol. Biol. Cell.* 28:1825–1832.
- Patel, G., N. Xu, ..., D. T. Tambe. 2019. Mechanical signaling in a pulmonary microvascular endothelial cell monolayer. *Biochem. Biophys. Res. Commun.* 519:337–343.
- Wang, J. H. C., and J.-S. Lin. 2007/01/03 2007. Cell traction force and measurement methods. *Biomech. Model. Mechanobiol.* 6:361–371. <https://doi.org/10.1007/s10237-006-0068-4>.
- Maskarinec, S. A., C. Franck, ..., G. Ravichandran. 2009. Quantifying cellular traction forces in three dimensions. *Proc. Natl. Acad. Sci. USA.* 106:22108–22113.
- Kong, H. J., T. R. Polte, ..., D. J. Mooney. 2005. FRET measurements of cell-traction forces and nano-scale clustering of adhesion ligands

- varied by substrate stiffness. *Proc. Natl. Acad. Sci. USA*. 102:4300–4305. <https://doi.org/10.1073/pnas.0405873102>.
6. Gayraud, C., and N. Borghi. 2016. FRET-based Molecular Tension Microscopy. *Methods*. 94:33–42. <https://doi.org/10.1016/j.ymeth.2015.07.010>.
7. Treppe, X., M. R. Wasserman, ..., J. J. Fredberg. 2009. Physical forces during collective cell migration. *Nat. Phys.* 5:426–430.
8. Dembo, M., and Y.-L. Wang. 1999. Stresses at the cell-to-substrate interface during locomotion of fibroblasts. *Biophys. J.* 76:2307–2316.
9. Kraning-Rush, C. M., S. P. Carey, ..., C. A. Reinhart-King. 2011. The role of the cytoskeleton in cellular force generation in 2D and 3D environments. *Phys. Biol.* 8, 015009.
10. Rape, A. D., W.-h. Guo, and Y.-l. Wang. 2011. The regulation of traction force in relation to cell shape and focal adhesions. *Biomaterials*. 32:2043–2051.
11. Indra, I., V. Undyala, ..., K. A. Beningo. 2011. An *in vitro* correlation of mechanical forces and metastatic capacity. *Phys. Biol.* 8, 015015.
12. Koch, T. M., S. Münster, ..., B. Fabry. 2012. 3D traction forces in cancer cell invasion. *PLoS One*. 7, e33476.
13. Kraning-Rush, C. M., J. P. Califano, and C. A. Reinhart-King. 2012. Cellular traction stresses increase with increasing metastatic potential. *PLoS One*. 7, e32572.
14. Marinković, A., J. D. Mih, ..., D. J. Tschumperlin. 2012. Improved throughput traction microscopy reveals pivotal role for matrix stiffness in fibroblast contractility and TGF- β responsiveness. *Am. J. Physiol. Lung Cell Mol. Physiol.* 303:L169–L180.
15. Cetera, M., G. R. Ramirez-San Juan, ..., S. Horne-Badovinac. 2014. Epithelial rotation promotes the global alignment of contractile actin bundles during *Drosophila* egg chamber elongation. *Nat. Commun.* 5:5511. <https://doi.org/10.1038/ncomms6511>.
16. Wang, S., K. Matsumoto, and K. M. Yamada. 2020. Reconstituting Stratified Epithelial Branching Morphogenesis by Engineering Cell Adhesion. Preprint at bioRxiv. <https://doi.org/10.1101/2020.06.24.165795>.
17. Barriga, E. H., K. Franze, ..., R. Mayor. 2018. Tissue stiffening coordinates morphogenesis by triggering collective cell migration *in vivo*. *Nature*. 554:523–527. <https://doi.org/10.1038/nature25742>.
18. Tambe, D. T., C. C. Hardin, ..., X. Treppe. 2011. Collective cell guidance by cooperative intercellular forces. *Nat. Mater.* 10:469–475.
19. Tambe, D. T., U. Croult, ..., J. J. Fredberg. 2013. Monolayer stress microscopy: limitations, artifacts, and accuracy of recovered intercellular stresses. *PLoS One*. 8, e55172.
20. Nier, V., S. Jain, ..., P. Marcq. 2016. Inference of Internal Stress in a Cell Monolayer. *Biophys. J.* 110:1625–1635. <https://doi.org/10.1016/j.bpj.2016.03.002>.
21. Nier, V., G. Peyret, ..., P. Marcq. 2018. Kalman Inversion Stress Microscopy. *Biophys. J.* 115:1808–1816. <https://doi.org/10.1016/j.bpj.2018.09.013>.
22. Oakes, P., S. Banerjee, ..., M. Gardel. 2014. Geometry regulates traction stresses in adherent cells. *Biophys. J.* 107:825–833.
23. Califano, J. P., and C. A. Reinhart-King. 2010. Substrate stiffness and cell area predict cellular traction stresses in single cells and cells in contact. *Cell. Mol. Bioeng.* 3:68–75.
24. Han, S. J., K. S. Bielawski, ..., N. J. Sniadecki. 2012. Decoupling substrate stiffness, spread area, and micropost density: a close spatial relationship between traction forces and focal adhesions. *Biophys. J.* 103:640–648.
25. Parker, K. K., A. L. Brock, ..., D. E. Ingber. 2002. Directional control of lamellipodia extension by constraining cell shape and orienting cell tractional forces. *Faseb. J.* 16:1195–1204.
26. Oakes, P. W., T. C. Bidone, ..., M. L. Gardel. 2018. Lamellipodium is a myosin-independent mechanosensor. *Proc. Natl. Acad. Sci. USA*. 115:2646–2651.
27. Ghosh, K., Z. Pan, ..., R. A. F. Clark. 2007. Cell adaptation to a physiologically relevant ECM mimic with different viscoelastic properties. *Biomaterials*. 28:671–679.
28. Kang, J., R. L. Steward, ..., K. M. Puskar. 2011. Response of an actin filament network model under cyclic stretching through a coarse grained Monte Carlo approach. *J. Theor. Biol.* 274:109–119.
29. Maruo, N., I. Morita, ..., S. Murota. 1992. IL-6 increases endothelial permeability *in vitro*. *Endocrinology*. 131:710–714. <https://doi.org/10.1210/endo.131.2.1639018>.
30. Marcos-Ramiro, B., D. García-Weber, and J. Millán. 2014. TNF-induced endothelial barrier disruption: beyond actin and Rho. *Thromb. Haemostasis*. 112:1088–1102. (in eng). <https://doi.org/10.1160/th14-04-0299>.
31. Huynh, J., N. Nishimura, ..., C. A. Reinhart-King. 2011. Age-related intimal stiffening enhances endothelial permeability and leukocyte transmigration. *Sci. Transl. Med.* 3:112ra122. (in eng). <https://doi.org/10.1126/scitranslmed.3002761>.
32. Urbano, R. L., C. Furia, ..., A. M. Clyne. 2017. Stiff Substrates Increase Inflammation-Induced Endothelial Monolayer Tension and Permeability. *Biophys. J.* 113:645–655. <https://doi.org/10.1016/j.bpj.2017.06.033>.
33. Beckers, C. M., N. Knezevic, ..., G. P. van Nieuw Amerongen. 2015. ROCK2 primes the endothelium for vascular hyperpermeability responses by raising baseline junctional tension. *Vascul. Pharmacol.* 70:45–54.
34. Najjar, S., S. Pahlajani, ..., D. Chong. 2017. Neurovascular Unit Dysfunction and Blood-Brain Barrier Hyperpermeability Contribute to Schizophrenia Neurobiology: A Theoretical Integration of Clinical and Experimental Evidence. *Front. Psychiatry*. 8:83. (in eng). <https://doi.org/10.3389/fpsy.2017.00083>.
35. Kim, S. U., and J. de Vellis. 2009. Stem cell-based cell therapy in neurological diseases: A review. *J. Neurosci. Res.* 87:2183–2200. <https://doi.org/10.1002/jnr.22054>.
36. Su, J. B. 2015. Vascular endothelial dysfunction and pharmacological treatment. *World J. Cardiol.* 7:719.
37. Deshpande, D. D., D. R. Janero, and M. M. Amiji. 2011. Therapeutic strategies for endothelial dysfunction. *Expert Opin. Biol. Ther.* 11:1637–1654.
38. Steward, R., Jr., D. Tambe, ..., J. J. Fredberg. 2015. Fluid shear, intercellular stress, and endothelial cell alignment. *Am. J. Physiol. Cell Physiol.* 308:C657–C664. <https://doi.org/10.1152/ajpcell.00363.2014>.
39. Stroka, K. M., and H. Aranda-Espinoza. 2011. Endothelial cell substrate stiffness influences neutrophil transmigration via myosin light chain kinase-dependent cell contraction. *Blood*. 118:1632–1640. <https://doi.org/10.1182/blood-2010-11-321125>.
40. Islam, M. M., and R. L. Steward, Jr. 2019. Perturbing endothelial biomechanics via Connexin 43 structural disruption. *JoVE*. 152, e60034.
41. Gui, L., and S. T. Wereley. 2002. A correlation-based continuous window-shift technique to reduce the peak-locking effect in digital PIV image evaluation. *Exp. Fluid.* 32:506–517.
42. Butler, J. P., I. M. Tolić-Nørrelykke, ..., J. J. Fredberg. 2002. Traction fields, moments, and strain energy that cells exert on their surroundings. *Am. J. Physiol. Cell Physiol.* 282:C595–C605.
43. Raj, A.. Unlocking the True Power of Support Vector Regression. <https://towardsdatascience.com/unlocking-the-true-power-of-support-vector-regression-847fd123a4a0#:~:text=Support%20Vector%20Regression%20is%20a,the%20maximum%20number%20of%20points>.
44. Lewis, M. 2007. Stepwise versus Hierarchical Regression: Pros and Cons. Online Submission.
45. Munevar, S., Y.-l. Wang, and M. Dembo. 2001. Traction force microscopy of migrating normal and H-ras transformed 3T3 fibroblasts. *Biophys. J.* 80:1744–1757.
46. Pasqualini, F. S., A. Agarwal, ..., K. K. Parker. 2018. Traction force microscopy of engineered cardiac tissues. *PLoS One*. 13:e0194706.

SubramanianBalachandar et al.

47. Banerjee, I., K. Carrion, ..., V. Nigam. 2015. Cyclic stretch of embryonic cardiomyocytes increases proliferation, growth, and expression while repressing Tgf- β signaling. *J. Mol. Cell. Cardiol.* 79:133–144.
48. Bakker, A. D., V. C. D. Silva, ..., J. Klein-Nulend. 2009. Tumor necrosis factor α and interleukin-1 β modulate calcium and nitric oxide signaling in mechanically stimulated osteocytes. *Arthritis Rheum.* 60:3336–3345.
49. Yoshie, H., N. Koushki, ..., A. J. Ehrlicher. 2019. High throughput traction force microscopy using PDMS reveals dose-dependent effects of transforming growth factor- β on the epithelial-to-mesenchymal transition. *JoVE.* 148, e59364.
50. Rokhzan, R., C. C. Ghosh, ..., R. Krishnan. 2019. Multiplexed, high-throughput measurements of cell contraction and endothelial barrier function. *Lab. Invest.* 99:138–145.
51. Berg, I. C., and G. H. Underhill. 2019. High throughput traction force microscopy for multicellular islands on combinatorial microarrays. *Bio. Protoc.* 9:e3418.
52. Hardin, C., K. Rajendran, ..., R. Krishnan. 2013. Glassy dynamics, cell mechanics, and endothelial permeability. *J. Phys. Chem. B.* 117:12850–12856.