



OPEN

Improving Prediction of Prostate Cancer Recurrence using Chemical Imaging

SUBJECT AREAS:

IR SPECTROSCOPY
PROGNOSTIC MARKERSJin Tae Kwak^{1,2,3}, André Kajdacsy-Balla⁴, Virgilia Macias⁴, Michael Walsh^{3,4}, Saurabh Sinha² & Rohit Bhargava^{3,5}Received
19 September 2014Accepted
3 February 2015Published
4 March 2015Correspondence and
requests for materials
should be addressed to
S.S. (sinhas@illinois.
edu) or R.B. (rxb@
illinois.edu)

¹Center for Interventional Oncology, National Institutes of Health, Bethesda, MD 20892, USA, ²Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA, ³Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA, ⁴Department of Pathology, University of Illinois at Chicago, Chicago, IL 60612, USA, ⁵Department of Bioengineering, Mechanical Science and Engineering, Electrical and Computer Engineering, Chemical and Biomolecular Engineering and University of Illinois Cancer Center, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

Precise Outcome prediction is crucial to providing optimal cancer care across the spectrum of solid cancers. Clinically-useful tools to predict risk of adverse events (metastases, recurrence), however, remain deficient. Here, we report an approach to predict the risk of prostate cancer recurrence, at the time of initial diagnosis, using a combination of emerging chemical imaging, a diagnostic protocol that focuses simultaneously on the tumor and its microenvironment, and data analysis of frequent patterns in molecular expression. Fourier transform infrared (FT-IR) spectroscopic imaging was employed to record the structure and molecular content from tumors prostatectomy. We analyzed data from a patient cohort that is mid-grade dominant – which is the largest cohort of patients in the modern era and in whom prognostic methods are largely ineffective. Our approach outperforms the two widely used tools, Kattan nomogram and CAPRA-S score in a head-to-head comparison for predicting risk of recurrence. Importantly, the approach provides a histologic basis to the prediction that identifies chemical and morphologic features in the tumor microenvironment that is independent of conventional clinical information, opening the door to similar advances in other solid tumors.

Prostate cancer is the second leading cause of deaths following lung cancer in US men and constitutes one-third of non-skin cancer diagnoses every year¹. One of the most acute needs in PCa management today is higher precision in prediction of clinical outcomes for more effective decision-making². Various prediction tools have been developed for this purpose and they largely rely on patients' clinical, pathologic and demographic data³. These data include age, PSA level, Gleason grade, and pathologic stage. The available tools include risk grouping⁴, look-up tables⁵, machine learning⁶, nomograms^{7,8}, and risk scoring^{9,10}. The performance of these predictive tools is more consistent and superior to manual decisions; hence, they are implemented in modern patient care. Currently, the widely-validated Kattan nomogram^{7,8} and CAPRA-S score¹⁰ are considered as the best performing tools for prostate outcome prediction after radical prostatectomy (RP). The accuracy of prediction is often measured using a receiver operating curve (ROC) and the area under the curve (AUC) to segment patients with poor and good outcomes. The modest overall accuracy of available tools is further devalued by the realization that the tools largely fail for the most common cohort in the modern era (mid-grade, confined disease with moderate PSA level) and there is a tendency to overestimate likelihood of disease recurrence for lower risk patients^{10,11}. Incorporating additional clinical variables has not significantly improved predictions¹². Likewise, integrating gene expression^{13,14}, immunohistochemistry^{15,16}, magnetic resonance imaging^{17–19} or tissue morphology^{20,21} data has met with limited success² and none has achieved widespread clinical acceptance.

While most of the diagnostic and prognostic efforts have focused on epithelial cells, an important role for the tumor microenvironment or stroma in cancer progression has also been reported^{15,20–22}. No methods are available, however, to make anecdotal and laboratory observations about the microenvironment into a practical clinical test. The challenges of measuring multiple cell types, via multiple, cumbersome steps, and associating the data manually from multiple sections and with clinical management has limited the potential of this promising avenue. Here, our first aim was to establish an imaging approach that promises to change this deficiency. We use chemical imaging, which is strongly emerging as a platform for combining spatial and molecular analysis. Fourier transform infrared (FT-IR) spectroscopic imaging²³, in particular, combines the spatial specificity of optical microscopy with the molecular selectivity of vibrational spectroscopy. The IR absorption spectrum at each



position (or pixel) is a quantitative record of chemical composition²⁴ and FT-IR imaging has been applied to study various cell types and different types of cancers^{25–30}, including that in the prostate^{31–33}. In particular, chemical changes in both epithelial cells³⁴ and cancer-associated stroma^{35,36} were recently reported using FT-IR spectroscopic imaging³⁷. While promise in imaging different cell types has been shown, the use of IR imaging for prognosis has not been previously proposed. In particular, the question of prostate cancer recurrence has not been addressed, which is the second aim of this study. We hypothesized that the chemical properties of the tissue, and especially of the stroma, may add additional information that can improve prediction. To test this hypothesis in the most pressing area of contemporary PCa need, we employed a recurrence-enriched and mid-grade dominant cohort³⁸ where recurrent cases and non-recurrent controls were matched for age at prostatectomy, race, Gleason score, and pathologic (pTNM) stage. This cohort not only provides the most challenging and demanding subset for modern prostate pathology but also facilitates discovering novel predictors of cancer recurrence which are independent of the conventional clinical parameters (age, race, Gleason score, and pathologic stage). This is an exceptionally challenging cohort that requires us to also change analytical methods, which was the third aim of our study. Biological variation has been shown to be a primary source of variation in a tissue³⁹, which is exacerbated with heterogeneity in cancer¹⁶. Hence, conventional approaches in analyzing chemical imaging data^{27,28,40} are inappropriate for this application. Therefore, we propose to use a new approach called “frequent pattern mining”⁴¹ to discover subsets, subsequences, or substructures in a dataset⁴² that are not apparent using simple dimensionality reduction. To make the final diagnosis or predict risk score, we adopt a Ranking-Support Vector Machine⁴³ such that a recurrent case is ranked at higher risk than a non-recurrent control. Finally, the fourth part of our study compares the new approach to the current gold standards - Kattan nomogram and CAPRA-S score - to determine if the approach could yield improved PCa outcome predictions. It is notable that the design of the study addresses a pressing need, utilizes new ideas to examine cancer by focusing on both the tumor and microenvironment and our approach is entirely compatible with other tests - whether MRI-based non-invasive assays or digital pathology^{44,45} on the same tissue specimens.

Results

Stromal IR features distinguish recurrence cases from non-recurrence controls. Five tissue microarrays (TMAs), including 186 patients, were imaged to provide patient and clinical diversity for the most difficult contemporary cohort of patients (predominantly mid-grade and organ confined disease). The characteristics of the patient set are detailed in Table 1 and data acquisition and analysis protocols are detailed in the methods section. We first classified the tissue into different cell types following previous protocols³² and compared the average IR absorption spectra (>500 pixels) of epithelium and two types of stroma (distal and adjacent to tumor) to discern differences between recurrent cases and non-recurrent controls. There was no significant difference between recurrent cases and non-recurrent controls for different cell types (Figure 1). Transitioning from average values to examining each patient, two separate sets of analyses were conducted - on epithelial cells and on stromal cells. Based on both population difference and individual variation, IR spectra were converted into 98 spectral metrics that are individually indicative of molecular content in tissue (Supplementary Table S1). The spectral features from each cell type were separately handled in our data analysis pipeline (Figure 2) since cell types greatly differ in morphology, chemistry, and function. From data for each cell type, we built a machine learning classifier that utilizes not only an individual IR metric but also the combination of IR metrics, designated as “patterns” (Supplementary Figure S1). The present approach to machine learning is to simply treat patterns from all pixels of a sample as equal - as assumption that does not hold if there are subclasses of pixels, outliers, or contaminated pixels. We recognize that a pattern could also identify a subset of pixels that may share (un)known biochemical functions or represent sub-cellular components within a sample. The subset of pixels, we hypothesize, has predictive power that may be lost in examining average values from all cells, even of the same type. The use of patterns can make the analysis more sensitive by not focusing on small differences in absorption that may arise from natural variation or sampling and make the analysis more robust by focusing on biochemical characteristics or metric expression. The frequency of a pattern, i.e., the occurrence of a subset of pixels, is used as a feature in our classification algorithm. Numerous patterns can be generated, but we do not know *a priori* which patterns may represent the

Table 1 | Clinical characteristics of patient cohort in TMAs. The patient characteristics are shown for the calibration dataset of 60 pairs and 82 patients without neoadjuvant or adjuvant therapy and the validation dataset of 66 patients. n and SD denote a number and standard deviation, respectively

	Calibration (n = 120)		Calibration (n = 82)		Validation (n = 66)	
	Cases (n = 60)	Controls (n = 60)	Cases (n = 35)	Controls (n = 47)	Cases (n = 38)	Controls (n = 28)
Age at prostatectomy, mean (SD)	61.6 (7.1)	62.9 (6.5)	62.2 (7.1)	63.1 (6.7)	63.1 (7.1)	63.8 (5.3)
Race, n (%)						
White	51 (85)	51 (85)	30 (85.7)	40 (85.1)	33 (86.8)	25 (89.3)
African American	9 (15)	9 (15)	5 (14.3)	7 (14.9)	4 (10.5)	3 (10.7)
Other					1 (2.6)	
Gleason sum, n (%)						
≤6	12 (20)	12 (20)	7 (20)	10 (21.3)	17 (44.7)	6 (21.4)
7(3 + 4)	35 (58.3)	35 (58.3)	21 (60)	28 (59.6)	16 (42.1)	15 (53.6)
7(4 + 3)	7 (11.7)	7 (11.7)	6 (17.1)	6 (12.8)	4 (10.5)	3 (10.7)
≥8	6 (10)	6 (10)	1 (2.9)	3 (6.4)	1 (2.6)	4 (14.3)
Pathological stage, n (%)						
T2a	7 (11.7)	7 (11.7)	4 (11.4)	7 (14.9)	2 (5.3)	2 (7.1)
T2b	33 (55)	33 (55)	18 (51.4)	30 (63.8)	26 (68.4)	15 (53.6)
T3a	19 (31.7)	19 (31.7)	13 (37.1)	10 (21.3)	10 (26.3)	11 (39.3)
T3b	1 (1.7)	1 (1.7)				
Serum PSA ng/mL, mean (SD)	10.6 (8.7)	8.6 (6.0)	10.4 (10.2)	8.7 (6.1)	9.5 (7.2)	8.5 (4.8)
Month to recurrence, mean (SD)	38.6 (29.0)		44.5 (33.1)		36.6 (28.7)	

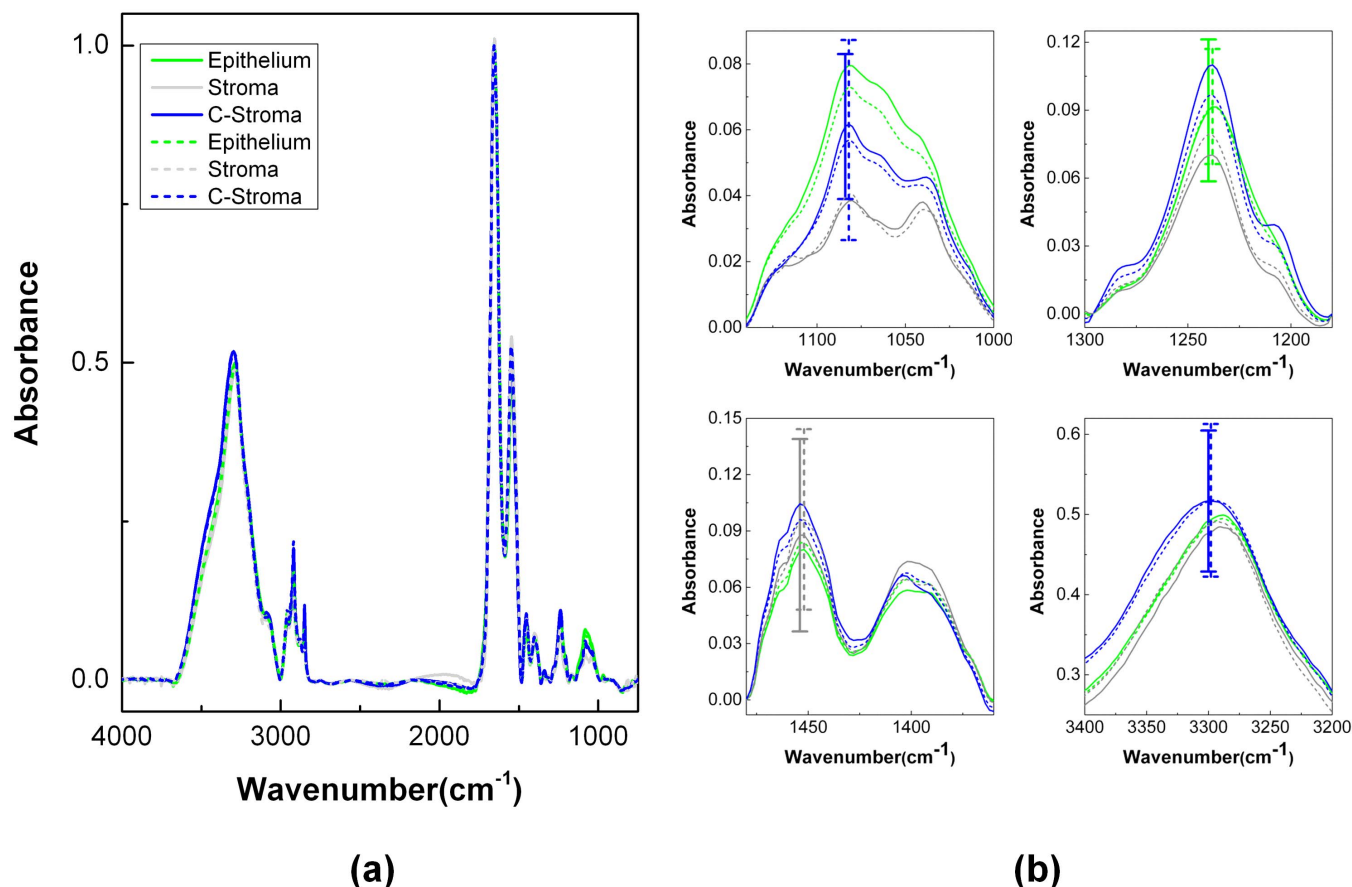


Figure 1 | IR Spectra. (a) IR spectra (4,000–720 cm^{-1}) of epithelium (green), distal stroma (grey; Stroma) and adjacent stroma (blue; C-Stroma) to tumor. (b) A close-up view of four different spectral areas (1,140–1,000 cm^{-1} , 1,300–1,180 cm^{-1} , 1,480–1,360 cm^{-1} , and 3,400–3,200 cm^{-1}) and error bars at the spectral peaks. Solid and dotted line represent recurrence case and non-recurrence control, respectively.

important biochemical functions or sub-cellular components. We adopted a pattern selection step⁴⁶ that automatically identifies the useful patterns from training data, before using such patterns in constructing the final classifier. The classification algorithm used in this study is a Ranking-Support Vector Machine (SVM)⁴³. Details of the classification scheme are provided in the Methods section.

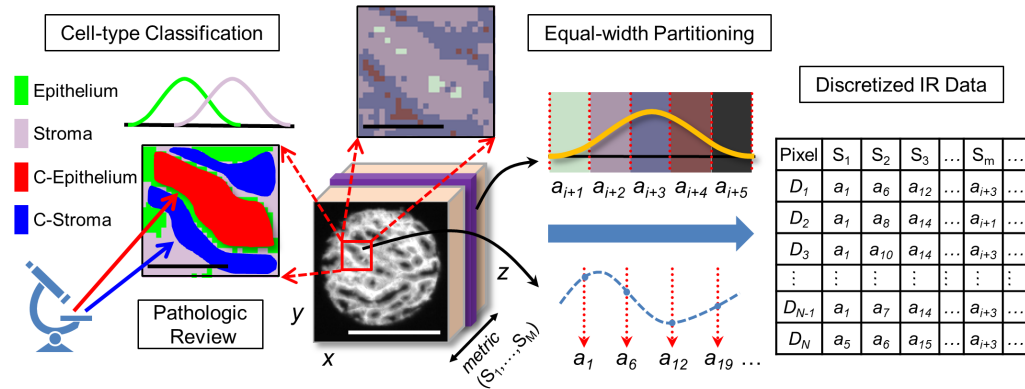
To ensure robustness, we tested the classifier in a variety of ways: by cross-validation on a calibration data set as well as by training on the calibration data set and testing on a separate validation dataset. The former uses a nested case-control study design, where a recurrence case is matched with a non-recurrence control, and the latter adopts a non-nested design. For both validations, the testing dataset and outcome were blinded to the classification model. We also note that multiple TMA slides are used, which includes variance from sample preparation and handling, with a blinded selection of patients into a calibration and a validation set. We detail these multiple modes of evaluation next. We first tested the classifier by K -fold cross-validation ($K = 10$) on the calibration dataset (120 patients, see Methods), using a nested design. In this test, the classifier was always presented with a case-control pair of patients that were matched by clinical descriptors (age at prostatectomy, race, Gleason score, and pathologic stage), and was asked to discriminate the case from the control. As is standard in cross-validation, the dataset was divided into K equal-sized partitions and one partition was left out as the “test partition”. The classifier was trained on the union of the remaining $K-1$ partitions and made to predict outcomes in the test partition. After repeating this K times with systematically different choices of the left-out partition, correct and incorrect predictions were summarized. The classifier based on epithelial data was not capable of distinguishing recurrent cases from non-recurrent controls ($\sim 50\%$

accuracy). On the other hand, the classifier based on stromal data discriminated cases from controls with $\sim 70\%$ accuracy ($P < 0.001$, Binomial test). Note that existing methods that rely on clinical variables to predict outcome are expected to perform close to a random classifier (50% accuracy) in this nested design sample set, since each case-control pair presented to the classifier was identical/matched in terms of many clinical variables.

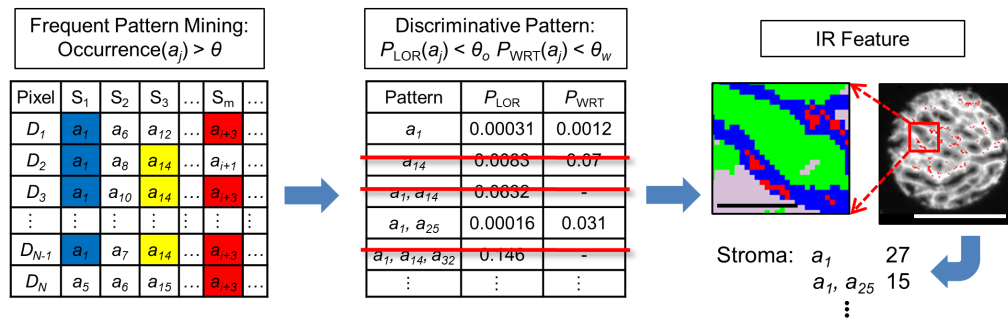
We next performed evaluations on the same data set as above but using a non-nested design, where the goal is to predict outcome of unknown individual patients (query). This was done with a modification of the classifier discussed above. For any test query, the algorithm finds the most similar recurrent and non-recurrent patients from the training dataset using clinical variables alone, and compares them to the query as detailed in Methods. The evaluations are done with K -fold cross validation as above, and in each “fold” the $K-1$ partitions other than the left out partition serve as training data. The predicted probability of recurrence (PPR) of the test patient based on these comparisons is computed using a logistic regression model. We call this PPR value the “IR Score” of the patient, as described in Methods. The sensitivity and specificity of the classifier are summarized using a receiver operating characteristic (ROC) curves and by the area under the ROC curve (AUC). We also calculate a 95% confidence interval (CI) by Bootstrap re-sampling for comparison. The classifier using data only from epithelial pixels could not appreciably distinguish the query as recurrent or non-recurrent (0.60 AUC, 95% CI: 0.50–0.69), but the classifier based on stromal data was effective with an AUC of 0.71 (95% CI: 0.61–0.81) in the calibration dataset (120 patients) in K -fold cross-validation ($K = 10$). After evaluating predictive performance on the calibration data set of 120 patients, we subjected the classifier to a



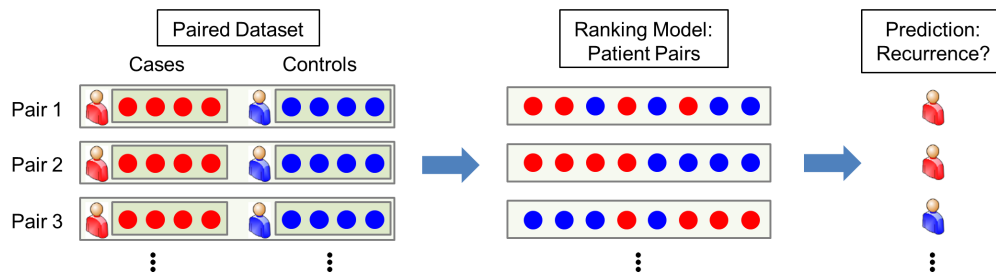
(a) Pre-processing: Digitization



(b) IR Feature Extraction



(c) Prediction Model for Patient Pairs



(d) Prediction Model for Individual Patients

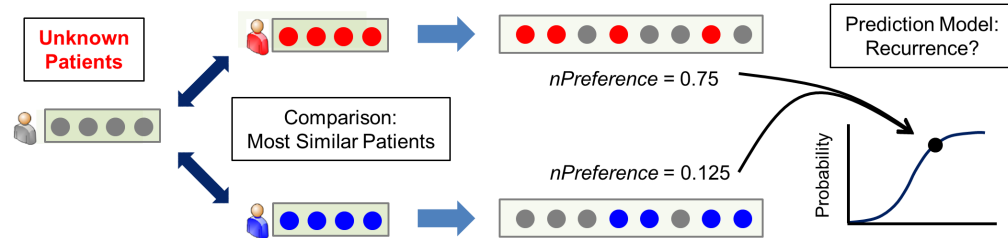


Figure 2 | Overview of Outcome Prediction. Given an IR image, (a) cell-types (Epithelium and Stroma) are identified by IR imaging and, by pathologic review, cancer epithelial cells (red; C-Epithelium) and adjacent stromal cells (blue; C-Stroma) are manually selected and discretized. Upon discretization, pixels in tissue contain a set of discrete values representative of the bins. (b) Applying frequent pattern mining and statistical tests, IR features are extracted. Using the features, prediction models for (c) patient pairs and (d) an individual patient are formed and predictions are made. C-Epithelium and C-Stroma denote the selected cancerous epithelium and stroma next to it, respectively. $nPreference$ is a normalized preference score for the unknown patient against the most similar patient. Black and white bars denote 100 μ m and 500 μ m, respectively.

separate, completely independent validation dataset of 66 patients. The IR Score of each patient in the validation dataset was recorded and an AUC of 0.57 (95% CI: 0.43–0.72) was obtained when using epithelium pixels. Using stromal pixels, an AUC of 0.71 (95% CI: 0.57–0.83) was observed, indicating that the test is robust.

IR Score outperforms popular existing tools in predicting outcomes. We compared the performance of the IR Score to the two

of the most commonly used clinical tools (Kattan nomogram and CAPRA-S score) for predicting outcomes after radical prostatectomy. Here, 82 patients, who have no record of neoadjuvant or adjuvant therapy, from the calibration dataset were employed since Kattan and CAPRA-S score are only applicable to such patients. As shown in Figure 3a, the IR Score (AUC = 0.74, 95% CI: 0.61–0.85) outperforms both Kattan (AUC = 0.60, 95% CI: 0.49–0.72) and CAPRA-S (AUC = 0.63, 95% CI: 0.50–0.75) scores when

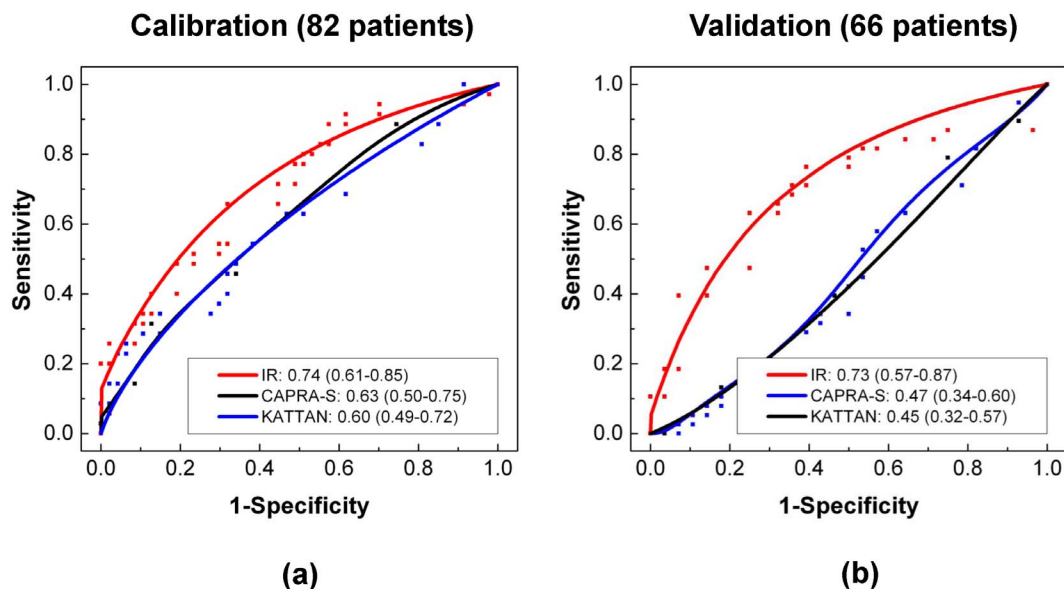


Figure 3 | Performance of Outcome Prediction. Outcome prediction of three models – CAPRA-S (black), Kattan nomogram (blue), and IR (red) – on (a) the calibration dataset and (b) the validation dataset. AUCs and 95% confidence intervals in parentheses are shown on the plots. Lines represent the smoothed ROC curves and rectangular dots denote the empirical sensitivity and 1-specificity points.

evaluated on the calibration dataset (82 patients). In a similar evaluation on the independent validation dataset (66 patients), Kattan (AUC = 0.45, 95% CI: 0.32–0.57) and CAPRA-S (AUC = 0.47, 95% CI: 0.34–0.60) scores performed far worse than the IR Score (AUC = 0.73, 95% CI: 0.57–0.87) as seen in Figure 3b. We note that these comparisons were from the individual query method only, since the nested case-control design is not amenable to better-than-random classification by Kattan or CAPRA-S scores, by definition.

We next compared the distribution of the patients according to the IR Score used by our algorithm for predicting recurrence. We compared the IR Score distributions of cases and controls to analogous distributions from both the Kattan nomogram and CAPRA-S score (Figure 4). PPR values represented by the IR Score had distinct distributions for recurrence cases versus non-recurrence controls ($P \approx 0.02$, Kolmogorov-Smirnov test). In contrast, the distributions of either the Kattan or the CAPRA-S scores were indistinguishable between the two classes ($P > 0.1$, Kolmogorov-Smirnov test). A majority of the patients, regardless of recurrence status, had very low PPR values using the Kattan nomogram. Most samples were assigned to either low-risk (score 0–2) or intermediate-risk (score 3–5) category by CAPRA-S score. This shows that the discriminative capability of the two widely-used tools in this challenging data set is very limited, and demonstrates the potential of the new approach proposed here to add to clinical decision-making.

Stromal IR features are independent predictors of cancer recurrence. The relationship between IR Score and recurrence was investigated in the context of conventional clinical variables (age at prostatectomy, Gleason grade, pathologic stage, and PSA level). In particular, we asked if the IR Score and clinical variables together can better predict recurrence than clinical variables alone can. If this is the case, then there must be additional information in the IR Score that is not present in clinical variables. By fitting a logistic regression model of recurrence using IR Score and the clinical variables as covariates, we estimated the strength of the association between IR Score and cancer recurrence, regardless of the clinical variables, on the calibration dataset (see Supplementary Information for details). IR Score was used as either a continuous or a categorical variable. As a continuous variable, the strength of the association with cancer recurrence was examined for a one-unit increase in IR Score as

fixing the clinical variables. As a categorical variable, patients were assigned to quartiles (1–4) by IR Score and, fixing the clinical variables, the relationship with cancer recurrence was compared between the patients in the lowest quartile and the patients in other quartiles (Table 2). For both types, a significant association was observed ($P < 0.001$ as a continuous variable and ~ 21 -fold differences between the lowest quartile and the highest quartile as a categorical variable). These indicate that higher IR Scores are significantly correlated with higher likelihood of recurrence, in ways that are not captured by clinical variables, i.e., independent from these variables. The association between Kattan and CAPRA-S scores with cancer recurrence was not significant ($P > 0.06$ as a continuous variable and ~ 2 - to 6-fold differences between the extreme quartiles as a categorical variable; Wald chi-square test) when tested.

IR Score adds independent predictive power to popular existing tools. As shown above, the relationship between IR Score and cancer recurrence is independent of clinical variables that the existing tools rely upon. To further test this hypothesis, we combined IR Score with Kattan and CAPRA-S scores as follows:

$$C_{score} = IR_{score} + xE_{score}.$$

C_{score} and IR_{score} denote combined score and IR Score respectively, while E_{score} is either Kattan score or CAPRA-S score and x is a weight for the scores ($0 < x \leq 1$). As changing a weight x from 0 to 1, the performance of the combined score C_{score} was measured by computing AUC. Increasing x , the performance of the combined score only marginally improved upon the performance of IR Score alone (even degraded capability in the combination of IR Score and Kattan score on the validation dataset) and the improvement was not statistically significant (Supplementary Figure S2). Combined with Kattan and CAPRA-S scores, the AUC of IR Score reached 0.75 (95% CI: 0.64–0.85) and 0.75 (95% CI: 0.63–0.85) respectively on the calibration dataset, whereas IR Score alone achieved AUC of 0.74 (95% CI: 0.61–0.85). Moreover, on the validation dataset the combined score led to degradation in classification performance. An AUC of 0.65 (95% CI: 0.50–0.79) and 0.74 (95% CI: 0.58–0.87) was attained by combining IR with Kattan and CAPRA-S scores respectively, while an AUC of 0.73 (95% CI: 0.57–0.87) was achieved by IR Score alone. Hence, the IR Score is truly independent of the existing tools.

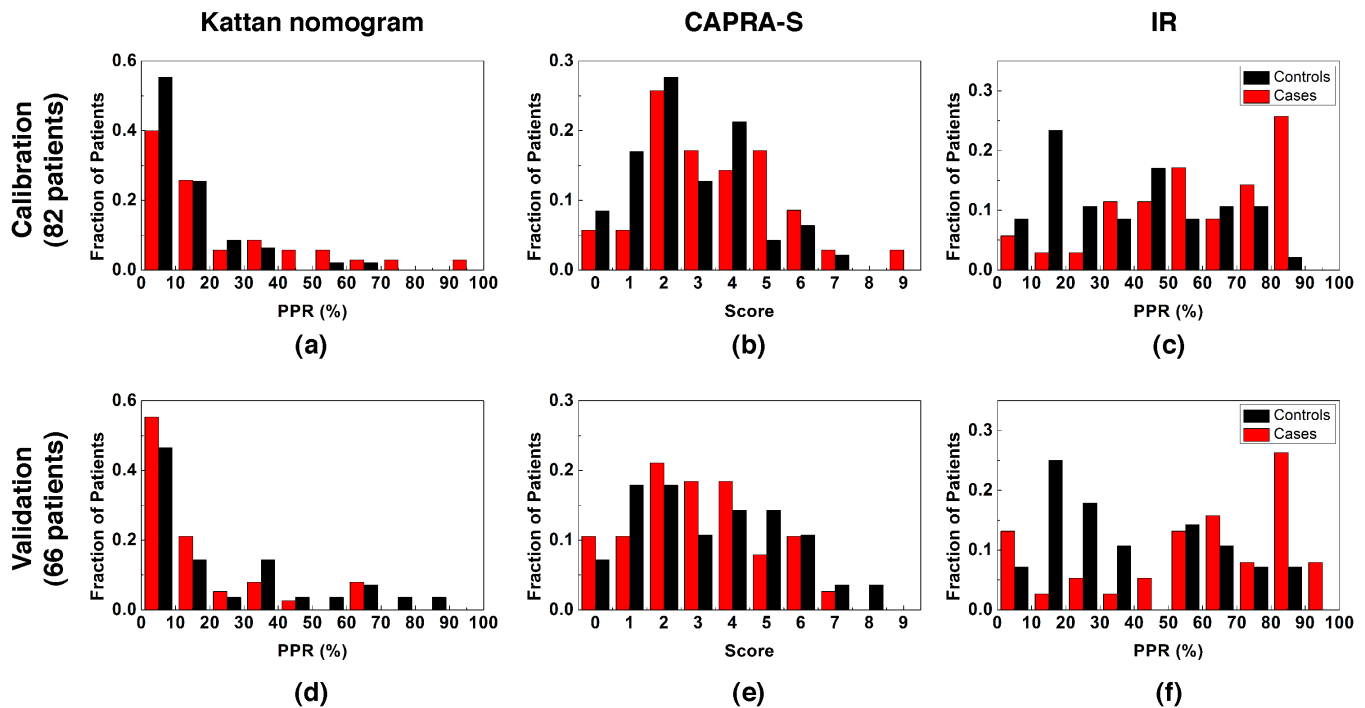


Figure 4 | Score distribution. Distribution of the recurrence cases (red) and non-recurrence controls (black) by (a)(d) Kattan (b)(e) CAPRA-S (c)(f) IR Scores. Distributions on (a)(b)(c) the calibration dataset and (d)(e)(f) the validation dataset are plotted. PPR represents the predicted probability of recurrence.

IR features correspond to specific functions in stroma. We examined the data to understand the underlying spectral and spatial basis of our successful classification. Examining the metrics, two prominent spectral regions were recognized. The peaks in $3000\text{--}3600\text{ cm}^{-1}$ range (Figure 1b) were present in all of the features; these are related to N-H stretching vibrational modes (in proteins) or O-H stretching vibration^{47,48}. About half of the features were related to the peaks in $990\text{--}1132\text{ cm}^{-1}$, which include C-O stretching vibrational mode at 1042 cm^{-1} due to oligosaccharides⁴⁹. The complete list of IR metrics useful in outcome prediction is available in Supplementary Table S1, indicated by a “Yes” in whether the feature was found to be useful. While the chemical content is elucidated, the spatial origin of our differences is also enlightening. Spectra can be used to find regions of the image most closely associated with prediction. Pixels containing the discriminating stromal features are identified and compared to their corresponding H&E images (Figure 5). The features appear in the regions where reactive stroma is present. More specifically, loose or myxoid stroma and fibroblasts are often observed around the pixels that were most useful in outcome prediction. Stromal reaction (or desmoplasia), which occurs in most solid human cancers⁵⁰, contains increased number of fibroblast and myofibroblasts and modifies extracellular matrix composition. We also found that the discriminating features are frequently present next to periacinar retraction (clefing), a separation of the gland from the adjacent stroma that has been associated with cancer recurrence⁵¹.

Discussion

We have presented a novel approach for precise predictions in prostate cancer using a combination of emerging IR imaging and a novel computational approach. Stromal features, not epithelial features, were found to be more significant in distinguishing biochemical recurrence cases from non-recurrence controls for both nested and non-nested case-control designs. While these results indicate that the stroma contains molecular signatures related to cancer progression, as has been shown in the past; here, however, we have not used any dyes, molecular stains or human interpretation. Automated methods that can comprehensively analyze the entire tumor microenvironment lead to a robust and accurate predictive model. This presents a new opportunity for combined molecular and imaging-based outcome predictions, as opposed to simply using clinical data or molecular biomarker techniques⁵². In a head-to-head comparison, IR-based prediction outperforms both Kattan nomogram and CAPRA-S score on multiple datasets. IR prediction is especially exciting as the patient cohort here is of clinically difficult cases, where nomograms are of limited utility. There were slight differences in the performance of Kattan and CAPRA-S scores for the calibration dataset and the validation dataset. This may be attributable to the differences in clinical characteristics of the patients (Table 1). For instance, in the validation dataset, differences in age and PSA level between recurrence cases and non-recurrence controls are smaller compared to those in the calibration dataset. More recurrence cases than non-recurrence controls (17 recurrence cases and 6 non-recurrence con-

Table 2 | Odds ratios for cancer recurrence by quartiles of IR score

Quartile	1 (low)	2	3	4 (high)	P trend
IR OR (95% CI)	1	6.5 (1.3–46.8)	10.8 (2.0–86.8)	21.3 (4.2–168.4)	0.020
CAPRA-S OR (95% CI)	1	3.2 (0.8–15.0)	1.3 (0.3–5.3)	6.1 (1.1–40.3)	0.270
KATTAN OR (95% CI)	1	0.8 (0.2–3.1)	1.5 (0.4–6.5)	1.9 (0.4–10.2)	0.111

OR and CI denote odds ratio and confidence interval, respectively.

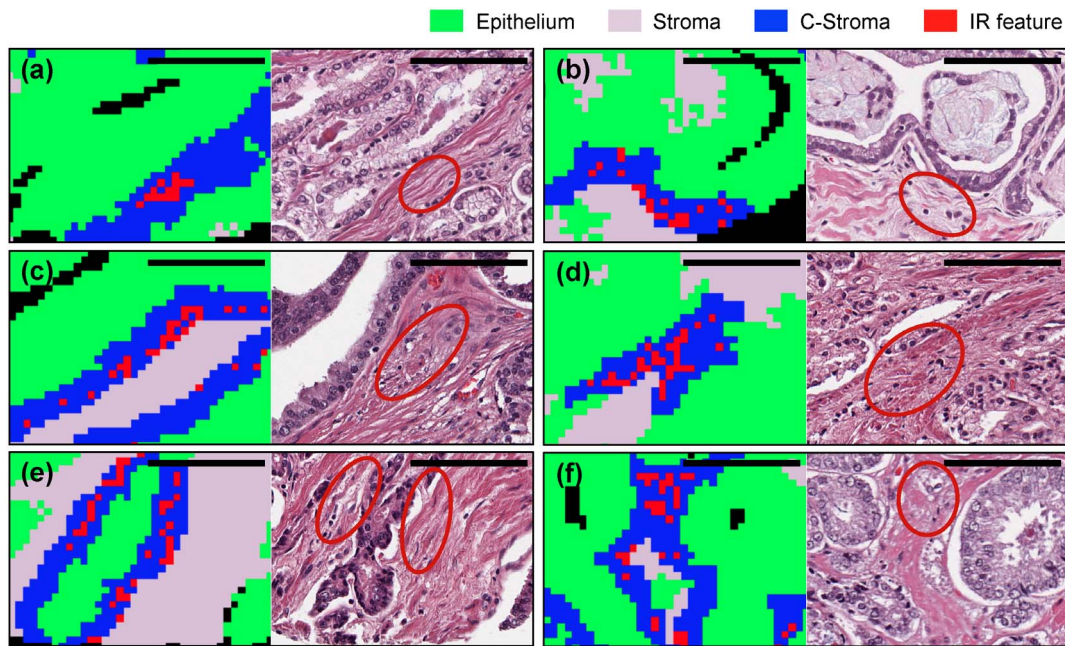


Figure 5 | Comparison of IR and H&E images. IR stromal features (red rectangles) and their corresponding areas (red circles) are marked on IR and H&E images, respectively. (a)(c)(e) Recurrence images and (b)(d)(f) non-recurrence images. C-Stroma represents stroma cells adjacent to cancer cells. Black scale bar denotes 100 μ m. IR features are related to (a)(b)(e)(f) loose or myxoid stroma, (b)(c)(d)(e)(f) fibroblasts, and (a)(d)(e)(f) retraction clefting.

trols) with lower Gleason sum (≤ 6) exist in the validation dataset. All of the 17 recurrence cases were designated as either lower-risk (70%) or intermediate-risk (30%) category by CAPRA-S score. The recurrence cases had relatively lower PPR values (88% had $< 10\%$ PPR) by Kattan nomogram. However, the performance of IR Score was consistent and the separation was evident by IR Score; at the level of PPR $\leq 30\%$, ~ 3 -fold more non-recurrence controls were observed than recurrence cases; 2- to 4-fold more recurrence cases were found at $\geq 70\%$ PPR. Overall, the patients, grouped into an intermediate (or lower) risk group by the conventional tools, can be further stratified by IR Score whereas both Kattan and CAPRA-S scores were not contributory. Since Kattan and CAPRA-S scores did not perform well, it is not surprising that the combined score with IR Score was not able to significantly improve upon IR Score alone. The synergy between IR Score and the two scores does not lie in a simple combination. We believe that the way to utilize IR Score is to add it to the existing tools as an additional factor and to improve the utility of the tools. This is especially true for the clinically difficult intermediate risk where Kattan and CAPRA-S methods fail, as demonstrated by the cohort here. The method presented is likely to prove an exciting addition to the power of Kattan and CAPRA-S methods and could substantially increase their predictive power.

In addition to Kattan and CAPRA-S scores, several imaging- and molecular-based techniques have been developed; however, the relationship between IR Score and these techniques is yet unclear. This is largely due to the non-availability yet of the data and/or images on the same set. The approach presented here is entirely compatible with these approaches, including pre-operative MRI-based approaches⁵³, downstream genomic tests⁵⁴, emerging mass spectrometric^{55,56} or even other intra-operative⁵⁷ or emerging vibrational spectroscopic⁵⁸ approaches. None of these techniques offers the unique microenvironment profiling and imaging capability of IR microscopy, using protein, nucleic acid and carbohydrate signals inherent in the spectroscopic capability used here. The presented data are both unique and complementary; hence, it will be desirable to examine if there is a synergy between IR Score and predictive features from other techniques. Finally, both the chemical and bio-

logical underpinnings of the discriminative stromal spectroscopic features uncovered by the approach here provide further avenues for examination. The features may be related to the molecules pertaining to (myo)fibroblasts, extracellular matrix composition or retraction clefting; for instance, chondroitin sulfate, known to be associated with prostate cancer progression⁵⁹. However, the specific molecules or chemical/biological causes relevant to the IR features cannot be identified without further biochemical assays such as immunohistochemistry, proteomics, or glycomics. In an emerging trend, obtaining higher spatial resolution IR imaging⁶⁰ could improve the recognition of the features yielding a finer classification model and, in turn, achieve better predictions of recurrence. In summary, the presented approach not only shows tremendous promise to directly address a major contemporary need in PCa management but can also help narrow the focus for further research studies. The same approach may be used to examine other questions of importance in contemporary cancer care, for example, the determination of the cohort in which adjuvant therapy may prove useful, for examining the metastatic potential of organ-confined disease and for the analysis of solid tumors in other organs.

Methods

Samples and data preparation. This study and protocols were conducted as approved by the University of Illinois Institutional Review Board (IRB) via protocol #06684. Informed consent was obtained from all subjects included in this study. We used a set of five tissue microarray slides (TMAs) constructed by the NIH-sponsored Cooperative Prostate Cancer Tissue Resource (CPCTR)³⁸. The TMA contain up to four cores each from 404 patients. Half the patients experienced biochemical recurrence after radical prostatectomy (cases) and the other half did not for at least 5 years after radical prostatectomy (controls). Cases and controls were matched for age at prostatectomy, race, Gleason score, and pathologic (pTNM) stage. $\sim 90\%$ of the patients are in the intermediate risk group using D'Amico risk grouping⁴, and 268 of the 404 patients have Gleason sum 7 – the most problematic and prevalent score. These TMAs, hence, are recurrence-enriched as well as mid-grade dominant, which are unique and challenging compared to other studies^{7–10} – but represent the most crucial subset where prediction capabilities are sorely desired. Further description of this outcomes-based TMA is available in Ref. 61.

For accurate diagnosis, tissues with insufficient malignant epithelial cells ($< 10\%$ of core) were omitted by combined pathologic review and IR-based cell-type classification³². Stromal cells in the vicinity of epithelial cells (within $\sim 50 \mu$ m) were also selected as we have previously noted using engineered tissue models³⁵ that stromal



reaction can be observed within a safe margin of 100 μm from the tumor. Finally, the study only included subjects who had at least two cores with both sufficient epithelial cancer cells and adjacent stromal cells. Hence, calibration dataset consists of 60 patient pairs, with a total of 380,685 and 333,167 epithelial pixels for the cases and controls, respectively (the “IR epithelium dataset”). Similarly, 101,576 and 95,925 stromal recurrence and non-recurrence pixels, respectively (the “IR stroma dataset”), were collected. For comparison with Kattan nomograms and CAPRA-S score, 82 patients (35 recurrence cases and 47 non-recurrence controls) who have no record of neoadjuvant or adjuvant therapy, were selected from the calibration dataset. Similarly, a completely independent “validation dataset” of 66 patients (38 recurrence cases and 28 non-recurrence controls) with 193,620 epithelial recurrence, 133,126 epithelial non-recurrence, 50,745 stromal recurrence, and 41,130 stromal non-recurrence pixels, respectively, was obtained. Clinical characteristics of the patient cohort are available in Table 1.

IR imaging data were acquired at a spatial pixel size of $6.25\mu\text{m} \times 6.25\mu\text{m}$, a spectral resolution of 4 cm^{-1} at an undersampling ratio of 2 using a Perkin-Elmer Spotlight imaging system. Each pixel’s interferogram was processed using NB-medium apodization and spectra truncated to $4,000\text{--}720\text{ cm}^{-1}$. Tissue samples were microtomed to a thickness of $\sim 5\mu\text{m}$. One section was placed on IR transparent BaF_2 slides, while consecutive sections were placed on a standard glass slide and stained with hematoxylin and eosin (H&E) for pathologic review. H&E stained images were acquired on an optical microscope at $40\times$ magnification. Since IR spectra are high-dimensional, we convert data into 98 spectral metrics (Supplementary Table S2). Metrics are defined by spectroscopic review and can be absorbance at a peak position, ratio between two peaks, and area and center of gravity of a spectral region³². Conversion to metrics not only reduces the dimensionality of the original data to a manageable size but also preserves the intrinsic and essential characteristics of the data while allowing interpretation of the results.

Chemical feature extraction. Due to biological variation and heterogeneity in tissue and cancer, signatures of cancer recurrence may not be apparent. Here, we hypothesize that the signatures of recurrence reside in a part of cells or tissues which share certain biochemical properties and can be identified by IR spectra. Using the combinations of discretized IR data, we attempt to find such signatures efficiently and effectively. Ranking approach is adopted to achieve an accurate and robust prediction of recurrence.

Discretization. Each IR metric is independently discretized by dividing the entire range of that metric into a number of equal-sized partitions or “bins” (Figure 2a), and a representative value is designated for each bin. The value of the IR metric for a pixel is then transformed to the appropriate discrete value representative of the bin containing the original value. Thus, each pixel d_i in a dataset $D = \{d_{i1}, d_{i2}, \dots, d_{in}\}$ (n pixels), originally a vector of IR metric values (k IR metrics), is now represented as a vector $\vec{d}_i = (d_{i1}, d_{i2}, \dots, d_{ik})$ where each dimension d_{ij} is the discretized value of a distinct IR metric j corresponding to pixel d_i . We used 20 bins to discretize each IR metric. Transformation of a continuous IR metric into a discrete format further bounds the joint probability distributions of IR metrics and helps mitigate the effect of variance by discarding the unnecessary precision.

Frequent pattern mining. A “pattern” is a specification of values for a certain subset of dimensions, e.g., a pattern $\phi = ((2, \phi_2), (5, \phi_5), (6, \phi_6))$ specifies that the second dimension has value ϕ_2 , the fifth dimension has value ϕ_5 , etc, and a vector \vec{d}_i is said to match this pattern if $d_{i2} = \phi_2$, $d_{i5} = \phi_5$, and $d_{i6} = \phi_6$. A frequent pattern is a pattern that is matched by many pixels (Figure 2b), i.e., the number of pixels matching the pattern is no less than $\theta|D|$, where θ is a user-specified minimum threshold and $|D|$ is the total number of pixels in the data set D (We set $\theta = 0.02$). Frequent patterns are discovered for each group of recurrence and non-recurrence subjects separately by applying FP-growth⁴¹. Using frequent patterns, we are now able to deal with any subset of pixels in an efficient and effective way.

Discriminative patterns. Many of these frequent patterns may not be useful for classification⁶². We obtain discriminative patterns by comparing the frequency of a pattern between two classes (recurring and non-recurring) in two stages: (Figure 2b; see Supplementary Information for details). In the first stage, we quantify the frequency of a pattern for all samples in each class and test if the frequency of a pattern is associated with cancer recurrence. In the second stage, we test if the pattern frequency of the individual subjects in one class (e.g., recurrence) is larger than that in the other class (e.g., non-recurrence). After the two stages, the frequency of the top m discriminative patterns ($m = 100$ is set here) in a tissue sample forms an IR feature for the classification algorithm.

Feature selection. We examine the IR features with respect to the training dataset and select a subset of them to classify the test dataset by adopting a two-stage feature selection scheme. In the first stage, we select a subset of the features that yield a higher relevance with class labels and a lower redundancy among them. In the second stage, the subset is refined by adding new feature(s) and/or removing the already selected feature(s). The detailed description of the feature selection procedure is available in Ref. 45.

Prediction models. We apply Ranking-Support Vector Machine (SVM)⁴³ (Figure 2c) for classification, which assigns relative ranks to any given set of samples. Formally

speaking, given two samples, x_i and x_j , the Ranking-SVM computes a function f on each sample and assigns ranks y_i and $y_j \in \mathcal{R}$ to the two instances by the rule:

$$y_i > y_j \Leftrightarrow f(x_i) > f(x_j) \quad (1)$$

Here, a higher rank indicates greater confidence in the sample being recurrent. Ideally, given a pair of recurrence patient p_1 and non-recurrence patient p_2 , the trained SVM should assign a higher rank to any sample of patient p_1 compared to that from patient p_2 . The Ranking-SVM algorithm is optimized on the training dataset to minimize the fraction of samples that are mis-ranked.

While the description above formulates a protocol and tests it for a case-control set, translation to a clinical unknown can be incorporated into the same framework by adopting a lookup methodology. We matched the nomogram input data to find the most similar recurrence case and non-recurrence control from the training dataset to an individual patient (query). The similarity between two patients is measured as the inverse of Euclidean distance between their clinical variables – age at prostatectomy, Gleason sum, and pathologic stages (see Supplementary Information for details). Hence, each query returns from the database a recurrent and a non-recurrent match. We rank the query against the two matches via Ranking-SVM. If the query is from a patient who will recur, it will be higher ranked than the non-recurrent sample from the database. If the query is from a non-recurrent patient, it will be ranked lower than the recurrent sample from the database. Quantifying the tendency in the ranks, a preference score for a patient p_1 against a patient p_2 is defined as follows:

$$\text{Preference}(p_1; p_2) = \sum_{s \in S_1} \text{ranking}(s; S_1 \cup S_2)^d \quad (2)$$

where S_i represents the set of cores samples from a patient p_i , $\text{ranking}(\bullet; S)$ denotes a rank of \bullet among the set of samples in S , and $d \geq 1$. Since the preference score is dependent on the number of samples of the two patients, a normalized preference score is computed:

$$n\text{Preference}(p_1; p_2) = \frac{\text{Preference}(p_1; p_2) - \text{Preference}^{\min}(p_1; p_2)}{\text{Preference}^{\max}(p_1; p_2) - \text{Preference}^{\min}(p_1; p_2)} \quad (3)$$

where $\text{Preference}^{\min}(p_1; p_2)$ and $\text{Preference}^{\max}(p_1; p_2)$ denote the minimum and maximum possible preference scores for the given two patients, respectively. Based on the two preference scores, we predict the probability of recurrence (“IR Score”) of the query patient using a logistic regression model (Figure 2d).

Kattan nomogram and CAPRA-S score. Kattan nomogram⁷⁸ is a graphical tool to predict up to 10-year probability of biochemical recurrence after radical prostatectomy. It is constructed based on the Cox proportional hazards regression model. It includes several clinical variables, for instance, PSA level, Gleason grade, lymph node invasion, seminal vesicle invasion, extracapsular extension, surgical margin status, and year of radical prostatectomy. Each variable is represented as a scale, and each scale position corresponds to a prognostic point that shows its prognostic significance in the model. Adding up the prognostic points of all the clinical variables, the total sum is mapped to the probability of biochemical recurrence.

CAPRA-S score¹⁰ also predicts 5-year biochemical recurrence after radical prostatectomy and uses the Cox proportional hazards regression model. It utilizes PSA level, Gleason score, surgical margin status, extracapsular extension, seminal vesicle invasion, lymph node invasion. Depending on a value of a clinical variable, it assigns a risk point up to 3 for PSA level and Gleason Score, 2 for surgical margin status and seminal vesicle invasion, and 1 for extracapsular extension and lymph node invasion. The minimum and maximum possible score is 0 and 12, respectively. The scores 0–2, 3–5, and ≥ 6 are considered to be a low-risk, intermediate-risk, and high-risk group, respectively.

Statistical Analysis. Statistical significance in discriminating pairs of recurrent cases and non-recurrent controls is determined by binomial test with a success probability of 0.5. ROC curves were smoothed prior to computing AUCs to compensate for the underestimation in the empirical AUCs and to estimate the true AUCs⁶³. Bootstrap re-sampling with 2000 repetitions is adopted to assess 95% confidence intervals (CI) of AUCs and statistically substantial differences between AUCs of the two ROC curves⁶⁴. Kolmogorov-Smirnov test is employed to examine the differences in score distributions between recurrent cases and non-recurrent controls. Wald chi-square statistic is used to measure the significance of individual predictors in a logistic regression model. The tests were performed using R software version 2.15.2.

1. Siegel, R., Naishadham, D. & Jemal, A. Cancer Statistics, 2012. *CA Cancer J Clin* **62**, 10–29 (2012).
2. Shariat, S. F., Kattan, M. W., Vickers, A. J., Karakiewicz, P. I. & Scardino, P. T. Critical review of prostate cancer predictive tools. *Future Oncol* **5**, 1555–1584 (2009).
3. Shariat, S. F., Karakiewicz, P. I., Margulis, V. & Kattan, M. W. Inventory of prostate cancer predictive tools. *Curr Opin Urol* **18**, 279–296 (2008).
4. D’Amico, A. V., Whittington, R., Malkowicz, S. et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA* **280**, 969–974 (1998).



5. Partin, A. W. *et al.* The Use of Prostate-Specific Antigen, Clinical Stage and Gleason Score to Predict Pathological Stage in Men with Localized Prostate-Cancer. *J Urol* **150**, 110–114 (1993).
6. Dayhoff, J. E. & DeLeo, J. M. Artificial neural networks - Opening the black box. *Cancer* **91**, 1615–1635 (2001).
7. Kattan, M. W., Wheeler, T. M. & Scardino, P. T. Postoperative nomogram for disease recurrence after radical prostatectomy for prostate cancer. *J Clin Oncol* **17**, 1499–1507 (1999).
8. Stephenson, A. J. *et al.* Postoperative nomogram predicting the 10-year probability of prostate cancer recurrence after radical prostatectomy. *J Clin Oncol* **23**, 7005–7012 (2005).
9. Cooperberg, M. R. *et al.* The University of California, San Francisco cancer of the prostate risk assessment score: A straightforward and reliable preoperative predictor of disease recurrence after radical prostatectomy. *J Urol* **173**, 1938–1942 (2005).
10. Cooperberg, M. R., Hilton, J. F. & Carroll, P. R. The CAPRA-S Score A Straightforward Tool for Improved Prediction of Outcomes After Radical Prostatectomy. *Cancer* **117**, 5039–5046 (2011).
11. Greene, K. L. *et al.* Validation of the Kattan preoperative nomogram for prostate cancer recurrence using a community based cohort: Results from cancer of the prostate strategic urological research endeavor (CaPSURE). *J Urol* **171**, 2255–2259 (2004).
12. Gallina, A. *et al.* Obesity does not predispose to more aggressive prostate cancer either at biopsy or radical prostatectomy in European men. *Int J Cancer* **121**, 791–795 (2007).
13. Stephenson, A. J. *et al.* Integration of gene expression profiling and clinical variables to predict prostate carcinoma recurrence after radical prostatectomy. *Cancer* **104**, 290–298 (2005).
14. Glinsky, G. V., Glinskii, A. B., Stephenson, A. J., Hoffman, R. M. & Gerald, W. L. Gene expression profiling predicts clinical outcome of prostate cancer. *J Clin Invest* **113**, 913–923 (2004).
15. Ayala, G. *et al.* Reactive stroma as a predictor of biochemical-free recurrence in prostate cancer. *Clin Cancer Res* **9**, 4792–4801 (2003).
16. Karnes, R. J. *et al.* The Ability of Biomarkers to Predict Systemic Progression in Men with High-Risk Prostate Cancer Treated Surgically Is Dependent on ERG Status. *Cancer Res* **70**, 8994–9002 (2010).
17. Cheng, G. C. *et al.* Clinical utility of endorectal MRI in determining PSA outcome for patients with biopsy Gleason score 7, PSA \leq 10, and clinically localized prostate cancer. *Int J Radiat Oncol Biol Phys* **55**, 64–70 (2003).
18. Shukla-Dave, A. *et al.* Prediction of Prostate Cancer Recurrence Using Magnetic Resonance Imaging and Molecular Profiles. *Clin Cancer Res* **15**, 3842–3849 (2009).
19. Wassberg, C. *et al.* The Incremental Value of Contrast-Enhanced MRI in the Detection of Biopsy-Proven Local Recurrence of Prostate Cancer After Radical Prostatectomy: Effect of Reader Experience. *Am J Roentgenol* **199**, 360–366 (2012).
20. Cordon-Cardo, C. *et al.* Improved prediction of prostate cancer recurrence through systems pathology. *J Clin Invest* **117**, 1876–1883 (2007).
21. Beck, A. H. *et al.* Systematic Analysis of Breast Cancer Morphology Uncovers Stromal Features Associated with Survival. *Sci Transl Med* **3** (2011).
22. Khamis, Z. I., Sahab, Z. J., Byers, S. W. & Sang, Q. X. Novel stromal biomarkers in human breast cancer tissues provide evidence for the more malignant phenotype of estrogen receptor-negative tumors. *J Biomed Biotechnol* **2011**, 723650 (2011).
23. Bhargava, R. Infrared Spectroscopic Imaging: The Next Generation. *Appl Spectrosc* **66**, 1091–1120 (2012).
24. Ellis, D. I. & Goodacre, R. Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and Raman spectroscopy. *Analyst* **131**, 875–885 (2006).
25. Argov, S. *et al.* Diagnostic potential of Fourier-transform infrared microspectroscopy and advanced computational methods in colon cancer patients. *J Biomed Opt* **7**, 248–254 (2002).
26. Miller, L. M. & Dumas, P. Chemical imaging of biological tissue with synchrotron infrared light. *Biochim Biophys Acta* **1758**, 846–857 (2006).
27. Petibois, C. & Deleris, G. Chemical mapping of tumor progression by FT-IR imaging: towards molecular histopathology. *Trends Biotechnol* **24**, 455–462 (2006).
28. Zagouri, F. *et al.* Female breast cancer in Europe: statistics, diagnosis and treatment modalities. *J Thorac Dis* **6**, 589–590 (2014).
29. Walsh, M. J., Reddy, R. K. & Bhargava, R. Label-Free Biomedical Imaging With Mid-IR Spectroscopy. *IEEE Sel Top Quant Electron* **18**, 1502–1513 (2012).
30. Kumar, S., Desmedt, C., Larsimont, D., Sotiriou, C. & Goormaghtigh, E. Change in the microenvironment of breast cancer studied by FTIR imaging. *Analyst* **138**, 4058–4065 (2013).
31. Gazi, E. *et al.* The combined application of FTIR microspectroscopy and ToF-SIMS imaging in the study of prostate cancer. *Faraday Discuss* **126**, 41–59; discussion 77–92 (2004).
32. Fernandez, D. C., Bhargava, R., Hewitt, S. M. & Levin, I. W. Infrared spectroscopic imaging for histopathologic recognition. *Nat Biotechnol* **23**, 469–474 (2005).
33. Bhargava, R. Towards a practical Fourier transform infrared chemical imaging protocol for cancer histopathology. *Anal Bioanal Chem* **389**, 1155–1169 (2007).
34. Hughes, C. *et al.* FTIR microspectroscopy of selected rare diverse sub-variants of carcinoma of the urinary bladder. *J Biophoton* **6**, 73–87 (2012).
35. Kong, R., Reddy, R. K. & Bhargava, R. Characterization of tumor progression in engineered tissue using infrared spectroscopic imaging. *Analyst* **135**, 1569–1578 (2010).
36. Holton, S. E., Walsh, M. J. & Bhargava, R. Subcellular localization of early biochemical transformations in cancer-activated fibroblasts using infrared spectroscopic imaging. *Analyst* **136**, 2953–2958 (2011).
37. Li, M. J., Hsu, H. S., Liang, R. C. & Lin, S. Y. Infrared microspectroscopic detection of epithelial and stromal growth in the human benign prostatic hyperplasia. *Ultrastruct Pathol* **26**, 365–370 (2002).
38. Kajdacsy-Balla, A. *et al.* Practical aspects of planning, building, and interpreting tissue microarrays: The Cooperative Prostate Cancer Tissue Resource experience. *J Mol Histol* **38**, 113–121 (2007).
39. Kwak, J. T., Reddy, R., Sinha, S. & Bhargava, R. Analysis of Variance in Spectroscopic Imaging Data from Human Tissues. *Anal Chem* **84**, 1063–1069 (2012).
40. DeSantis, C. E. *et al.* Cancer treatment and survivorship statistics, 2014. *CA Cancer J Clin* **64**, 252–271 (2014).
41. Han, J., Pei, J. & Yin, Y. Mining frequent patterns without candidate generation. *SIGMOD Rec* **29**, 1–12 (2000).
42. Graefen, M. *et al.* Validation study of the accuracy of a postoperative nomogram for recurrence after radical prostatectomy for localized prostate cancer. *J Clin Oncol* **20**, 951–956 (2002).
43. Joachims, T. Training linear SVMs in linear time. *Proc 12th ACM SIGKDD international conference on Knowledge Discovery and Data mining*, Philadelphia, PA, United States, 217–226, ACM (doi: 10.1145/1150402.1150429).
44. Waldon, J. *et al.* A comparison of cancer statistics in New Zealand and Australia: 1996–2007. *N Z Med J* **127**, 20–29 (2014).
45. Kwak, J. T., Hewitt, S. M., Sinha, S. & Bhargava, R. Multimodal microscopy for automated histologic analysis of prostate cancer. *BMC Cancer* **11**, 62 (2011).
46. Iwamoto, M., Nakamura, F. & Higashi, T. Estimated life expectancy and risk of death from cancer by quartiles in the older Japanese population: 2010 vital statistics. *Cancer Epidemiol* **38**, 511–514 (2014).
47. Chalmers, J. M. & Griffiths, P. R. *Handbook of vibrational spectroscopy*. (J. Wiley, New York, 2002).
48. Skorniyakov, I. V., Tolstorozhev, G. B. & Butra, V. A. Infrared Absorption Spectra of Human Malignant Tumor Tissues. *J Appl Spectrosc* **75**, 420–425 (2008).
49. Severcan, F., Kaptan, N. & Turan, B. Fourier transform infrared spectroscopic studies of diabetic rat heart crude membranes. *Spectroscopy* **17**, 569–577 (2003).
50. Gogoire, M. & Lieubeau, B. The role of fibroblasts in tumor behavior. *Cancer Metastasis Rev* **14**, 339–350 (1995).
51. Tomas, D. *et al.* Extensive retraction artefact predicts biochemical recurrence-free survival in prostatic carcinoma. *Histopathology* **58**, 447–454 (2011).
52. Paris, P. L. *et al.* A Group of Genome-Based Biomarkers That Add to a Kattan Nomogram for Predicting Progression in Men with High-Risk Prostate Cancer. *Clin Cancer Res* **16**, 195–202 (2010).
53. Parivar, F. *et al.* Detection of locally recurrent prostate cancer after cryosurgery: evaluation by transrectal ultrasound, magnetic resonance imaging, and three-dimensional proton magnetic resonance spectroscopy. *Urology* **48**, 594–599 (1996).
54. Klein, E. A. *et al.* A 17-gene Assay to Predict Prostate Cancer Aggressiveness in the Context of Gleason Grade Heterogeneity, Tumor Multifocality, and Biopsy Undersampling. *Eur Urol* **66**, 550–560 (2014).
55. Eberlin, L. S. *et al.* Ambient mass spectrometry for the intraoperative molecular diagnosis of human brain tumors. *Proc Natl Acad Sci USA* **110**, 1611–1616 (2013).
56. Eberlin, L. S. *et al.* Molecular assessment of surgical-resection margins of gastric cancer by mass-spectrometric imaging. *Proc Natl Acad Sci USA* **111**, 2436–2441 (2014).
57. Kong, K. *et al.* Diagnosis of tumors during tissue-conserving surgery with integrated autofluorescence and Raman scattering microscopy. *Proc Natl Acad Sci USA* **110**, 15189–15194 (2013).
58. Freudiger, C. W. *et al.* Label-free biomedical imaging with high sensitivity by stimulated Raman scattering microscopy. *Science* **322**, 1857–1861 (2008).
59. Ricciardelli, C. *et al.* Elevated stromal chondroitin sulfate glycosaminoglycan predicts progression in early-stage prostate cancer. *Clin Cancer Res* **3**, 983–992 (1997).
60. Nasse, M. J. *et al.* High-resolution Fourier-transform infrared chemical imaging with multiple synchrotron beams. *Nat Methods* **8**, 413–U458 (2011).
61. Ananthanarayanan, V. *et al.* Subcellular localization of p27 and prostate cancer recurrence: automated digital microscopy analysis of tissue microarrays. *Hum Pathol* **42**, 873–881 (2011).
62. Hong, C., Xifeng, Y., Jiawei, H. & Chih-Wei, H. Discriminative Frequent Pattern Analysis for Effective Classification. *Proc ICDE 2007 Istanbul, Turkey*, 716–725, IEEE (doi: 10.1109/icde.2007.367917).
63. Robin, X. *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC bioinformatics* **12**, 77 (2011).

Acknowledgments

Support from the National Institutes of Health via grant 1R01CA138882 is gratefully acknowledged. We would like to thank Prof. Peter Gann of University of Illinois at Chicago for fruitful discussions and several valuable suggestions.



Author contributions

Conception and design: J.K., A.B., S.S. and R.B. Development of methodology: J.K., S.S. and R.B. Acquisition of data: J.K., A.B., V.M. and M.W. Analysis and interpretation of data: J.K., A.B., V.M., M.W., S.S. and R.B. Writing, review, and/or revision of the manuscript: J.K., A.B., V.M., M.W., S.S. and R.B. Study supervision: S.S. and R.B.

Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Kwak, J.T. *et al.* Improving Prediction of Prostate Cancer Recurrence using Chemical Imaging. *Sci. Rep.* 5, 8758; DOI:10.1038/srep08758 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>