Radiomics in CT imaging for Longitudinal Analysis in Oncology: Emerging Complements, Advantages and Limitations

Paula Castro Ramírez
p.castroramirez@students.uu.nl

MSc Medical Imaging, Utrecht University,
University Medical Center Utrecht (UMCU), Image Sciences Institute

***Abstract*—Precision medicine plays a crucial role for cancer diagnosis and treatment planning. With significant advancements in the medical imaging field, new techniques are being developed to characterize diseases, predict treatment outcomes, and determine survival rates. In this context, the collection of image data over time known as longitudinal imaging has become increasingly popular in clinical oncology for its prognosis potential. Radiomics is a widely used tool to extract features from such images to analyze them. While traditional radiomic approaches such as texture or shape analysis have proven to be effective, they may lack sensitivity to certain changes, occasionally oversimplifying the complexity of the data under study. Furthermore, these methods are sometimes not stable enough to noise or image artifacts, and do not generalize well, making it challenging to compare scans over time or across different studies. In this paper, we present a review of the use of radiomics in longitudinal oncologic studies, specially focusing on its applications in CT imaging. To this end, this study aims to explore three main alternative methods to traditional radiomics: for feature extraction, topological data analysis (TDA), including persistent homology, and geodesic geometry; and for data analysis, Cox and joint statistical modeling. By capturing and analyzing more complex features, these methods offer new valuable insights into disease progression, making them strong candidates for the development of more accurate treatment planning and prognostic models in clinical oncology.**

***Index Terms*—Radiomics, Computed Tomography, Longitudinal imaging, Texture analysis, Topology, Persistent Homology, Geodesic Geometry, Statistical Analysis, Joint models.**

I. Introduction

Significant advancements in the medical imaging field have led to the development of new and more refined methods to evaluate and analyze medical images. By accurately deciphering the clinical meaning beneath such images at an early stage, diseases could be detected and treated with a much greater precision [1]. Radiomics is a technique that extracts high-dimensional quantitative information invisible to the naked human eye by means of sophisticated and, sometimes, considerably complex mathematical algorithms [2]. Countless radiomic features can be calculated, creating lists that could be used to create correlations between image characteristics and the condition under study. Radiomics presents a novel approach to gather the information from the images, helping the progress of clinical decision-support systems, response to treatments, and the prediction of disease evolution or even survival rates[3][4].

Although radiomics has been applied in a variety of diseases, it has found its most extended applications in oncology. Specifically, radiomics plays a crucial role in precision medicine for cancer diagnosis, treatment and prognosis. Computed tomography (CT) has emerged as a key modality in this field [5], where there has been a growing use of longitudinal images to track tumor evolution over time. Contrary to using single CT images at a certain moment, by acquiring a series of scans at multiple time points, longitudinal studies allows for a more continuous monitoring of the disease. Longitudinal data presents multiple advantages to the field as it includes further information to static scans, offers a much deeper understanding about such tumor, and potentially admits a superior analysis of heterogeneity changes, tumor dynamics, and overall progression of the disease [6].

Radiomics can potentially optimize treatment planning and monitoring of tumor dynamic factors by analysing quantitative parameters over time. In the context of longitudinal image analysis, traditional radiomic approaches have proven to be of help by capturing valuable characteristics. For instance, texture analysis examines intensity pixel variations and extract distinct patterns from the images [7], and shape analysis can describe geometric properties to characterize specific lesions [8]. Nevertheless, such methodologies tend to be sensitive to noise, image artifacts, and may not capture the complete behavior of the tumor microenvironment, occasionally oversimplifying the complexity of the time-dependent data and misreading tumor progression [9][10]. In addition, new statistical modeling alternatives have emerged to analyze time-dependent data, encouraging researchers to reevaluate currently used models like the Cox proportional hazards model for its multiple limiting assumptions that hardly hold for this type of datasets [11]. Given the lack of consensus among the research community for analyzing images over time, and the limitations of the currently used methodologies, the purpose of this literature review is to give an overview of the field of radiomics in evaluating longitudinal images, highlighting possible radiomic feature extraction alternatives to traditional analytical methods, statistical models, challenges, and future perspectives in this growing field.

The literature review is organized as follows: (II) the search strategy followed; (III) a brief theoretical introduction of the radiomics workflow, focusing on its applications in CT imag-

ing; (IV) an introduction to radiomics in longitudinal image characterization, together with the description of alternatives for feature extraction and data analysis with real applications in the clinical setting; (V) discussion of the current challenges of the methodology described, and some future directions of radiomics field; (VI) and a short conclusion.

II. Search Strategy

For this literature review, different research strategies were followed for paper inclusion. The base search strategy was as follows: *(radiomic*) AND (oncology OR cancer) AND ((computed tomography) OR CT OR (medical imaging))*.

Depending on the methodology described, the search strategy was modified as needed including the required key words for each section. Therefore, for Section III, we included *AND (classification OR prognosis OR (treatment planning))*; for Section IV-A we included *AND (longitudinal OR follow-up OR (over time)) AND ((topological data analysis) OR TDA OR (persistent homology) OR (topolog*))*; for Section IV-B we included *AND (longitudinal OR follow-up OR (over time)) AND (geodesic* OR (geodesic geometry))*; and for Section IV-C, we included *AND (longitudinal OR follow-up OR (over time)) AND ((joint model*) OR (survival prediction) OR Cox)*.

The minimum publication date inclusion was set to 2014, and the search was used in Google Scholar and PubMed. To analyze the query results, papers were organized from most relevant to least relevant, rejecting those corresponding to systematic reviews, reviews, books, and from other imaging modalities.

III. Radiomics workflow

Radiomics relies on a series of steps to extract the required features from medical images and analyze them. The radiomics workflow is summarized in Figure 1.

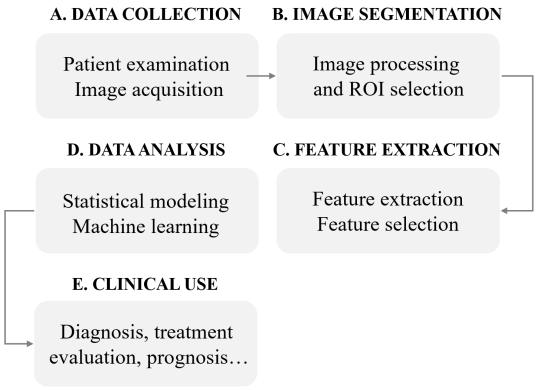

Fig. 1. Summary of the steps to be followed in a radiomics workflow. Abbreviations: Region of Interest (ROI)

A. Data collection

The first step in a radiomic study is the collection of the required images. Although this might appear like a simple step, it involves multiple challenges concerning acquisition parameters such as slice thickness, dose modulation, reconstruction algorithms, or scanning protocols [5]. Standardized imaging protocols have proven to be essential in minimizing unwanted variability within the datasets under study. Defeudis *et al.* [12] investigated how CT standardization between multiple centers affected radiomics workflow and radiomic feature calculation, concluding that inter-center normalization is necessary and key for obtaining accurate results.

B. ROI selection and image segmentation

The next step involves the pre-processing of the images, centered in segmenting the region of interest (ROI), i.e. delineating the region of the image where the radiomic features will be extracted. Currently, there exists multiple image segmentation tools, from manual to completely automatic ones. While manual delineation has been the gold standard method for image segmentation for many years, it may be prone to bias and inconsistencies coming from the clinicians performing the task, as well as being an extremely time-consuming process. Because of that, automatic and semiautomatic methods have gained interest, reducing computational time and variability [1][13].

C. Feature extraction, feature selection and dimensionality reduction

To effectively analyze the ROI delineated, features should be extracted and selected. Feature extraction consists of the calculation of multiple quantitative features (properties of the tissue under study) to uncover previously unseen characteristics on the images. As shown in Table I, a vast number of features can be derived from CT images. Some classical approaches are statistical methods such as histogram descriptors to study the spatial arrangement of gray-level intensity values [14], or texture analysis to examine the intensity pixel variation to characterize distinct patterns of different structures [15]; mathematical morphology methods for shape and size characterization [16]; or model-based methods, modelling texture probabilistically [17], geometrically or using basis functions [7]. Such variations are generally related to underlying qualities of the dataset under inspection, making feature selection a key step in numerous fields. Such properties have proven to be significantly important in the medical imaging field for diverse tasks such as classification, segmentation, or object detection [18]; however, not all of them are appropriate for the specific goals of each investigation.

Feature selection is the process of reducing the number of extracted features to achieve an optimal analysis. Dimensionality reduction can be accomplished through a variety of methods, but a universally accepted standard methodology has not yet been established for excluding the redundant and irrelevant features [1]. Among the most common possibilities of feature extraction methods in CT imaging we can find principal component analysis (PCA) [25], LASSO feature selection [26], hierarchical clustering [27], random forest [28], etc. In the critical reflection from Timmeren *et al.* [1] they

TABLE I
FOUR EXAMPLES OF MAJOR CLASSES OF RADIOMIC FEATURES.
SUMMARY EXTRACTED FROM [19]

Radiomic Feature Class	Definition	Radiomic Feature Examples
Morphological features	Physical properties related to shape and volume [16]	Surface area, surface-to-volume ratio, sphericity [20], etc.
Statistical features	Histogram based features (First-order features) [14]	Mean, standard deviation, kurtosis, skewness, energy, entropy, uniformity, and variance, etc.
	Texture based features (Higher-order features)	Gray level co-occurrence matrix (GLCM) [21]
		Gray level run length matrix (GLRL) [22]
		Neighbourhood gray-tone difference matrix (NGTDM) [23], etc.
Regional features	Variation intensity between regions	Sub-regional partition of groups of voxels into clusters [24]
Model-based features	Characterize shape complexity over scale ranges	Fractal dimension [17]

presented a summary of widely used steps for dimensionality reduction. It first involved a reproduciblility analysis, followed by the calculation of an "importance variable", data layout, cluster correlation, and final feature selection.

D. Data analysis

The final goal of radiomics is to build models and improve clinical decision-making systems for diagnosis, prognosis, and treatment planning, among others. Because of that, once the most relevant features have been selected, the last step in the workflow is to analyze them. Many tools can be used for different purposes [28]. For instance, in classification tasks, logistic regression [29] is the statistical model most widely used, while other commonly applied methods include random forest [30], support vector machine (SVM) [31], k-nearest neighbour [32], and neural networks [33]. Moreover, apart from classifying diseased and not diseased patients, radiomics can also predict treatment response and survival outcomes. For prognosis studies, the most commonly used models are Cox's proportional hazard models [34], alongside regression [35], random forest [36], and SVM models [37].

Applications of Radiomics in Computed Tomography

Radiomics has become an emerging field in CT over the past decades. Several studies have focused on proving its utility, ranging from identifying critical risk factors to classifying various stages of different types of cancer.

For instance, Zeng *et al.* [38] identified different risk factors for the development of brain metastases (BM) in patients with stage III non-small cell lung cancer. They extracted from the delineated gross tumor volume multiple features, including first-order statistical features, texture features and morphological features, and developed three different BM models to determine the positive prognostic value of age and cell carcinoma type. Moreover, in another study, Aerts *et al.* [39] related underlying tumor phenotypic patterns with radiomic features, proving the prognostic quality of feature extraction in lung and head-neck cancer patients. They developed a radiomics signature to capture tumor heterogeneity and demonstrated once more its prognostic power, revealing a real possibility of improving clinical decision-support systems.

In addition to predictive models, multiple studies also focused on classification tasks. For example, Gitto *et al.* [40] classified cartilaginous tumors by training a machine learning (ML) model using radiomic features, potentially serving as guidance in preoperative tumor characterization.

IV. RADIOMICS IN LONGITUDINAL IMAGES

The usage of longitudinal imaging has become increasingly important in clinical oncology due to its rapid advancements in personalized medicine and treatment therapies. With the study of time-dependent scans, additional temporal information is introduced, enriching the analysis, but also including more complexity to it. With these images, more precise decisions could be made at specific time points to maximize the potential of such images, relating the current state of the disease with the predictions of the condition's future status with the help of mathematical and computational procedures [41][42].

In that context, delta-radiomics was introduced, assessing longitudinal changes in radiomic properties and highlighting variations of the extracted features over time. This allowed for a more accurate evaluation of changes related to external variables such as immunotherapies or targeted treatments that may have contributed to morphological changes invisible to the naked eye [41][43]. Nevertheless, limitations such as lack of validation, generalizability, and absence of professional background on the topic still appear [44]. Multiple advanced analytical methods exist to capture features containing texture, shape and size information, and associate such longitudinal markers to occurrences at different time points. This review focuses on the potential usage of topological data analysis (TDA) [45], persistent homology [46], geodesic geometry analysis [47] for feature extraction, and the latter statistical modeling for data analysis.

A. Topological data analysis and persistent homology

As introduced in Section III-C, numerous features can be derived from medical images, yet careful consideration is required to ensure the selection of the most appropriate ones for the specific task at hand. Specifically focusing on texture analysis, challenges remain in effectively capturing global

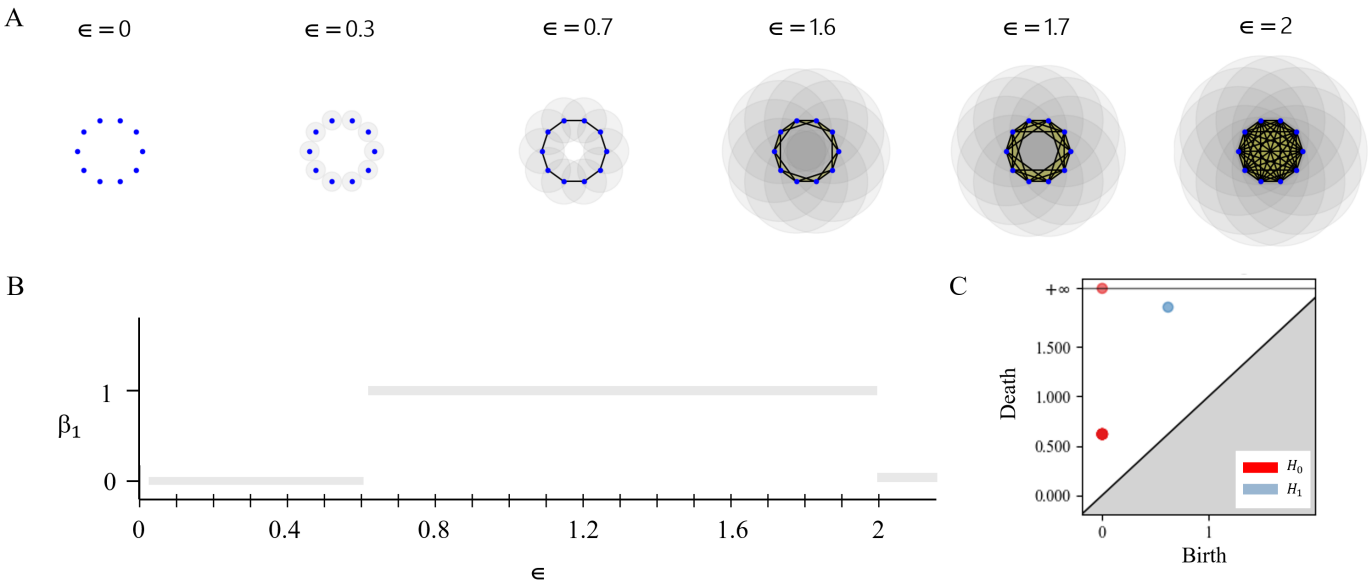

Fig. 2. A: Example of the evolution of simplicial complexes formation in a cloud of N=10 points sampled from a circle; B: corresponding barcode for β_1, which refers to *Betti number* 1 that is the total number of 1-dimensional holes (loops); and C: corresponding persistence diagram, where H_0 refers to zero-level homology features (connected components) and H_1 refers to one-level homology features (loops). Birth corresponds to the value ϵ when the feature appears for the first time, while death corresponds to the value ϵ when the feature disappears. Image adapted from [48].

and local behaviors in a reproducible way as methodologies greatly vary between studies [49]. They may fail to accurately discriminate complex patterns related to changes in time as features tend to be local, sometimes missing more complex contextual information [50].

Topological data analysis

TDA seeks to link spatial properties of a dataset in a more global and connected manner, characterizing structures as shapes based on topological features. This method, instead of focusing on individual points, connects them to study how they relate to each other. By doing so, TDA extracts valuable information that remain unchanged even if the data is distorted, identifying for instance how separated are different groups (connected components), if there are any gaps (voids), or if more complex arrangements (higher dimension structures) appeared [45]. TDA provides different techniques to tackle some of the main limitations in the modern data analysis field like the high dimensionality problem, the lack of simultaneous local and global information, or the absence of methods to capture multivariate shapes [50][48].

Let us consider a cloud of N points sampled from a circle. In TDA, point clouds are frequently perceived as simplicial complexes, mathematical structures used to describe shapes. The units that conform simplicial complexes are called simplexes, which can be of different forms depending on their connectivity as shown in Figure 3. A subset of data points $(k+1)$ is known as k-simplex. In other words, a single data point is a 0-simplex, a line formed between two data points a 1-simplex, a triangle formed between three data points a 2-simplex, and so on [46].

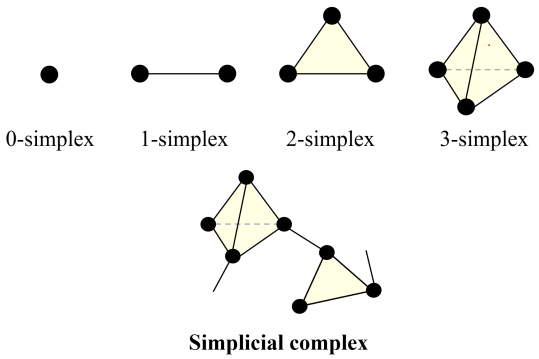

Fig. 3. Simplexes summary: 0-simplex, 1-simplex, 2-simplex, 3-simplex, with an example of a simplicial complex.

Persistent Homology

Persistent homology is a method in TDA that describes the aforementioned extracted structures in a more intuitive manner by characterizing the features in the dataset at different scales [46]. This method makes use of homology, a principle that uses algebraic computations to relate different connectivity properties. Persistent homology looks at how features appear (birth) and disappear (death) by modifying ϵ, a scale parameter that describes different radii sizes around the data points [48]. As ϵ increases, so does the radii around the data, eventually connecting points to each other to form complex arrangements, such as holes, voids, etc.

To evaluate and study the appearance and disappearance of such structures, the evolution need to be tracked based on ϵ. Structures are counted in groups called *Betti numbers* as

β_{th}, where β_0 is the total number of connected components, β_1 is the total number of 1-dimensional holes (loops), β_2 is the total number of 2-dimensional voids (cavities), etc. This tracking process is known as filtration, a procedure that offers a deeper understanding of the dataset by registering the variations in shape and texture across distances to devise the most informative features. Topological features that are kept constant in the filtration process are known as persistent, and are considered to be the most significant [48].

Figure 2.A shows a visual representation of the evolution of simplexes as ϵ increases in a simple point cloud extracted from a circle. This specific dataset presents an example of a circular feature β_1 that persists across multiple values of ϵ, from $\epsilon = 0.7$ to $\epsilon = 2$. To summarize such evolution, persistent homology is often represented in barcodes (Figure 2.B), where each bar starts where the feature appears for the first time and ends where the feature disappears; or alternatively, in persistence diagrams (Figure 2.C) where the features are presented in an Euclidean plane such that *(x,y) = (birth, death)*. The x-axis represents when the feature appears for the first time (birth at $\epsilon = 0.7$), and the y-axis when the feature disappears (death at $\epsilon = 2$). These birth and death of features can be extrapolated to real applications. An example could be the study of the presence of vessel bumps, where $|death - birth|$ will relate to the prominence of the bulge, i.e. the persistence of the extracted homological feature corresponding to the bump [51].

Applications of TDA and Persistent Homology in the clinical setting

Topology analysis and persistent homology have been widely used to tackle diverse research questions. In a study from Somasundaram *et al.* [52], topological features were extracted from CT lung scans by examining pixel intensity values. They implemented a TDA analysis called Cubical complexes, successfully capturing and analyzing topological characteristics as the image was thresholded at different pixel values. Moreover, they also proposed a novel output visualization plot by displaying 0-dimensional topological features against distinct Hounsfield unit values, which they later related with survival models. They demonstrated the real practical use of topology which, together with traditional radiomics, could greatly improve clinical decisions.

Moreover, in a study from Tanabe *et al.* [53], persistent homology was used to detect and analyze 3D physiological abnormalities on chest CT images. Following a similar reasoning as in the previous study from Somasundaram *et al.* [52], they determined the lifetime of features by modifying threshold values. Persistent homology was applied on 3D segmentations, obtaining the correspondent three persistence diagrams, one for each dimension (H_0 for connected components, H_1 for holes, and H_2 for voids). The results confirmed that topological analysis can successfully assess challenging structural modifications in CT imaging, expanding possibilities for future longitudinal data analysis of the topic.

Persistent homology has also been applied for tumor classification. For instance, Belchi *et al.* [54] demonstrated that TDA can greatly improve results obtained using standard features, such as percentage area or volume calculation, by extracting quantitative features derived from persistent homology. They highlighted its superior disease classification and potential use for personalized medicine. However, they also stated that additional research needs to be done to evaluate the direct application in longitudinal studies. In a recent research performed by Vandaele *et al.* [55], it has also been reported that TDA has proved superiority in classifying histology, progression, and treatment response.

In addition to those findings, Igbal *et al.* [56] presented state of the art results for detection and classification of SARS-CoV-2 using persistent homology, achieving a precision and accuracy of 99.4%. Their study emphasized that TDA is an appropriate tool to analyze CT data, even outperforming modern deep learning algorithms. TDA generates more interpretable features, needs less training data and attains comparable quantitative results to traditional radiomic approaches.

B. Geometric, shape, and geodesic analysis

We have previously discussed that topology captures global features that remain unchanged under continuous deformations, extracting shape and texture features from the images; however, geometric information is also crucial to achieve a correct and complete analysis. Traditional morphological features have been widely used to extract geometric properties such as volume, surface area, and size [57][58][59]. Despite this, some barriers continue to exist, specially when dealing with longitudinal data. Classical morphological features might fail to precisely describe small shape changes due to its simplified methodology. For instance, volume features provide one single number to measure the size of the tumor oversimplifying its shape, and potentially missing crucial information about its heterogeneity, curvature, or local changes. More sophisticated methods need to be included to extract additional geometric information and provide a finer dynamic shape analysis.

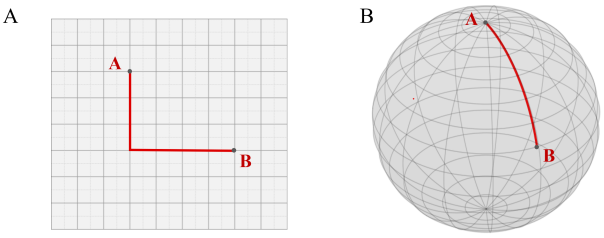

Fig. 4. Geodesics examples. A: Riemannian manifold corresponding to an Euclidean space, a 2D plane; B: Riemannian manifold on a 3D surface, a sphere. Geodesics displayed correspondingly in red from points A to B in each example.

Geodesic geometry

To extrapolate geometric analysis to more complex surface analysis, lets define a Riemannian manifold as a smooth continuous space where Riemannian metrics can be defined. These Riemannian metrics allow researchers to define geometric concepts in surfaces such as tumors [60] and characterize more advanced features. Given two points A and B in a graph, a geodesic is defined as the shortest path between them. In other words, a geodesic distance is the minimum number of edges one have to travel to connect two points.

An example of such shortest trajectories is presented in Figure 4. The simplest scenario is displayed in Fig 4.A, where the Riemannian manifold corresponds to a Euclidean space. In this 2D space, the geodesic distance is reduced to straight lines. Moreover, Fig 4.B shows a similar scenario in a 3D sphere, where the geodesics are not longer straight, but segments of great circles.

Applications of Geodesic Geometry and Structural Analysis in the clinical setting

Such definition of geodesic distance in Riemannian manifolds can be extrapolated to longitudinal studies, successfully overcoming some of the previously mentioned limitations encountered when analyzing the data with traditional morphological features. Numerous papers have used geodesic geometry methodology for multiple applications.

Liu *et al.* [61] applied geodesic geometry to enclose a very specific region of the colon called haustral loop and simplify its geometrical complexity to extract anatomical landmarks. More specifically, they extracted longitudinal geodesics from the colon surface, located paired points to identify the long and short geodesics, and generated loops from each paired point. By leveraging geodesic methodology, they identified approximately 92% of all haustral loops, demonstrating its potential use for computed-aided diagnosis.

In another study, Hong *et al.* [62] experimented with the usage of a hierarchical geodesic models on synthetic longitudinal data to identify shape changes on a S^2 Riemannian manifold. By implementing multivariate models for intercept and slope, and a univariate geodesic model to outline changes for each individual subject, they constructed a multi-geodesic framework. They showed promising outcomes, confirming the hypothesis that time-dependent anatomical changes are closely related to geodesic covariates. Moreover, on that same line of research, Fishbaugh *et al.* [63] developed a hierarchical geodesic model using real longitudinal data. Such model included two types of geodesics, the first type representing subject trajectories, and the second type capturing population trajectories, i.e. reflecting deviations of the individual subject trajectories from the mean.

Han *et al.* [64] performed several studies on the usage of mixed-effects models to trace shape changes in longitudinal data. Although they analyzed their models on 4D right ventricular data and not in CT scans, they also validated their geodesic model implementation with synthetic longitudinal data, concluding that geodesic polynomials of the 3^{rd} order fitted the data points much better than linear geodesic models. In subsequent work, Han *et al.* [65] mentioned that including extra demographic and clinical information about the patients could further improve the prognosis of the models under research.

Lastly, the integration of multiple analytical methods has proven to be useful in longitudinal data analysis. Kaji *et al.* [66] evaluated CT data recorded on a 5 years period to study the pathological physiology of lung diseases. They combined persistent homology and geodesic distance computations to characterize the shape and topology of the trachea, and compared the results to conventional CT lung metrics such as volumes and ratios. As a conclusion, they confirmed the superiority of this novel methodology for more precise morphology estimation and greater accuracy in risk evaluation.

C. Statistical modeling for data analysis

As explained in Section III, once the radiomic features have been extracted, the final goal is to analyze them. This subsection will focus on statistical modeling for survival analysis in longitudinal studies.

Time-dependent Cox models vs Joint models

Analyzing temporal dynamics within datasets has always been a challenge in research, requiring accurate and precise methodologies to capture the progression and behavior of features over time. A common method involves the usage of Cox regression models with time-dependent covariates, however, more sophisticated approaches have been developed in the past years, overcoming some of the limitations of the aforementioned model. One of the most extended alternatives is the usage of Joint models [11].

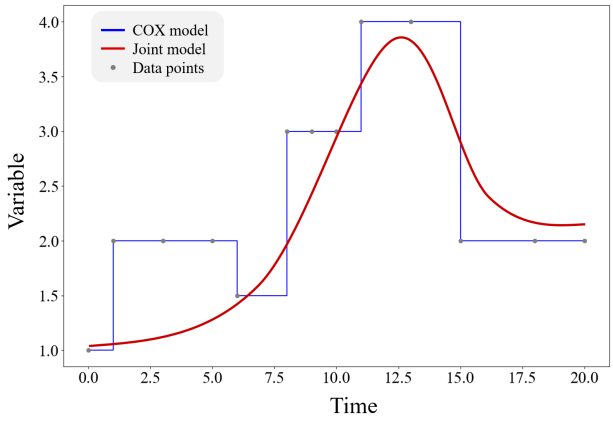

Fig. 5. Example of a collection of data over time for an unknown variable. Data fitted with a Cox model (blue line) and with a Joint model (red line). Data points represented with gray dots.

TABLE II
SUMMARY OF APPLICATIONS OF RADIOMICS, TOPOLOGICAL DATA ANALYSIS, PERSISTENT HOMOLOGY, GEODESIC GEOMETRY, AND STATISTICAL MODELING IN CT IMAGING.

Application area	Reference	Data type	Year	Main Purpose
Radiomics in CT	Aerts *et al.* [39]	Chest and Head-and-neck CT	2014	Study radiomic features for capturing tumor phenotypic differences.
	Gitto *et al.* [40]	CT or PET-CT	2021	Development of a machine learning network for tumor classification.
	Zeng *et al.* [38]	Contrast enhancement CT	2023	Investigate clinical risk factors identification for brain metastasis.
TDA and persistent homology	Belchi *et al.* [54]	Chest CT	2018	Quantitative feature extraction for precise disease classification.
	Somasundaram *et al.* [52]	Chest CT	2021	Study useful clinical correlates extraction and survival prognosis.
	Tanabe *et al.* [53]	Chest CT	2021	Study physiological abnormalities influence based on structural changes.
	Iqbal *et al.* [56]	Chest CT	2021	Topology analysis for prompt detection and classification of SARS-CoV-2.
	Vandaele *et al.* [55]	Chest CT	2023	Topology analysis for accurate tumor histology classification.
Geodesic geometry	Liu *et al.* [61]	CT colonography	2017	Study of geometric structures for structural complexity reduction and registration.
	Hong *et al.* [62]	Synthetic longitudinal data	2019	Development of hierarchical multi-geodesic model for longitudinal analysis.
	Fishbaugh *et al.* [63]	CT longitudinal data	2023	Development of hierarchical geodesic model for subject-specific study progression.
	Han *et al.* [64]	4D Right Ventricular Data	2023	Development of hierarchical geodesic polynomial models for longitudinal analysis.
	Han *et al.* [65]	4D Right Ventricular Data	2024	Combination of subject-specific shape models for longitudinal shape analysis.
	Kaji *et al.* [66]	CT longitudinal data	2024	Study of tubular structures by applying persistent homology and geodesics.
Statistical modeling	Baart *et al.* [11]	Echocardiographic data	2021	Critical comparison of limitations of Cox and Joint modeling in survival analysis.
	Paez *et al.* [67]	CT longitudinal data	2023	Usage of Joint models to predict future cancer probabilities.
	Balbi *et al.* [68]	CT longitudinal data	2024	Delta radiomics. Utilized Cox modeling for survival prediction.

Time-dependent Cox models are commonly build based on the Cox proportional hazards model [69] by extending the already existing methodology with an assumption. Since researchers only have access to data at hospital visits, a hypothesis has to be made about the value of the measurement in the intervals between observations. Time-dependent Cox models assume such value to be constant, constructing a step-function as illustrated in Figure 5 by the blue line. In the graph, measurements are fixed after observations until a new measurement (gray dot) is made. Although this approach might seem straightforward, two main limitations arise. Firstly, the assumed behavior overlooks possible measurement errors, potentially leading to results that deviate from real clinical observations. The more separated in time the measurements are, the larger the assumed error becomes. Secondly, the model does not consider the timing of the measurements, either the patient's disease status, as patients usually visit the hospital when their condition worsens, nor the context of data collection, including how and under what conditions the values were extracted [11].

As a more appropriate method for analyzing longitudinal data, joint models appeared [70]. Contrary to Cox models, joint models develop individual regression models to characterize the behavior of the individual measurements between hospital visits, rather than assuming them constant. Joint models use mixed-effects models, integrating fixed and random effects [11]. While the fixed effects describe the average trajectory of the feature under study over time, random effects account for the particular deviations from the common path of each measurement [71][72][73]. This behavior can be seen in Figure 5 by the red line. As all the measurements made on the patient are endogenous, i.e. depend on events at earlier time points, the corresponding longitudinal process is calculated taking as a principle the joint distribution of both effects. This approach addresses the main limitation presented in time-dependent Cox models by calculating the whole feature trajectory instead of assuming continuity between measurements. Moreover, joint models also admit extra extensions such as the integration of survival sub-models to tackle the possible time imbalance related to visit times, the combination of multiple features in the same model, or the possibility of computing dynamic predictions including new information as it becomes available [11].

Applications of Statistical Joint and Cox Modeling in the clinical setting

Survival risk prediction is one of the main targets when analyzing oncologic data, behavior that can be modeled using different methodologies. For instance, Paez *et al.* [67] classified pulmonary nodules by using a deep learning algorithm, and predicted cancer risk by fitting longitudinal data with a mixed-effects model. Focusing on their second achievement, they developed a joint model considering the random slope effect to define the association between patient measurements, the random intercept effect to characterize the variation difference between each patient measurements, and the interaction between class groups and time. Overall, they confirmed the superior predictive ability of joint models in estimating cancer probability.

Moreover, Balbi *et al.* [68] evaluated the influence of longitudinal variations in radiomic features and inflammatory indices compared to single time-point measurements on cancer survival prediction. To develop such delta-radiomics study, they defined three different Cox based models. They concluded that the model including delta-radiomics significantly outperformed the baseline-radiomics model in terms of prognostic evaluation performance.

In another study, Baart *et al.* [11] directly compared the prognostic ability of time-dependent Cox models with joint

modeling. Although the study was carried out using echocardiographic images and not CT scans, it gave a clear overview of the main differences of both models when examining longitudinal data. Due to the multiple limiting hypothesis made for the Cox longitudinal variant, they proposed the joint models as the superior alternative, greatly helping to develop more personalized treatments and monitoring.

V. Discussion

Radiomics is a powerful tool that offers the possibility of analyzing vast amounts of data, with the ultimate goal of optimizing treatment planning and advancing personalized medicine. However, there are currently multiple challenges and gaps in knowledge, specially regarding the analysis of time-dependent datasets.

In this literature review we have revised the radiomics workflow for medical images, and presented an analysis of its use in longitudinal CT examinations. More specifically, we have given an overview of three alternative analytical methods to traditional radiomics that could potentially capture texture, shape and size features with higher precision and accuracy. Moreover, we have also summarized the current clinical applications of radiomics in longitudinal data inspection, with a special focus on the work that has been done utilizing topology, persistent homology and geodesic geometry for feature extraction, and statistical models such as COX and joint models for data analysis.

A. Current challenges

One of the main challenges when analysing longitudinal data relies on the selection of the most appropriate features, i.e. selection of the features that are most informative about changes over time. Extracting such information is not straightforward, and still there has not yet been a consensus in the research community on a standardized methodology for addressing this issue. The alternative analytical methods described in this paper have the potential to tackle real problems on traditional radiomics such as difficulties when detecting complex patterns oversimplifying shapes and sizes, sometimes missing global and local information, or complications finding stable features over time. Despite that, these methods come with their own limitations. Let's address the specific topics introduced in the review one by one.

TDA and persistent homology remain to be state-of-the-art approaches when analyzing medical images, however, there is still not enough validation in the field. Limited access to other studies makes it challenging for these alternatives to interpret biological meanings and be reproducible for different image types. Most research in the last decade has primarily focused on the study of physiological changes on lung CT scans, leaving a gap in references for other diseases. There is a need for more extensive investigation to apply TDA in other research areas such as longitudinal studies. In that line of work, TDA and persistent homology have shown potential to characterize time-dependent data in cloud points [74][75] but, while multiple papers have mentioned the possibility [54][53], no studies using longitudinal CT scans were identified utilizing such methodologies in this review.

Moreover, topological tools might also require specific parameter modifications [50], such as the choice of scale parameter epsilon ϵ, or the decision up to which Betti number (β_0, β_1, β_2, etc.) to calculate. Computations can easily become highly complex as the number of simplicial complexes (nodes, lines, triangles, etc.) increase exponentially with certain parameters. These choices could pose an important limitation because of the insufficient literature and standard practices, making it difficult to achieve the most optimal results of the methodology.

Moving on to geodesic geometry analysis, the most noticeable limitation would be computational complexity. A strong foundation of mathematical, statistical, and Riemannian geometry is required to correctly integrate geodesics into the radiomics workflow. Geodesics have to be applied based on several mathematical rules. For instance, to be able to use minimization algorithms, surfaces have to be continuous, the surface under investigation has to be tractable, that is, it has to be defined by a discrete number of finite points creating a mesh, vertices must be pre-computed in order to allow for the approximation of the derivatives that fall in between vertices, etc. In addition to that, medical image segmentations frequently result in very intricate shapes, making the process of finding the most accurate parameterization an extremely difficult task [8].

In longitudinal studies, geodesic geometry has proven to be of great use to track shape changes across time [62][63][66]. However, it may not always be the most optimal methodology for all case studies. For instance, when dealing with relatively simple datasets, the inherent complexity of the geodesic calculations can become an obstacle, potentially introducing unnecessary computational load and complicating the analysis interpretability. In those cases, more simple and efficient alternative approaches exist, such as PCA [76][25], or graph-based methods [77].

Lastly, we must discuss some of the limitations of the statistical model under consideration for data analysis in this review. While joint modeling provides a consolidated approach for analyzing time-to-event data, its usage also presents challenges that go from computational demands to model decisions. As mentioned in Section IV-C, these models describe the non-linearity of the dataset by adding extra terms to the methodology. One of the main limitations come with the selection that has to be made regarding those terms. Splines, quadratic, or higher order polynomials [11] can be included depending on the focus of the research. Moreover, statistical models may also sensitive be to missing data, an issue that should also be considered and addressed accordingly. For instance, the data can be missing completely at random (MCAR), missing at random (MAR), or not missing at random (MNAR) [78].

B. Future perspectives

Based on the current literature and evolving methodologies, the future of radiomics in longitudinal studies is very promising. In this review, we have explored some alternative approaches to traditional feature extraction methods focusing on TDA, homology and geodesics. In addition to that, substantial efforts have been made to enhance radiomics results through the integration of artificial intelligence (AI). For instance, in a study from Nasief *et al.* [79], they developed a machine learning model with CT-derived delta-radiomic features based on a feed forward network to find clinical relationships between the previously calculated features and potential pathological reactions. Also, in another paper from Farina *et al.* [80], a deep learning CNN-based was used to gather high-level spatial characteristics related to immunotherapy response using transfer learning.

Moreover, variational autoenconders (VAEs) have also gained attention for their potential application in disease progression. VAEs are deep learning models that can encode large amounts of data (high dimensional data) into a low-dimensional latent space. This allows them to capture intrinsic features of the progression of the disease, study such process in a more precise way, and trace possible changes that were not apparent using conventional delta-radiomic methodologies. In an article from Sauty *et al.* [81], they successfully developed a VAE architecture to translate complex longitudinal image data to a lower dimensional space with the aim of relating patient variations with time. Similarly, Li *et al.* [82] described a beta-VAE that encoded lesion sizes from lung CT scans into a latent space, being able to predict pathological phases of the disease.

Although these studies have already been published, their findings serve as a solid foundation for the introduction of emerging AI concepts into radiomic workflows, integrating some methodologies that are still in the early stages of exploration and development. As they continue to evolve, their future routine incorporation in the clinical practice holds significant potential for enhancing oncology treatment decisions, improving treatment planning, and polishing prognostic models.

VI. Conclusion

In this paper, we have presented an overview of the use of radiomics in longitudinal studies, specially focusing on its applications in CT imaging. Although radiomics have proven to be a very valuable tool for analyzing medical images, commonly used feature extraction and data analysis tools present several limitations when closely examining time-dependent data. As an alternative, three main methodologies have been described in this review: persistent homology and geodesic geometry for feature extraction, and joint statistical modeling for data analysis. Moreover, some future steps regarding radiomics and deep learning networks have also been addressed. Despite of the challenges, the integration of more sophisticated topological, shape and statistical methods in radiomics can offer extra valuable insights into disease progression, making them strong candidates for the development of more accurate treatment planning and prognostic models in clinical oncology.

References

[1] Janita E. van Timmeren et al. "Radiomics in medical imaging—"how-to" guide and critical reflection". In: *Insights into Imaging* 11 (1 Dec. 2020), p. 91. ISSN: 1869-4101. DOI: 10.1186/s13244-020-00887-2.

[2] Roberta Fusco et al. "Radiomics in medical imaging: pitfalls and challenges in clinical management". In: *Japanese Journal of Radiology* 40 (9 Sept. 2022), pp. 919–929. ISSN: 1867-1071. DOI: 10.1007/s11604-022-01271-4.

[3] Philippe Lambin et al. "Predicting outcomes in radiation oncology—multifactorial decision support systems". In: *Nature Reviews Clinical Oncology* 10 (1 Jan. 2013), pp. 27–40. ISSN: 1759-4774. DOI: 10.1038/nrclinonc.2012.196.

[4] Zhenyu Liu et al. "The Applications of Radiomics in Precision Diagnosis and Treatment of Oncology: Opportunities and Challenges". In: *Theranostics* 9 (5 2019), pp. 1303–1322. ISSN: 1838-7640. DOI: 10.7150/thno.30309.

[5] Geewon Lee, So Hyeon Bak, and Ho Yun Lee. "CT Radiomics in Thoracic Oncology: Technique and Clinical Applications". In: *Nuclear Medicine and Molecular Imaging* 52 (2 Apr. 2018), pp. 91–98. ISSN: 1869-3474. DOI: 10.1007/s13139-017-0506-5.

[6] Atanu Bhattacharjee. "Longitudinal Data Analysis". In: *Bayesian Approaches in Oncology Using R and OpenBUGS* (2020). URL: https://api.semanticscholar.org/CorpusID:3058205.

[7] Maryam Khaksar Ghalati et al. "Texture Analysis and Its Applications in Biomedical Imaging: A Survey". In: *IEEE Reviews in Biomedical Engineering* 15 (2022), pp. 222–246. ISSN: 1937-3333. DOI: 10.1109/RBME.2021.3115703.

[8] Alexander M Bronstein, Michael M Bronstein, and Ron Kimmel. *Numerical geometry of non-rigid shapes*. Springer Science & Business Media, 2008.

[9] Alfonso Reginelli et al. "Radiomics as a New Frontier of Imaging for Cancer Prognosis: A Narrative Review". In: *Diagnostics* 11 (10 Sept. 2021), p. 1796. ISSN: 2075-4418. DOI: 10.3390/diagnostics11101796.

[10] Elisa Scalco and Giovanna Rizzo. "Texture analysis of medical images for radiotherapy applications". In: *The British Journal of Radiology* 90 (1070 Jan. 2017). ISSN: 0007-1285. DOI: 10.1259/bjr.20160642.

[11] Sara J. Baart et al. "Joint Modeling of Longitudinal Markers and Time-to-Event Outcomes: An Application and Tutorial in Patients After Surgical Repair of Transposition of the Great Arteries". In: *Circulation: Cardiovascular Quality and Outcomes* 14 (11 Nov. 2021). ISSN: 1941-7713. DOI: 10.1161/CIRCOUTCOMES.120.007593.

[12] Arianna Defeudis et al. "Standardization of CT radiomics features for multi-center analysis: impact of software settings and parameters". In: *Physics in Medicine Biology* 65 (19 Sept. 2020), p. 195012. ISSN: 1361-6560. DOI: 10.1088/1361-6560/ab9f61.

[13] Philippe Lambin et al. "Radiomics: the bridge between medical imaging and personalized medicine". In: *Nature Reviews Clinical Oncology* 14 (12 Dec. 2017), pp. 749–762. ISSN: 1759-4774. DOI: 10.1038/nrclinonc.2017.141.

[14] A. Azoulay et al. "Morphological imaging and CT histogram analysis to differentiate pancreatic neuroendocrine tumor grade 3 from neuroendocrine carcinoma". In: *Diagnostic and Interventional Imaging* 101 (12 Dec. 2020), pp. 821–830. ISSN: 22115684. DOI: 10.1016/j.diii.2020.06.006.

[15] William Henry Nailon. "Texture analysis methods for medical image characterisation". In: *Biomedical imaging* 75 (2010), p. 100.

[16] Sumijan Sumijan et al. "Hybrids Otsu method, feature region and mathematical morphology for calculating volume hemorrhage brain on CT-scan image and 3D reconstruction". In: *TELKOMNIKA (Telecommunication Computing Electronics and Control)* 15.1 (2017), pp. 283–291.

[17] Bidyut Baran Chaudhuri and Nirupam Sarkar. "Texture segmentation using fractal dimension". In: *IEEE Transactions on pattern analysis and machine intelligence* 17.1 (1995), pp. 72–77.

[18] P. S. P. Wang (eds.) C. H. Chen L. F. Pau. "The Handbook of Pattern Recognition and Computer Vision (2nd Edition)". In: vol. 2. 1998, pp. 207–248.

[19] Geewon Lee et al. “Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical management: State of the art”. In: *European Journal of Radiology* 86 (Jan. 2017), pp. 297–307. ISSN: 0720048X. DOI: 10.1016/j.ejrad.2016.09.005.

[20] Yuichi Hayakawa and Takashi Oguchi. “Evaluation of gravel sphericity and roundness based on surface-area measurement with a laser scanner”. In: *Computers Geosciences* 31 (6 July 2005), pp. 735–741. ISSN: 00983004. DOI: 10.1016/j.cageo.2005.01.004.

[21] P Mohanaiah, P Sathyanarayana, and L GuruKumar. “Image texture feature extraction using GLCM approach”. In: *International journal of scientific and research publications* 3.5 (2013), pp. 1–5.

[22] H-H Loh, J-G Leu, and Ren C Luo. “The analysis of natural textures using run length features”. In: *IEEE Transactions on Industrial Electronics* 35.2 (1988), pp. 323–328.

[23] Usman Mahmood et al. “Investigating the robustness neighborhood gray tone difference matrix and gray level co-occurrence matrix radiomic features on clinical computed tomography systems using anthropomorphic phantoms: evidence from a multivendor study”. In: *Journal of computer assisted tomography* 41.6 (2017), pp. 995–1001.

[24] Wei-Ying Ma and Bangalore S Manjunath. “Texture features and learning similarity”. In: *Proceedings CVPR IEEE computer society conference on computer vision and pattern recognition*. IEEE. 1996, pp. 425–430.

[25] Damian J. J. Farnell and Peter Claes. “Initial Steps towards a Multilevel Functional Principal Components Analysis Model of Dynamical Shape Changes”. In: *Journal of Imaging* 9 (4 Apr. 2023), p. 86. ISSN: 2313-433X. DOI: 10.3390/jimaging9040086.

[26] Ramakrishnan Muthukrishnan and R Rohini. “LASSO: A feature selection technique in predictive modeling for machine learning”. In: *2016 IEEE international conference on advances in computer applications (ICACA)*. Ieee. 2016, pp. 18–20.

[27] Dino Ienco and Rosa Meo. “Exploration and reduction of the feature space by hierarchical clustering”. In: *Proceedings of the 2008 siam international conference on data mining*. SIAM. 2008, pp. 577–587.

[28] Gary Ge and Jie Zhang. “Feature selection methods and predictive models in CT lung cancer radiomics”. In: *Journal of Applied Clinical Medical Physics* 24 (1 Jan. 2023). ISSN: 1526-9914. DOI: 10.1002/acm2.13869.

[29] Mutlu Gülbay. “A radiomics-based logistic regression model for the assessment of emphysema severity”. In: *Tuberk Toraks* 71 (3 Sept. 2023), pp. 290–298. ISSN: 04941373. DOI: 10.5578/tt.20239710.

[30] Fen-hua Zhao et al. “Predictive Efficacy of a Radiomics Random Forest Model for Identifying Pathological Subtypes of Lung Adenocarcinoma Presenting as Ground-Glass Nodules”. In: *Frontiers in Oncology* 12 (May 2022). ISSN: 2234-943X. DOI: 10.3389/fonc.2022.872503.

[31] B. Pushpa et al. “Liver fat analysis using optimized support vector machine with support vector regression”. In: *Technology and Health Care* 31 (3 May 2023), pp. 867–886. ISSN: 09287329. DOI: 10.3233/THC-220254.

[32] K. Murphy et al. “A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification”. In: *Medical Image Analysis* 13 (5 Oct. 2009), pp. 757–770. ISSN: 13618415. DOI: 10.1016/j.media.2009.07.001.

[33] Sovanlal Mukherjee et al. “Radiomics-based Machine-learning Models Can Detect Pancreatic Cancer on Prediagnostic Computed Tomography Scans at a Substantial Lead Time Before Clinical Diagnosis”. In: *Gastroenterology* 163 (5 Nov. 2022), 1435–1446.e3. ISSN: 00165085. DOI: 10.1053/j.gastro.2022.06.066.

[34] Akifumi Nishida et al. “Clinicoradiological outcomes of 33 cases of surgically resected pulmonary pleomorphic carcinoma: correlation with prognostic indicators”. In: *European Radiology* 26 (1 Jan. 2016), pp. 25–31. ISSN: 0938-7994. DOI: 10.1007/s00330-015-3811-3.

[35] Yufeng Zhu et al. “Establishment and validation of prognosis model for patients with cerebral contusion”. In: *BMC Neurology* 21 (1 Dec. 2021), p. 463. ISSN: 1471-2377. DOI: 10.1186/s12883-021-02482-4.

[36] Dezhi Shan et al. “Non-Contrasted CT Radiomics for SAH Prognosis Prediction”. In: *Bioengineering* 10 (8 Aug. 2023), p. 967. ISSN: 2306-5354. DOI: 10.3390/bioengineering10080967.

[37] Xinghua Xu et al. “Prognostic prediction of hypertensive intracerebral hemorrhage using CT radiomics and machine learning”. In: *Brain and Behavior* 11 (5 May 2021). ISSN: 2162-3279. DOI: 10.1002/brb3.2085.

[38] Haiyan Zeng et al. “The Association of Gross Tumor Volume and Its Radiomics Features with Brain Metastases Development in Patients with Radically Treated Stage III Non-Small Cell Lung Cancer”. In: *Cancers* 15 (11 May 2023), p. 3010. ISSN: 2072-6694. DOI: 10.3390/cancers15113010.

[39] Hugo J. W. L. Aerts et al. “Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach”. In: *Nature Communications* 5 (1 June 2014), p. 4006. ISSN: 2041-1723. DOI: 10.1038/ncomms5006.

[40] Salvatore Gitto et al. “CT radiomics-based machine learning classification of atypical cartilaginous tumours and appendicular chondrosarcomas”. In: *EBioMedicine* 68 (June 2021), p. 103407. ISSN: 23523964. DOI: 10.1016/j.ebiom.2021.103407.

[41] Dalit Engelhardt and Franziska Michor. “A Quantitative Paradigm for Decision-Making in Precision Oncology”. In: *Trends in Cancer* 7 (4 Apr. 2021), pp. 293–300. ISSN: 24058033. DOI: 10.1016/j.trecan.2021.01.006.

[42] Bibhas Chakraborty and Erica E.M. Moodie. *Statistical Methods for Dynamic Treatment Regimes*. Springer New York, 2013. ISBN: 978-1-4614-7427-2. DOI: 10.1007/978-1-4614-7428-9.

[43] Geewon Lee et al. “Measurement Variability in Treatment Response Determination for Non–Small Cell Lung Cancer”. In: *Journal of Thoracic Imaging* 34 (2 Mar. 2019), pp. 103–115. ISSN: 0883-5993. DOI: 10.1097/RTI.0000000000000390.

[44] Valerio Nardone et al. “Delta radiomics: a systematic review”. In: *La radiologia medica* 126 (12 Dec. 2021), pp. 1571–1583. ISSN: 0033-8362. DOI: 10.1007/s11547-021-01436-7.

[45] Vsevolod Salnikov, Daniele Cassese, and Renaud Lambiotte. “Simplicial complexes and complex systems”. In: *European Journal of Physics* (July 2018). DOI: 10.1088/1361-6404/aae790.

[46] Shafie Gholizadeh, Armin Seyeditabari, and Wlodek Zadrozny. “Topological Signature of 19th Century Novelists: Persistent Homology in Text Mining”. In: *Big Data and Cognitive Computing* 2 (4 Oct. 2018), p. 33. ISSN: 2504-2289. DOI: 10.3390/bdcc2040033.

[47] Jürgen Jost and Jeurgen Jost. *Riemannian geometry and geometric analysis*. Vol. 42005. Springer, 2008.

[48] Richard M. Levenson et al. “Advancing Precision Medicine: Algebraic Topology and Differential Geometry in Radiology and Computational Pathology”. In: *Laboratory Investigation* 104 (6 June 2024), p. 102060. ISSN: 00236837. DOI: 10.1016/j.labinv.2024.102060.

[49] Meghan G. Lubner et al. “CT Texture Analysis: Definitions, Applications, Biologic Correlates, and Challenges”. In: *RadioGraphics* 37 (5 Sept. 2017), pp. 1483–1503. ISSN: 0271-5333. DOI: 10.1148/rg.2017170056.

[50] Yara Skaf and Reinhard Laubenbacher. “Topological data analysis in biomedicine: A review”. In: *Journal of Biomedical Informatics* 130 (June 2022), p. 104082. ISSN: 15320464. DOI: 10.1016/j.jbi.2022.104082.

[51] Dario Arnaldo Domanin et al. “Persistence diagrams for exploring the shape variability of abdominal aortic aneurysms”. In: *arXiv preprint arXiv:2402.07966* (2024).

[52] Eashwar Somasundaram et al. “Persistent homology of tumor CT scans is associated with survival in lung cancer”. In: *Medical Physics* 48 (11 Nov. 2021), pp. 7043–7051. ISSN: 0094-2405. DOI: 10.1002/mp.15255.

[53] Naoya Tanabe et al. “A homological approach to a mathematical definition of pulmonary fibrosis and emphysema on computed tomography”. In: *Journal of Applied Physiology* 131 (2 Aug. 2021), pp. 601–612. ISSN: 8750-7587. DOI: 10.1152/japplphysiol.00150.2021.

[54] Francisco Belchi et al. “Lung Topology Characteristics in patients with Chronic Obstructive Pulmonary Disease”. In: *Scientific Reports* 8 (1 Mar. 2018), p. 5341. ISSN: 2045-2322. DOI: 10.1038/s41598-018-23424-0.

[55] Robin Vandaele et al. “Topological data analysis of thoracic radiographic images shows improved radiomics-based lung tumor histology prediction”. In: *Patterns* 4 (1 Jan. 2023), p. 100657. ISSN: 26663899. DOI: 10.1016/j.patter.2022.100657.

[56] Sohail Iqbal et al. “Classification of COVID-19 via Homology of CT-SCAN”. In: *arXiv preprint arXiv:2102.10593* (2021).

[57] Bartjan de Hoop et al. “Pulmonary Ground-Glass Nodules: Increase in Mass as an Early Indicator of Growth”. In: *Radiology* 255 (1 Apr. 2010), pp. 199–206. ISSN: 0033-8419. DOI: 10.1148/radiol.09090571.

[58] Elias Munoz et al. “3D-Morphomics, Morphological Features on CT scans for lung nodule malignancy diagnosis”. In: (July 2022).

[59] Stefano Diciotti et al. “The *LoG* Characteristic Scale: A Consistent Measurement of Lung Nodule Size in CT Imaging”. In: *IEEE Transactions on Medical Imaging* 29.2 (2010), pp. 397–409. DOI: 10.1109/TMI.2009.2032542.

[60] Jürgen Jost. *Riemannian Geometry and Geometric Analysis*. Springer Berlin Heidelberg, 2011. ISBN: 978-3-642-21297-0. DOI: 10.1007/978-3-642-21298-7.

[61] Yongkai Liu et al. "Haustral loop extraction for CT colonography using geodesics". In: *International Journal of Computer Assisted Radiology and Surgery* 12 (3 Mar. 2017), pp. 379–388. ISSN: 1861-6410. DOI: 10.1007/s11548-016-1497-x.

[62] Sungmin Hong et al. "Hierarchical Multi-geodesic Model for Longitudinal Analysis of Temporal Trajectories of Anatomical Shape and Covariates". In: Med Image Comput Comput Assist Interv, 2019, pp. 57–65. DOI: 10.1007/978-3-030-32251-9_7.

[63] James Fishbaugh et al. "Modeling Longitudinal Optical Coherence Tomography Images for Monitoring and Analysis of Glaucoma Progression". In: Shape Med Imaging, 2023, pp. 236–247. DOI: 10.1007/978-3-031-46914-5_19.

[64] Ye Han et al. "Hierarchical Geodesic Polynomial Model for Multilevel Analysis of Longitudinal Shape". In: 2023, pp. 810–821. DOI: 10.1007/978-3-031-34048-2_62.

[65] Ye Han et al. "Geodesic-based Predictive Shape Modeling on Right Ventricles of Patients with Hypoplastic Left Heart Syndrome". In: (June 2024).

[66] Shizuo Kaji et al. "Quantification of Airway Structures by Persistent Homology". In: *IEEE Transactions on Medical Imaging* 43 (8 Aug. 2024), pp. 2758–2768. ISSN: 0278-0062. DOI: 10.1109/TMI.2024.3376683.

[67] Rafael Paez et al. "Longitudinal lung cancer prediction convolutional neural network model improves the classification of indeterminate pulmonary nodules". In: *Scientific Reports* 13 (1 Apr. 2023), p. 6157. ISSN: 2045-2322. DOI: 10.1038/s41598-023-33098-y.

[68] Maurizio Balbi et al. "Longitudinal Changes of CT-radiomic and Systemic Inflammatory Features Predict Survival in Advanced Non–Small Cell Lung Cancer Patients Treated With Immune Checkpoint Inhibitors". In: *Journal of Thoracic Imaging* (Aug. 2024). ISSN: 0883-5993. DOI: 10.1097/RTI.0000000000000801.

[69] D. R. Cox. "Regression Models and Life-Tables". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 34 (2 Jan. 1972), pp. 187–202. ISSN: 1369-7412. DOI: 10.1111/j.2517-6161.1972.tb00899.x.

[70] Maria Sudell, Ruwanthi Kolamunnage-Dona, and Catrin Tudur-Smith. "Joint models for longitudinal and time-to-event data: a review of reporting quality with a view to meta-analysis". In: *BMC Medical Research Methodology* 16 (1 Dec. 2016), p. 168. ISSN: 1471-2288. DOI: 10.1186/s12874-016-0272-6.

[71] Dean Follmann and Margaret Wu. "An approximate generalized linear model with random effects for informative missing data". In: *Biometrics* (1995), pp. 151–168.

[72] R. Henderson. "Joint modelling of longitudinal measurements and event time data". In: *Biostatistics* 1 (4 Dec. 2000), pp. 465–480. ISSN: 14654644. DOI: 10.1093/biostatistics/1.4.465.

[73] Graciela Muniz Terrera et al. "Joint Modeling of Longitudinal Change and Survival". In: *GeroPsych* 24 (4 Dec. 2011), pp. 177–185. ISSN: 1662-9647. DOI: 10.1024/1662-9647/a000047.

[74] Nalini Ravishanker and Renjie Chen. "Topological Data Analysis (TDA) for Time Series". In: *arXiv e-prints* (2019), arXiv–1909.

[75] Nalini Ravishanker and Renjie Chen. "An introduction to persistent homology for time series". In: *Wiley Interdisciplinary Reviews: Computational Statistics* 13.3 (2021), e1548.

[76] Siamak Ardekani et al. "Computational Method for Identifying and Quantifying Shape Features of Human Left Ventricular Remodeling". In: *Annals of Biomedical Engineering* 37 (6 June 2009), pp. 1043–1054. ISSN: 0090-6964. DOI: 10.1007/s10439-009-9677-2.

[77] Beniamin Di Veroli et al. "Graph-Theoretic Automatic Lesion Tracking and Detection of Patterns of Lesion Changes in Longitudinal CT Studies". In: 2023, pp. 106–115. DOI: 10.1007/978-3-031-43904-9_11.

[78] Inês Sousa. "A Review on Joint Modelling of Longitudinal Measurements and Time-To-Even". In: *REVSTAT-Statistical Journal* 9.1 (2011), pp. 57–81.

[79] Haidy Nasief et al. "A machine learning based delta-radiomics process for early prediction of treatment response of pancreatic cancer". In: *npj Precision Oncology* 3 (1 Oct. 2019), p. 25. ISSN: 2397-768X. DOI: 10.1038/s41698-019-0096-z.

[80] Benito Farina et al. "Integration of longitudinal deep-radiomics and clinical data improves the prediction of durable benefits to anti-PD-1/PD-L1 immunotherapy in advanced NSCLC patients". In: *Journal of Translational Medicine* 21 (1 Mar. 2023), p. 174. ISSN: 1479-5876. DOI: 10.1186/s12967-023-04004-x.

[81] Benoît Sauty and Stanley Durrleman. "Progression Models for Imaging Data with Longitudinal Variational Auto Encoders". In: 2022, pp. 3–13. DOI: 10.1007/978-3-031-16431-6_1.

[82] Yiheng Li et al. "A 3D lung lesion variational autoencoder". In: *Cell Reports Methods* 4.2 (2024).

Layman's Summary

Radiomics in CT imaging for Longitudinal Analysis in Oncology: Emerging Complements, Advantages and Limitations

For the past decade, significant advancements in the medical imaging field have permitted the development of more advanced methods to study the characteristics beneath medical images. Radiomics is a technique that uses complex mathematical algorithms to extract and analyse information that could be invisible to the naked human eye. This approach allows the calculation of many different features to find connections between the image data and the studied condition, helping physicians make diagnoses, track responses to treatments, or predict the survival of patients. Radiomics has been mostly applied to oncology, where longitudinal computed tomography (CT) scans are especially useful. Longitudinal studies utilize repeated images at multiple time points to track how the disease changes, giving a clearer picture on how the condition might progress. Traditional radiomic feature extraction approaches such as texture or shape analysis have the potential to monitor those changes. However, they may not always capture accurately the complete tumor behavior because they can be sensitive to noise or image artifacts, i.e. errors or distortions in the images. Moreover, current statistical methods are also being reevaluated for their multiple limiting assumptions that hardly hold for these type of images. Since the research community has not yet found a standard methodology for analysing images over time, and the limitations of the currently used strategies, this review aims to give an overview of the field of radiomics for studying longitudinal CT images, highlighting possible feature extraction and data analysis alternatives to traditional methods, challenges, and future directions in the field.

Feature extraction is the process of calculating quantitative features from images. In this review we focus on the potential use of topological data analysis (TDA), persistent homology and geodesic geometry. TDA is a method that extracts structures (features) by connecting individual points, identifying for example how separated are different groups, if there are any gaps between them, or if more complex arrangements are present. Persistent homology is a method in TDA that tracks how those features appear and disappear based on a scale parameter, allowing for a deeper understanding of the structures and texture beneath the images. On the other hand, geodesic geometry is a type of geometric analysis that describes shapes in complex surfaces, defining the shortest path from one point to another. Geodesics include additional geometric details, which can help provide a more detailed analysis of how shapes change over time.

Data analysis is the process of studying the extracted features and developing systems to improve medical decisions for diagnosis, disease prediction, or treatment planning. In this paper we discuss joint models. Joint statistical models describe how features change over time by using mixed-models to predict the trajectory of the features individually, allowing for more accurate and precise estimations.

Although the discussed methods have demonstrated to be strong candidates for feature extraction and data analysis of longitudinal CT images, they have limitations. For instance, TDA and persistent homology do not have enough validation and background studies in the medical imaging field for time dependent data, geodesic geometry is computationally complex as it requires a strong foundation of mathematical principles, and joint models involve making several challenging decisions. In conclusion, despite the challenges, the integration of more sophisticated and complex methodologies for feature extraction and data analysis offer extra valuable knowledge into disease progression, making them strong candidates for the improvement of treatment planning and prediction models in clinical oncology.