HHS Public Access

Author manuscript

*Cancer Prev Res (Phila)*. Author manuscript; available in PMC 2020 February 01.

# Three-dimensional nanoscale nuclear architecture mapping of rectal biopsies detects colorectal neoplasia in patients with inflammatory bowel disease

**Shikhar Uttam, PhD**[1,*], **Jana G. Hashash, MD**[2,*], **Justin LaFace, BS**[3], **David Binion, MD**[2], **Miguel Regueiro, MD**[4], **Douglas J Hartman, MD**[5], **Randall E Brand, MD**[2,6], **Yang Liu, PhD**[2,3,6]

[1]Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, Pennsylvania.

[2]Division of Gastroenterology, Hepatology and Nutrition, Department of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania.

[3]Biomedical Optical Imaging Laboratory, Departments of Medicine and Bioengineering, University of Pittsburgh, Pittsburgh, Pennsylvania.

[4]Department of Gastroenterology, Hepatology and Nutrition, Cleveland Clinic, Cleveland, Ohio.

[5]Department of Pathology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.

[6]University of Pittsburgh Hillman Cancer Center, Pittsburgh, Pennsylvania.

## Abstract

Patients with inflammatory bowel disease (IBD) colitis are at an increased risk of developing colorectal cancer and are currently recommended to undergo extensive annual or biennial colonoscopy, a costly and invasive procedure. Most surveillance colonoscopies are negative with no existing objective measures for assessing their risk of developing cancer. We have recently developed a less invasive, cost-effective and objective method to assess cancer risk by detecting the presence of colonic neoplasia via three-dimensional (3D) nanoscale nuclear architecture mapping (nanoNAM) of normal-appearing rectal biopsies. To establish its translational relevance, we prospectively recruited 103 patients with IBD colitis undergoing surveillance colonoscopy and measured sub-microscopic alterations in aberrant intrinsic nuclear architecture of epithelial cells from normal-appearing rectal biopsies with nanoNAM. The results were correlated with the histologic diagnoses from all random biopsies obtained during initial and follow-up colonoscopy within 3 years. Using nanoNAM-based structural characterization as input features into a soft-margin-based nu-SVM risk classifier, we show that nanoNAM detects colonic neoplasia with area under the curve of 0.87±0.04, sensitivity of 0.81±0.09, and specificity of 0.82±0.07 in the independent validation set. Additionally, projecting nanoNAM features onto a 2-sphere reveals low-risk and high-risk IBD colitis patients existing on separate hemispheres. Finally, we show that

Correspondence: Shikhar Uttam, PhD, Department of Computational and Systems Biology, University of Pittsburgh, Room 3088, 3064 Biomedical Science Tower 3, 3501 Fifth Avenue, Pittsburgh, PA 15260. shf28@pitt.edu, Yang Liu, PhD, University of Pittsburgh, University of Pittsburgh Hillman Cancer Center, Suite 2.32e, 5117 Centre Ave, Pittsburgh, PA 15232. liuy@pitt.edu.
*These two authors contribute the work equally.

**Conflict of Interest:** The technology used in this manuscript is associated with US Patents 9,983,399 and 10,156,479 owned by University of Pittsburgh.

this ability to assess cancer risk translates to clinically-relevant estimation of individual-patient likelihood of being truly at risk. We demonstrate the potential of nanoNAM to identify IBD patients at higher risk of developing cancer from normal-appearing rectum tissue, which may aid clinicians in personalized IBD colitis patient surveillance.

## INTRODUCTION

Inflammatory bowel diseases (IBD), including Crohn's disease and ulcerative colitis (UC), affect about 1.6 million Americans (1). IBD colitis patients are at an increased risk for dysplasia and colorectal cancer (CRC). Chronic colonic inflammation predisposes patients to dysplasia and subsequent progression to neoplasia. Neoplastic progression in IBD colitis is hypothesized to follow the inflammation–dysplasia–carcinoma sequence with progressive accumulation of oncogenic traits in clonal populations modulated by flexible epigenetic mechanisms in the background of intestinal inflammation, resulting in aberrant epithelial morphology and adenoma formation (2). Given the increased risk for developing CRC, the American Gastroenterology Association (AGA) and American Society for Gastrointestinal Endoscopy (ASGE) guidelines recommend that IBD colitis patients should undergo extensive colonoscopy surveillance with multiple random biopsies (4-quadrant biopsies every 10 cm throughout the entire involved colon), repeated every 1–2 years (3). Chromoendoscopy with targeted biopsies is also endorsed (4), although its use is limited due to interobserver variability, requisite expertise, higher cost and prolonged procedure. However, despite their higher risk, most IBD colitis patients will not develop CRC, and a majority of these surveillance procedures are negative (5). Furthermore, prediction of risk of developing cancer based on established risk factors is far from accurate, with no objective measures currently existing to assess cancer risk for individual colitis patients. Given the high cost, invasiveness, poor patient compliance and inefficiencies associated with frequent colonoscopies and biopsies, there is an urgent need for a less-invasive, objective, cost-effective and clinically-applicable risk-surveillance approach to triage IBD colitis patients into low- and high-risk groups, where close surveillance is performed on high-risk patients and lengthened surveillance interval is applied to low-risk patients.

Histologic evaluation of colonic tissue biopsies remains the clinical gold standard for assessing the presence of dysplasia/neoplasia amongst patients with IBD colitis, with microscale architectural changes in the epithelial cell nuclei due to spatial arrangement of chromatin and other nuclear components (6) being among the most striking features. It is well established that the remodeling of nuclear architecture plays an important role in epigenetic dysregulation of cellular processes during tumorigenesis (7). We have recently developed *nano*scale *n*uclear *a*rchitecture *m*apping (nanoNAM) approach that measures sub-microscopic manifestations of aberrant nuclear architecture. It is a cost-effective imaging approach that uses standard clinically-obtained formalin-fixed, paraffin-embedded (FFPE) tissue section and is capable of quantifying nanoscale-sensitive, depth-resolved, sub-microscopic alterations in the *intrinsic* architecture of epithelial cell nuclei (8,9). Unlike conventional histopathology that indirectly examines microscopic nuclear architecture through visualization of spatial distribution of the chromogenic contrast within the stained tissue sample, nanoNAM directly quantifies nuclear architecture of unstained tissue samples

through its three-dimensional (3D) refractive index distribution. The physics behind light-matter interaction (10) reveals that biological samples have a biophysical representation (11) as specimens defined by their refractive index distribution (12). This refractive index distribution quantifies the optical density of the sample, and hence its intrinsic nuclear architecture. nanoNAM captures depth-resolved properties of this nuclear optical density with nanoscale sensitivity (9). We emphasize that these properties are differential in nature, in that, they measure mean and entropic *change* in sample optical density. In a retrospective study, we showed that nanoNAM-based nuclear architecture characterization of histologically normal-appearing tissue from the colonic mucosa detected high-risk IBD colitis patients (13) who progressed into high-grade dysplasia or adenocarcinoma.

Built upon our previous work, this prospective study focuses on characterizing 3D nanoscale-sensitive intrinsic nuclear architecture of normal-appearing rectal biopsies to detect the presence of neoplasia throughout the entire colon and assess the risk in developing colorectal cancer in patients with IBD colitis. It is based on the concept of field cancerization, also known as field effect or field defect, that hypothesizes that lesions located at a site in the more proximal colon alter the genetic and epigenetic milieu at a distant site like the rectum (14). This phenomenon has been well established in CRC (15,16). We aim to demonstrate that by detecting the intrinsic nanoscale changes of nuclear architecture from the histologically normal-appearing cell nuclei of the rectal mucosa, we can distinguish low- and high-risk IBD colitis patients.

## METHODS

### Nanoscale nuclear architecture mapping (nanoNAM) principle and instrument

The nanoNAM approach builds on the well-established principle behind Spectral-domain Optical Coherence Tomography (SD-OCT) and the underlying physics has been described in detail previously (8,9). Figure 1A illustrates the general principle. The backscattered light from within the unstained rectal tissue sample and the reference reflected-light from the single-layer dielectric coating (~20% reflection, ~80% transmission) at the sample-substrate interface form a common-path interferometer. This common-mode interference allows weak back-scattered light from within the sample to “ride” the reference signal, which is then captured on the imaging sensor as a spectral interference signal. We have shown that within cell nuclei where the nuclear architecture has slowly varying refractive index without a strong interface, these back-scattered signals carry *intrinsic* nuclear architectural information (9,13). nanoNAM accesses this information through the Fourier phase mapping of the spectral interference signal to compute, with nanoscale sensitivity, two specific depth-resolved properties of the nuclear architecture: (a) depth-resolved alteration in average nuclear architectural density, and (b) depth-resolved localized heterogeneity in the density of the nuclear architecture (8). The physical principle and detailed mathematical derivation of these structural properties accessed via nanoNAM have been described previously (8,9) (see also Supplementary Methods and Supplementary Fig. S1 for further details). The depth resolution of these properties is due to the implicit coherent windowing, a Fourier consequence of the spectral bandwidth of the light source (17). This windowing ensures that

the nanoNAM-based measurements of intrinsic nuclear architectural properties are not confounded by variation in the overall tissue section thickness (8,18).

The nanoNAM optical system has been previously described (9,13). In brief, it is a tri-modal optical system with a two-stage workflow. The three optical modalities include (1) depth-resolved nanoNAM on the unstained sample (Supplementary Fig. S2A); (2) bright-field imaging of H&E-stained sample (Supplementary Fig. S2B); and (3) diffraction phase imaging of unstained and stained samples for image registration (Supplementary Fig. S2C). Such workflow is to facilitate reliable identification of cell nuclei from low-contrast unstained nanoNAM images. Following nanoNAM on unstained tissue sample, standard H&E-staining is performed to ensure correct identification of epithelial nuclei and integration with pathology information. The diffraction-phase images performed on both unstained and stained samples are used to register these two images.

## Patient selection

This study was approved by the Institutional Review Board of the University of Pittsburgh following the US Common Rules. The investigators who were involved in patient recruitment obtained informed written consent from the subjects. We prospectively recruited 103 patients with long-standing UC or Crohn's colitis who were scheduled for a standard surveillance colonoscopy at the University of Pittsburgh Medical Center (Table 1 details patient clinical and demographic information along with histological evaluation of surveillance biopsies and any follow-up surveillance biopsies within 3 years, if available). All of these patients underwent CRC surveillance either with chromoendoscopy and targeted biopsies or via colonoscopy with 4-quadrant random biopsies every 10cm throughout their colon, per AGA and ASGE IBD colitis surveillance protocol (3). Patients with active malignancy anywhere in their body and those patients who were scheduled to undergo colonoscopy for accessing disease activity and response to treatment were excluded from the study. In addition to the standard-of-care surveillance biopsies, 2–3 pinch biopsies from normal-appearing rectal mucosa were obtained at the beginning of each colonoscopy procedure. If chromoendoscopy was performed, the biopsies were obtained before the dye was sprayed. The rectal biopsies were immediately fixed in 10% formalin and then subsequently processed with a standard tissue histology protocol (i.e. paraffin embedding, sectioning at 5 μm, and coverslip) for nanoNAM. The person who performed data acquisition and nuclei segmentation was blinded to the clinical data. This was easily implemented without bias because the rectal tissue was histologically normal-appearing.

## Histopathologic diagnosis and categorization of risk groups

The histopathology of each patient's random colonic biopsies was reviewed by our expert gastrointestinal pathologist (DH). All rectal biopsies analyzed by nanoNAM showed no active inflammation with good tissue preservation, with just a few showing mild inflammation. Based on the final pathology results from all random biopsies taken throughout the colon, patients were categorized as low risk or high risk. Low-risk patients were those whose pathology of the surveillance biopsies throughout the entire colon showed no dysplasia or CRC at the time of the evaluation, while the high-risk patients were those who either had any dysplasia, including tubular adenoma, indefinite for dysplasia, low-grade

dysplasia, high-grade dysplasia or adenocarcinoma anywhere in the colon at the time of initial evaluation or during the follow-up surveillance colonoscopy biopsies within 3 years.

The rationale for the above low- and high-risk IBD patient grouping was motivated by the clinical need to assess each patient's risk from easily accessible rectum and determine whether a subsequent full colonoscopy was needed or not. In this IBD patient surveillance setting, although adenoma is not traditionally considered high-risk in the context of aggressive colonic lesions, any patient with adenoma or more advanced neoplastic lesions deserves a closer follow-up. Consequently, we included within the high-risk cohort those patients with adenoma/low-grade dysplasia, high-grade dysplasia, or adenocarcinoma. As a result, follow-up colonoscopy would be limited to this high-risk cohort, thereby limiting the number of negative surveillance procedures. We note that inclusion of non-advanced adenoma emphasizes the ability of nanoNAM to assess the risk of entire spectrum of dysplasia from normal-appearing rectal biopsies.

## nanoNAM processing and patient features

The nanoNAM-based imaging of an epithelial cell nucleus generates a volumetric data-cube $\phi_{nanoNAM}(x,y,z_{opl})$ describing the 3D intrinsic nuclear architecture of the cell. Its 3D visualization is shown in Fig. 1B. Here the $(x,y)$ coordinates specify the en face location in the physical space, while $z_{opl}$ is the depth coordinate in the optical space, a consequence of the optical density based biophysical characterization of the cell. Details are provided in the Supplementary Methods.

We performed nanoNAM on ~500 normal-appearing epithelial cells from the rectal mucosa of each IBD colitis patient, with corresponding ~500 $\phi_{nanoNAM}(x,y,z_{opl})$ data-cubes per patient. We summarized the intrinsic nuclear architecture information within these data-cubes at the patient level following a two-step procedure. First, we defined *mean* and the *entropy* functions, each of which took a data-slice $\phi_{nanoNAM}(\cdot,\cdot,z_{opl})$ at optical depth $z_{opl}$ from the data-cube $\phi_{nanoNAM}(x,y,z_{opl})$ for each epithelial cell nucleus, and as shown in Fig. 1C, respectively computed the mean and entropy of the intrinsic nuclear architecture for positive values within that data-slice. The positivity ensures an injective correspondence between increase (or decrease) in $\phi_{nanoNAM}(\cdot,\cdot,z_{opl})$ values with increase (or decrease) in depth-resolved alteration in average nuclear architectural density and depth-resolved localized heterogeneity in the density of the nuclear architecture (8). The resulting optical-depth dependent $\phi_{mean}(z_{opl})$ and $\phi_{entropy}(z_{opl})$ values respectively quantify the average intrinsic architecture and the distribution of that architecture at optical depth $z_{opl}$. We performed the operations at discrete optical depths with a step-size of 45 nm in each epithelial cell nucleus resulting in depth-dependent mean and entropy vectors (Fig. 1D) summarizing the properties of intrinsic architecture of each nucleus at each optical depth. We column-wise arranged these vectors from all epithelial cells to obtain the mean and entropy matrices summarizing the depth-dependent intrinsic-architecture at the patient level, as illustrated in Fig. 1E.

Second, to obtain overall quantitative nuclear architecture summary for each patient, we summarized their mean and entropic intrinsic nuclear architectural matrices using three matrix norms – the 1-norm, the infinity-norm and the spectral-norm. In matrix theory, these

matrix norms respectively capture the maximal values of the matrix along the columns, rows and eigen-directions of the matrix. In our context, the columns correspond to patients' cell nuclei, while the rows correspond to the sample optical depth. The eigen-directions are the directions within the intrinsic nuclear architectural space defined by the depth-resolved nuclear architecture of all the patient's cell nuclei that sequentially capture spectral directions that span this space. As a result, for nanoNAM, the 1-, infinity- and spectral-norms for the mean and entropy patient matrices, respectively capture the dominant nuclear architectural mean and entropy across all (a) patients' cell nuclei, (b) optical depths, and (c) the optical density space defined by all the patients' cell nuclei. (Details are provided in the Supplementary Methods.) Thus, the matrix norms provide us three different *views* of two different – mean and entropy – nuclear architectural properties related to alterations in the average depth-resolved density of the patients' cell nuclei, and the heterogeneity of these alterations. Together they summarize the nuclear architecture of the patient via a six-dimensional (6D) vector of nanoNAM-based nuclear architectural properties. We refer to these properties as spectral-nanoNAM mean, 1-nanoNAM mean, infinity-nanoNAM mean, spectral-nanoNAM entropy, 1-nanoNAM entropy, and infinity-nanoNAM entropy.

We note that the above two-step procedure did not apply any ad-hoc heuristics in its computations and did not use any empirical thresholds to include or exclude data. We also emphasize that the norms compute the dominant values of the nuclear architecture from their depth-resolved mean and entropy values. Consequently, they summarize from three different perspectives patient-level dominant alterations in the average cell nuclei density and the heterogeneity of these alterations.

## Support vector machine-based risk prediction

We considered the six-dimensional nanoNAM-based patient characterization as a point in a 6D Euclidean vector space. The collection of these points for the IBD colitis patient cohort describes our nanoNAM patient-data constellation. We next considered training a support vector machine-based classifier to test the ability of nanoNAM-based measurements to partition the patient constellation based on patient risk, and therefore, serve as a risk prediction model.

Support vector machines (19) are among the best binary supervised-learning classifiers with deep foundation in statistical learning theory (20). We trained a soft-margin support vector machine using nu-parameterization to classify the cancer risk of IBD colitis patients based on the 6D nanoNAM-based patient characterization of their intrinsic nuclear architecture. Soft-margin-based learning optimized the trade-off between classification ability and noise in the data, while its nu-parametrization allows this trade-off to be bounded between zero and one (21). The patient cohort was split into mutually exclusive training and validation sets. The split was performed via stratified sampling such that both training and validation sets respectively contained 50% of the relative proportion of the low- and high-risk IBD colitis patients. The stratification ensured that the relative mix of patient risk in the two sets was similar. The nu-SVM classifier was trained using leave-one-out cross-validation and was tested on the independent validation set to compute the sensitivity and specificity of risk prediction via computation of receiver operator characteristics (ROC). Stratified

bootstrapping, based on the stratified sampling procedure described above, was employed to ensure that our training and testing-based risk-prediction model was robust. Specifically, 500 pairs of bootstrapped training-testing pairs were generated, and nu-SVM-based risk prediction model was trained on each of the training sets and tested on the corresponding validation sets. From the resulting distribution of 500 ROC curves, confidence band indicating the confidence interval in which 95% of the curves lay was computed (22).

We further studied clinical translational relevance of the nanoNAM-based cancer-risk prediction model by using its ROC-based performance to compute the likelihood of truly assessing patient-risk at the individual patient level via characterization of nuclear architecture of histologically-normal appearing rectal tissue biopsies. Additionally, we also computed the individual effect-size and corresponding power of each of the six nanoNAM properties to identify high-risk IBD colitis patients among the patients included in the prospective study.

## RESULTS

### nanoNAM measurements quantify risk of cancer progression in IBD colitis patients

We computed the six nanoNAM properties for both low- and high-risk patients. Figure 2 shows representative images of 3D nanoNAM measurements and their mean projection along the sample depth – z-axis – for both a low- (Figs. 2, A–E) and a high-risk (Figs. 2, F–J) patient. The pseudo-color images of mean projection map clearly show a higher value (redder) of nanoNAM properties in the high-risk patient. To ascertain the cohort-level behavior for each of these properties, we next computed the kernel density estimate of the probability distribution for each of these properties in the low-risk and high-risk patients. Figures 3(A–F) shows the probability distribution for each of the properties for both low- and high-risk patients. It can be clearly seen that the probability distribution for all six properties shows a rightward shift in high-risk patients with respect to low-risk patients. This rightward shift in nanoNAM properties of high-risk IBD colitis patients, corresponding to increased alterations in mean nuclear density and associated heterogeneity, suggests that epithelial cell nuclei from high-risk patients have higher irregularity and size variability when compared to low-risk patients, even though their epithelial cell nuclei appear histologically normal. The p-values corresponding to this rightward shift at the 0.05 significance level for the six nanoNAM properties are shown in Fig. 3G. They indicate the significance of all but one nanoNAM property in detecting this shift in high-risk patients. To ensure that the statistical significance of this ability is not confounded by sample size, we computed the effect size of each of the properties to obtain their respective power to correctly identify IBD colitis patients at a higher risk of developing cancer. As can be seen in Fig. 3G, four of the six properties have high power to reject low-risk classification of IBD patients if their real risk is high. These results strongly suggest that nanoNAM-based quantification of nuclear architecture can not only detect the field effect of colorectal carcinogenesis in IBD colitis patients but also has the potential to be used to identify patients at a higher cancer risk. We note that since patients undergoing surveillance colonoscopy are ulcerative or Crohn’s colitis patients in remission without inflammation in their rectum, inflammation does not serve as a confounding factor in our study (see Supplementary

Methods and Supplementary Fig. S3 for further details.) We also analyzed the effect of diverse orientations and sectioning of cell nuclei across the tissue sample on the NAM parameters, which shows no bias (see Supplementary Methods and Supplementary Fig. S4 for further details).

## Projection of nanoNAM properties on a 2-sphere reveals low-risk and high-risk IBD colitis patients existing on separate hemispheres

To visualize the nature of risk captured by nanoNAM in IBD colitis patients and probe if it has an underlying biophysical structure, we used the six mean and entropic nanoNAM properties to construct a set of three meta properties. Each meta property was respectively defined by the ratio of one of the norm-based entropic architectural property and the corresponding norm-based mean architectural property. Since we considered three norms, we computed three ratios. These ratios between entropy of the intrinsic architecture distribution and its mean have distinct similarity to the coefficient of variation, which is the ratio between standard deviation of a distribution and its mean. Note that entropy, like standard deviation, is a measure of the spread of the distribution, albeit a more comprehensive one. Therefore, these coefficient-of-variation-like meta properties capture the mean-corrected spread of the intrinsic nuclear architecture across patient cell nuclei, (1-norm), optical depths (infinity-norm), and the nuclear-architectural space defined by all the patient cell nuclei (spectral-norm). We respectively refer to these three meta properties as CV-1, CV-infinity, and CV-spectral nuclear architectural properties. We represented these properties as points in a 3D Euclidean space and projected them onto a unit 2-sphere, the 2D surface of a unit sphere centered at zero, as shown in Fig. 4A. As can be seen, the patients in the two risk groups approximately lie on two separate hemispheres. This visualization interestingly suggests the possible existence of biophysical low- and high-risk states associated with the spread of the distribution of mean alterations and heterogeneity of nuclear architecture from normal-appearing epithelial cell nuclei of IBD colitis patients captured by nanoNAM-based measurements. Our future work will build on this possible existence of risk-states to assess how nanoNAM could help longitudinally surveil patients with evidence-based adjustment of their surveillance interval.

## nanoNAM measurements predict risk of cancer progression

We finally used the 6D nanoNAM properties as features to train a soft-margin support vector machine risk-classifier using nu-parameterization and a polynomial kernel of degree 3. As described in the Methods section we used stratified bootstrapping of low- and high-risk patients to test prediction performance of the risk-classifier. Figure 4B shows the resulting ROC on the independent validation sets with 95% confidence interval. We also computed three ROC-based prediction metrics of the nanoNAM-based prediction model: area under ROC of 0.87±0.04, sensitivity of 0.81±0.09 and specificity of 0.82±0.07. The sensitivity and specificity values corresponded to the point on the ROC curve that jointly maximized both. These metrics demonstrate the promise of nanoNAM-based quantification of field-effect to predict cancer risk of individual IBD colitis patients from their histologically normal-appearing biopsies of the rectal mucosa. (See Supplementary Fig. S5A for training set performance.)

Clinical translation relevance of nanoNAM-based risk-classification for an individual IBD colitis patient, however, not only depends on the overall prediction accuracy but also on how the classification should be interpreted for that patient. Figure 4C shows the clinically-relevant interpretation of nanoNAM-based testing for a patient via a negative and positive likelihood ratio plot corresponding to the ROC shown in Fig. 4B. It shows that nanoNAM-based risk prediction has a positive likelihood ratio of about 17.77, with a corresponding negative likelihood ratio of about 0.55. Thus, if nanoNAM-based risk-classifier labels a patient as high-risk, it increases the odds of the patient being high-risk by a factor of 17.77, a large effect. In contrast, if nanoNAM testing classifies a patient as low-risk, then the pre-nanoNAM test odds of a patient decrease by 0.55, a moderate effect. These results indicate that patients identified by nanoNAM as high risk are indeed truly at risk with a high probability. At the same time specificity of 0.82 indicates that the probability of nanoNAM falsely identifying a low-risk patient as high risk is low. Taken together, the results strongly suggest that nanoNAM-based quantification has capability to identify and triage high-risk IBD patients.

## Quantitative feature analysis of conventional pathology does not distinguish high-risk from low-risk cohort

Currently, the most widely used approach to aid conventional pathology – the gold-standard for assessing risk of cancer progression – is quantification of microscale H&E-stained histology images to reveal non-obvious patterns that correlate with pathologic or clinical evidence. We, therefore, compared our results with both conventional pathology and its quantitative analysis. First, our expert pathologist (DH) reviewed all patient slides and did not identify any histologic features differentiating high- from low- risk IBD colitis patients. Second, using bright-field images of H&E stained tissue samples, we extracted a comprehensive set of microscopic image features to evaluate whether they could identify high-risk patients. We focused on morphological, textural and statistical features and extracted four distinct features for each of the two risk groups. Specifically, we extracted, area, elongation, compactness, and perimeter as the four morphological features; energy, contrast, homogeneity, and correlation as textural features; and mean, standard deviation, skewness and kurtosis as statistical features. Together, these features and their derivatives are extensively used in computational pathology-based image analysis. As illustrated in Figs. 5A–C, none of the twelve features showed a separation between the two groups at the significance level of 0.05. To complete the analysis, we used the twelve features to train a nu-SVM classifier in a manner identical to nanoNAM and the resulting validation set ROC with 95% confidence interval computed using stratified bootstrapping is shown in Fig. 5D. (See Supplementary Fig. S5B for training set performance.) As can be seen the ROC has a poor trade-off between sensitivity and specificity, while its area under curve of 0.76±0.05 is significantly lower than 0.87±0.040 for nanoNAM. Moreover, the increase in the bootstrapped ROC confidence interval for these microscale image features indicates a less robust risk classifier when compared with nanoNAM, thereby illustrating that (a) the variation in H&E staining is a fundamental confounding factor in the computation of these features with resulting variability in risk-prediction performance; and (b) microscale features are not always sensitive to detecting sub-microscopic manifestations of aberrant nuclear architecture that appear histologically normal to pathologists.

## DISCUSSION

Despite being at an increased risk for developing CRC, the majority of IBD colitis patients will not develop cancer (23). However, current guidelines for colonoscopy surveillance in IBD colitis patients are imprecise and invasive. The results of this prospective study with 103 patients suggest that nanoNAM has the potential to detect the field effect of CRC in IBD colitis patients and identify patients at higher cancer risk based on characterizing the intrinsic nuclear architecture of histologically normal-appearing cells from rectal biopsies. Although still considered somewhat invasive, biopsies from the rectal mucosa can be performed with a simple flexible sigmoidoscopy or even without an endoscopic procedure that does not require bowel preparation or sedation as needed for a full colonoscopy. Moreover, the ability of nanoNAM-based analysis in identifying IBD colitis patients at a higher risk of developing cancer, together with its specificity of separating those IBD colitis patients that are at lower risk would help triage these patients in CRC surveillance based on their individual risk of developing CRC. We note that our proposed approach is not a replacement for a full colonoscopy that remains a gold standard for the detection of polyps and cancer. This risk-based personalized surveillance strategy may potentially maximize efficiency by identifying a small subset of truly high-risk patients for subsequent full colonoscopy, thus reducing cost and improving patient care.

Since being first hypothesized by Slaughter et al. in 1953 (24), the evidence for field effect (or field cancerization) has been accumulating in colorectal and many other cancers (25). Although research on elucidating the underlying mechanisms of field cancerization is ongoing, the potential clinical significance of utilizing field cancerization in improving cancer risk assessment is widely recognized (25,26). Our less-invasive yet clinically relevant approach to surveilling IBD patients at risk of developing colorectal cancer adds to the growing evidence on the utility of field effect in improving early cancer detection and risk assessment.

The nanoNAM approach interrogates the 3D structural changes using the Fourier phase of the spectral interference between light back-reflected from within the cell nuclei and the common-mode reference. It characterizes the intrinsic nuclear architecture by measuring the depth-resolved alteration in the average local density of the cell nuclei, and the localized heterogeneity in these alterations. The common-mode setup provides robustness to noise, thereby allowing nanoNAM to detect sub-microscopic abnormalities of nuclear architecture with nanoscale sensitivity of ~1–2 nm before their manifestation at the microscopic scale (9,13).

Our nanoNAM approach has the potential to be translated to the clinic because it fits into the standard clinical workflow: 1) It only requires FFPE tissue samples prepared from rectal biopsies that are easily obtained in a less-invasive manner than colonoscopy. 2) It is cost-effective with minimal running cost or consumable biochemical reagents. 3) It is relatively fast – imaging a slide (sampling 500–1000 epithelial cell nuclei only takes about 40 minutes. 4) Finally, its computational burden is minimal as the data-processing can be performed on a standard personal computer.

Projecting nanoNAM properties onto the surface of a unit sphere reveals an interesting dichotomization of the low- and high-risk spaces into two distinct hemispheres. The transition from one risk-space to the other may suggest the existence of an underlying “path” for assessing risk progression in IBD patients. We intend to pursue this as part of our future work with more extensive testing with well-controlled cancer progression models.

Unlike chromogenic or immunofluorescence-based imaging approaches, nanoNAM uses unstained samples. This is required for nanoNAM, as it directly characterizes properties of the intrinsic biophysical structure of the cell nucleus defined by the nuclear refractive index. Staining would introduce error in such measurements because of its indirect effect on nuclear refractive index through the Kramers-Kronig relations (10). Although post-hoc staining correction is an option, it is either a statistical correction using a regression model, or a model-based correction with strong assumptions about smoothness and asymptotic decay properties of the stain absorption spectrum (27). Consequently, it is difficult to completely remove the effect of stain at every location of the cell nucleus. Using unstained sample allows us to avoid stain-induced alteration of nanoNAM measurements resulting in accurate characterization of 3D-intrinsic nuclear architecture.

Imaging modalities based on quantitative phase imaging principles do provide nanoscale sensitive measurements (28,29). Unlike nanoNAM, however, they measure optical thickness of the sample-section that can potentially be confounded by variations in sample-section thickness. Other physical-optics-based modalities such as confocal light absorption and scattering (CLASS) microscopy (30) and partial-wave spectroscopy (PWS) (31) also detect structural changes with nanoscale sensitivity. PWS, for example, measures interference between scattering from within the cell and the strong reflection from the cell surface – with its varying surface topography – to quantify the disorder strength of cellular mass density (32), while CLASS models sub-cellular size distributions within the cell by averaging light back-scattered from within a confocal volume. Unlike these modalities, nanoNAM uses a stable and sample-independent reference signal to obtain the spectral interference signal from which it extracts 3D depth-resolved intrinsic nuclear architectural properties after mapping it to the optical space via Fourier mapping. Imaging modalities, such as optical coherence phase microscopy (33,34), are closer to nanoNAM in their principle, but they measure nanoscale shifts at strong interfaces at the sample surface, or within the sample such as blood vessel boundaries, and do not characterize intrinsic architecture within the cell nuclei, where the nuclear architecture presents slowly varying refractive index without a strong interface.

Currently, conventional pathology remains the clinical gold standard for assessing the risk of an IBD colitis patient for colorectal cancer. We view nanoNAM as a low-cost extension of pathology to the sub-microscopic realm, to aid clinicians and pathologists better assess the cancer risk of IBD colitis patients. Combined, with our earlier results based on mouse models and a retrospective study of IBD colitis patients, this study using prospectively recruited IBD patients provides further evidence for the ability of nanoNAM to detect the presence of colonic neoplasia and assess cancer risk in IBD colitis patients from histologically normal-appearing cells from the easily accessible rectum. Despite the promise, we aware that our relatively small size, especially in the high-risk group, of the

current patient cohort requires further validation and more detailed subgroup analysis of patients in the high-risk group using a larger patient study. Our future work will focus on using nanoNAM in a multi-center study to ensure a successful translation to the clinic and to investigate the underlying biophysical basis of the abnormal nuclear architecture to assess cancer risk.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

**Grant Funding:** This work is supported by National Institute of Health Grant Number R01EB016657, R01CA185363, R01CA232593 and Broad Medical Research Foundation.

## REFERENCES

1. Molodecky NA, Soon IS, Rabi DM, Ghali WA, Ferris M, Chernoff G, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. Gastroenterology. 2012;142:46–54.e42. [PubMed: 22001864]
2. Waldner MJ, Neurath MF. Mechanisms of Immune Signaling in Colitis-Associated Cancer. CMGH 2015 page 6–16.
3. Shergill AK, Lightdale JR, Bruining DH, Acosta RD, Chandrasekhara V, Chathadi KV., et al. The role of endoscopy in inflammatory bowel disease. Gastrointest Endosc. 2015;81:1101–21. [PubMed: 25800660]
4. Carballal S, Maisterra S, López-Serrano A, Gimeno-García AZ, Vera MI, Marín-Garbriel JC, et al. Real-life chromoendoscopy for neoplasia detection and characterisation in long-standing IBD. Gut. 2018;67:70 LP–78. [PubMed: 27612488]
5. Mattar MC, Lough D, Pishvaian MJ, Charabaty A. Current Management of Inflammatory Bowel Disease and Colorectal Cancer Gastrointest Cancer Res. Melville, NY: International Society of Gastrointestinal Oncology; 2011;4:53–61. [PubMed: 21673876]
6. Zink D, Fischer AH, Nickerson JA. Nuclear structure in cancer cells. Nat. Rev. Cancer 2004 page 677–87. [PubMed: 15343274]
7. Shen H, Laird PW. Interplay between the cancer genome and epigenome. Cell. 2013 page 38–55. [PubMed: 23540689]
8. Uttam S, Liu Y. Fourier phase in Fourier-domain optical coherence tomography. J Opt Soc Am A Opt Image Sci Vis 2015;32:2286–306. [PubMed: 26831383]
9. Uttam S, Liu Y. Fourier phase based depth-resolved nanoscale nuclear architecture mapping for cancer detection Methods. Academic Press; 2018;134–51.
10. Jackson JD. Classical Electrodynamics Third Edition. John Wiley Sons, INC 1999.
11. Liu PY, Chin LK, Ser W, Chen HF, Hsieh C-M, Lee C-H, et al. Cell refractive index for cell biology and disease diagnosis: past, present and future. Lab Chip. 2016;16:634–44. [PubMed: 26732872]
12. Hwu Y, Tsai WL, Chang HM, Yeh HI, Hsu PC, Yang YC, et al. Imaging cells and tissues with refractive index radiology. Biophys J. 2004;87:4180–7. [PubMed: 15465870]
13. Uttam S, Pham HV., LaFace J, Leibowitz B, Yu J, Brand RE, et al. Early prediction of cancer progression by depth-resolved nanoscale mapping of nuclear architecture from unstained tissue specimens. Cancer Res. 2015;75:4718–27. [PubMed: 26383164]
14. Slaghter DP, Southwick HW, Smejkal W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. Cancer. 1953;6:963–8. [PubMed: 13094644]

15. Alberts DS, Einspahr JG, Krouse RS, Prasad A, Ranger-Moore J, Hamilton P, et al. Karyometry of the colonic mucosa. Cancer Epidemiol Biomarkers Prev. 2007;16:2704–16. [PubMed: 18086777]
16. Lochhead P, Chan AT, Nishihara R, Fuchs CS, Beck AH, Giovannucci E, et al. Etiologic field effect: Reappraisal of the field effect concept in cancer predisposition and progression. Mod Pathol. 2015;28:14–29. [PubMed: 24925058]
17. Fercher AF, Drexler W, Hitzenberger CK, Lasser T. Optical coherence tomography - Principles and applications. Reports Prog Phys. 2003;66:239–303.
18. Uttam S, Bista RK, Staton K, Alexandrov S, Choi S, Bakkenist CJ, et al. Investigation of depth-resolved nanoscale structural changes in regulated cell proliferation and chromatin decondensation. Biomed Opt Express. 2013;4:596–613. [PubMed: 23577294]
19. Cristianini N, Shawe-Taylor J. An introduction to Support Vector Machines. Cambridge Univ. Press 2000.
20. Vapnik VN. Statistical Learning Theory. Wiley-Interscience 1998.
21. Pai-Hsuen C, Chih-Jen L, Bernhard S. A tutorial on ν-support vector machines. Appl Stoch Model Bus Ind. 21:111–36.
22. Macskassy SA, Provost F, Rosset S. ROC Confidence Bands: An Empirical Evaluation. Proc 22Nd Int Conf Mach Learn New York, NY, USA: ACM; 2005 page 537–44.
23. Beaugerie L, Itzkowitz SH. Cancers complicating inflammatory bowel disease. N Engl J Med. 2015;372:1441–52. [PubMed: 25853748]
24. Slaughter DP, Southwick H, Smejkal W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. Cancer. 1953;6:963–968. [PubMed: 13094644]
25. Curtius K, Wright NA, Graham TA. An evolutionary perspective on field cancerization. Nat Rev Cancer. 2018;18:19–32. [PubMed: 29217838]
26. Braakhuis BJ, Tabor MP, Kummer JA, Leemans CR, Brakenhoff RH. A genetic explanation of slaughter's concept of field cancerization: Evidence and clinical implications. Cancer Res. 2003;63:1727–30. [PubMed: 12702551]
27. Uttam S, Bista RK, Hartman DJ, Brand RE, Liu Y. Correction of stain variations in nuclear refractive index of clinical histology specimens. J Biomed Opt. 2011;16:116013. [PubMed: 22112118]
28. Popescu G Quantitative phase imaging of cells and tissues. McGraw-Hill; 2011.
29. Cotte Y, Toy F, Jourdain P, Pavillon N, Boss D, Magistretti P, et al. Marker-free phase nanoscopy. Nat Photonics. 2013;7:113–7.
30. Itzkan I, Qiu L, Fang H, Zaman MM, Vitkin E, Ghiran IC, et al. Confocal light absorption and scattering spectroscopic microscopy monitors organelles in live cells with no exogenous labels. Proc Natl Acad Sci. 2007;104:17255–60. [PubMed: 17956980]
31. Subramanian H, Pradhan P, Liu Y, Capoglu IR, Li X, Rogers JD, et al. Optical methodology for detecting histologically unapparent nanoscale consequences of genetic alterations in biological cells. Proc Natl Acad Sci. 2008;105:20118–23. [PubMed: 19073935]
32. Gladstein S, Damania D, Almassalha LM, Smith LT, Gupta V, Subramanian H, et al. Correlating colorectal cancer risk with field carcinogenesis progression using partial wave spectroscopic microscopy. Cancer Med. 2018;7:2109–20. [PubMed: 29573208]
33. Joo C, Akkin T, Cense B, Park BH, de Boer JF. Spectral-domain optical coherence phase microscopy for quantitative phase-contrast imaging. Opt Lett. 2005;30:2131–3. [PubMed: 16127933]
34. Choma MA, Ellerbee AK, Yang C, Creazzo TL, Izatt JA. Spectral-domain phase microscopy. Opt Lett. 2005;30:1162. [PubMed: 15945141]

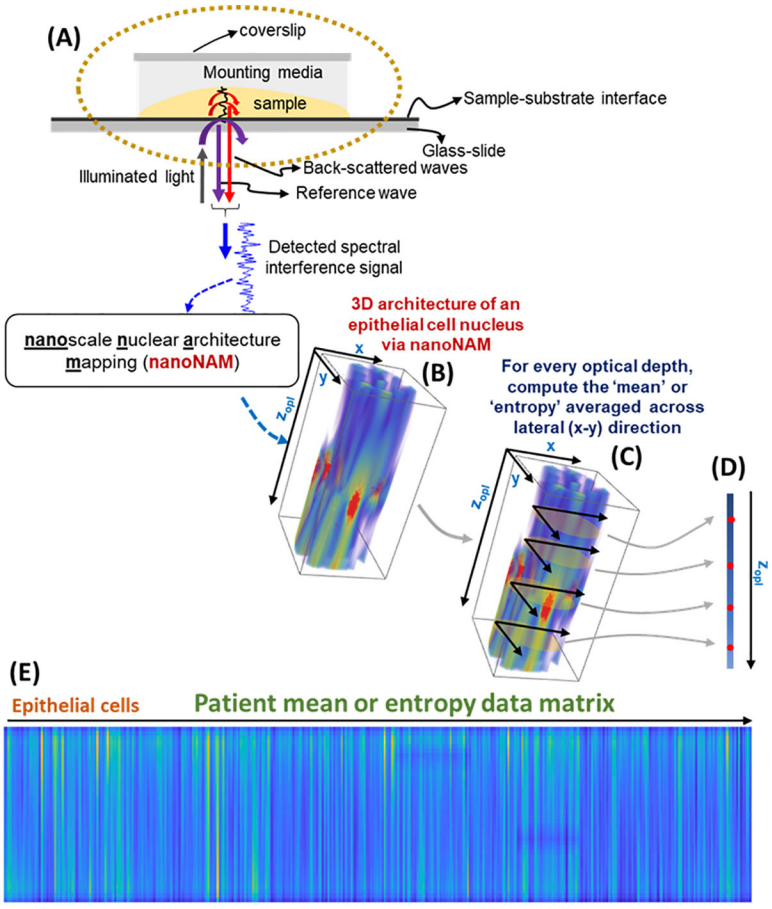



**Figure 1.**

Nanoscale nuclear architecture mapping (nanoNAM) principle. (**A**) Spectral interference between the back-scattered waves from within the sample (cell nucleus) and the reference wave from the sample and glass-slide interface is captured on the imaging sensor. (**B**) Fourier phase of the Fourier-mapped spectral interference signal is used by nanoNAM to obtain 3D nuclear architecture that jointly captures depth-resolved nanoscale alterations in the average density of the cell nuclei, and the heterogeneity of these alterations. (**C**) This 3D nuclear architecture is sampled every 45 nm along its depth, and the mean and entropy properties of these nuclear architectural slices are computed to obtain (**D**) depth-dependent mean and entropy vectors. (**E**) These vectors are computed for all patient cell nuclei and arranged column-wise to obtain mean and entropy data matrices respectively.

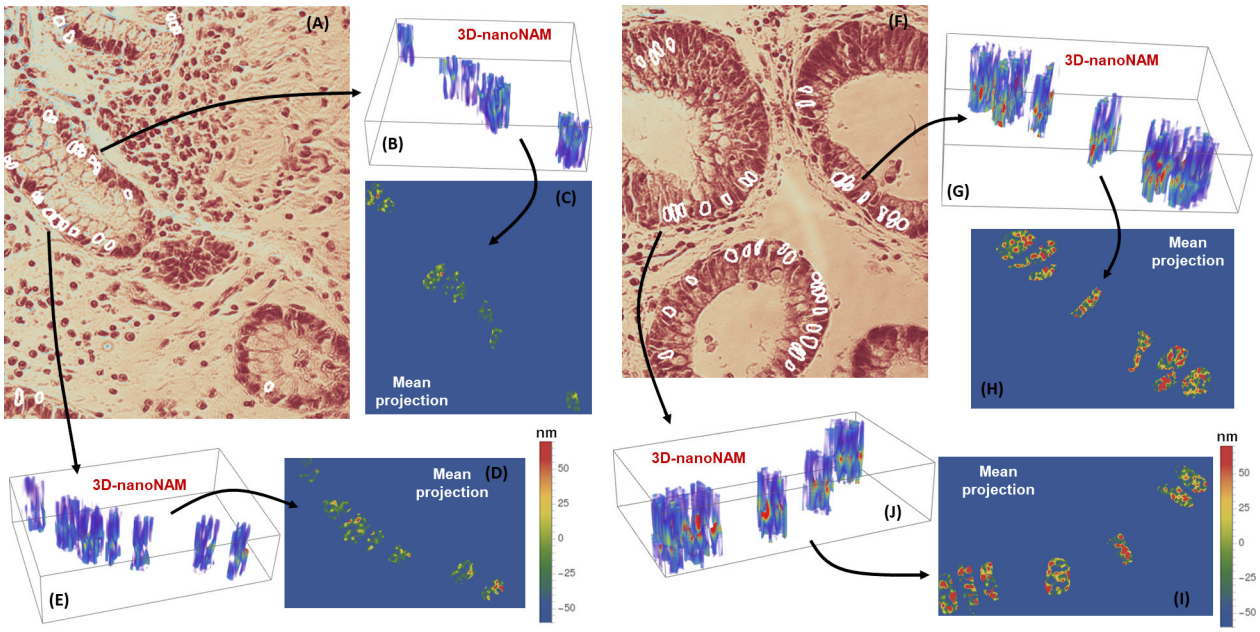
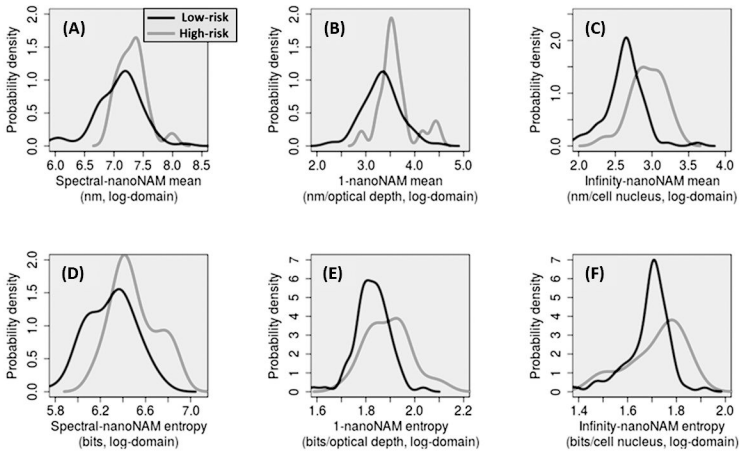



**Figure 2.**

nanoNAM-based visualization of 3D nuclear architecture of histologically normal appearing epithelial cell nuclei from rectal biopsies of (**A-E**) low-risk and (**F-J**) high-risk IBD colitis patients. (**A and F**) H&E stained tissue sample with segmented cell-nuclei of (**A**) low-risk and (**F**) high-risk patients. (**B and E**) nanoNAM-based 3D nuclear architecture of the unstained cell nuclei corresponding to the segmented nuclei in the stained tissue sample from low-risk patients. (**G and J**) depict the same for high-risk patients. In all cases, nanoNAM was performed on unstained tissue samples co-registered with the stained samples. (**C and D**) Mean projection of 3D nanoNAM-based nuclear architecture of cell nuclei form low-risk patients along the axial depth. (**H and I**) show the mean projection for high-risk patients. Although at microscopic scale the cell nuclei from both low- and high-risk patients appear histologically normal, nanoNAM-based sub-microscopic characterization shows increased alterations in mean nuclear density and associated heterogeneity in high-risk patients (redder values) compared to low-risk patients.




(G)

| | p-value | | effect size | | power | |
|---|---|---|---|---|---|---|
| | Mean | Entropy | Mean | Entropy | Mean | Entropy |
| **Spectral-nanoNAM** | 0.0010624 | 1.05e-05 | 0.573756 | 0.7958753 | 0.6513359 | 0.9028961 |
| **1-nanoNAM** | 0.0036342 | 0.0044166 | 0.7155077 | 0.8494736 | 0.8337823 | 0.9354492 |
| **infinity-nanoNAM** | 3.2e-06 | 0.1981355 | 1.1081969 | 0.3642601 | 0.9950144 | 0.3199077 |

**Figure 3.**

Probability distribution of six nanoNAM properties characterizing the intrinsic nuclear architecture of IBD colitis patients. For ease of visualization the properties have been depicted in the log domain. (**A**) Spectral-nanoNAM mean, (**B**) 1-nanoNAM mean, (**C**) infinity-nanoNAM mean, (**D**) spectral-nanoNAM entropy, (**E**) 1-nanoNAM entropy, and (**F**) infinity-nanoNAM entropy. All six properties show a rightward-shift in their values for high-risk patients when compared to low-risk patients. (**G**) *p-values* demonstrating the significance of the shift at the 0.05 significance level. *Effect size* capturing the size of the shift in the each of the six properties for high-risk patient group with respect to the low-risk group. *Power* of the each of the individual properties to correctly identify high-risk patients, when the patient is truly at risk for the prospective patient cohort recruited for this study.

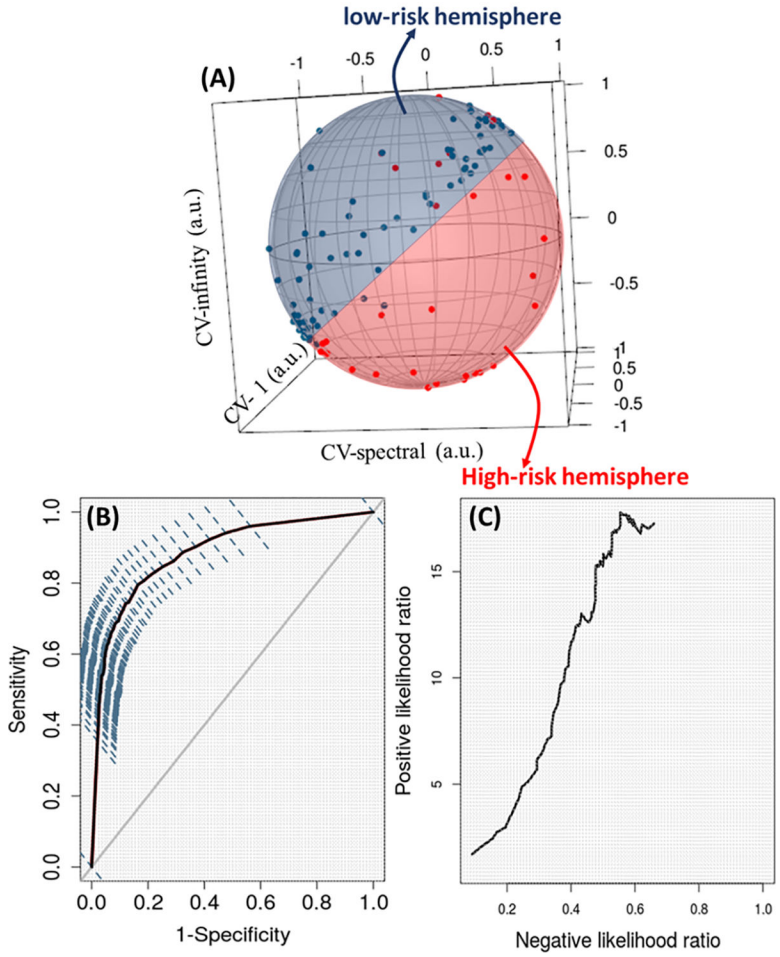



**Figure 4.**

Performance of nanoNAM-based prediction of cancer risk and associated risk spaces. (**A**) Projection of CV-spectral, CV-1 and CV-infinity meta nuclear architectural properties onto a unit 2-sphere. The low- and high-risk patients lie on two separate hemispheres of the 2-sphere. (**B**) Receiver operator characteristics (ROC) of nu-SVM based risk classification on the independent validation sets generated using stratified bootstrapping, with six nanoNAM-based nuclear architectural properties being the input features. The 95% confidence interval of the ROC is also shown. (**C**) Negative and positive likelihood ratios of the risk-classifier depicting its ability to assess the respective odds of a patient being truly at low- or high-risk of developing cancer.

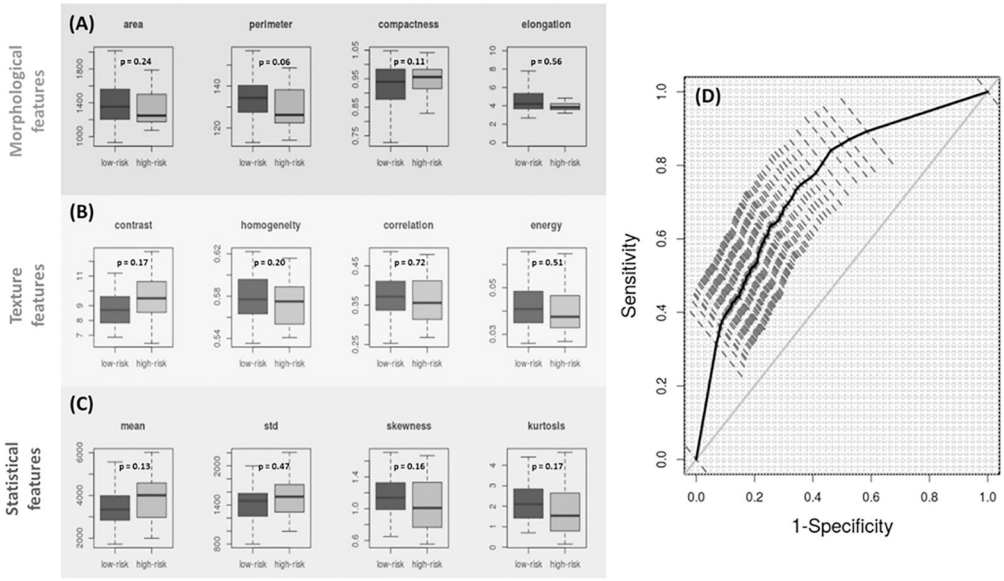



**Figure 5.**

Boxplots of (**A**) Morphological (area perimeter, compactness, elongation), (**B**) texture (contrast, homogeneity, correlation and energy) and (**C**) statistical (mean, standard deviation, skewness and kurtosis) microscopic features of histologically normal appearing cell nuclei from rectal biopsies of low-risk (dark gray) and high-risk (light gray) IBD colitis patients. (p-values have been indicated.) (**D**) Receiver operator characteristics (ROC) of nu-SVM based risk-classification based on the same independent validation sets used in Fig. 4, with the morphological, texture and statistical features forming the input features.

**Table 1.**

**Patient information.**

A total of 103 IBD patients who underwent surveillance colonoscopy were recruited, of whom 20 were identified as high risk while 83 were low risk. Of the 20 high risk patients, 6 had colonic biopsies indefinite for dysplasia, 2 had sessile serrated adenomas, 10 had tubular adenomas, while 2 harbored high-grade dysplasia. The table compares the clinical and endoscopic outcomes of low- and high-risk IBD patients along with their demographics.

| | High Risk (n=20) | Low Risk (n=83) | P-value |
|---|---|---|---|
| Mean age (yrs) ± SD | 52.5 ± 14.6 | 50.3 ± 14.4 | 0.557 |
| Gender (Male:Female) | 10:10 | 48:35 | 0.526 |
| Disease duration (yrs) ± SD | 15.6 ± 14.8 | 11.5 ± 7.5 | 0.252 |
| Ethnicity (White:AA:Other) | 19:0:1 | 78:3:1 | 0.166 |
| Smoking (Current:Never:Prior use) | 0:16:4 | 5:55:22 | 0.387 |
| PSC (Yes:No) | 3:17 | 3:80 | 0.051 |
| Pseudopolyps (Yes:No) | 1:19 | 15:68 | 0.147 |
| Prior history of dysplasia (Yes:No) | 8:12 | 15:68 | 0.035 |
| Family history of IBD (Yes:No) | 3:14 | 14:64 | 0.977 |
| Family history of colon cancer (Yes:No) | 3:14 | 10:68 | 0.600 |
| Future dysplasia (Yes:No) | 6 (HGD found in 2 patients during the 3-year study period; colorectal cancer found in the same two patients after the study was completed):11 | 8:63 | 0.015 |