A transcriptomic model to predict increase in fibrous cap thickness in response to high-dose statin treatment: Validation by serial intracoronary OCT imaging

Kipp W. Johnson, Benjamin S. Glicksberg, Khader Shameer, Yuliya Vengrenyuk, Chayakrit Krittanaawong, Adam J. Russak, Samin K. Sharma, Jagat N. Narula, Joel T. Dudley, Annapoorna S. Kini

Article history:
Received 15 February 2019
Received in revised form 15 April 2019
Accepted 3 May 2019
Available online 22 May 2019

ABSTRACT

Background: Fibrous cap thickness (FCT), best measured by intravascular optical coherence tomography (OCT), is the most important determinant of plaque rupture in the coronary arteries. Statin treatment increases FCT and thus reduces the likelihood of acute coronary events. However, substantial statin-related FCT increase occurs in only a subset of patients. Currently, there are no methods to predict which patients will benefit. We use transcriptomic data from a clinical trial of rosuvastatin to predict if a patient’s FCT will increase in response to statin therapy.

Methods: FCT was measured using OCT in 69 patients at (1) baseline and (2) after 8–10 weeks of 40 mg rosuvastatin. Peripheral blood mononuclear cells were assayed via microarray. We constructed machine learning models with baseline gene expression data to predict change in FCT. Finally, we ascertained the biological functions of the most predictive transcriptomic markers.

Findings: Machine learning models were able to predict FCT responders using baseline gene expression with high fidelity (Classification AUC = 0.969 and 0.972). The first model (elastic net) using 73 genes had an accuracy of 92.8%, sensitivity of 94.1%, and specificity of 91.4%. The second model (KTSP) using 18 genes has an accuracy of 95.7%, sensitivity of 94.3%, and specificity of 97.1%. We found 58 enriched gene ontology terms, including many involved with immune cell function and cholesterol biometabolism.

Interpretation: In this pilot study, transcriptomic models could predict if FCT increased following 8–10 weeks of rosuvastatin. These findings may have significance for therapy selection and could supplement invasive imaging modalities.

© 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Acute coronary events most commonly result from atherosclerotic plaque rupture with luminal thrombosis and vessel occlusion. The plaques vulnerable to rupture demonstrate several distinct histomorphological signatures: Fibrous cap thickness (FCT) is the most important determinant of plaque vulnerability, followed by measures of the extent of plaque inflammation and necrotic core size [1,2]. FCT is best assessed in vivo by optical coherence tomography (OCT) [3–5]. However, since OCT is an invasive intravascular imaging test, it is not routinely performed. FCT is known to increase in response to statin (HMG-CoA reductase inhibitor) therapy. This increase in FCT is indicative of plaque stabilization. However, observable increases in FCT occur in only a subset of patients following statin therapy, and it would be important to identify these subjects without an invasive imaging modality [6].

Plaque morphology and composition cannot generally be identified by traditional coronary artery angiography, and numerous imaging modalities, such as intravascular ultrasound (IVUS) [7], OCT [8], and near infrared spectroscopy (NIRS) have been employed in clinical practice [9,10]. Of all the imaging strategies, only OCT imaging is capable of

---

**Abbreviations:** FCT, fibrous cap thickness; OCT, optical coherence tomography; Hmg-CoA, 3-hydroxy-3-methyl-glutaryl-coenzyme A; KTSP, K top-scoring pairs.

⁎ Corresponding author.

E-mail address: annapoorna.kini@mountsinai.org (A.S. Kini).
Acute myocardial infarction (“heart attack”) results from the rupture of atherosclerotic plaques in the coronary arteries, with subsequent aggregation of platelets and formation of a clot that blocks the flow of blood through the artery, ultimately leading to tissue death in the heart.

One of the key determinants of plaque rupture in the coronary arteries is the thickness of the fibrous cap that forms on top of plaque. When fibrous cap thickness (FCT) is larger, the plaque is more stable and less likely to rupture, and thus less likely to cause myocardial infarction. Drugs such as the statins function to increase FCT and thus reduce the risk of cardiovascular events. However, it is known that statins only increase fibrous cap thickness in a subset of patients. Not every patient prescribed a statin will have a beneficial increase in FCT.

Unfortunately, accurate measurement of FCT is difficult in living patients. The best method to measure FCT is called optical coherence tomography, which requires threading an imaging wire into the coronary arteries to directly image the atherosclerotic plaque. Thus, it would be beneficial if a test could be performed that would be able to predict whether a patient’s FCT will respond to statins without performing this invasive procedure.

**Research in context**

**Evidence before this study**

Lipid lowering agents, most prominently statins, increase the FCT of atheromas [13–15]. Statins primarily inhibit the rate-limiting enzyme in cholesterol biosynthesis, HMG-CoA reductase, and also demonstrate pleiotropic effects [6,16]. The degree of FCT thickening (and therefore protection) caused by statins is positively correlated with the statin dose [17]. Unfortunately, not all patients respond equivalently to statin therapy, and the poor responders experience greater atheroma progression and higher rates of adverse cardiovascular outcomes [18,19].

Using OCT imaging, we also observed that only a subset of patients increased the FCT with statin therapy [6]. Usually the statin response is measured by changes in serum lipid levels following the initiation of statin therapy. However, even though there is likely some correlation, serum lipid levels do not directly predict the change in plaque FCT [20]. It is therefore important to non-invasively distinguish responders from non-responders [21], so that non-responders may seek other therapeutic options. We present a predictive model for FCT response to statin therapy utilizing imaging, clinical data, and gene expression (transcriptomics) from the previously described patients included in the Yellow II clinical trial [6].

**2. Materials and methods**

**2.1. Ethics statement**

Written informed consent was obtained from all patients. This study was approved by the Institutional Review Board of the Icahn School of Medicine at Mount Sinai.

**2.2. Study design and procedure**

We analyzed patient data from the Yellow II clinical trial which has been previously described [6]. Briefly, 85 patients with stable multi-vessel coronary artery disease underwent percutaneous coronary intervention (PCI) for an obstructive lesion. Lesions had greater than or equal to 70% stenosis as measured via angiographic visual assessment and fractional flow reserve. In the same procedure, a second “non-culprit” lesion was imaged with OCT imaging, and peripheral blood mononuclear cell (PBMC) gene expression was characterized by microarray. Trial patients were then assigned to 40 mg daily rosuvastatin therapy. Full trial inclusion criteria and details are deposited at clinicaltrials.gov under the clinical trial identifier NCT01567826. Notably, 69 patients had previously been exposed to lower-intensity statin therapy, primarily atorvastatin ($n = 29$) and simvastatin ($n = 23$). At eight to ten weeks post-enrollment the patients underwent a staged PCI, the non-culprit lesion was re-imaged with OCT, and PBMC gene expression was again characterized by microarray. We ultimately obtained complete gene expression and OCT imaging data for 69 of these patients.

**2.3. Microarray preparation and gene expression data**

PBMCs were isolated from fasting blood samples by density gradient centrifugation using FicollPaque, and mRNA was isolated using the TRIzol protocol. We used the Illumina HT-12 v4 microarray for gene expression profiling. We chose PBMCs because they are readily obtainable and gene expression signatures are robust across tissues [22]. All data preparation was conducted in R version 3.3.3 [23]. Quality control and normalization was accomplished with the lumi Bioconductor package [24,25] with a variance-stabilizing transformation applied to background-adjusted probe intensities [26]. Normalization was performed with the quantile method. We corrected for batch effects resulting from staged microarray preparation using the ComBat parametric empirical Bayes technique in the R package sva [27]. Probe-to-gene mapping was conducted with the nurID mapping scheme implemented in lumi. The gene expression values are deposited in the NIH Gene Expression Omnibus (GEO) database with ID GSE86216.

**2.4. FCT imaging by OCT**

OCT images were obtained with the C7–XR OCT device (St. Jude Medical, Minneapolis, Minnesota) with continuous intracoronary contrast injection. The OCT catheter was placed at least 10 mm distal to measuring FCT. Noninvasive approaches such as computed tomography angiography [11] are able to identify the extent of necrotic core and expansive remodeling of the lesion sites, and magnetic resonance imaging is able to identify lipid-rich plaques [12]. However, the resolution of the non-invasive imaging modalities is not sufficient for the assessment of FCT. It is therefore important to develop biomarker surrogates of FCT that could distinguish patients who would or would not benefit from statin therapy.

Lipid lowering agents, most prominently statins, increase the FCT of thin-cap fibroatheromas [13–15]. Statins primarily inhibit the rate-limiting enzyme in cholesterol biosynthesis, HMG-CoA reductase, and also demonstrate pleiotropic effects [6,16]. The degree of FCT thickening (and therefore protection) caused by statins is positively correlated with the statin dose [17]. Unfortunately, not all patients respond equivalently to statin therapy, and the poor responders experience greater atheroma progression and higher rates of adverse cardiovascular outcomes [18,19].
the target lesion, after which each OCT pullback imaged 54 mm of the vessel. OCT images were evaluated offline by an independent core laboratory (Cardiovascular Research Foundation, New York, New York) and analyzed at 1-mm intervals according to previously validated criteria for plaques characterization with the St. Jude Medical review workstation. The FCT of each lipid-rich plaque was measured three times at each time point at its thinnest point (6 measurements in total). The average value was then calculated for both baseline and follow-up time points for each patient (mean baseline and follow FCT). Finally, we calculated the change in FCT between baseline and follow-up timepoints by taking the absolute value of the difference. Compared to baseline, the location of the thinnest portion of the fibrous cap at follow-up remained the same in 65 of 69 patients (94.2%).

2.5. Responder identification

Each patient was identified as either a statin responder or non-responder depending upon whether FCT increased or did not increase from baseline time point to follow-up time point (Fig. 1b). We dichotomized this measurement because it is clinically meaningful to ascertain whether a patient’s FCT will change in response to statin therapy, and thus whether statin therapy benefits a given patient. For example, in

![OCT Imaging of FCT](image)

![FCT Response Prediction](image)

![Study Workflow](image)

![Model Sensitivity Testing](image)

**Fig. 1.** Central figure illustrating the components of study. (a) An example OCT image of an atherosclerotic plaque, before and after 8–12 weeks of high intensity rosuvastatin therapy. (b) The study workflow. Blood transcriptomics and OCT imaging were performed at 69 patients at baseline and follow-up periods. 35 patients had increased FCT (responders), and 34 patients did not (non-responders). (c) Predictive modeling of FCT response. We combined clinical variables and transcriptomic data and used two machine learning methods to predict responder type. (d) Graphical explanation of extensive sensitivity testing. We iteratively performed the strategy depicted in panel (c) upon randomly selected subsets of our dataset in order to understand the variability of the results.
2014, Kim et al. published a machine learning model using the expression levels of 100 genes to predict how LDL cholesterol levels in lymphoblastic cell lines changes following statin therapy [21]. This model for LDL cholesterol was able to classify the highest 15% of responders from the lowest 15% of responders with area under the ROC curve (AUC) up to 0.98 [28]. However, the model’s utility diminished as they attempted to differentiate more moderate responders from moderate non-responders, with an AUC only modestly better than a random classifier (AUC < 0.6) for fully dichotomized data. Thus, for this initial step we chose to predict the binary outcome of response to statin therapy with increased FCT or not, instead of attempting to model the quantitative increase in cap thickness.

2.6. Statistical analysis

We used the glmnet [29] package in R [23] version 3.3.3 to fit elastic net regularized generalized linear models using a logistic link function via penalized maximum likelihood to transcription data. The elastic net mixes the regression penalty of ridge regression with that of the least absolute shrinkage and selection operator (LASSO) regression. Elastic net coefficient estimates are computed by minimizing Eq. (1).

\[
\beta = \text{argmin} \left( |y-X\beta|^2 + \lambda_1|\beta|^2 + \lambda_2|\beta||_1 \right)
\]  

(1)

\[
\min(\beta_0, \beta) = \left[ \frac{1}{N} \sum_{i=1}^{N} y_i (\beta_0 + \beta_i x_i) \right] - \log \left(1 + e^{\beta_0 + \beta_i x_i} \right) + \lambda \left(1 - \alpha \right) |\beta||_1 / 2 + a |\beta||_1
\]  

(2)

To fit K top scoring pair (KTSP) classifiers, we used the switchbox [30] Bioconductor R package. KTSP is an extension of the original Top Scoring Pairs (TSP) classification algorithm [31]. Briefly, given a matrix of expression values X with dimension genes x samples where columns 1 to N1 are samples belonging to outcome class 1 and columns N1+1 to N2 are samples of class 2, KTSP finds the best pair of genes i and j which maximize the value \( \Delta_{ij} \) as given in Eq. (3).

\[
\Delta_{ij} = |P_{ij}(1) - P_{ij}(2)|
\]  

(3)

where \( P_{ij}(1) \) and \( P_{ij}(2) \) are given by Eqs. (4) and (5):

\[
P_{ij}(1) = \text{Pr}(x_i < x_j | \text{class} = 1) = \frac{1}{N_1} \sum_{n=1}^{N_1} I(x_{in} < x_{jn})
\]  

(4)

\[
P_{ij}(2) = \text{Pr}(x_i < x_j | \text{class} = 2) = \frac{1}{N_2} \sum_{n=N_1+1}^{N_2} I(x_{in} > x_{jn})
\]  

(5)

For the extension to K pairs of genes, the next highest top-scoring pairs of genes are added to the classifier for a given range of K. The addition of genes is penalized by choosing the value of K which produces a classifier that maximizes \( \tau = \hat{\delta}_0 / \hat{\delta} \), where \( \hat{\delta}_0 \) is sum of K pairwise gene scores as given in Eq. (3) and \( \hat{\delta} \) is the maximum-likelihood estimator of the variance [32].

2.7. Incorporation of clinical variables into predictive models

To assess whether inclusion of clinical variables may help to alter or improve predictive performance, we conducted two additional analyses. First, we built an elastic net regression model as described above composed of only clinical variables. Second, we built another elastic net regression model as described above including both clinical variables as well as the transcriptomic data from the original elastic net model. A full list of clinical variables is included in the Supplemental Material.

2.8. Gene function analysis

We sought to ascertain the biological function of genes found to be predictive of FCT response by characterizing their Gene Ontology Enrichments. We used the TopGO R package to perform weighted Fisher Exact tests. We used the TopGO “weight01” algorithm to account for the hierarchical nature of Gene Ontology functional group assignments and to protect against false discovery from multiple hypothesis testing.

3. Results

3.1. FCT responder prediction

Clinical values for our patient cohort are available in Table 1. The mean increase in FCT for rosuvastatin-responders was 36.3 ± 69.8 μm. The mean change in FCT for rosuvastatin non-responders was −4.41 ± 7.05 μm. A plot demonstrating the distribution of FCT values is available in the Supplementary Material. LDL cholesterol and total cholesterol levels were not significantly different between responder and non-responders, either at baseline, follow-up, or when examining change in lipid levels from baseline to follow-up (Supplementary Material, Table S3). Using transcriptomic data to predict FCT response, we obtained a final model with leave-one-out-cross-validation (LOOCV) area under the receiver operating characteristic curve of 0.975. We thus could classify individuals as FCT statin responders or non-responders with high fidelity. The elastic net model using 73 genes had an accuracy of 92.8%, sensitivity of 94.1%, and specificity of 91.4%. Similarly, the KTSP classifier could discriminate between responders and non-responders with high performance, obtaining LOOCV accuracy of 95.7%, sensitivity of 94.3%, and specificity of 97.1% (Fig. 2a). Notably, this classifier required only 18 genes. Fig. 3 provides a visual demonstration of how well this small number of genes divide responders and non-responders.

| Table 1 | Clinical variables of individuals in dataset, stratified by Responder/Non-responder type. |
|---------|---------------------------------------------------------------|
|         | Non-responder (n = 34)                  | Responder (n = 35)       | Total (n = 69)  | P-Value |
| Gender  | Female                                      | Male                        |
| Age     | 9 (26.5%)                                  | 25 (73.5%)                  | 34 (47.1%)      | 0.489   |
| Weight  | 62.9 (11.2)                                | 65.0 (10.6)                 | 64.3 (10.6)     | 0.11    |
| Smoking | 16 (47.1%)                                 | 31 (44.9%)                  | 47 (44.9%)      | 0.913   |
| Systolic BP (mmHg) | 145 (22.2) | 137 (24.8) | 141 (23.7) | 0.17   |
| Diastolic BP (mmHg) | 73.3 (12.6) | 69.5 (10.2) | 71.4 (11.5) | 0.17   |
| Total Cholesterol (mg/dL) | 147 (42.0) | 146 (35.6) | 147 (38.6) | 0.93   |
| LDL Cholesterol (mg/dL) | 39.9 (10.6) | 41.9 (13.9) | 40.9 (12.3) | 0.52   |
| Hemoglobin A1c (%) | 6.80 (1.58) | 6.93 (1.48) | 6.89 (1.52) | 0.48   |

*P* values for continuous variables computed with the two-sample *t*-test. *P* values for categorical variables computed with the Chi-square test of independence.
3.2. Sensitivity testing

When constructing predictive models, the potential for overfitting the training dataset is of high concern, especially when sample sizes are small. We conducted extensive cross validation and sensitivity testing to characterize the stability of our predictive models to patient sampling (Fig. 1d). Briefly, our strategy for sensitivity testing was to (1) randomly split the data in half; (2) build a predictive model on half of the data, using LOOCV to select the most robust model; (3) test this model on the remaining held out 50% of data to obtain a true test set validation of the model. We then repeated steps 1–3 1000 times to gain insight into the sensitivity of this model-building procedure and distributions of the model statistics (Fig. 2b and c).

We could predict FCT responder status in a held-out test set with high discrimination. The median elastic net AUC was 0.969, and the median KTSP AUC was 0.972. Our sensitivity analysis revealed that even the lower-performing models still performed with high accuracy on the held-out testing set: 97% of the elastic net models had AUCs of better than 0.886, and 97% of the KTSP models had AUCs better than 0.946.

3.3. Incorporation of clinical variables into predictive models

After discovering that the transcriptome-only model provided excellent results for predictive models, we explored whether clinical variables could either augment or replace our transcriptome-based predictive models. We included 49 different clinical and laboratory parameters. A full table of clinical variables is given in the supplementary material Table S2. We followed identical training procedures to those described above for the transcriptome-only model. An elastic net trained on transcriptome plus clinical data produced identical results described above for the transcriptome-only model. An elastic net trained on transcriptome plus clinical data produced identical results to the elastic net trained on transcriptome data alone. The elastic net training algorithm did not select any clinical variables for inclusion in the final predictive model, instead choosing the same 73 genes. An elastic net trained on the clinical variables only and then used to predict FCT response performed very poorly (AUC = 0.16, Accuracy = 0.33, Sensitivity = 0.09, Specificity = 0.57) (Supplementary Fig. 3). In fact, the very poor AUC of this model mean that is performed well below random chance (AU = 0.5). Theoretically, simply flipping each of this model’s predictions would have made it perform moderately well (AUC 0.83). However, it would be statistically unfair to reverse model predictions after looking at the outcome of the trained model on the test set.

3.4. Gene ontology term enrichment

Our elastic net regression model utilized 73 gene expression values (Table 2). The final KTSP model used only 18 genes (nine pairs of top-scoring genes) (Table 3). 12 of the 18 genes from the KTSP algorithm were present in the elastic net regression selected genes (Fig. 4). We performed Gene Ontology (GO) biological process enrichment testing upon significant elastic net covariates to understand biological processes which may regulate the fibrous cap thickening response to statin therapy. GO terms are a controlled vocabulary of curated genes and gene products mapped to specific biological processes, molecular functions, and cellular components. We found 58 significant GO term enrichments at \( p < .05 \), including many involved with (1) immune cell function (GO terms: locomotion involved in locomotory behavior; T cell homeostasis; positive regulation of MHC class I biosynthesis; positive regulation of NK T cell activation; negative regulation of T cell mediated cytotoxicity; monocyte activation), and (2) cholesterol metabolism (cholesterol catabolic process; cholesterol import; C21-steroid hormone metabolic process; SREBP signaling pathway), and (3) protein modification (protein phosphorylation; protein adenyllation; protein K11-linked ubiquitination; N-terminal

Fig. 2. Predictive Model Receiver Operating Characteristic Curves. The receiver operating characteristic (ROC) curves for the elastic net and KTSP pairs predictive models are shown in (a). ROC scores were computed for KTSP by dividing the number of potential votes (i.e. gene pairs) in the classifier as the predicted probability. Sensitivity testing using elastic net (b) and KTSP (c) showed performance is highly robust to sampling error.
4. Discussion

In this post-hoc analysis of the Yellow-II study, we demonstrate that gene expression (when characterized via microarray with statistical models capable of dealing with high-dimensional gene expression data) can be used to reliably predict whether a given patient’s fibrous cap thickness will increase with statin treatment. Our study is one of the first attempts to use predictive modeling techniques applied to transcriptomic data to predict plaque response to lipid lowering therapy. Transcriptomic profiling has been used for many years in cardiovascular research to identify genes differentially expressed between cases and controls in a variety of diseases [33], or somewhat more recently to analyze putative biological interactions using methods such as weighted gene coexpression networks [34,35]. We attempted to, instead of simply identifying statistically differentially expressed genes across a condition, use gene expression data to predict outcomes. In fact, statistically significant variables (e.g., differentially expressed genes) are often poorly predictive variables [36], and often explain only a fraction of the variance in the outcome. To accurately predict an outcome, one must instead explain a significant fraction of the outcome variance, which is a more difficult task. In the present study, we explain much of the variance in outcome as demonstrated by our high prediction statistics (AUCs on the order of 0.95). For clinically applicable personalized medicine, a physician would be interested in a test which is predictive of an outcome for an individual patient, such as the test we present here.

Second, we demonstrate that the genes which are most predictive of FCT response to statin therapy are those broadly involved in immune cell function, cholesterol biosynthesis and metabolism, and cellular protein modification processes such as phosphorylation and adenylation. A number of researchers have linked fibrous cap production and degradation to inflammatory processes [37]. The same statin used in the YellowII trial—rosuvastatin—was demonstrated to reduce major cardiovascular events in the JUPITER study with elevated levels of the nonspecific inflammatory marker c-reactive protein (CRP) [38] but presumably normal cholesterol levels. Indeed, we have previously demonstrated in this cohort an inverse relationship between CRP levels and FCT [6,39]. Through MRI-PET/CT and mathematical fluid-structure-interaction modeling it was observed that in mice that cap inflammation can lead to increased cap strain, especially in the presence of thin-caps and hypertension [40]. The CANTOS trial of the Interleukin-1β (IL-1β) monoclonal antibody successfully demonstrated that selective anti-
inflammatory therapy can reduce the incidence of adverse cardiovascular events [41]. Other recent work has focused upon the role of IL-1β-modulated neutrophils in the atherosclerotic plaque [42], or the roles of neutrophil extracellular traps [43,44]. Another recent paper has demonstrated that neutrophil reprogramming can modulate chronic inflammatory processes play a role in the immunologic response to gram-negative bacteria [54]. The intracellular calcium-binding protein calbindin 2 (CALB2) plays a role in a number of cellular functions, such as the buffering of intracellular calcium, and has been implicated in vascular smooth muscle cell adhesion [55], cardiogenic differentiation [56], cellular response to ischemia [57], steroidogenesis [58]. Finally, we were interested in the negative correlation between expression of steroidogenic acute regulator-related lipid transfer domain protein (STARD5) and FCT response. This family of proteins, which modulate intracellular lipid and cholesterol transport, have been investigated for their roles in cholesterol and lipid metabolism [59,60] as well as their function in atherogenic lipid phenotypes [61] and fibroblast function [62]. This gene also has been identified as an emerging cardiovascular disease risk biomarker [63].

Table 2

| Covariate | Coefficient | Covariate | Coefficient | Coefficient | Coefficient |
|-----------|-------------|-----------|-------------|-------------|-------------|
| (Intercept) | -90.652 | ZSCAN12L1 | -0.047 | WDR23 | 0.322 |
| RNF113A | -1.155 | GSDM1 | -0.045 | CBX5 | 0.365 |
| MEI1 | -1.042 | RN6U-15 | -0.041 | XRCC5 | 0.397 |
| PBX2 | -0.711 | DTX3L | -0.026 | EIF2A | 0.401 |
| STARD5 | -0.587 | SPATA13 | 0.006 | NOL7 | 0.424 |
| TIMM117B | -0.568 | HSPH1 | 0.012 | FMOD | 0.477 |
| RNAM4ATAC | -0.557 | CLK1 | 0.016 | CGFRF27 | 0.496 |
| FCD | -0.455 | SMAP1 | 0.089 | MEF2A | 0.549 |
| HCST | -0.408 | TIPN1 | 0.104 | CSE1L | 0.577 |
| CALB2 | -0.401 | TMEM189-UBE2V1 | 0.112 | SNORD57 | 0.633 |
| TP53TG1 | -0.399 | SLC46A3 | 0.118 | HSP90 | 0.635 |
| DHX37 | -0.393 | KIAA0020 | 0.118 | AQP12B | 0.656 |
| USP48 | -0.393 | PPNP3CB | 0.126 | TMEM183A | 0.661 |
| FOXO4 | -0.299 | XCR1 | 0.13 | TAOK1 | 0.73 |
| HNRNPCUL2 | -0.29 | UBE2D3 | 0.148 | CI40RF88 | 0.74 |
| MRPS16 | -0.268 | CXYF3B | 0.157 | CREBBP | 0.801 |
| MT1G | -0.26 | C190RF12 | 0.181 | SFRS2 | 0.829 |
| RNU6-1 | -0.222 | NRC2C2 | 0.204 | MBLN3 | 0.837 |
| AURKAIP1 | -0.204 | SR140 | 0.223 | YARS2 | 0.878 |
| FNP2 | -0.133 | GTF2A1 | 0.236 | CN1D6 | 0.981 |
| UBE28 | -0.129 | ZRANB2 | 0.236 | RBAX | 1.012 |
| SLC35A3 | -0.108 | SEC11A | 0.253 | KCTD7 | 1.069 |
| LOC148413 | -0.074 | ZKSCAN1 | 0.273 | LOC729402 | 1.118 |
| AKR1D1 | -0.066 | TP53RK | 0.307 | IRX5 | 2.525 |
| ZNF264 | -0.053 | GNL2 | 0.313 | |

Table 3

| Gene Pair | Score |
|-----------|-------|
| TAOK1, MEI1 | 0.884 |
| ZRANB2, PMIPCA | 0.884 |
| GN1C, AP1M1 | 0.884 |
| PPIG, PBX2 | 0.855 |
| ATOH7, RBAK | 0.826 |
| LOC729402, CALB2 | 0.797 |
| CMO RF38, TNP2O | 0.797 |
| TIMM117B, OSBP | 0.796 |
| TP53TG1, SNORD57 | 0.796 |

Fig. 4. Genes Shared Between Elastic Net and KTSP predictive models. Venn diagram showing the overlap of genes included in the elastic net and KTSP algorithms. 12 of 18 KTSP genes were also selected by the elastic net algorithm.
4.1. Study limitations

Our study must be interpreted in the light of its limitations. The major limitation to this study is the limited sample size, essentially due to the invasive nature of repeated OCT imaging. We are exploring options for external validation, which would require a second clinical trial. However, we are comforted by the extensive cross validation efforts we undertook to test that we are not overfitting our model to the dataset. Furthermore, the gene expression results have been previously validated in the Yellow II study [6]. Second, there are currently few transcriptome-based tests utilized in cardiovascular medicine, and likely none used in routine clinical cardiology. In this measure, cardiovascular disease has lagged behind other fields such as oncology, which have rapidly embraced transcriptome characterization via microarray and RNA sequencing [64]. However, there is no fundamental reason that transcriptomic medicine will not ultimately prove to be useful in cardiovascular medicine, especially as future efforts toward precision cardiology increase [65].

5. Conclusions

Taken altogether, we believe our results represent a step toward enabling statin therapy recommendation based upon the patient's transcriptome in chronic cardiovascular disease. OCT is a precise and highly useful technique with a large body of existing literature, and it would require further validation studies before a transcriptomic approach could be used with confidence in the clinic. However, our findings here are promising: a non-invasive, blood-based test appears to detect with high sensitivity and specificity whether an individual patient in our sample will benefit from high intensity statin therapy. Since OCT is associated with systemic cardiovascular disease outcomes and risk of myocardial infarction, development of a non-invasive diagnostic test for patient response is potentially of high clinical interest. We envision that predicted non-responders to statin therapy may benefit from additional intervention, such as earlier introduction of PCSK9 inhibitors.

Funding sources

No specific funding sources exist for this project. Kipp Johnson is partially supported by the Medical Scientist Training Program grant 2T32GM007280-41 to the Icahn School of Medicine at Mount Sinai (National Institute of General Medical Sciences (NIGMS), National Institute of Health (NIH), United States). No funders had any role in study design, data collection, data analysis, interpretation, writing of the report.

Declaration of interests

Dr. Annapoorna Kini received grants from AstraZeneca (Gaithersburg, MD) to conduct the initial Yellow II clinical trial of high-intensity rosuvastatin therapy. Dr. Shameer Khader is an employee of AstraZeneca. No other authors have any financial conflicts of interest to disclose.

Author contributions

All authors contributed to study design and data analysis. Data collection was performed by Yuliya Vengrenyuk, Samin Sharma, Jagat Narula, and Annapoorna Kini. The first draft of the article was written by Kipp Johnson, Benjamin Glicksberg, Khader Shameer, Yuliya Vengrenyuk, Jagat Narula, Joel Dudley, and Annapoorna Kini. All authors contributed to subsequent revisions and drafts of the paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.embio.2019.05.007.

References

[1] Narula J, Nakano M, Virmani R, et al. Histopathologic characteristics of atherosclerotic coronary disease and implications of the findings for the invasive and noninvasive detection of vulnerable plaques. J Am Coll Cardiol 2013;61(10):1041–51.
[2] Burke AP, Farb A, Malcom GT, Liang YH, Virmani R. Coronary risk factors and plaque morphology in men with coronary disease who died suddenly. N Engl J Med 1997;336(18):1276–82.
[3] Osuka F, Joner M, Prati F, Virmani R, Narula J. Clinical classification of plaque morphology in coronary disease. Nat Rev Cardiol 2014;11(7):379–89.
[4] Feldman MD, Milner TE. Defining plaque classification with optical coherence tomography. JACC Cardiovasc Imaging 2017;11(11):1677–8. https://doi.org/10.1016/j.jcmg.2017.11.007.
[5] Kini AS, Vengrenyuk Y, Yoshimura T, et al. Fibrous cap thickness by optical coherence tomography in vivo. J Am Coll Cardiol 2017;69(6):644–57.
[6] Kini AS, Vengrenyuk Y, Shameer K, et al. Intracoronary imaging, cholesterol efflux, and transcriptomics after intensive statin treatment: the YELLOW II study. J Am Coll Cardiol 2017;69(6):628–40.
[7] Yamagishi M, Terashima M, Awan K, et al. Morphology of vulnerable coronary plaque: insights from follow-up of patients examined by intravascular ultrasound before an acute coronary syndrome. J Am Coll Cardiol 2000;35(1):106–11.
[8] Jang IK, Bouna BE, Kang DL, et al. Visualization of coronary atherosclerotic plaques in patients using optical coherence tomography: comparison with intravascular ultrasound. J Am Coll Cardiol 2002;39(4):604–9.
[9] Gardner CM, Tan H, Hull EL, et al. Detection of lipid core coronary plaques in autopsy specimens with a novel catheter-based near-infrared spectroscopy system. JACC Cardiovasc Imaging 2008;1(5):338–48.
[10] Kini AS, Baber U, Kovicic JC, et al. Changes in plaque lipid content after short-term intensive versus standard statin therapy: the YELLOW trial (reduction in yellow plaque by aggressive lipid-lowering therapy). J Am Coll Cardiol 2013;62(1):21–9.
[11] Sato A, Hoshi T, Kakefuda Y, et al. In vivo evaluation of fibrous cap thickness by optical coherence tomography for positive remodeling and low-attenuation plaques assessed by computed tomography angiography. Int J Cardiol 2015;182:419–25.
[12] Sun J, Zhao SQ, Burla N, et al. Carotid plaque lipid content and fibrous cap status predicting systemic CV outcomes: the MBI substudy in AIM-HIGH. JACC Cardiovasc Imaging 2017;10(3):241–9.
[13] Ekn SM, Jin H, Chernogubova E, et al. MicroRNA-21 enhances fibrous cap stability in advanced atherosclerotic lesions. Circ Res 2017;120(4):533–44.
[14] Wang Z, Che YS, Soeda T, et al. Three-dimensional morphological response of lipid-rich coronary plaques to statin therapy: a serial optical coherence tomography study. Coran Artery Dis 2016;27(5):350–6.
[15] Park SJ, Kang SJ, Ahn JM, et al. Effect of statin treatment on modifying plaque composition: a double-blind, randomized study. J Am Coll Cardiol 2016;67(15):1772–83.
[16] Oesterle A, Laufs U, Liao JK. Pleiotropic effects of statins on the cardiovascular system. Circ Res 2017;120(1):229–43.
[17] Komukai K, Kubo T, Kitabata H, et al. Effect of atorvastatin therapy on fibrous cap thickness in coronary atherosclerotic plaque as assessed by optical coherence tomography: the EASY-FTI study. J Am Coll Cardiol 2014;64(21):2207–17.
[18] Kataoka Y, Sr John J, Wolski K, et al. Atheroma progression in hyporesponders to statin therapy. Arterioscler Thromb Vasc Biol 2015;35(4):990–5.
[19] Wei KI, Zhang LK. Interactions between CYP3A5 and POR polymorphisms and lipid lowering response with atorvastatin. Clin Drug Invest 2015;35(9):583–91.
[20] Kataoka Y, Yamada M, Puri R, et al. Plaque microstructures in patients with coronary artery disease who achieved very low low-density lipoprotein cholesterol levels. Atherosclerosis 2015;242(2):490–5.
[21] Trompet S, Postmus I, Slagboom PE, et al. Non-response to (statin) therapy: the importance of distinguishing non-responders from non-adherers in pharmacogenetic studies. Eur J Clin Pharmacol 2016;72(4):431–7.
[22] Dudley JT, Tibshirani R, Deshpande T, Butte AJ. Disease signatures are robust across tissues and experiments. Mol Syst Biol 2009;5:307.
[23] R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2017.
[24] Du P, Kibbe WA, Lin SM. null: a universal naming scheme of oligonucleotides for illumina, affymetrix, and other microarrays. Bioil Direct 2007;2:16.
[25] Du P, Kibbe WA, Lin SM. lumic: a pipeline for processing Illumina microarray. Bioinformatics 2008;24(11):1547–8.
[26] Lin SM, Du P, Huber W, Kibbe WAock, Model-based variance-stabilizing transformation for Illumina microarray data. Nucleic Acids Res 2008;36(2):e11.
[27] Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 2012;28(6):828–33.
[28] Kim K, Bolotin E, Theusch E, Huang H, Medina MW, Krauss RM. Prediction of LDL cholesterol response to statin using transcriptomic and genetic variation. Genome Biol 2014;15(9):460.
[29] Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010;33(1):1–22.
[30] Ahari B, Fertig EJ, German RT, Marinchioni L, switchBox: an R package for k-top scoring pairs classifier development. Bioinformatics 2015;31(2):273–4.
Gomez D, d’Avignon C, Naiman DQ, Winslow RL. Classifying gene expression profiles from pairwise mRNA comparisons. Stat Appl Genet Mol Biol 2004;3:19.

Afzali B, Braga-Neto UM, Geman D. Rank discriminants for predicting phenotypes from RNA expression. Ann Appl Stat 2014;8(3):1469–91.

Mendelson MM, Marioni RE, Johanes R, et al. Association of body mass index with DNA methylation and gene expression in blood cells and relations to cardiometabolic disease: a Mendelian randomization approach. PLoS Med 2017;14(1):e1002215.

Miao L, Yin RX, Pan SL, Yang S, Yang DZ, Lin WX. Weighted gene co-expression network analysis identifies specific modules and hub genes related to hyperlipidemia. Cell Physiol Biochem 2018;48(3):1151–63.

Chen J, Yu L, Zhang S, Chen X. Network analysis-based approach for exploring the potential diagnostic biomarkers of acute myocardial infarction. Front Physiol 2016;7:615.

Lo A, Chernoff H, Zheng T, Lo SH. Why significant variables aren’t automatically good predictors. Proc Natl Acad Sci U S A 2015;112(45):13892–7.

Ridker PM, Narula J. Will reducing inflammation reduce vascular event rates? JACC Cardiovasc Imaging 2018;11(2 Pt 2):317–9.

Ridker PM, Danielson E, Fonseca FA, et al. Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. N Engl J Med 2008;359(21):2195–207.

Chamaria S, Johnson KW, Vengrenyuk Y, et al. Intracoronary imaging, cholesterol efflux, and transcriptomics after intensive statin treatment in diabetes. Sci Rep 2017;7(1):7001.

Tang D, Yang C, Huang S, et al. Cap inflammation leads to higher plaque cap strain and lower cap stress: an MRI-PET/CT-based FSI modeling approach. J Biomech 2017;50:121–9.

Ridker PM, Everett BM, Thuren T, et al. Antiinflammatory therapy with canakinumab for atherosclerotic cardiovascular disease. N Engl J Med 2017;377(12):1119–31.

Tall AR, Westerterp M. Inflammasomes, neutrophil extracellular traps, and cholesterol. J Lipid Res 2019;60(4):721–7. https://doi.org/10.1194/jlr.S091280.

Soehnlein O, Steffens S, Hidalgo A, Weber C. Neutrophils as protagonists and targets in chronic inflammation. Nat Rev Immunol 2017;17(4):248–61.

Doring Y, Soehnlein O, Weber C. Neutrophil extracellular traps in atherosclerosis and atherothrombosis. Circ Res 2017;120(4):736–43.

Geng S, Zhang Y, Lee C, Li N. Novel reprogramming of neutrophils modulates inflammation resolution during atherosclerosis. Sci Adv 2019;5(2):eaav2309.

Gomez D, Bayls RA, Durgin BG, et al. Interleukin-1beta has atheroprotective effects in advanced atherosclerotic lesions of mice. Nat Med 2018;24(9):1418–29.

Stone NJ, Robinson JC, Lichtenstein AH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2014;63(25 Pt B):2889–934.

Nasu K, Terashima M, Habara M, et al. Impact of cholesterol metabolism on coronary plaque vulnerability of target vessels: a combined analysis of virtual histology intravascular ultrasound and optical coherence tomography. JACC Cardiovasc Interv 2013;6(7):746–55.

Ozaki Y, Tanaka A, Komukai K, et al. High-density lipoprotein cholesterol level is associated with fibrous cap thickness in acute coronary syndrome. Circ J 2013;77(12):2982–9.