RESEARCH ARTICLE

Radiomics on spatial-temporal manifolds via Fokker–Planck dynamics

Jack B. Stevens[1]　|　**Breylon A. Riley**[1]　|　**Jihyeon Je**[2]　|　**Yuan Gao**[3]　|
Chunhao Wang[1,4]　|　**Yvonne M. Mowery**[4,5,6]　|　**David M. Brizel**[4,5]　|　**Fang-Fang Yin**[1,4]　|
Jian-Guo Liu[7,8]　|　**Kyle J. Lafata**[1,2,4,9,10]

[1]Medical Physics Graduate Program, Duke University, Durham, North Carolina, USA

[2]Department of Electrical and Computer Engineering, Duke University Pratt School of Engineering, Durham, North Carolina, USA

[3]Department of Mathematics, Purdue University, West Lafayette, Indiana, USA

[4]Department of Radiation Oncology, Duke University School of Medicine, Durham, North Carolina, USA

[5]Department of Head and Neck Surgery & Communication Sciences, Duke University School of Medicine, Durham, North Carolina, USA

[6]Department of Radiation Oncology, UPMC Hillman Cancer Center/University of Pittsburgh, Pittsburgh, Pennsylvania, USA

[7]Department of Mathematics, Duke University, Durham, North Carolina, USA

[8]Department of Physics, Duke University, Durham, North Carolina, USA

[9]Department of Radiology, Duke University School of Medicine, Durham, North Carolina, USA

[10]Department of Pathology, Duke University School of Medicine, Durham, North Carolina, USA

Correspondence
Kyle J Lafata, Medical Physics Graduate Program, Duke University, 2424 Erwin Rd Suite 301, Durham, NC 27705, USA.
Email: kyle.lafata@duke.edu

Abstract

Background: Delta radiomics is a high-throughput computational technique used to describe quantitative changes in serial, time-series imaging by considering the relative change in radiomic features of images extracted at two distinct time points. Recent work has demonstrated a lack of prognostic signal of radiomic features extracted using this technique. We hypothesize that this lack of signal is due to the fundamental assumptions made when extracting features via delta radiomics, and that other methods should be investigated.

Purpose: The purpose of this work was to show a proof-of-concept of a new radiomics paradigm for sparse, time-series imaging data, where features are extracted from a spatial-temporal manifold modeling the time evolution between images, and to assess the prognostic value on patients with oropharyngeal cancer (OPC).

Methods: To accomplish this, we developed an algorithm to mathematically describe the relationship between two images acquired at time $t = 0$ and $t > 0$. These images serve as boundary conditions of a partial differential equation describing the transition from one image to the other. To solve this equation, we propagate the position and momentum of each voxel according to Fokker–Planck dynamics (i.e., a technique common in statistical mechanics). This transformation is driven by an underlying potential force uniquely determined by the equilibrium image. The solution generates a spatial-temporal manifold (3 spatial dimensions + time) from which we define dynamic radiomic features. First, our approach was numerically verified by stochastically sampling dynamic Gaussian processes of monotonically decreasing noise. The transformation from high to low noise was compared between our Fokker–Planck estimation and simulated ground-truth. To demonstrate feasibility and clinical impact, we applied our approach to ^{18}F-FDG-PET images to estimate early metabolic response of patients ($n = 57$) undergoing definitive (chemo)radiation for OPC. Images were acquired pre-treatment and 2-weeks intra-treatment (after 20 Gy). Dynamic radiomic features capturing changes in texture and morphology were then extracted. Patients were partitioned into two groups based on similar dynamic radiomic feature expression via k-means clustering and compared by Kaplan–Meier analyses with log-rank tests ($p < 0.05$). These results were compared to conventional delta radiomics to test the added value of our approach.

Results: Numerical results confirmed our technique can recover image noise characteristics given sparse input data as boundary conditions. Our technique was able to model tumor shrinkage and metabolic response. While no

delta radiomics features proved prognostic, Kaplan–Meier analyses identified nine significant dynamic radiomic features. The most significant feature was Gray-Level-Size-Zone-Matrix gray-level variance ($p = 0.011$), which demonstrated prognostic improvement over its corresponding delta radiomic feature ($p = 0.722$).

Conclusions: We developed, verified, and demonstrated the prognostic value of a novel, physics-based radiomics approach over conventional delta radiomics via data assimilation of quantitative imaging and differential equations.

1 | INTRODUCTION

Radiomics is a high-throughput computational technique that enables the extraction and analysis of quantitative features from radiological images.[1–3] Radiomic features provide measures of different morphological, topological, and/or textural characteristics of an image, which can lead to new diagnostic and/or prognostic value.[4,5] In particular, delta radiomics aim to capture quantitative changes in serial, time-series image representation. This technique is based on calculating the relative change in radiomic features between two acquisition time points.[6] Delta radiomic features have shown improved prognostic value compared to radiomic features of a single time point.[7,8]

There are numerous applications of delta radiomics throughout the literature.[8–14] However, despite the potential advantages of incorporating multiple images, an often-understudied aspect of delta radiomics is accurate and reliable time interval analysis. Time interval analysis is the characterization of quantitative changes in image metrology over time.

Clinical image acquisition is temporally sparse, which results in disease characterization based on radiomic feature differences between a limited number of images. For example, delta radiomics calculated via two images acquired before and after treatment assumes a linear relationship between radiomic feature change and treatment response. This is potentially an oversimplification of the problem that may not be valid for complex diseases such as cancer, where tumors likely exhibit non-linear time dynamics.[15]

To address this limitation, we propose a novel approach to delta radiomics based on the assimilation of stochastic differential equations and quantitative radiomic analysis. Our approach is based on the Fokker–Planck[16–18] equation, that is, a partial differential equation describing the probability distribution of a stochastic process. Essentially, we first use Fokker–Planck dynamics to model non-linear behavior between time separated images, from which radiomic features are calculated from a pseudo-continuous function. In general, Fokker–Planck dynamics mathematically describe countless *many-body* problems in physics (e.g., the diffusion of pollutants through the atmosphere,[19] electron transport in semiconductors,[20] calcium absorption in bones,[21] and the random walk of stars and black holes due to the gravitational force of nearby stellar bodies[22]). Thus, we consider Fokker–Planck a reasonable mathematical framework for conducting time interval analysis between images to drive delta radiomics.

In this paper, we first provide a theoretical framework and numerical validation of our method. We then apply our method to a characterize the early metabolic response of patients undergoing definition radiation therapy for oropharyngeal head and neck cancer, where dynamic radiomic features on PET imaging are calculated via Fokker–Planck dynamics and compared to their classical delta radiomic analogues.

2 | METHODS

2.1 | Theory

2.1.1 | Equilibrium-driven deformation via a Fokker–Planck algorithm

We developed an equilibrium-driven deformation algorithm (EDDA) to simulate the temporal evolution between two tomographic images, $\rho_o(x, y, z) \in \mathbb{R}^3$ and $\rho_t(x, y, z) \in \mathbb{R}^3$, acquired at $t = 0$ and $t > 0$, respectively. Given two images, one initial (ρ_o) and one final (ρ_t), the process of generating motions (i.e., extrapolated intermediate images) between them is known as "inbetweening auto-animation[23]". Instead of using complicated kinematic equations to evolve each object within the starting image, this work used an algorithm based on Fokker–Planck dynamics known as equilibrium-driven deformation.[24,25] The fundamental principle behind this algorithm is the definition of a potential function that drives the time dynamics between ρ_o and ρ_t. This potential function is uniquely defined by ρ_t, which acts as the equilibrium state under Fokker–Planck dynamics. By using Fokker–Planck dynamics as an approach to time interval analysis, it may capture the complex and
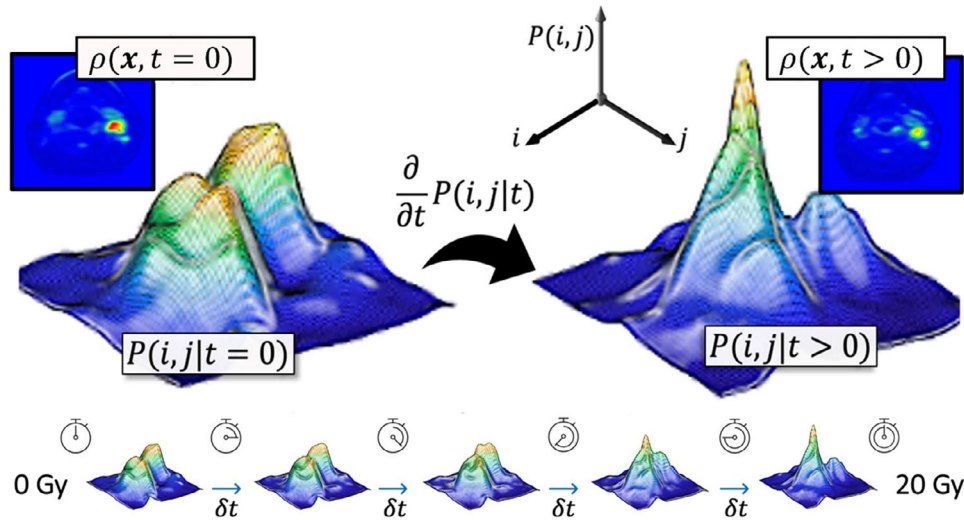
FIGURE 1 Illustration of the Fokker–Planck image-inbetweening algorithm. The functions $\rho(\boldsymbol{x}, t = 0)$ and $\rho(\boldsymbol{x}, t > 0)$ represent the initial and equilibrium boundary conditions, respectively. $P(i,j|t)$ is the corresponding probability density function. The algorithm generates inbetween states between 0 and 20 Gy shown by the lower panel.

often poorly described dynamics associated with the biological processes of tumor treatment response.

First, let $\Omega \subset \mathbb{R}^3$ be a closed subset on $\mathbb{R}^3$. We define the initial and equilibrium images on Ω as ρ_0 and ρ_t, respectively. We then assume that the final image on Ω is described by the equilibrium density function $\pi(x) : \Omega \to R$. Then, the time-evolution of the gray-level value of the image, ρ, is described by the Fokker–Planck equation,

$$\partial_t \rho = \Delta \rho + \nabla \cdot (\rho \nabla \phi) = \nabla \cdot \left(\pi \nabla \left(\frac{\rho}{\pi} \right) \right) \quad (1)$$

where ϕ represents a given energy landscape. One can show by direct computation that the equilibrium is given by $\pi \propto e^{-\phi}$. This is analogous to the concept of the Gibbs measure in statistical mechanics.[26] The initial data, ρ_0, satisfies,

$$\int \rho_0 \, dx = \int \pi \, dx. \quad (2)$$

In this context, the images ρ_o and ρ_t are interpreted as no-flux boundary conditions of Equation (1), that is, they obey

$$n \cdot \nabla \left(\frac{\rho}{\pi} \right) = 0 \text{ on } \partial\Omega, \quad (3)$$

where n is the normal vector to the boundary surface $\partial\Omega$. Physically, ρ represents gray-level intensities (pixel values) of the equilibrium image ρ_t. The evolution between ρ_0 and ρ_t is driven by a unique underlying potential force defined by the equilibrium (Figure 1). For tomographic images on $\mathbb{R}^3$, Equation (1) generates a set of 3D matrices which contain pixel values at each spatial

location. Each matrix corresponds to a given time point between the initial and equilibrium image. Therefore, we defined the spatial-temporal manifold as the 4D array that describes the time evolution of a 3D spatial volume according to Equation (1). Equation (1) essentially projects image data on $\mathbb{R}^3$ to time-series data, defined as a tensor object on $\mathbb{R}^4$. Mathematically, we define this manifold Ψ as the following tensor object,

$$\Psi \subset \mathbb{R}^4 = \left\{ (x, y, z, t) \, | x, y, z \in \mathbb{R}, t \in \mathbb{R}^+ \right\}. \quad (4)$$

To numerically solve Equation (1), we apply an EDDA using a finite-volume method on structured grids, as implemented by Gao et al.[23] Finite-volume discretization methods can be applied to many conservation laws, with a particular advantage being local conservativity of fluxes.[27] This is useful in problems such as image-inbetweening, where fluxes are crucial in defining the boundary conditions.

2.1.2 | Dynamic radiomic feature extraction

As described in Section 2.1.1, Equation (1) produced a 4-dimensional spatial-temporal manifold. This manifold was then mapped back into image space, which produced a tensor object representing the time evolution between the initial image and the equilibrium image. High-throughput radiomics features were extracted using an in-house pipeline,[28] validated both through standard Image Biomarker Standardization Initiative (IBSI) benchmarking[4] and externally, using digital bar phantoms.[29] This feature extraction generated a dynamic radiomic feature space, capturing the time
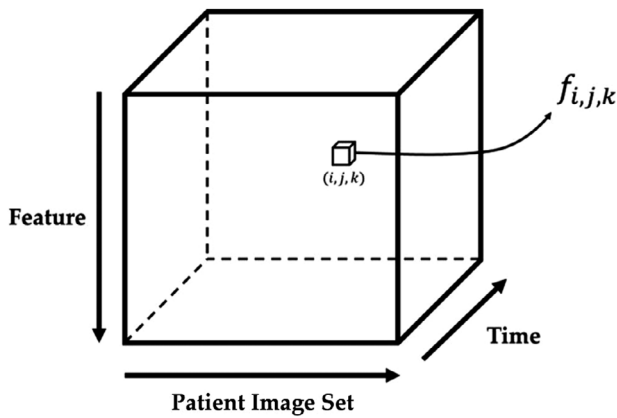
FIGURE 2 Visualization of the dynamic radiomic feature space extracted from the spatial-temporal manifold generated via Fokker–Planck dynamics.

evolution of the radiomic features. Mathematically, the feature space can be represented by a tensor object, $\mathcal{F}$, given a set of p features for q image sets (i.e., pairs of initial and equilibrium images) across r time points,

$$\mathcal{F} = \left(f_{i,j,k}\right) \in \mathbb{R}^{p \times q \times r} \tag{5}$$

where, the coordinates (i, j, k) represent the ith radiomic feature observed for the jth image set at the kth time step. As an illustrating example, if each image set corresponded to a given patient, Figure 2 shows the structure of the resulting dynamic feature space.

This feature space encodes variation of radiomic features across different patients and across time. By construction, this feature space allows for analysis along all three dimensions, that is, features for all patients at a given time, variations of single features across time for all patients etc.

2.1.3 | Numerical validation on 2D random Gaussian processes

To experimentally verify the technique outlined in Section 2.1.1, the algorithm was used to model monotonically decreasing noise generated by randomly sampling 2D Gaussian processes. By stochastically sampling 2D Gaussian processes of different full-width-half-max values, images of various noise levels were generated, which served as the ground-truth for the experiment. In this experiment we consider "noise evolution", that is, the transition between images of high noise value to images of low noise value, rather than "time evolution".

All steps of the evolution were zero-mean centered Gaussian noise acting on the same 128×128 random matrix of pixel intensities ranging from 0 to 1. Each step had a non-linear narrowing of the noise distribution full width at half maximum (FWHM) ranging from 0.8 to 0.0003 spanning 14 measurements.

The Fokker–Planck method (Equation (1)) was then applied, using only the images with maximum and minimum noise as the initial and equilibrium boundary conditions, respectively, according to Equation (3). To compare the Fokker–Planck technique to the experimental ground-truth, image energy and entropy were calculated and plotted against noise level. These features were calculated via the image intensity histograms and the standard IBSI definitions,[30]

$$E = \sum_{i=1}^{N_g} p_i^2 \tag{6}$$

$$S = -\sum_{i=1}^{N_g} p_i \log_2 p_i, \tag{7}$$

where N_g is the number of discretized intensity bins and p_i is ratio of counts in the ith intensity bin to the total number of pixels in the image.

2.2 | Proof-of-concept application of Fokker–Planck dynamics to patients treated for oropharyngeal cancer

Figure 3 shows the overall workflow of our patient study, as described in the following section.

2.2.1 | Clinical trial design and PET image acquisition

In this work, patient data was acquired from a single-institution, prospective clinical study (NCT01908504) conducted at Duke University Medical Center. Patients underwent curative intensity modulated radiation therapy (IMRT) for a positive diagnosis of oropharyngeal cancer (OPC). Radiation treatment was delivered via two distinct fractionation schemes: a total dose of 70 Gy in 35 fractions, and a total dose of 67.5 Gy in 30 fractions. Most patients received concurrent chemotherapy according to well-established clinical procedures (modified bolus cisplatin at 20 mg/m^2 over Days 1−5 and 29−33, weekly cisplatin, or weekly docetaxel). For the purposes of this analysis, exclusion criteria were as follows: (i) a diagnosis of p16-positive head and neck squamous cell carcinoma with unknown primary; (ii) any prior surgical removal/resection of the primary tumor; (iii) multiple synchronous primary tumors; and (iv) any neoadjuvant chemotherapy before definitive radiation treatment.

At the conclusion of radiation treatment, patients were examined every 2−3 months via fiberoptic laryngoscopy, diagnostic imaging (PET/CT and additional imaging as necessary) and biopsy of recurrence-indicating lesions. Recurrence-free survival (RFS) was used as the main
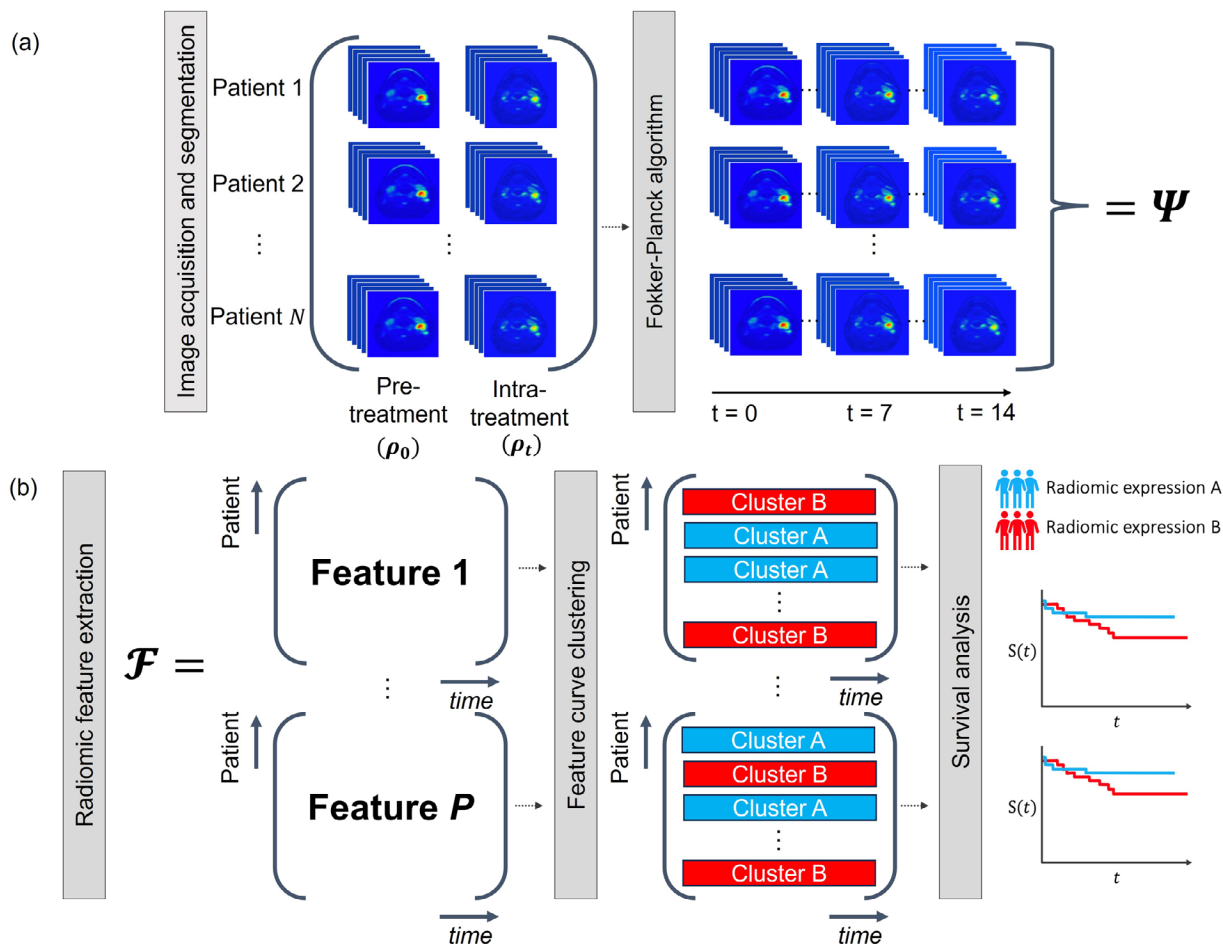
FIGURE 3 Retrospective patient study workflow. (a) Pre- and intra-treatment PET/CT images are acquired for N patients and the gross tumor volume (GTV) is manually segmented. Images serve as boundary conditions to the Fokker–Planck equation (ρ_o and ρ_t) and are input into the algorithm, generating simulated images between the boundary conditions (i.e., the spatial-temporal manifold Ψ defined in Equation (4). (b) Radiomic features are extracted from the time series data, generating the feature space $\mathcal{F}$, defined in Equation (5). Patient-specific feature trajectories are clustered using a k-means algorithm into 2 groups and tested for prognostic significance, according to Equation (9).

endpoint for the study and defined as the time between conclusion of radiation treatment and detection of recurrent/residual local, regional, and/or distant disease. After the final follow-up patients were censored, and median follow-up time was computed via the reverse Kaplan–Meier approach.[31]

For each patient, two sets of ^{18}F-FDG-PET/CT images were taken. The first set was obtained before patients underwent radiation treatment, and the second set was obtained after 2 weeks (20 Gy) of treatment. The same scanner (Siemens Biograph mCT PET/CT, Siemens Medical Solutions, Knoxville, TN) was used for each patient. Imaging protocols were standardized across the patient according to the prospective clinical trial design. Depending on the weight of the patient, 8−15 mCi of FDG activity was injected following 4 h of fasting, with the time between injection and imaging kept constant across pre-treatment and intra-treatment imaging for each patient. For the PET imaging protocol, a 54 cm *field-of-view* with a 400 × 400 matrix

size and 2 mm slice thickness. For the CT imaging protocol, an extended 65 cm *field-of-view* was used, with a 512 × 512 matrix size and 3 mm slice thickness. PET images were reconstructed via ordered subset expectation maximization (OSEM) with time-of-flight (TOF) correction and attenuation correction (from CT). CT images were reconstructed using filtered back-projection.

Following image acquisition, a radiation oncologist manually delineated the gross tumor volumes at the primary disease location (GTVp). This was done on the CT images and then transferred to the PET images. Images were re-sampled to an isotropic resolution of 1.17 mm via tri-cubic spline interpolation and subsequently re-binned to a dynamic range of 64 gray levels, as recommended by the IBSI. Image registration and verification was done following a method established by prior work.[32] Intra-treatment PET/CT images were registered to pre-treatment images first by rigid bony structure alignment, followed by local soft tissue
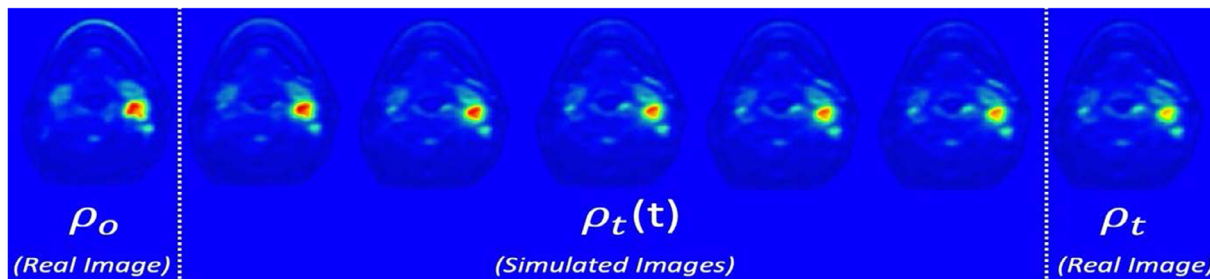
FIGURE 4 Inbetween PET/CT images of a given patient generated by Fokker–Planck evolution from two clinical images, ρ_0 and ρ_t.

adjustment using the deformable multi-pass registration algorithm within the Velocity software (Varian Medical Systems, Palo Alto, CA, United States).

2.2.2 | Construction of a spatial-temporal manifold

The pre-treatment and intra-treatment PET/CT images defined the initial and equilibrium conditions for Equation (1), which obey the boundary conditions prescribed by Equation (3). As an illustrating example, consider Figure 4, which shows a single 2D slice of a PET/CT image volume (for a representative patient) at different timepoints across the 2-week interval, with the real and simulated images indicated.

PET/CT imaging data is generally sparse and shows mostly uniform uptake with small regions of much higher uptake, which define the tumor. Therefore, a minimum bounding box around the initial gross tumor volume was defined using contours drawn by physicians. These minimum bounding boxes define the initial and equilibrium images on which we applied the Fokker–Planck method.

To evaluate convergence of the Fokker–Planck solution to the equilibrium image, the following maximum error metric was used,

$$M(t) = \max\{|\rho_t(i,j,k,t) - \rho_t(i,j,k,nt)|\} \quad (8)$$

where, $nt = 10000$ is the final timestep.

Fokker–Planck dynamics follow a non-linear time evolution determined by Equation (1). The differences between an image at time t and the equilibrium image (i.e., time nt) decrease very quickly at the start of the evolution and then gently converge at larger timesteps. To capture this non-linearity, timestep values were sampled at equal intervals along the y-axis, which gave a scalar vector of timesteps representing equivalent change in maximum error. By sampling the original spatial-temporal manifold at these calcu-

lated timesteps, a new four-dimensional tensor, $\overline{\rho_t}$ was defined.

Given the well-studied importance of image texture as a biomarker for metabolic heterogeneity and subsequent recurrence-free survival,[1] it was necessary to describe changes in the tumor shape and volume across the evolution. As there is no "gold-standard" method for thresholding tumor volume in PET/CT images, an absolute standardized uptake value (*SUV*) of 2.5 was used.[33] Given that the transformation into Fokker–Planck space (i.e. probability density functions) does not preserve the physical meaning of pixel intensity values, an absolute *SUV* threshold of 2.5 was scaled by the SUV_{max} of the initial image, thereby defining a tumor-specific relative threshold.[33] This threshold allowed for generation of binary masks at each timestep and therefore calculation of intensity masks and subsequent extraction of radiomic features, as introduced in Section 2.1.2.

2.2.3 | Dynamic radiomic feature analysis

For each patient included in this clinical application, we applied Equation (1) to generate a fourth-order tensor representing the time-evolution of a spatial volume surrounding the primary tumor bed. This formed the (patient-specific) spatial-temporal manifold from which we extracted dynamic radiomic features.

We performed feature extraction via the method outlined in Section 2.1.2. For each dynamic radiomic feature, we partitioned the curves (representing patient-specific time evolution of a given feature) using the following k-means clustering algorithm: Given a dynamic radiomic feature $\boldsymbol{f}$, we have a set of N feature vectors $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_N)$, where $N = 57$ is the number of patients. Therefore, each feature vector $\boldsymbol{f}_i$ represents the time evolution of feature $\boldsymbol{f}$ for the ith patient. To compare patients that responded to treatment with those that did not, the algorithm then partitions the set of feature vectors into $k = 2$ sets (or clusters), $\boldsymbol{S} = \{\text{Group 1, Group 2}\}$. This is accomplished by minimizing the within-cluster sum of squares (WCSS), which is
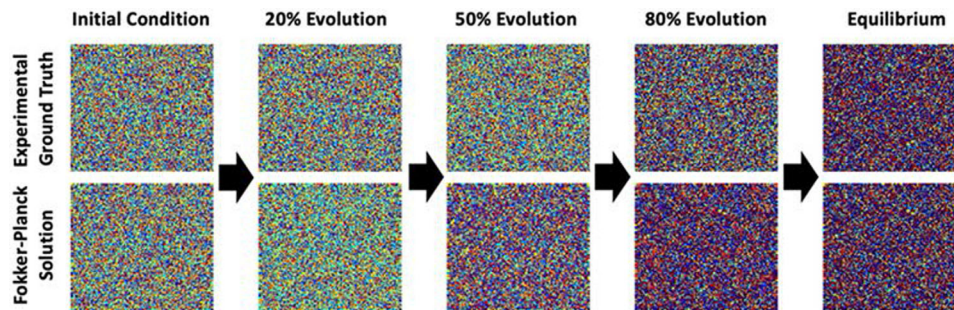
FIGURE 5 Noisy images generated via random sampling of 2D Gaussian processes. Top: ground truth images. Bottom: images estimated via the Fokker–Planck algorithm.
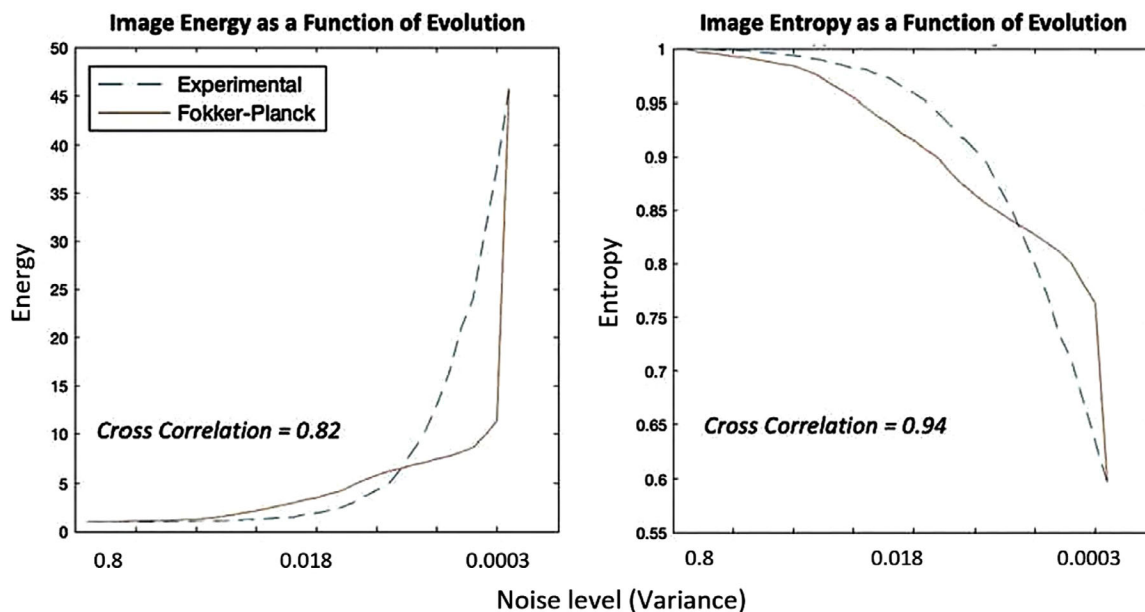

FIGURE 6 Energy and entropy versus noise level (variance) of images with monotonically decreasing noise generated by random 2D Gaussian processes.

equivalent to the variance. Hence, the objective is to find

$$\arg\min_{S} \sum_{i=1}^{k} \sum_{\boldsymbol{f} \in S_i} ||\boldsymbol{f} - \boldsymbol{\mu_i}||^2, \qquad (9)$$

where $\boldsymbol{\mu_i}$ is the mean of all points in cluster i, S_i denotes cluster i, and $||.||$ is the standard L^2 norm. This algorithm was used to identify those patients with intrinsically similar feature curves.

To compare with traditional delta radiomics, we calculated another feature space by taking the difference between the final and initial feature values and patients were partitioned based on their median feature value. Kaplan–Meier analyses were performed to test the prognostic value of the radiomic encoding. Log-rank tests[34] were used to test for differences in patient partitions, where a p-value less than or equal to 0.05 was considered statistically significant. This analysis was constructed in MATLAB (Math-

Works, Natick, MA) using the *MatSurv* package.[35] To evaluate the potential added value of our technique, dynamic radiomic features computed via Fokker–Planck were compared to their classical delta radiomic counterparts.

3 | RESULTS

3.1 | Numerical validation

Numerical results confirmed that the Fokker–Planck method (introduced in Section 2.1.1) can recover image noise characteristics given sparse input data as boundary conditions. Figure 5 depicts the noise evolution for the experimental ground truth images and images generated via Fokker–Planck.

Figure 6 shows plots of image energy and entropy as noise level, that is, variance of the Gaussian distribution, is decreased.

As expected, image energy increases as noise level decreases, as the image becomes more homogeneous. The converse is observed for image entropy, again as expected. The calculated cross-correlations of 0.82 and 0.94 for energy and entropy, respectively, demonstrate that the image series generated by the Fokker–Planck method was consistent with the experimental ground truth.

3.2 | Patient study

3.2.1 | Patient characteristics

Relevant characteristics of the patient dataset are summarized in Table 1.

Median follow up time was 40.9 months. 82% ($n = 47$) of patients were HPV-positive or p16-positive. Tumor staging was performed via the AJCC 8th edition.

3.2.2 | Fokker–Planck trajectories

Implementation of the Fokker–Planck method via Equation (1) generated a 4D tensor, $\overline{\rho_t}$, which consisted of simulated 3D images sampled from the manifold. To visualise the structure of the data, consider Figure 7, which shows an illustrating example of a 2D slice of the 3D time evolution for the primary GTV images of a single patient. A qualitative inspection of the image series reveals reduction in FDG uptake as a function of time and accumulation of therapeutic dose, as expected. This reduction is highly non-linear across both spatial and temporal dimensions, and complex textural and morphological changes are observed across the 2-week treatment period. By implementing the thresholding procedure, a volumetric representation of tumor shrinkage as a response to treatment can be generated. An example of this is shown in Figure 8 and demonstrates the expected shrinkage of the primary tumor volume across the treatment period.

3.2.3 | Association of dynamic radiomics with treatment response

Beginning with simple, first order image features, normalized image energy and entropy changes across the treatment period of a single representative patient are shown in Figure 9.

Kaplan–Meier analysis identified nine prognostic dynamic radiomic features where the corresponding delta radiomic feature was not prognostic. Summary statistics for these features are shown in Table 2.

Figure 10 shows Kaplan–Meier survival curves for a representative radiomic feature, in this case Gray Level Size Zone Matrix (GLSZM) gray-level variance.

TABLE 1 Characteristics of the patient dataset used in this work.

Parameter	Total ($n = 57$)
Sex	
Male	46 (80.7%)
Female	11 (19.3%)
Patient age (years)	
Median (max—min)	59.3 (77.9 – 39.8)
Primary tumor (T)	
T0	1 (1.8%)
T1	12 (21.1%)
T2	22 (38.6%)
T3	9 (15.8%)
T4	4 (7%)
T4a	8 (14%)
T4b	1 (1.8%)
Regional lymph nodes (N)	
N0	5 (8.8%)
N1	3 (5.3%)
N2a	2 (3.5%)
N2b	30 (52.6%)
N2c	14 (24.6%)
N3	3 (5.3%)
STAGE	
I	1 (1.8%)
II	1 (1.8%)
III	3 (5.3%)
IVa	48 (84.2%)
IVb	4 (7%)
Chemotherapy	
Yes	53 (93%)
No	4 (7%)
Smoking status	
≤ 10 packs per year	34 (59.6%)
> 10 packs per year	23 (40.4%)
Recurrence/residual disease	
Yes	16 (28.1%)
No	41 (71.9%)

4 | DISCUSSION

Our algorithm was able to estimate the time evolution of PET images of patients with HNSCC throughout a 2-week treatment period. We observed non-linear changes in the texture of the images, which is to be expected given the non-linear changes in tumor size/shape during radiation treatment. Furthermore, there was apparent spatial heterogeneity in the evolution of metabolic uptake. Future work will implement a spatial clustering approach to interrogate potential
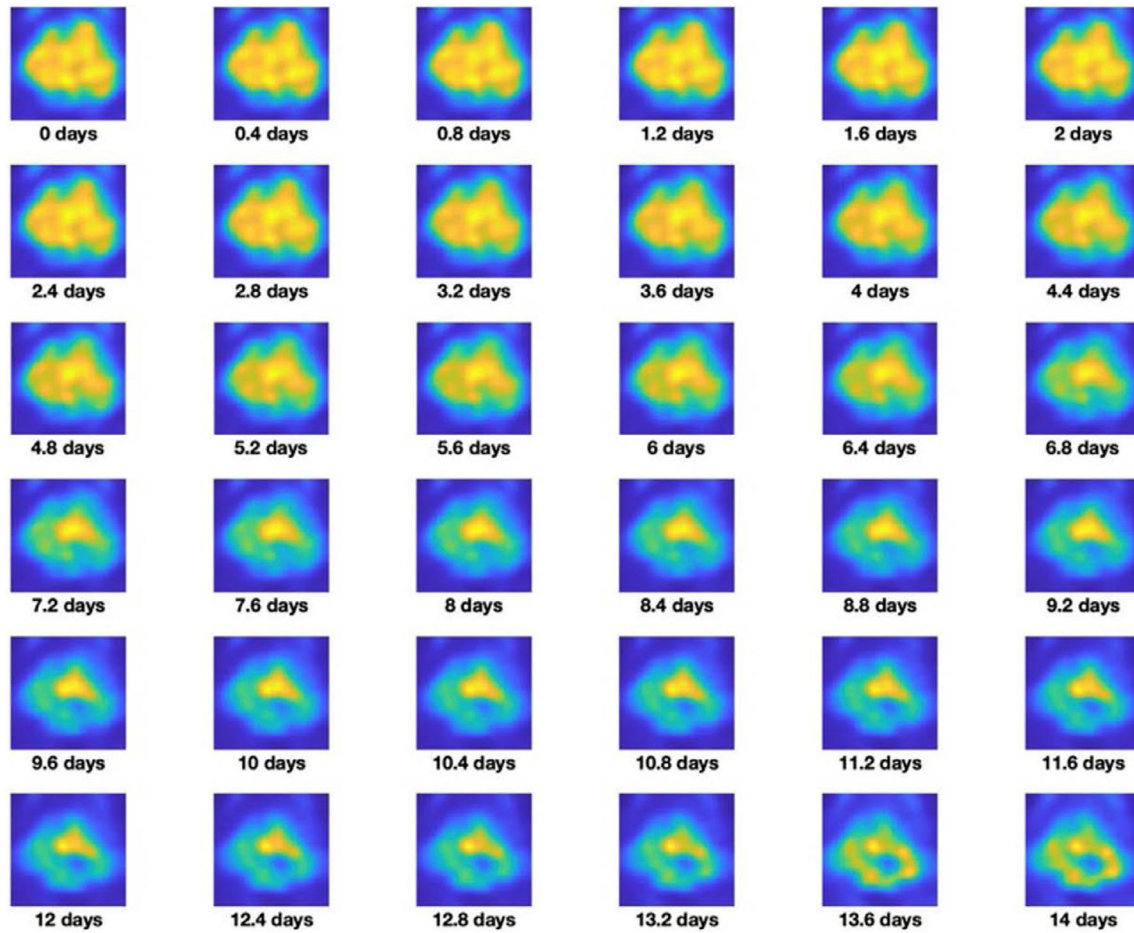
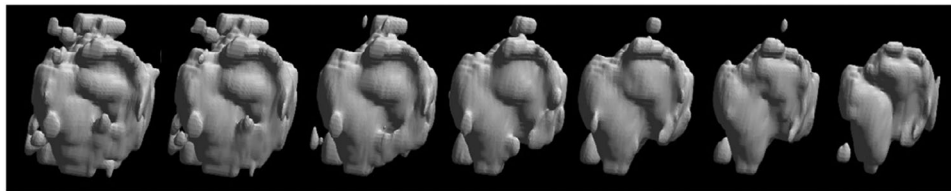**FIGURE 7** Temporal evolution for a single axial slice of a tumor volume across the 2-week treatment period.

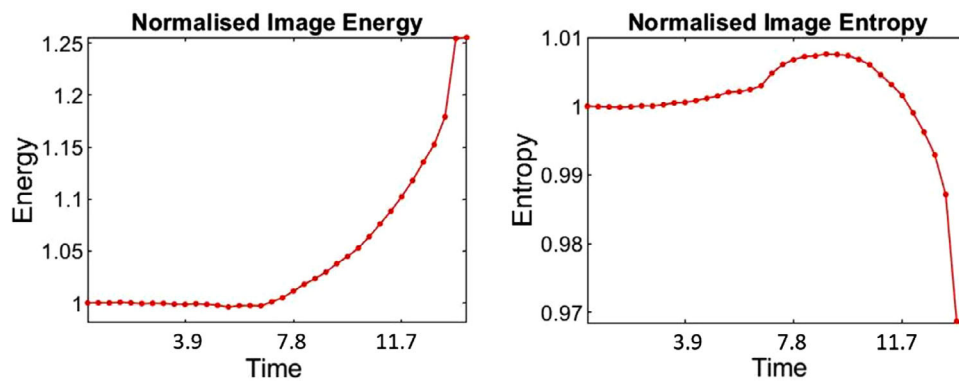**FIGURE 8** Volumetric visualization of tumor shrinkage in response to treatment.

FIGURE 9 Time evolution (in days) of image energy and entropy for a representative patient during the 2-week treatment period.

TABLE 2 Summary statistics for significant dynamic features showing lack of significance in delta radiomic features.

| Feature	p-value (dynamic)	p-value (delta)	Hazard ratio (dynamic)	Hazard ratio (delta)
GLRLM Gray-level non-uniformity	0.0479	0.0938	0.31	2.94
GLRLM Run-length non-uniformity	0.0479	0.103	0.31	2.87
GLRLM Low gray level run emphasis	0.0179	0.397	0.125	1.69
GLRLM Run-length variance	0.0429	0.152	0.314	2.55
GLSZM Gray-level variance	0.0113	0.722	0.174	0.806
MORPHOLOGY Volume	0.0479	0.0832	0.31	3.04
MORPHOLOGY Surface area	0.0162	0.377	0.26	1.73
MORPHOLOGY Compactness 1	0.0218	0.669	0.264	1.29
MORPHOLOGY Perimeter	0.0162	0.155	0.26	2.53
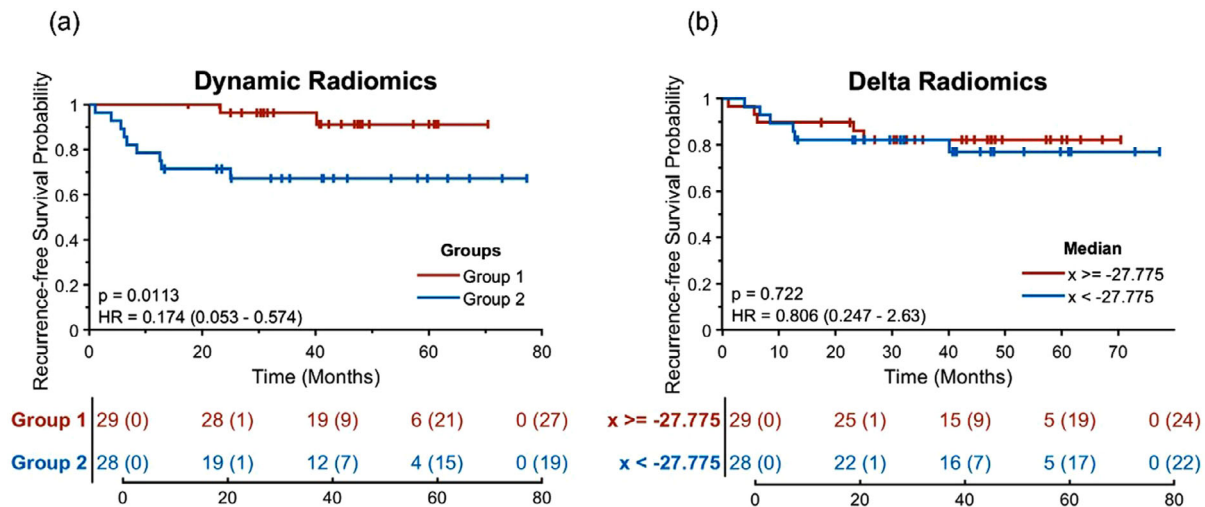
FIGURE 10 Kaplan–Meier curves for the radiomic feature x = GLSZM Gray-level variance. (a) Dynamic radiomics features are partitioned into two groups via k-means clustering of the patient-specific feature curves. This feature is prognostically significant ($p = 0.01$) and the curves show clear separation between group 1 and group 2. (b) Delta radiomics features are partitioned into two groups using the median feature value. This feature is not prognostically significant ($p = 0.72$) and shows no clear separation between groups. Risk tables are shown below each plot and the number of events (and censored patients) is displayed for a set of time points.

heterogeneity in the biological response of the tumor due to treatment.

We aim to map the regions identified by the clustering to specific areas in the tumor bounding box, thereby identifying regions of tumor shrinkage. We hypothesize that the formation of these spatial tumor habitats is potentially due to underlying biological phenomena, such as aerobic vs sub-hypoxic regions of disease and/or spatial differences in tumor radiosensitivity. Further work to characterize and interrogate these phenomena will provide further biological insight and advance our understanding of treatment response of oropharyngeal tumors.

The key finding of our study was the prognostic significance of dynamic radiomics over traditional delta radiomics. Specifically, we observed significance in dynamic texture and morphology features, but did

not see any prognostic benefit in the corresponding delta radiomic features. This supports our hypothesis that the application of delta radiomics in this dataset may remove useful radiomic signal and affect prognostic value of these features. This work has demonstrated that the application of delta radiomics should be more closely studied, and further techniques to boost radiomic signal should be investigated.

This is a valuable insight, given the prevalence of delta radiomics in the literature. It has been used as an analytical tool for a wide variety of applications and disease sites. Fave et al. investigated changes in, and prognostic value of, delta radiomics features extracted from CT images of patients undergoing treatment for non-small cell lung cancer (NSCLC). To a similar end, Delgadillo et al. examined applications of delta radiomics to prostate cancer radiotherapy, with an

emphasis on features extracted from multiple MRIs pre- and post-treatment.[36] Furthermore, Wang et al. applied a delta radiomics analysis to study acute normal tissue (pulmonary) toxicity following radiation treatment for oesophageal cancer. PET and CT-based radiomics is often used to quantify metabolic response of patients undergoing definitive radio/chemotherapy for head and neck squamous cell carcinoma. Applications include tumor segmentation,[37] predictive/prognostic studies,[38] and normal tissue response to radiation therapy.[39,40] Delta radiomics has also been applied outside the field of radiation therapy, including treatment of metastatic melanoma via immunotherapy[41,42] and differentiation of radiation-induced necrosis and cancer tissue in treatments using Gamma Knife radiosurgery.[43] Review articles by Nardone et al., Kothari et al., Spohn et al., and Wang, H. et al. provide exhaustive analyses on the various applications of delta radiomics.

Another key advantage of our technique is the breadth of analysis that can be performed on the radiomic feature space. By generating another dimension of input data, the feature space becomes a higher-dimensional object. Therefore, it becomes possible to analyze the data across additional dimensions and various methods can be used to reduce the dimensionality of the feature space such that it can be compared to traditional delta radiomics. In our case, we primarily implemented a k-means clustering technique, as we hypothesized that this would most effectively capture intrinsic differences between the feature curves. However, other methods could be used, many of which apply deep learning tools.[44] These methods would provide a valuable extension to the work done here.

A key hypothesis underpinning this work stated that the lack of prognostic significance of delta radiomics features (obtained from pre- and intra-treatment PET imaging) arises, at least in part, due to the coarse-grained nature of delta radiomics.[45] There are, however, other potential reasons for this lack of signal. One such reason is the lack of 3D radiation dosimetry data in the original study. One cannot expect to fully realize the behavior of the physical system without including such information, as the biological perturbation induced by radiation therapy is the most significant driver of changes between pre- and intra-treatment imaging data. For our data, dosimetry modeling was not a main concern, as patients were given a uniform dose under well-defined prescriptions outlined in the clinical trial. Furthermore, prior work has demonstrated that dosimetry is a shallow feature in this dataset, and that all necessary useful information comes directly from the PET/CT images.[46] However, if our algorithm were to be applied to other disease sites, this phenomenon would need to be accounted for, as we could no longer guarantee dosimetric homogeneity.

One strategy to incorporate radiation dose information is implementation of a biologically guided deep learning model for post-radiotherapy outcome prediction. Our lab has developed and applied this method to the OPC dataset studied in this work.[32,46] The use of deep learning tools alone was common to assess treatment response using the pre-treatment images and dose distribution information as inputs (e.g., Wang et al.).

To extend this analysis, Ji et al. developed a novel partial differential equation, based on a reaction-diffusion model, to incorporate spatial radiation dose information. Then, a 7-layer encoder-decoder-based convolutional neural network was trained to generate post-radiation PET images and break them down into constituent parts related to each of the terms in the biological model. This represents integration of both deep learning models and analytical techniques that encode biological information. The study was able to generate the post-treatment images and break them down into the constituent mathematical components of the model. Gamma tests indicated good agreement between the generated images and ground truth images. This analysis is particularly relevant to this work, as a possible extension of our analysis is to model the effects of the radiation dose as a heat bath in our statistical mechanical framework. Further, encoding biological information directly via partial differential equations represents a key step in improving the explainability of analyses such as ours.

There is an important distinction to be made when defining the term "equilibrium" in this work. The equilibrium state of the Fokker–Planck equation, which generates the evolution of the images, is defined as the intra-treatment PET/CT images, i.e., after 20 Gy of radiation therapy. This is clearly different from the biological equilibrium, as the tumor will continue to evolve dynamically throughout the remaining treatment period and beyond. It is therefore important to emphasize that the equilibrium referred to in this work is not biological, rather it is statistical and defined a priori by our algorithm. However, we do not consider this a major limitation of the work—indeed we make no claim that our algorithm directly simulates biology. Rather, we consider this work an application of statistical mechanics and thus our aim was to describe the macroscopic behavior of a system without direct modeling of the microscopic environment.

This work sought to integrate a traditional radiomics approach with techniques from applied mathematics and statistical mechanics to generate a new formalism. Integration of radiomic analysis and other analytical tools that are mechanistically informed may increase both generalization and interpretation.[46–50] For example, radiomics-boosted deep learning models have been developed for diverse applications, such as COVID-19 pneumonia detection via chest radiographs,[51] post-resection survival prediction of

patients with glioblastoma[50] and identification of radionecrosis following stereotactic radiosurgery (SRS) for brain metastases.[48] In each case, integration of radiomics and deep learning approaches serves to improve interpretability of deep learning models otherwise described as "black boxes". Additionally, inclusion of time-series information, such as cell-free DNA samples (acquired via liquid biopsies[52]) into these types of studies further enhances the robustness of their results. These types of analyses, collectively known as "data fusion", and the quantification of fusion quality are exciting new frontiers in computational imaging research.[53]

Despite the prognostic value and numerical validation of our algorithm, a key limitation of this work is lack of biological validation of our algorithm. By construction, we attempted to model and analyze data for which no ground-truth images exist. While we did not explicitly aim to simulate biology, the question of validation is nonetheless of paramount importance. We addressed this question first via a numerical validation on 2D Gaussian processes. We showed good agreement between first-order features (energy and entropy) extracted from images generated by our algorithm and the corresponding features extracted from the ground truth images. This suggests that our algorithm was able to capture noise characteristics of the images to a reasonable degree.

Ultimately, this key limitation lies in the lack of independent, validation data. The proof-of-concept patient study outlined in Section 2.2 is a secondary analysis of a prospective clinical trial (NCT01908504). As such, enlargement of this dataset is not yet a feasible validation strategy. However, the methodology outlined in Section 2.1 was developed with small sample sizes in mind. By using the PDE model based on Equation (1), we were able to generate unique solutions for each patient in the dataset. Hence, our method does not suffer from the traditional pitfalls of overfitting that are common in analyses based on machine learning techniques.[1]

To further validate our method, we require quasi-continuous ground truth datasets of time-series imaging. One possible example includes 4DCT/4D-MRI images of patients acquired during radiation therapy treatment planning.[54,55] These would provide the necessary ground truth images on which our algorithm could be validated. This would also necessitate the adaptation of our algorithm to other disease sites, which introduces further challenges as described above.

Mouse models could serve as another potential strategy for validation. Specifically, our lab has plans to conduct an animal trial wherein genetically engineered mice[56] are treated for HNSCC and images are acquired across multiple length scales (radiological and pathological). These images would then serve as a ground truth for our algorithm, with the significant advantage of similarity with our study regarding disease type and imaging modality.

5 | CONCLUSION

In this work, we successfully developed a novel, physics-based analytical framework, which we define as dynamic radiomics, that integrated radiomics with partial differential equations. We were able to validate our method using a numerical experiment and compared first order features calculated using our method with ground truth, which showed good agreement. We then estimated time evolution between ^{18}F-FDG-PET images of patients undergoing definitive radiation therapy for OPC. Survival analysis demonstrated added prognostic value of our technique over traditional delta radiomics. Our study lays the groundwork for deeper and more nuanced analysis of the implications of delta radiomics and the integration of radiomic analysis with diverse data sources and techniques.

CONFLICT OF INTEREST STATEMENT
The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT
Research data are stored in an institutional repository and deidentified data will be shared upon request to the corresponding author through a data use agreement.

REFERENCES

1. Lafata KJ, Wang Y, Konkel B, Yin FF, Bashir MR. Radiomics: a primer on high-throughput image phenotyping. *Abdom Radiol (NY)*. 2021;47(9):2986-3002. doi:10.1007/s00261-021-03254-x

2. Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges. *Magn Reson Imaging*. 2012;30(9):1234-1248. doi:10.1016/j.mri.2012.06.010

3. Yip SS, Aerts HJ. Applications and limitations of radiomics. *Phys Med Biol*. 2016;61(13):R150-166. doi:10.1088/0031-9155/61/13/R150

4. Zwanenburg A, Vallieres M, Abdalah MA, et al. The Image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*. 2020;295(2):328-338. doi:10.1148/radiol.2020191145

5. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012;48(4):441-446. doi:10.1016/j.ejca.2011.11.036

6. Nardone V, Reginelli A, Grassi R, et al. Delta radiomics: a systematic review. *Radiol Med*. 2021;126(12):1571-1583. doi:10.1007/s11547-021-01436-7

7. Capobianco E, Deng J. Radiomics at a glance: a few lessons learned from learning approaches. *Cancers*. 2020;12(9):2453. doi:10.3390/cancers12092453

8. Fave X, Zhang LF, Yang JZ, et al. Delta-radiomics features for the prediction of patient outcomes in non-small cell lung cancer. *Sci Rep-Uk*. 2017;7(1):588. doi:10.1038/s41598-017-00665-z

9. Carles M, Fechter T, Radicioni G, et al. FDG-PET radiomics for response monitoring in non-small-cell lung cancer treated with radiation therapy. *Cancers*. 2021;13(4):814. doi:10.3390/cancers13040814

10. Kothari G, Korte J, Lehrer EJ, et al. A systematic review and meta-analysis of the prognostic value of radiomics based models

in non-small cell lung cancer treated with curative radiotherapy. *Radiother Oncol*. 2021;155:188-203. doi:10.1016/j.radonc.2020.10.023

11. Spohn SKB, Bettermann AS, Bamberg F, et al. Radiomics in prostate cancer imaging for a personalized treatment approach—current aspects of methodology and a systematic review on validated studies. *Theranostics*. 2021;11(16):8027-8042. doi:10.7150/thno.61207

12. Delgadillo R, Ford JC, Abramowitz MC, Dal Pra A, Pollack A, Stoyanova R. The role of radiomics in prostate cancer radiotherapy. *Strahlenther Onkol*. 2020;196(10):900-912. doi:10.1007/s00066-020-01679-9

13. Wang L, Gao ZH, Li CM, et al. Computed tomography-based delta-radiomics analysis for discriminating radiation pneumonitis in patients with esophageal cancer after radiation therapy. *Int J Radiat Oncol*. 2021;111(2):443-455. doi:10.1016/j.ijrobp.2021.04.047

14. Wang HX, Zhou Y, Li L, Hou WX, Ma XL, Tian R. Current status and quality of radiomics studies in lymphoma: a systematic review. *Eur Radiol*. 2020;30(11):6228-6240. doi:10.1007/s00330-020-06927-1

15. Lowengrub JS, Frieboes HB, Jin F, et al. Nonlinear modelling of cancer: bridging the gap between cells and tumours. *Nonlinearity*. 2010;23(1):R1-R9. doi:10.1088/0951-7715/23/1/r01

16. Fokker AD. The median energy of rotating electrical dipoles in radiation fields. *Ann Phys-Berlin*. 1914;43(5):810-820.

17. Planck M. An essay on statistical dynamics and its amplification in the quantum theory. *Sitzber K Preuss Aka*. 1917;324-341.

18. Iram S, Dolson E, Chiel J, et al. Controlling the speed and trajectory of evolution with counterdiabatic driving. *Nat Phys*. 2021;17(1):135-142. doi:10.1038/s41567-020-0989-3

19. Medved A, Davis R, Vasquez PA. Understanding fluid dynamics from langevin and Fokker-Planck equations. *Fluids*. 2020;5(1):40. doi:10.3390/fluids5010040

20. Beck H, Nettel S. Electron-transport in disordered semiconductors. *Phys Lett A*. 1984;105(6):319-323. doi:10.1016/0375-9601(84)91007-7

21. Wang X, Weinberg SH, Hao Y, Sobie EA, Smith GD. Calcium homeostasis in a local/global whole cell model of permeabilized ventricular myocytes with a Langevin description of stochastic calcium release. *Am J Physiol-Heart C*. 2015;308(5):H510-H523. doi:10.1152/ajpheart.00296.2014

22. Cohn H, Kulsrud RM. Stellar distribution around a black-hole—numerical-integration of Fokker-Planck equation. *Astrophys J*. 1978;226(3):1087-1108. doi:10.1086/156685

23. Gao Y, Jin GZ, Liu JG. Inbetweening auto-animation via Fokker-Planck dynamics and thresholding. *Inverse Probl Imag*. 2021;15(5):843-864. doi:10.3934/ipi.2021016

24. Gao Y, Li TJ, Li XG, Liu JG. Transition path theory for langevin dynamics on manifolds: optimal control and data-driven solver. *Multiscale Model Sim*. 2023;21(1):1-33. doi:10.1137/21m1437883

25. Gao Y, Liu JG, Wu N. Data-driven efficient solvers for Langevin dynamics on manifold in high dimensions. *Appl Comput Harmon A*. 2023;62:261-309. doi:10.1016/j.acha.2022.09.003

26. Georgii H-O. *Gibbs Measures and Phase Transitions*. De Gruyter Studies in Mathematics. In: de Gruyter W, ed. 1988;xiv:525.

27. Eymard R, Gallouët T, Herbin R. Finite volume methods. In: Ciarlet PG, Lions JL, eds. *Handbook of Numerical Analysis*. Elsevier; 2000:713-1018.

28. Lafata KJ. Novel identification of radiomic biomarkers with Langevin annealing. Duke University; 2018.

29. Chang Y, Lafata K, Wang C, et al. Digital phantoms for characterizing inconsistencies among radiomics extraction toolboxes. *Biomed Phys Eng Express*. 2020;6(2):025016. doi:10.1088/2057-1976/ab779c

30. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun*. 2014;5:4006. doi:10.1038/ncomms5006

31. Xue X, Agalliu I, Kim MY, et al. New methods for estimating follow-up rates in cohort studies. *BMC Med Res Methodol*. 2017;17(1):155. doi:10.1186/s12874-017-0436-z

32. Wang C, Liu C, Chang Y, et al. Dose-distribution-driven PET image-based outcome prediction (DDD-PIOP): a deep learning study for oropharyngeal cancer IMRT application. *Front Oncol*. 2020;10:1592. doi:10.3389/fonc.2020.01592

33. Im HJ, Bradshaw T, Solaiyappan M, Cho SY. Current methods to define metabolic tumor volume in positron emission tomography: which one is better? *Nucl Med Molec Imag*. 2018;52(1):5-15. doi:10.1007/s13139-017-0493-6

34. Bland JM, Altman DG. The logrank test. *BMJ*. 2004;328(7447):1073. doi:10.1136/bmj.328.7447.1073

35. CreedJ MatSurv: Survival analysis and visualization in MATLAB. *J Open Source Software*. 2020;5:1830.

36. Pollack A, Chinea FM, Bossart E, et al. Phase I trial of MRI-guided prostate cancer lattice extreme ablative dose (LEAD) boost radiation therapy (vol 107, pg 305, 2020). *Int J Radiat Oncol*. 2020;108(1):328-328. doi:10.1016/j.ijrobp.2020.05.061

37. Yu H, Caldwell C, Mah K, Mozeg D. Coregistered FDG PET/CT-based textural characterization of head and neck cancer for radiation treatment planning. *Ieee T Med Imaging*. 2009;28(3):374-383. doi:10.1109/Tmi.2008.2004425

38. El Naqa I, Grigsby PW, Apte A, et al. Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. *Pattern Recogn*. 2009;42(6):1162-1171. doi:10.1016/j.patcog.2008.08.011

39. Scalco E, Fiorino C, Cattaneo GM, Sanguineti G, Rizzo G. Texture analysis for the assessment of structural changes in parotid glands induced by radiotherapy. *Radiother Oncol*. 2013;109(3):384-387. doi:10.1016/j.radonc.2013.09.019

40. Scalco E, Moriconi S, Rizzo G. Texture analysis to assess structural modifications induced by radiotherapy. *Ieee Eng Med Bio*. 2015:5219-5222.

41. Wang ZL, Mao LL, Zhou ZG, et al. Pilot study of CT-based radiomics model for early evaluation of response to immunotherapy in patients with metastatic melanoma. *Front Oncol*. 2020;10:1524. doi:10.3389/fonc.2020.01524

42. Basler L, Gabrys HS, Hogan SA, et al. Radiomics, tumor volume, and blood biomarkers for early prediction of pseudoprogression in patients with metastatic melanoma treated with immune checkpoint inhibition. *Clin Cancer Res*. 2020;26(16):4414-4425. doi:10.1158/1078-0432.Ccr-20-0020

43. Zhang ZJ, Yang JZ, Ho A, et al. A predictive model for distinguishing radiation necrosis from tumour progression after gamma knife radiosurgery based on radiomic features from MR images (vol 28, pg 2255, 2018). *Eur Radiol*. 2018;28(8):3570-3571. doi:10.1007/s00330-017-5276-z

44. Lara-Benitez P, Carranza-Garcia M, Riquelme JC. An experimental review on deep learning architectures for time series forecasting. *Int J Neural Syst*. 2021;31(3):2130001. doi:10.1142/S0129065721300011

45. Lafata KJ, Chang Y, Wang C, et al. Intrinsic radiomic expression patterns after 20 Gy demonstrate early metabolic response of oropharyngeal cancers. *Med Phys*. 2021;48(7):3767-3777. doi:10.1002/mp.14926

46. Ji H, Lafata K, Mowery Y, et al. Post-Radiotherapy PET image outcome prediction by deep learning under biological model guidance: a feasibility study of oropharyngeal cancer application. *Front Oncol*. 2022;12:895544. doi:10.3389/fonc.2022.895544

47. Yang Z, Hu Z, Ji H, et al. A neural ordinary differential equation model for visualizing deep neural network behaviors in multi-parametric MRI-based glioma segmentation. *Med Phys*. 2023;50(8):4825-4838. doi:10.1002/mp.16286

48. Zhao J, Yang Z, Hu Z, et al. A radiomics-integrated deep-learning model for identifying radionecrosis following brain metastasis stereotactic radiosurgery. *Medical Physics*. 2022;49(6):E457-E457.

49. Hu ZS, Yang ZY, Lafata KJ, Yin FF, Wang CH. A radiomics-boosted deep-learning model for COVID-19 and non-COVID-19 pneumonia classification using chest x-ray images. *Medical Physics*. 2022;49(5):3213-3222. doi:10.1002/mp.15582

50. Hu Z, Yang Z, Zhao J, et al. A deep learning model with radiomics analysis integration for glioblastoma post-resection survival prediction. *Medical Physics*. 2022;49(6):E364-E364.

51. Hu Z, Yang Z, Lafata KJ, Yin FF, Wang C. A radiomics-boosted deep-learning model for COVID-19 and non-COVID-19 pneumonia classification using chest x-ray images. *Med Phys*. 2022;49(5):3213-3222. doi:10.1002/mp.15582

52. Lafata KJ, Corradetti MN, Gao J, et al. Radiogenomic analysis of locally advanced lung cancer based on CT imaging and intratreatment changes in cell-free DNA. *Radiol Imaging Cancer*. 2021;3(4):e200157. doi:10.1148/rycan.2021200157

53. Wang Y, Li X, Konanur M, et al. Towards optimal deep fusion of imaging and clinical data via a model-based description of fusion quality. *Med Phys*. 2023;50(6):3526-3537. doi:10.1002/mp.16181

54. Cai J, Chang Z, Wang Z, Paul Segars W, Yin FF. Four-dimensional magnetic resonance imaging (4D-MRI) using image-based respiratory surrogate: a feasibility study. *Med Phys*. 2011;38(12):6384-6394. doi:10.1118/1.3658737

55. Li G, Citrin D, Camphausen K, et al. Advances in 4D medical imaging and 4D radiation therapy. *Technol Cancer Res Treat*. 2008;7(1):67-81. doi:10.1177/153303460800700109

56. Lee CL, Mowery YM, Daniel AR, et al. Mutational landscape in genetically engineered, carcinogen-induced, and radiation-induced mouse sarcoma. *JCI Insight*. 2019;4(13):e128698. doi:10.1172/jci.insight.128698