


Università degli Studi di Napoli Federico II

Ph.D. Program in

Computational and Quantitative Biology

XXXVI Cycle

Thesis for the Degree of Doctor of Philosophy

Computational Approaches to Support Clinical Decision on Different Scales of Biomedical Images in Oncology

by

Valentina Brancato

Advisors: Dr. Marco Aiello, Prof. Francesco Isgrò

Co-advisor: Prof. Giuseppe Aceto

Scuola Politecnica e delle Scienze di Base

Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione

To my beloved family,
Mom, Dad, Ciullo, Andrea, and...
With you by my side, I am never alone.

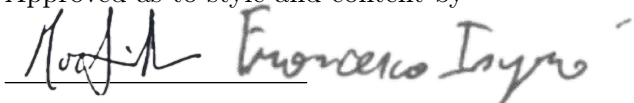
Computational Approaches to Support Clinical Decision on Different Scales of Biomedical Images in Oncology

Ph.D. Thesis presented
for the fulfillment of the Degree of Doctor of Philosophy
in Computational and Quantitative Biology
by

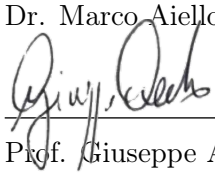
Valentina Brancato

March 2024

Approved as to style and content by



Dr. Marco Aiello, Prof. Francesco Isgrò, Advisors



Prof. Giuseppe Aceto, Co-advisor

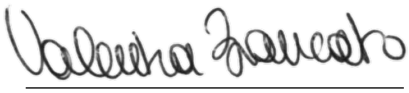
Università degli Studi di Napoli Federico II
Ph.D. Program in Computational and Quantitative Biology
XXXVI cycle - Chairman: Prof. Michele Ceccarelli

https://cqb.dieti.unina.it/

Candidate's declaration

I hereby declare that this thesis submitted to obtain the academic degree of Philosophiæ Doctor (Ph.D.) in Computational and Quantitative Biologyis my own unaided work, that I have not used other than the sources indicated, and that all direct and indirect sources are acknowledged as references.
Parts of this dissertation have been published in international journals and/or conference articles (see list of the author's publications at the end of the thesis).

Napoli, March 9, 2024



Valentina Brancato

Abstract

In the era of precision medicine, the integration of heterogeneous data across multiple scales is crucial for advancing cancer diagnosis and prognosis. Despite significant advancements in diagnostic techniques and computational analysis led to a big data production, these data are often characterized by a high heterogeneity and complexity, belonging to different domains, and their integration remains an open challenge. This PhD thesis explores a possible solution through Computational Decision Support Systems employing AI and data analytics for aggregating, structuring, and comprehending biomedical data, with a focus on medical imaging, that have historically played a crucial role in cancer screening, diagnosis, staging, and therapeutic response monitoring. The thesis first emphasizes the urgency of organizing complex biomedical data belonging to different diagnostic domains, framing digital biobanks as a multifactorial solution capable of containing curated and standardized imaging data, along with clinical, molecular, and pathologic data. Based on this foundation, the thesis delves into the development of computational and statistical tools for analyzing biomedical images across different imaging scales. The correlation between radiomic and pathomic features was explored, particularly in the context of Glioblastoma Multiforme. Preliminary findings revealed intriguing cross-scale relationships, offering a nuanced understanding of tumor heterogeneity and impacting diagnostic, prognostic, and therapeutic considerations. The concept of a "virtual biopsy" emerges, representing a transformative shift in diagnostic methods by relying on advanced analyses of radiological images. Furthermore, the thesis demonstrates the practical applications of pathomics in cancer diagnosis, exemplified through studies in breast and prostate cancers. The pathomic approach proved valuable in quantifying tumor-infiltrating lymphocytes in breast cancer and improving Gleason grading in prostate cancer. These applications showcase the transformative potential of pathomics in enhancing diagnostic precision and treatment strategies, positioning it as a key player in the future of oncology. In summary, this PhD

thesis makes significant contributions to the field by establishing a foundation for advancing precision medicine through computational approaches in oncological decision-making. The proposed model for digital biobanks, the exploration of cross-scale relationships in glioblastoma, and the practical applications of pathomics collectively contribute to the development of comprehensive tools that have the potential to revolutionize cancer diagnosis and decision support. This work serves as a compelling call to action for the scientific community to embrace these innovative approaches and drive positive changes in clinical practice, ultimately improving patient outcomes in oncology.

Keywords: Medical Imaging, Clinical Decision Support Systems, Machine Learning, Oncology, Radiomics, Digital Pathology, Pathomics.

Sintesi in lingua italiana

Nell'era della medicina di precisione, l'integrazione di dati eterogenei su diverse scale è cruciale per avanzare nella diagnosi e prognosi del cancro. Nonostante significativi progressi nelle tecniche diagnostiche e nell'analisi computazionale abbiano portato a una produzione di big data, tali dati sono spesso caratterizzati da un'alta eterogeneità e complessità, appartengono a diversi domini diagnostici, e la loro integrazione rimane una sfida aperta. Questa tesi di dottorato propone una soluzione attraverso Sistemi Computazionali di Supporto alle Decisioni che impiegano l'Intelligenza Artificiale e l'analisi dei dati per aggregare, strutturare e comprendere dati biomedici, con un focus sulle immagini mediche, che hanno sempre giocato un ruolo cruciale nello screening, nella diagnosi, nella stadiazione e nel monitoraggio della risposta terapeutica in ambito oncologico. La tesi sottolinea innanzitutto l'urgenza di organizzare dati biomedici complessi appartenenti a diversi domini diagnostici, inquadrando le biobanche digitali come soluzione multifattoriale in grado di contenere dati di imaging curati e standardizzati, insieme a dati clinici, molecolari e patologici. Sulla base di questa premessa, la tesi approfondisce lo sviluppo di strumenti computazionali e statistici per analizzare immagini biomediche su diverse scale di imaging. È stata esplorata la correlazione tra descrittori numerici radiomici e pathomici, in particolare nel contesto del Glioblastoma Multiforme. Le scoperte preliminari hanno rivelato interessanti relazioni trasversali tra le scale, offrendo una maggiore comprensione dell'eterogeneità tumorale ed influenzando considerazioni diagnostiche, prognostiche e terapeutiche, facendo emergere l'importanza del concetto di "biopsia virtuale". Inoltre, la tesi dimostra le applicazioni pratiche della patomica nella diagnosi del cancro, esemplificate da studi su carcinoma mammario e prostatico. L'approccio patomico si è rivelato prezioso nella quantificazione dei linfociti infiltranti il tumore nel carcinoma mammario e nel miglioramento della classificazione di Gleason nel carcinoma prostatico. Queste applicazioni mostrano il potenziale trasformativo della pathomica nel migliorare la precisione diagnostica e le strategie di trattamento, posizionandola come un attore chiave nel futuro dell'oncologia. In sintesi, questa tesi di dottorato apporta significativi contributi al campo stabilendo una base per avanzare nella medicina di precisione attraverso approcci computazionali nel decision-making in ambito oncologico. Il modello proposto

per le biobanche digitali, l'esplorazione delle correlazioni tra dati estratti da immagini a differenti scale di imaging, e le applicazioni pratiche della pathomica contribuiscono collettivamente allo sviluppo di strumenti completi con il potenziale di rivoluzionare la diagnosi del cancro e il supporto decisionale. Questo lavoro si presenta come un appello avvincente alla comunità scientifica affinché adotti questi approcci innovativi e guidi cambiamenti positivi nella pratica clinica, migliorando alla fine gli esiti dei pazienti in oncologia.

Parole chiave: Immagini mediche, Sistemi di supporto alle decisioni cliniche, Apprendimento Automatico, Oncologia, Radiomica, Patologia Digitale, Patomica.

Acknowledgements

The completion of this thesis has been made possible through the support provided by the University of Napoli Federico II. The university's commitment to academic excellence and research has played a pivotal role in facilitating my academic journey.
I extend my sincere gratitude to IRCCS SYNLAB SDN for their collaboration and support in the field of medical imaging. The opportunities afforded by IRCCS SYNLAB SDN have significantly contributed to my professional development and the advancement of scientific knowledge.
I wish to express my appreciation to my tutors and co-tutor, Dr. Marco Aiello, Prof. Francesco Isgrò, and Prof. Giuseppe Aceto for their guidance, expertise, and encouragement throughout this endeavor.

Contents

List of Acronyms

The following acronyms are used throughout the thesis.

ABRF Association of Biomolecular Resource Facilities

AGGC22 Automated Gleason Grading Challenge 2022

AI Artificial Intelligence

ADC Apparent Diffusion Coefficient

AUC Area Under the ROC curve

BBMRI-ERIC Biobanking and BioMolecular resources Research Infrastructure-European Research Infrastructure Consortium

BC Breast Cancer

BF Bayes Factor

BRAINS Brain Images of Normal Subjects

CAMDA Critical Assessment of Microarray Data Analysis

CDB Comprehensive Digital Biobank

CDISC	Clinical Data Interchange Standards Consortium
CDSS	Clinical Decision Support System
CNN	Convolutional Neural Network
COINS	Collaborative Informatics and NeuroImaging Suites
CPTAC	Clinical Proteomic Tumor Analysis Consortium
CRDC	Cancer Research Data Commons
CT	Computed Tomgraphy
CV	Cross-Validation
DCE-MRI	Dynamic Contrast-Enhanced Magnetic Resonance Imaging
DICOM	Digital Imaging and Communications in Medicine
DL	Deep Learning
DNA	Deoxyribonucleic Acid
DT	Decision Tree
DWI	Diffusion Weighted Imaging
DWT	Discrete Wavelet Transform
EIBALL	European Imaging Biomarkers Alliance
ESR	European Society of Radiology
FAIR	Findable, Accessible, Interoperable and Reusable
FDA	Food and Drug Association
FDR	False Discovery Rate

FFPE	Formalin-Fixed Paraffin-Embedded
FGED	Functional Genomics Data Society
FSD	Fourier Shape Descriptor
GAGH	Global Alliance for Genomics and Health
GBM	Glioblastoma Multiforme
GIAB	Genome in a Bottle
GDPR	General Data Protection Regulation
GLCM	Gray Level Co-occurrence Matrix
GLDM	Gray Level Dependence Matrix
GLRLM	Gray Level Run Length Matrix
GLSZM	Gray Level Size Zone Matrix
HER2+	Human Epidermal Growth Factor Receptor 2
HL7	Health Level 7
HE	Hematoxylin and Eosin
HTAN	Human Tumor Atlas Network
IBSI	Image Biomarker Standardization Initiative
IDC	Imaging Data Commons
ISBER	International Society for Biological and Environmental Repository
ISO	International Organization for Standardization

JSON JavaScript Object Notation

KNN K-Nearest Neighbors

LASSO Least Absolute Shrinkage and Selection Operator

LBP Local Binary Pattern

LDA Linear Discriminant Analysis

LONI Laboratory of Neuro Imaging

MAQC MicroArray Quality Control Project

MIABIS Minimum Information About BIobank Data Sharing

MIAME Minimum Information About a Microarray Experiment

MIGS Minimum Information About a Genome Sequence

MIMS Minimum Information About a Metagenomic Sequence

MINSEQE Minimum Information About a high throughput SEQuencing experiment

MDTA Material and Data Transfer Agreement

ML Machine Learning

MLP Multi-layer Perceptron

MRI Magnetic Resonance Imaging

NDA National Institute of Mental Health (NIMH) Data Archive

NGS Next-Generation Sequencing

NGTDM Neighboring Gray Tone Difference Matrix

NIfTI Neuroimaging Informatics Technology Initiative

NIST National Institute of Standards and Technology

NTO Normalization and Transformation Ontology

OMOP-CDM Observational Medical Outcomes Partnership-Common Data Model

PACS Picture Archiving and Communication System

PCa Prostate Cancer

PET Positron Emission Tomography

QIBA Quantitative Imaging Biomarkers Alliance

RF Random Forest

RNA Ribonucleic acid

ROC Receiver Operating Characteristic

ROI Region Of Interest

RQS Radiomics Quality Score

SMOTE Synthetic Minority Oversampling Technique

SNP Single Nucleotide Polymorphism

SOP Standard Operating Procedure

SR Structured Reporting

SVS Aperio ScanScope Virtual Slide

T1C Post-Contrast T1

TIGER Tumor Infiltrating lymphocytes in breast cancER

TILs Tumor-Infiltrating Lymphocytes

TCIA The Cancer Imaging Archive

TCGA the Cancer Genome Atlas

TIFF Tagged Image File Format

TNBC Triple-Negative Breast Cancer

TME Tumor Microenvironment

VCF Variant Call Format

VOI Volume Of Interest

WG Working Group

WMA World Medical Association

WSI Whole-Slide Image

XGB eXtreme Gradient Boosting

XNAT Extensible Neuroimaging Archive Toolkit

XTENS eXTensible Environment for NeuroScience

List of Figures

List of Tables

Chapter 1

Introduction

1.1 General Context

Cancer is a complex and heterogeneous disease, characterized by multiple factors that contribute to its development and progression [155]. In the era of precision medicine, the ability to anticipate and steer cancer diagnosis and prognosis, as well as obtain a comprehensive understanding of the disease, requires integrating information across multiple scales, from the patient to the molecular level [155, 56, 244]. Clinical imaging, pathology and next-generation sequencing (NGS) represent the most advanced diagnostic domains, providing the most significant amount of information for clinical research and decision-making in oncology [66]. This information can surely offer a complete view of complex biological phenomena at different scales (such as histological imaging and clinical imaging) and different characteristics (genotypic and phenotypic) of complex biological phenomena, making it critical to make the data accessible and usable to advance scientific research and technological progress.

Within the healthcare system, the rapid development of both modern diagnostic techniques and computational analysis methods has led to an explosion of data production, much of which is heterogeneous, belonging to different diagnostic domains, and often extremely complex [32]. Moreover, even when these data are made available, they are rarely integrated, and there have been limited reported advancements in leveraging computational methods to explore the research potential that arises

from integrating data across different diagnostic domains. Here, Clinical Decision-Support Systems (CDSS), powered by Artificial Intelligence (AI) and data analytics techniques can have enormous potential to aggregating, structuring, and understanding data, turning them into a new generation of diagnostic and prognostic models that can support clinical decision process [66, 168]. In fact, the ongoing development of CDSS in clinical practice and the establishment of interdisciplinary workflows will enable holistic medical approaches to improve the diagnosis and treatment of individual patients, creating new opportunities for the understanding of complex tumor characteristics, driving personalized diagnosis and treatment [289, 321]. Nevertheless, the potential of computational techniques often goes unrealized in biomedical contexts, where the availability of suitable datasets is critical. Focusing on the challenges mentioned above, the next sections will introduce the different imaging domains covered by this work, both from the point of view of data management and computational analysis.

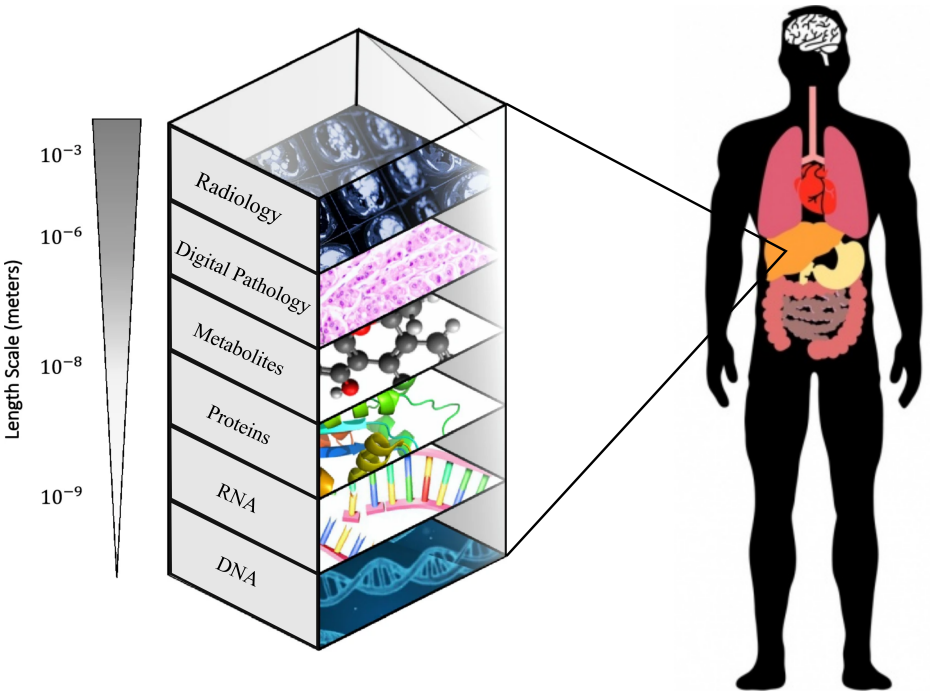


Figure 1.1. Information across multiple scales. Adapted from [179]

1.2 Medical Imaging at Different Imaging Scales

Medical images have historically played a crucial role in cancer screening, diagnosis, staging, and therapeutic response monitoring. Physicians constantly rely on medical imaging inputs to formulate patient management plans, especially within the context of modern oncological guidelines, where patients are stratified into increasingly complex subgroups based on biological, clinical, histopathological and radiological parameters [231]. Medical imaging is a broad term including multiple techniques used to view the human body with the aim to diagnose, monitor, or treat medical conditions. Common imaging techniques include ultrasound, X-ray, computed tomography (CT), magnetic resonance imaging (MRI), and histopathology. Modern oncology and medical imaging are now inseparable. In this thesis, the scope of medical images is limited to radiology images (in par-

ticular MRI) and histopathology whole-slide images (WSI).

The digitization process of radiology departments around the world has been ongoing for the past decades. This development is more recent in pathology, where the transition to digital pathology is ongoing in many countries and healthcare organizations. The workload of both radiologists and pathologists has increased in recent years due to increased volume of the generated data and increased complexity and granularity of diagnoses. Many argue that digital diagnostic tools may assist in lightening the load by automating tasks and reducing the reading time per case. The radiology and pathology specialties are similar in the clinical workflow but different in what types of images they produce. The radiologist and pathologist alike assess images that form the base for diagnosis and treatment suggestions. This subsection aims to give an overview of the clinical pipeline of acquisition of both pathology and radiology images as well as their specific characteristics.

1.2.1 Digital Pathology

Digital pathology is a general term that refers to the assemblage of digital workflow and imaging solutions that are geared towards creating a digital image-based practice environment in which a WSI or other digital image is acquired, managed, interpreted and searched for specific content [28]. The digital pipeline for preparing tissue is typically a sequential process with multiple steps and is, in large, the same as when the sample is reviewed under a microscope. The resected tissue arrives from surgery or biopsy, where the sample has been placed in formalin to preserve the tissue and prevent decay (fixation). The first step in the lab is to determine which parts of the tumor to section (gross examination). Once relevant pieces are sectioned, they are dried and put in a paraffin embedding to make cutting easier (embedding). Once the sample has been embedded, micrometer-thin sections are cut using a special tool called microtome, and are put on glass slides (sectioning). Human cells are practically translucent in their natural state. Therefore, the sample is stained to increase the contrast between relevant cell structures (staining). In the digital pipeline, there is a final step of scanning the glass slides in high magnification so they can be reviewed on a computer (scanning). The final review and diagnosis are traditionally made using a light microscope, but the digitization allows

reviewing on a computer using a specific viewer (review) [52].

The most common staining is hematoxylin and eosin (H&E), where hematoxylin stains cell nucleus purple, and eosin stains extra-cellular matrix and cytoplasm pink [120]. As not all structures are visualized with H&E, special stains are sometimes required. One special technique for imaging cellular structures is immunohistochemistry (IHC) [206]. IHC utilizes antibodies together with a staining agent that attaches to antigens in cells and tissue such that the presence or absence of specific antigens can be detected. This technique has become increasingly common in classifying and detecting cancer tumor cells, and may provide important prognostic information, helping specialize treatment. The pathologist's task is to examine the sample in high-resolution, both on a macro and on a micro level, to diagnose and give treatment guidelines. Different cellular appearances indicate the type of disease (such as cancer or bacterial infection) and the extent and aggressiveness of the malignancy. For many cancer diagnoses, the pathologist determines the size of the tumor and how close it is to the resection line (if the tumor is close, it means that parts of the tumor may not have been removed), and if the tumor has spread to nearby tissue or lymph nodes. Additional staining using IHC may further characterize the tumor [206]. The results of several smaller tasks give an overall description of the tumor, such that treatment can be personalized, preventing both under- and over-treatment of the patient. This thesis only relates to histopathology, the study of disease in cells and tissue, and does not include related disciplines such as cytology, bacteriology or hematology.

1.2.1.1 Whole Slides Images

Whole slide imaging is the process of creating a single high-resolution digital image from a glass slide, such that the slides can be reviewed digitally. A whole slides image is a high-resolution, large format scan of human tissue. WSIs are also called microscope slides, as they represent an enlarged image of cells and subcellular structure. Such slides are obtained from making multiple scans of different areas of a tissue sample and aggregating them into a large image [214]. Glass slides are scanned at high resolution, resulting in giga-pixel-sized images. In light microscopy, magnification has been the convention to denote the size of structures, which differs from resolution. Despite that digitization of images results

in a resolution (measured in microns per pixel (mpp)) rather than a magnification, the level of magnification is often still reported. For example, many scanners scan at "40x", which denote 400-times magnification, corresponding to (approximately) 0.25 mpp. Figure 1.2 shows an example of a WSI at different resolutions. Compared to the traditional microscopic analysis of histology, WSI provides a number of advantages. First of all, WSI usually come in giant resolutions like 100,000 x 100,000 pixels or even bigger, which gives a lot of clarity when visually inspecting tissue areas. Then, WSI is a digital scan that covers much larger areas all at once and is easier to zoom in/zoom out, compared to a microscopic view. Moreover, the image can be sent between the healthcare experts as a digital copy. Additionally, WSI scans can't deteriorate like specimens on a glass slide. However, in terms of operability, WSI analysis is not so great due to the time and effort needed during the manual examination. In real world practice, a doctor would sit before the image scanning by eye, zooming in and out for hours. In the case of manual examination, a whole slide should be examined to spot suspicious cell structures or mutations, because skipping some areas may result in medical error. WSI files are written to permanent storage (such as hard disks) in specific file formats. Unfortunately, due to a lack of standardization by vendors, a wide variety of different WSI file formats exist, and these are frequently proprietary formats. Most, but not all, of the vendor formats are based around the "tiled pyramid" paradigm. In the "tiled pyramid" approach, the WSI file contains the base image layer (the full resolution scan) as well as a number of pre-calculated zoom levels which are typically scaled by a factor of $1/4$ or $1/2$ at each level [214, 338].

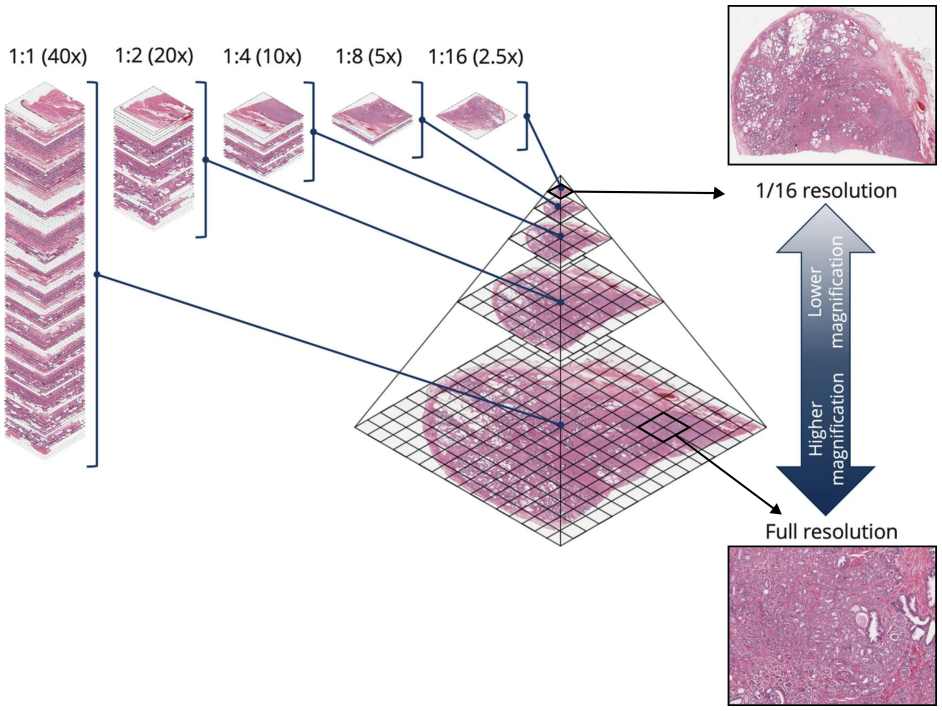


Figure 1.2. Whole Slide Image pyramid structure. Each pyramid level represents a different zoom level. The base level represents the full resolution image. The lowest magnification image is at the apex of the pyramid. Adapted from [214]

1.2.2 Clinical Imaging (Radiology)

Modern oncology and clinical imaging are now inseparable. Management of cancer involves many technical aspects that can be separated into four areas: screening, diagnosis, treatment and/or monitoring, and follow up. Clinical imaging is interwoven into every stage, and as a result, each offers fertile ground for imaging biomarker research which seeks to improve cancer management [286, 144]. Medical imaging reveals the hidden anatomical structures and biological functions of the human body. It is an incredible advantage to peer past the skin to assess a condition without an invasive surgical procedure for the patient, and imaging techniques are now indispensable in modern medicine as a result. There are a

variety of distinct image acquisition modalities that utilise different physical phenomenon to produce a functional or anatomical image, including high energy x-rays, radioactive decay, and magnetic resonance. Human with cancer or other diseases who need immediate treatment requires clinical imaging to help clinicians make correct diagnosis and select optimal treatment options. Clinical imaging modalities play an important role in all aspects of disease treatment, including screening, detection, diagnosis, image-guided biopsy, treatment planning, outcome assessment and follow-up examinations. Due to the complexity of cancer, no single modality can provide a complete overview, hence the prevalence of multimodal approaches that combine imaging techniques for a more complete tumour characterisation. Multimodal imaging is the merging of single modalities to provide complementary information. Usually, it involves a composite of functional and anatomical approaches [294]. For a particular aliment, a clinician utilises the modality (or modalities) that best maximises the image clarity needed for assessment whilst minimising any potential side effects or discomfort for the patient. The collection of modalities at their disposal spans a broad measurement scale, ranging from microscopic analysis of cell structure, to macroscopic whole-body assessment of organ and bone. Many of these modalities have potential for quantitative characterisation of disease using radiomic techniques, with a clear application in oncology as focused on in this body of work. Particularly, these include: Computed Tomography (CT), Positron Emission Tomography (PET), Single Photon Emission Computed Tomography (SPECT), Magnetic Resonance Imaging (MRI), and Ultrasound (US). This thesis utilises MRI, which is one of these key quantitative modalities and is prevalent in cancer management together with CT and PET, and as such it is introduced in more detail in the following paragraph. The following sections provide an introduction to MRI, which is, together with CT and PET, one of the major modalities utilised in the field of radiomics, and will be examined in this thesis.

1.2.2.1 Magnetic Resonance Imaging

MRI has emerged as a cornerstone in diagnostic imaging, employing a non-invasive approach that relies on the principles of magnetic fields and radio waves. In contrast to traditional imaging techniques using ionizing radiation, MRI showcases unparalleled capabilities in visualizing soft tis-

sues with remarkable detail and without harmful side effects. This imaging modality has become indispensable in the realm of oncology, where precision and comprehensive understanding are paramount [129, 173]. One facet of MRI that plays a pivotal role in oncological investigations is Diffusion-Weighted Imaging (DWI). DWI is founded on the principles of molecular motion within tissues, specifically the movement of water molecules. By capturing and analyzing this diffusion, DWI provides crucial insights into cellular density and microstructural variations. In the context of oncology, DWI becomes an invaluable tool for characterizing tumors. The ability to identify regions of restricted diffusion aids in mapping tumor boundaries, evaluating heterogeneity, and contributing significantly to treatment planning. The nuanced information gleaned from DWI enhances our understanding of tumor architecture and assists in tailoring therapeutic strategies to the individual characteristics of each case [50, 218]. Dynamic Contrast-Enhanced (DCE) MRI represents another sophisticated dimension of oncological imaging. DCE-MRI involves the introduction of contrast agents into the bloodstream, typically gadolinium-based, to accentuate vascular and perfusion dynamics within tissues. This technique offers real-time visualization of blood flow, capillary permeability, and tissue enhancement. In the oncological context, DCE-MRI is instrumental in unraveling the intricate details of tumor vascularity and microvascular permeability, providing crucial information about angiogenesis, that is a hallmark of tumor development. The ability to monitor treatment responses and predict outcomes based on the evolving vascular characteristics of tumors underscores the significance of DCE-MRI in guiding clinical decisions and optimizing therapeutic interventions [233]. In conclusion, the synergy of MRI, DWI, and DCE-MRI in the field of oncology represents a sophisticated and powerful arsenal for comprehensive tumor assessment. As the landscape of medical imaging continues to evolve, these techniques stand at the forefront of advancements, offering clinicians unparalleled insights into the complexities of cancer, ultimately contributing to improved patient care and outcomes. Until now, there are still differences in diagnostic opportunities between clinical imaging on the one side and digital pathology on the other. In clinical imaging, non-invasive techniques such as X-ray and ultrasound are used to define tumor infiltrates. The applied techniques lead to macroscopic images that enable the measurement of

structures in the resolution of cm and mm. Additionally, biological processes can be visualized with the help of tracers. In contrast, pathological examinations are based on invasively obtained biopsy material. Tumor characteristics are defined at a microscopic scale, ranging from mm to µm. In pathology, immunohistochemistry and immunofluorescence provide high specificity and sensitivity to detect and characterize a positive cell within a population of thousands of negative cells. Commonly applied methods are light microscopy, confocal laser, spinning disc laser microscopy, and light sheet microscopy. Even higher resolutions are possible by electron microscopy and cryo-electron tomography. The information of the cell can be additionally enriched by molecular features through the detection of proteins (immunohistochemistry), RNAs in situ hybridization, and DNA mutation analysis (deep sequencing, methylation, and others). Apart from the different opportunities that arise from these medical sub-disciplines, they also differ in the challenges that need to be overcome [212]. Imaging techniques in radiology and nuclear medicine face up with lower resolution than microscopic images in pathology. On the other hand, the high resolutions biopsy material has the downside of a restricted and limited area that can be examined. There is also a risk of clinical complications in pathology due to the invasiveness of tissue extirpation. The resolution of clinical images is much lower than that of pathology. Clinically collected volumes using conventional CT typically have an in-plane resolution of 0.5 mm per pixel, slightly lower for MRI at 1.5 mm per pixel, compared to 0.5 - 0.25 microns (0.0005 mm) per pixel in clinically collected WSI. However, recent advances in CT technology enable higher resolution images. These types of photon-counting machines are just starting to be used clinically at the time of writing. With increased resolution, these images close the gap between radiology and pathology, and introduces interesting research opportunities, which may lead to better diagnostic possibilities. Moreover, another notable distinction lies in the image dimensions encountered in these two medical imaging modalities. While radiology and nuclear medicine typically produce images with manageable file sizes, WSIs in pathology can be exceptionally large. WSI, capturing entire tissue sections at high resolution, can easily reach gigabytes in size. Managing and processing such extensive data sets pose significant challenges, from storage requirements to computational demands. The digitalization of pathology, though trans-

formative, has encountered obstacles in terms of data handling, hindering the seamless integration and analysis of vast datasets. The digitalization of pathology is a more recent development, and the comprehensive integration of molecular and pathological data is an ongoing area of research [205]. Additionally, the lack of large publicly available datasets in pathology hampers the development and validation of computational models, limiting progress in automated analysis techniques. Of note, both types of images at different scales can play a pivotal role in CDSS. This concept will be better highlighted in the following paragraph.

1.2.3 Computational Image Analysis

Historically, qualitative semantic features were used to describe tumor morphology observed in the patient images at different imaging scales. These descriptions were based on a scoring system relying on visual assessment. Semantic features were shown in the literature to have correlations with stage, prognosis, and even response prediction. Nevertheless, this approach suffered from limitations due to its dependence on subjective scoring and the limited sensitivity of the human eye. Radiological images generated by modern imaging modalities (e.g. CT, MRI, and PET), along with digital pathology images generated by digital pathology slide scanners, are primarily quantitative in nature. This characteristic is harnessed to extract numerical descriptors and generate mineable data exploiting computational algorithms. These numerical features can be used to objectively characterize tumor morphology, thus transforming the commonly used diagnostic and histopathologic examinations into personalized-medicine tools that are independent from the typical subjective interpretation that characterize the common clinical practice. In the realm of biomedical research and clinical decision-making, the synergy between computational methodologies and advanced imaging technologies has revolutionized our ability to extract intricate information from radiological and histopathological digital images. Two pioneering approaches, radiomics and pathomics, unravel the inherent complexity of these images, employing sophisticated computational algorithms for feature extraction. These extracted features serve as the foundation for predictive modeling of clinically relevant endpoints. Hand-crafted radiomics and pathomics represent a frontier in computational image analysis. In this approach,

domain expertise guides the selection and crafting of features that are deemed most pertinent to the intricacies of radiological or histopathological images. This methodical curation of features allows for a nuanced and tailored analysis, capturing specific aspects of tumor heterogeneity or cellular morphology that may be of paramount clinical significance. Handcrafted radiomics, for instance, involves the deliberate selection of features such as intensity, texture, and shape descriptors from radiological images. Similarly, in the realm of pathomics, features like nuclear morphology, spatial arrangements, and cellular textures are carefully chosen to illuminate distinctive aspects of tissue characteristics.

1.2.3.1 Radiomics (computational analysis of radiological images)

On a radiological scale, radiomics is a new frontier of medicine based on the extraction of numerical descriptors from radiologic images that are imperceptible to the human eye and are able to describe the intratumoral heterogeneity [65, 257, 69]. This could be useful in anticipating diagnosis and improving stratification and follow-up of cancer patients. Moreover, concerning intratumoral heterogeneity, radiomics is often referred as a sort of «virtual biopsy», since it allows to enrich the traditional diagnostic radiologic workflow with more information not detectable by human eye and associated with processes not included in radiologic workflow (e.g., bioptic outcomes such as cellular grading by histopathology inspection, gene expression from genomic analysis) [211, 306, 229, 276]. In general, a typical radiomics study aims to develop a clinically relevant model from imaging data and other clinical measures using machine learning techniques to categorize patients based on anticipated outcomes [182]. Noteworthy is the recent division of radiomics into two distinct domains, with this thesis primarily focusing on the first domain: the potential of hand-crafted features to support clinical decision. These features are mathematically defined to quantify specific aspects of an image region. The second domain pertains to the surge of deep-learning (DL) techniques, utilizing neural networks (e.g., convolutional neural networks (CNNs)) trained to autonomously learn relevant patterns directly from imaging data [130]. This project addresses key challenges within the first domain, with a comprehensive discussion on both domains. The traditional radiomics workflow initially outlined five

major stages that are shown in Figure 1.3: 1) Acquisition, 2) Segmentation, 3) Feature extraction, 4) Feature selection and 5) Model development [182]. However, this simplification belies the complexity of the pipeline needed to transform raw imaging into potentially valuable data. Each stage encompasses numerous sub-stages with unique challenges to address and overcome. Particularly crucial are the steps involved in processing an image between acquisition and feature extraction, as they significantly impact the consistency of extracted features. Before delving into image analysis, obtaining a medical image is a prerequisite. Medical imaging, acquired through scanners like CT, PET, or MRI, involves complex protocols and machine settings, impacting contrast and image quality. Ongoing technical advancements in hardware and software, such as refined reconstruction algorithms, contribute to continuous improvement in imaging acquisition. Understanding the impact of acquisition parameters on feature extraction is crucial for effective radiomics. The second macro-step involves the segmentation, namely the spatial outlining of a region of interest (ROI) or volume of interest (VOI) in an image for subsequent analysis. In oncology, VOIs often define tumor lesions or organs-at-risk susceptible to treatment. Manual contouring, considered the ground truth, is challenging and susceptible to inter and intra-reader variability, especially in 3D tumor volumes segmented slice by slice on the axial plane. To address these challenges, semi-automatic and automatic segmentation methods, such as region growing, clustering, and thresholding algorithms, are recommended in radiomics workflows for their increased robustness and reproducibility. Accurate contour automation not only enhances efficiency but also allows radiation experts to focus on more technical tasks, though specialist approval remains essential. The next step after segmenting the tumor, is to extract radiomics features. Feature extraction involves quantitatively describing the characteristics of a segmented VOI in an image. Types of features to extract depend on the problem at hand. Morphological, statistical, first-order histogram families, along with texture features, contribute to the extensive feature set used in radiomics studies. This thesis explores several feature families. Image processing techniques, including re-segmentation, interpolation, discretization, and filtering, are integral to feature extraction. However, a lack of reporting on implemented techniques impedes the ability to validate promising radiomics models. Selected fea-

tures will then go through a feature selection process, where redundant and non-relevant features are excluded. Once features are selected, radiomics aims to develop prognostic or predictive models based on them, and the methodologies of which vary considerably based on the study's hypothesis. Examples include models for diagnosis, overall survival, tumor aggression, prediction of distant metastases, treatment response, and linking tumor genetics to the imaging tumor phenotype (Radiogenomics studies).

Figure 1.3. Basic components of a radiomic workflow.

1.2.3.2 Pathomics (computational analysis of digital pathology images)

On a histopathological scale, it is well-recognized that the interpretation and evaluation of biological samples at microscopy characteristic of traditional pathology, is complicated and characterized by subjective and qualitative processes [205]. As previously highlighted, digital pathology is a general term that refers to the process of digitizing histopathology, immunohistochemistry or cytology slides using whole-slide scanners, along with the interpretation, management, and analysis of these digitized WSIs using computational approaches (computational pathology) [160, 315]. The digital pathology slides can be stored in a centralized repository, enabling remote access for manual review by a pathologist or automated evaluation by a data algorithm. Computational pathology utilizes advanced computational methods to extract valuable information from high-resolution WSIs that can be correlated with phenotypic features in different types of malignancies in association with the traditional histopathologic evaluation performed by pathologists. The advent of digital pathology gave rise to a recent discipline that is known as pathomics.

This area of research aims to quantitatively characterize cells and tissues obtained by examined samples, as radiomics does for radiological images [134]. The pathomic analysis result is different from a pathology report. The pathology report includes what the pathology sees, while pathomics could allow us to obtain a quantitative and reproducible tissue characterization. Common examples of pathomics applications include spatial characterization of tumor and stromal regions, shapes and textures of nuclei, classifications of cell types, quantitative characterization of lymphocytic infiltration, and efforts to quantitatively estimate the number of cells - labeled with different kinds of biomarkers in immunohistochemistry testing. The pathomics analysis workflow consists of four main steps: the selection of ROIs, color normalization, and the extraction and analysis of pathomics features. Figure 1.4 illustrates a typical pathomics workflow. The process initiates with the acquisition of high-resolution WSIs, capturing microscopic details of tissue samples. Subsequently, image preprocessing is imperative, involving the selection of ROIs, gridding, and tiling to manage vast datasets efficiently. Color normalization techniques are then applied to standardize color variations across images, ensuring consistency in subsequent analyses. Following preprocessing, the extraction of pathomics features becomes central to the workflow. These features can range from hand-crafted features, capturing morphological and textural attributes, to more sophisticated deep learning features, leveraging convolutional neural networks for automated feature extraction. Finally, the extracted features serve as the foundation for model building, where advanced algorithms, including machine learning or deep learning models, are employed to discern patterns, classify tissue types, and predict diagnostic outcomes.

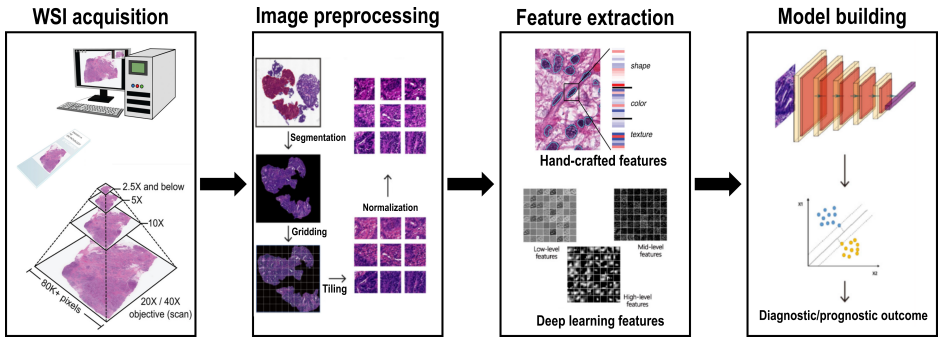


Figure 1.4. Typical pathomics workflow. After collecting and scanning pathological images, the ROI (region of interest) is manually or automatically labeled. Secondly, deep learning features or hand-crafted features are extracted from these images after a series of images pre-processing such as ROI segmentation, gridding, tile extraction, and color normalization. Finally, meaningful features are analyzed by machine learning or deep learning algorithms and classified or predicted according to different tasks.

The promises arising from the analysis of data at different imaging scales holds promises for integrative diagnostic and cell therapy applications, providing a holistic understanding of potential neoplastic changes. It would surely improve diagnostics and molecular knowledge about cancer, for diagnostic, prognostic and therapeutic purposes, and this would have direct implications in the clinical decision-making process [266, 66, 196]. The integration of data at different scales introduces heightened complexity, extending beyond the traditional scope of hypothesis testing and model validation. CDSS are poised to navigate this complexity by developing robust, high-performance algorithms and innovative inferential tools tailored for clinical use. Moreover, the integration of data at different imaging scales could be very useful for the validation of the radiomic approach in clinical practice as a virtual biopsy [229, 306, 194]. This integration could be also favored by the usage of digital biobanking, which is the infrastructure needed to organize and share imaging data curated and standardized, together with clinical, molecular and pathologic data. This is very important for interoperability in healthcare, and to allow a complete approach to cancer disease both from a clinical point of view and also to make these data usable from ML-powered CDSS systems for predicting clinical results

[92, 316, 212]. Finally, the genotype-phenotype correlations would also be favored [198]. At the same time, there are many challenges to face before addressing this integration, such as the lack of shared reference standards concerning data storage, the absence of an agreement on analysis procedures, and limitations in feature reliability and reproducibility that affect both radiomics and pathomics [204, 127, 55, 304].

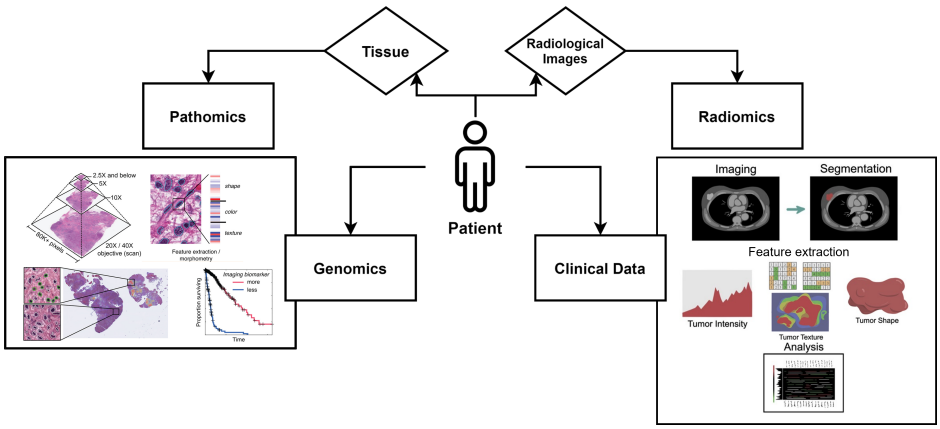


Figure 1.5. Computational Image Analysis at different imaging scales

1.3 Research Objectives and Thesis Contributions

Considering the outlined context, the focus of this thesis is twofold. The first goal is to address the challenges associated with integrating and organizing data across various diagnostic domains in the field of oncology. Simultaneously, the thesis delves into the development of methodologies for extracting and analyzing multiscale imaging data. A primary challenge addressed in the thesis is the integration of heterogeneous and complex biomedical data. The increasing volume of data, driven by advancements in acquisition and computational methods, poses a challenge that the thesis aims to overcome. To this end, contributions were made through the development of procedures for standardizing and harmonizing heterogeneous data from diagnostic imaging, histopathology, and genetics, thereby facilitating the collection of numerical descriptors from patients with oncological diseases. This process is intrinsically linked to the organization of

these data in digital biobanks containing curated and standardized imaging data, along with clinical, molecular and pathologic data in order. This comprehensive organization can contribute to provide a complete approach to cancer disease, promoting scientific research and technological development. The second objective concerns the supervised analysis of multi-scale data to support the clinical decision process. The research activities involved are the implementation and application of processing pipelines for radiomics and pathomics features extraction, respectively from radiological and histopathological images of patients with oncological diseases. Computational and statistical methods are implemented to foster the integration of radiomic domain with other diagnostics domains (pathomic, genomic, clinical) and to investigate the power of radiomics and pathomics numerical descriptors to support CDSS. In particular, the thesis aims at contributes to a deeper comprehension of tumor heterogeneity at different scales by exploring the correlation between radiomic and pathomic features, with the aim to provide a holistic view of cancer from microscopic details to larger-scale images. The thesis thoroughly explores and underscores the applications of pathomics in cancer diagnosis, with the aim to highlight the potential of pathomic information to support clinical decision-making. Integrating pathomic information into the diagnostic workflow is seen as a promising avenue for refining tumor characterization and improving diagnosis accuracy. These contributions and challenges can collectively drive advancements in understanding and managing of oncological diseases through the integration of cross-scale data sources and the development of computational tools.

1.4 Thesis Structure and Overview

The rest of the thesis is structured as follows.

Chapter 2 will emphasize the critical importance of proper data organization, particularly within biobanks. It will discuss the need for clear and shared standards for collecting, storing, and managing data related to radiology, pathology, and genetics. This forms the foundation upon which all subsequent analyses will rely, as disorganized or poorly managed data could compromise the quality and reliability of future AI applications in oncology.

Chapter 3 the development of computational and statistical tools designed for analyzing and integrating data across different scales of biomedical images exploring the two central pillars of radiomics and pathomics. The practical application of these approaches and their potential for supporting the clinical decision process for oncological pathologies will be highlighted.

Chapter 4 will discuss the results of this thesis with a summary of contributions and open perspectives.

Chapter 2

Data Integration and Organization

This chapter delves into the first primary objective of this thesis: integration and organization of data from different domains. The focal point of this part of investigation lies in the development of procedures aimed at standardizing and harmonizing data derived from diagnostic imaging, histopathology, and genetics within the realm of oncological diseases. As introduced in the previous chapter, the advancement in acquisition and computational methods has generated a vast amount of biomedical data which help characterize individual differences in patients and thus can transform the clinical practice into personalized-medicine, independent from the subjective interpretation. However, the huge amount of heterogeneous information belonging to diagnostic domains needs to be available and suitable to promote and support scientific research and technological development, supporting the effective adoption of the precision medicine approach in clinical practice. Digital biobanks can catalyze this process, facilitating the sharing of curated and standardized imaging data, clinical, pathological and molecular data, crucial to enable the development of a comprehensive and personalized data-driven diagnostic approach in disease management and fostering the development of computational predictive models. In this context, the aim of this chapter is to frame this perspective, first by evaluating the state of standardization of individual diagnostic domains and then by identifying challenges and proposing a

possible solution towards an integrative approach that can guarantee the suitability of information that can be shared through a digital biobank. The proposed integrative approach was based on the proposal of new models of standardization/harmonization of each diagnostic domain and of an integration among multidisciplinary domains, together with the harmonization/standardization concerning the generation of numerical descriptors associated with each single domain. An overview of the concept of biobanks, together with challenges associated with both traditional and digital biobanks is provided. Then, how a comprehensive digital biobank could be implemented and standardized, including the identification of key use cases and scenarios is provided. The ultimate goal is to leverage digital biobanks as source of multifactorial information to develop predictive models to support medical decision making towards precision medicine. Part of this chapter was published in [72]

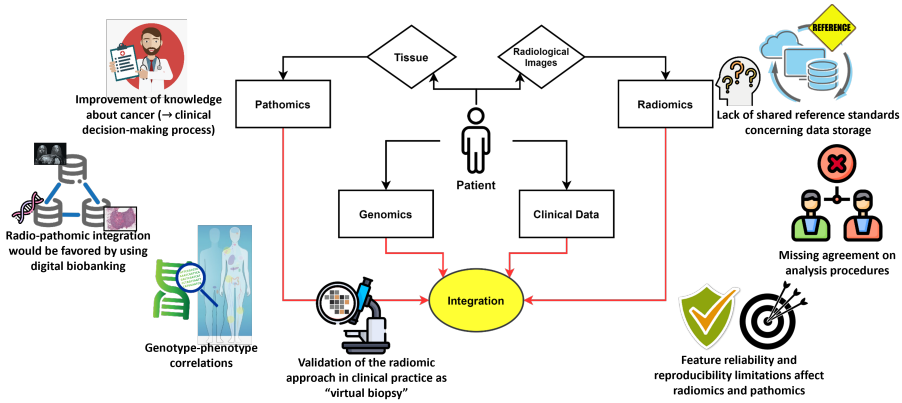


Figure 2.1. Pros and cons associated with the integration of data at different imaging scales.

2.1 Comprehensive banking of biomedical data: towards a common standardization.

2.1.1 Introduction

Within the healthcare system, the rapid development of modern diagnostic techniques has resulted in an explosion of data production, much of which is heterogeneous and belongs to different domains. Clinical imaging, pathology and next-generation sequencing (NGS) represent the most advanced sectors, bringing the most significant amount of information for clinical research and decision-making, especially in the field of oncology. This information can offer a complete view of the complex biological phenomena at different scales (such as histological and clinical imaging) and different characteristics (genotypic and phenotypic) of diseases, making it valuable to release curated raw data [32, 146, 337].

The concept of biobank precisely addresses this need. In fact, the guiding principles of the "Biobank Act" are the promotion of trustworthy, equal access to data and samples, protection of privacy, acceleration of innovation activities and exposing biobank activities to public scrutiny [282]. The aim of the biobanking is to ensure the availability of qualitatively annotated biological samples for planning research programs, and to foster innovative and personalized approach to disease treatment and diagnosis [208]. The more well-characterized, high-quality samples and associated data are available through biobanks, the faster research will progress and impact the today's healthcare. As biobanks are important sources for the provision of research-ready samples as well as associated data, they can face a dual bottleneck of data harmonization and curation [92]. These aspects are interconnected, and both can directly affect the biobank suitability. Variations associated with collecting, processing, and storing procedures make it extremely hard to extrapolate or to merge data from different domains (i.e., imaging, pathology, and molecular profiling). For instance, if one institution uses a different coding system for diagnoses compared to another, merging this data without proper harmonization could lead to erroneous conclusions or missed insights. Another example concerns genomic research that often involves data collected from various laboratories using different techniques and platforms. Harmonizing ge-

netic data requires aligning genetic markers, normalizing gene expression values, and reconciling discrepancies in annotations. Without proper harmonization, comparing data across studies becomes unreliable, impacting the accuracy of genetic associations or findings. Inconsistencies in data collection or storage methodologies could compromise the validity of longitudinal studies investigating disease progression or treatment outcomes. It is easy to introduce invisible bias, leading to irreproducible findings. Therefore, the standardization and harmonization of biobanking practices are of paramount importance [92]. Another fundamental aspect to be explored is the generation of numerical descriptors associated with each single domain, that is a crucial aspect for sharing digital information and passes through the definition of robust data curation and processing procedures. Dealing with standardized and harmonized procedures is a fertile ground for both developing -omics studies (e.g., radiomics, pathomics, genomics), as well as for the exploration of the potential links between -omics quantitative data and clinical outcomes of patients with specific diseases, primarily cancer [66]. In particular, the challenges associated with integrating data from diverse domains are multifaceted and can significantly impact the overall quality and interpretability of integrated datasets. One major obstacle stems from the heterogeneity in data formats across these domains, as each field often adopts distinct formats. Additionally, biological variability, inherent in living systems, manifests differently across domains. Therefore, integrating data across-domains requires careful consideration of biological variations to ensure that observed patterns are genuine and not artifacts of the integration process. Furthermore, differences in data resolutions present another hurdle. While imaging data might possess high spatial resolution, molecular data may operate at the molecular or genomic level. Integrating datasets with varying resolutions necessitates meticulous consideration to prevent loss of information or misinterpretation during the integration process. All these challenges presuppose having high-quality numerical descriptors for each domain.

These considerations highlight the need to put together different diagnostic and clinical domains in a comprehensive manner through a digital approach, while promoting data sharing and biobank sustainability. The concept of digital biobank, namely ecosystem of readily accessible, structured, and annotated datasets that can be dynamically queried and

analyzed [217], can be envisioned as companion infrastructure to support dynamic data access, processing and visualization of the growing data capital in research and healthcare. The digital biobank serves as a backbone structure for integrating diagnostic imaging, pathology, and NGS to allow a comprehensive approach. It should also be considered as a tool for the biomarkers discover and validation in order to define multifactorial precision medicine systems supporting decision making in the medical field [92]. On this premise, the aim of this work is to investigate procedures aiming at standardization and harmonization of data associated with diagnostic imaging, histopathology and NGS. This is directly linked to the potential data management in a digital biobank to allow a complete approach to diseases and also to make these data usable from artificial intelligence (AI) algorithms [124].

The study aims to address the following specific objectives:

- To review the current data standardization and harmonization initiatives, concerning all the domains and shared in common practice;
- To propose a comprehensive digital biobank (CDB) approach that integrates data from different domains. The approach is designed to address the identified standardization and harmonization needs, and could serve as a valuable tool for clinical decision-making in the field of precision medicine.

Figure 2.2 shows an overview of data types included in a CDB. Following the linear sequence for the generation of numerical descriptors, we represented (from the right side) all the data concerning storage and acquisition of a biological (e.g., tissue, blood, etc.) or digital (diagnostic or pathology image) sample together with the data provided by the reporting, curation, and processing procedures. On the other hand, following the vertical line, the integration of the diagnostic domain in each step of the sample lifecycle is achieved. The goal of a comprehensive approach should be the definition and application of specific standards for each kind of data and their integration across the domains. Various sources of information at different scales, e.g., clinical, imaging, pathology, molecular, and all the associated semantic, semi-quantitative and quantitative (of which -omics-) metrics, can support the clinical decision.

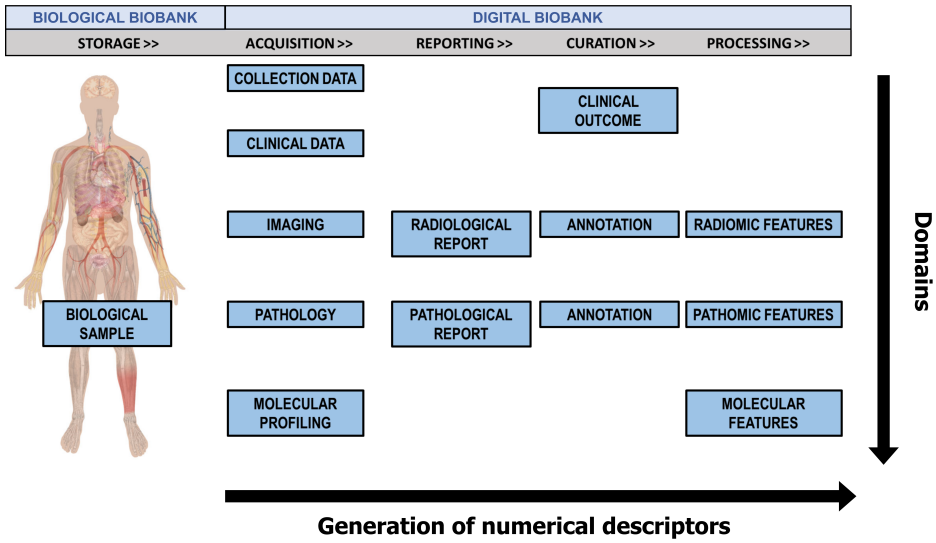


Figure 2.2. Overview of data included in a comprehensive digital biobank according to the generation of numerical descriptors (e.g. radiomic features extracted from radiological images, pathomic features from digital pathology images, as well as molecular features from molecular profiling) during the sample lifecycle (horizontal increasing arrow) and to the integration of different domains (vertical descending arrow).

In the first section the concept of biobank, along with challenges related to both traditional and digital biobanks will be overviewed. In the second section the focus will be on how a CDB could be implemented and standardized, including the identification of key use cases and scenarios addressing specific clinical questions. While always keeping in mind the necessary security and privacy regulations in accessing and retrieving biobank data [106, 281], the missing aspects to envision workable solutions will be tackled. Furthermore, the study aims to emphasize the role of biobanks as a tool for predictive research, interpreting them as source of multifactorial information to set up predictive models supporting decision making in the healthcare system. By evaluating the quality of quantitative features and using the reproducible procedures of the biobank, a standalone system can be established to generate predictive results starting from domain-specific data.

2.1.1.1 Biobanks' overview

From its first use in the scientific literature in 1996 [197], the term "biobank" has been defined in many different ways [274]. Nowadays, there is a consensus to define a biobank as a repository for the storage and retrieval of structured collections of human biological samples and associated data for present and future research use [238]. Data derived from different sources, such as bodily fluids, tissues, skin cells and other biological samples, through genomic and molecular processing, are associated with the medical records and potentially accessible by the researchers. With the growth of the biobanking field, many different types of data have been incorporated in these repositories, from the historical and annotated pathological information associated with the patient clinics to the new data coming from the advent of -omics science (genomics, transcriptomics, proteomics, metabolomics) [92]. International Organization for Standardization (ISO) standards and Standard Operating Procedures (SOPs) are integral aspects of the definition of a biobank.

The importance of standardization becomes evident in numerous challenges that arise when it is lacking in biobanking practices. For instance, without standardized protocols for sample collection and handling, variations in preservation methods and transportation procedures can compromise the quality of stored samples, impacting the reliability of subsequent analyses. In the absence of standardized data management procedures, maintaining accurate and comprehensive records becomes challenging, hindering the reproducibility of studies and traceability of sample characteristics. Ethical and legal compliance faces hurdles due to the absence of uniform guidelines, leading to uncertainties in obtaining informed consent and ensuring participant privacy. Quality control and assurance suffer when standard measures are lacking, risking issues such as contamination and mislabeling. The accessibility and sharing of samples are impeded without standardized practices, hindering global collaborations. Long-term storage conditions and data format standardization become critical for sample stability and integration of diverse datasets, respectively. In essence, adherence to ISO standards and SOPs is essential to address these standardization-related challenges by providing a framework for consistent practices and uphold the fundamental princi-

ples of sample quality, data integrity, and ethical conduct in biobanking. Standardization promotes transparency, reproducibility, and interoperability, ultimately enhancing the reliability and impact of biomedical research [67, 45, 239]. SOPs ensure the correct implementation of essential biobanking components (anonymization, samples' acquisition, transport, preparation, analysis, storage conditions and terms of sharing). Regarding the use of SOPs, a series of technical specifications relating to pre-analytical procedures for human samples developed by ISO/TC 212 are available to biobanks. The final purpose is the standardization of pre and post analytical procedures, analytical performances, laboratory safety, reference systems and quality assurance. Unfortunately, although promoted by international networks, this process is still not homogeneous and adopted by all biobanks. Furthermore, it is necessary to share an internationally accepted and implemented ISO standard for biobanks, as the recently introduced ISO 20387:2018 [101]. It specifies the general competence, impartiality, and consistent functioning requirements for biobanks, including quality control requirements, to ensure biological materials and associated data of proper quality. This document applies to all organizations that conduct biobanking procedures for research and development purposes, including the management of biological samples for the study of circulating tumor markers. The global adoption of the ISO 20387:2018 standard for biobanking faces multifaceted challenges rooted in the diverse nature of biobanks and the substantial commitment required for accreditation. One key challenge involves the shift from the ISO 9001:2015 certification scheme to the more comprehensive ISO 20387:2018 accreditation, necessitating a transition from a focus on operational aspects to governance and management considerations. The extensive documentation and rigorous internal processes required for accreditation demand a significant investment of time, skilled resources, and financial commitment, ranging from EUR 15–25 K. Additionally, the varied sizes and capabilities of biobanks introduce challenges related to the implementation of standardized Quality Management Systems, infrastructure upgrades, and personnel training. Overcoming these challenges requires the development of flexible guidelines, financial support mechanisms for resource-limited biobanks, and a concerted effort to raise awareness globally. Strategies should also include fostering international collaboration, creating networks for the exchange of

best practices, and emphasizing the long-term benefits of ISO 20387:2018 accreditation in terms of improved sample quality, stakeholder confidence, and research reproducibility. Overall, a carefully planned approach, involving stakeholders at different levels, is crucial for the successful and widespread adoption of the ISO 20387:2018 standard across the global biobanking community [101].

2.1.1.1.1 Beyond traditional biobanks As described above, biobanks were originally intended as a collection of biological samples and associated clinical information. The progressive digitization process has made it possible to digitally archive an enormous number of images and several types of data. In 2014, the Imaging Biobanks Working Group (WG) of the Research Committee was established by the European Society of Radiology (ESR). It defined imaging biobanks as "organized databases of medical images, and associated imaging biomarkers, shared among multiple researchers, linked to other biorepositories" [117]. It is evident that an imaging biobank is not simply a system of archiving and transmitting images as are the PACS (Picture Archiving and Communication System) systems used in the hospitals. An imaging biobank not only allows the storage and retrieval of medical images and associated metadata, but the added value is that these data are linked to the imaging biomarkers, and to clinical, molecular, biological and genomic data. Imaging biomarkers are defined as characteristics extracted from the images of an individual that can be objectively measured and function as indicators of a normal biological process, a disease, or a response to a therapeutic intervention. Imaging biomarkers are complementary to conventional radiological readings to detect a specific disease or lesion, to quantify its biological situation, to evaluate its progression, to stratify phenotypic abnormalities and to assess the treatment response [105, 231]. This connection could be necessary for the researchers to find an association between phenotype and genotype [227, 200], to design and validate new imaging biomarkers, as well as to understand their biological significance, which may be a crucial point in precision medicine [36]. Another important difference is the organization of data and the way they can be retrieved. Whilst a PACS can be defined patient-based, in the sense that a query is based on the search for single-patient examinations, an imaging biobank can be defined

population-based or disease-oriented as patients having a common disease are grouped and a user can query a specific study or disease [227]. The architectures for the creation of medical imaging and molecular imaging biobanks must incorporate advanced high-performance computing capabilities to allow high-throughput processing. Nowadays, institutions gather a whole spectrum of mostly digital information, including social, clinical, imaging and pathological records together with genomic profiles. Consequently, modern biobanking is shifting its focus from sample-driven to data-driven strategies. This implies that data management and integration has become a major component of contemporary biomedical research. In particular, the transition to digital biobanks enabled the amalgamation of diverse datasets, necessitating sophisticated data management and integration techniques. This shift has implications for the accessibility of data, requiring robust systems for efficient storage, retrieval, and analysis. Additionally, the move towards digital platforms facilitates collaborative research and accelerates advancements in precision medicine by enabling the exploration of associations between various data types [223].

2.1.1.1.2 Biobanking networks and research infrastructures In the initial stages of accreditation of a biobank, it is essential to be recognized by the own institution and to relate to the regional and national authorities responsible for managing public health. Biobanking networks aim to connect biobanks together to standardize institutional recognition procedures and coordinate the sharing of common strategies at European level. The biobanking field underwent a huge development with the fostering of such networks. The Biobanking and BioMolecular resources Research Infrastructure (BBMRI)-European Research Infrastructure Consortium (ERIC) represents a reference as biobanking infrastructure. The network was designed to operate across European countries with the aims of improving interoperability and giving quality management services to biobankers and researchers. Today the BBMRI-ERIC includes 20 countries and one international organization, making it one of the widest biobanks network [4]. Furthermore, the network has recently introduced a tool, the BBMRI Negotiator, to facilitate data sharing and collaboration among different biobanks, making data and materials rapidly and widely available to researchers. An online catalog has been established for the collection

and presentation of data describing the majority of European biobanks [323]. During the state of emergency caused by the COVID-19 pandemic, BBMRI promoted the organized collection and safe sharing of COVID-19 patient samples, data and images. The collections were registered in a dedicated and publicly available directory.

Of note, the COVID-19 emergency has highlighted crucial lessons for future biobanking strategies, particularly emphasizing the need for agile implementation tools alongside established standards. The recognition that the existence of standards doesn't guarantee immediate applicability underscores the importance of developing tools that facilitate the swift adoption and interpretation of standards. Aiello et al. [33] highlighted that, although the Digital Imaging and Communications in Medicine (DICOM) is the format used in the clinical acquisition routine, its limited adoption for the release of COVID-19 CT public datasets may indicate that the actual emergency conditions enhance the difficulty in finding suitable tools during emergencies, leading researchers to resort to alternative, more manageable formats. Another important lesson is the imperative need for quality assurance, traceability, and financial investment in biobanking. The urgency of the pandemic has underscored the critical role of resilient infrastructure and well-trained personnel to ensure the safety and accuracy of procedures. Future strategies must prioritize these aspects to effectively handle emergencies [45].

Among the other existing infrastructures, the International Society for Biological and Environmental Repository (ISBER) is a global biobank organization that creates opportunities for networking, education and innovation. ISBER provides a community for harmonizing approaches to emerging challenges in repositories, as well as promoting ideas for creating new solutions and sharing best practices [14]. Moreover, there are networks that bring together disease or pathology-oriented biobanks, such as for example EuroBioBank [9]. It is a unique network of 25 rare disease biobanks located in 11 countries. These biobanks store and distribute quality DNA, cell and tissue samples for scientists conducting research on rare diseases .

2.1.1.1.3 Imaging biobanks projects A recent systematic review of existing image repositories shows that of the 54 selected biobanks containing images (of which 61.1% disease-oriented and 38.9% population-based)

a relatively small proportion can be classified as imaging biobanks [125].

An example of one of the largest and most comprehensive worldwide biobanks is the United Kingdom (UK) biobank, a large-scale biomedical database and research resource, containing in-depth genetic, imaging studies and health information from half a million UK participants [22].

It is worth mentioning two registered European biobanks that have established organ and / or pathology-based images collections: the Central Biobanking facility at the Erasmus MC (Netherland) [8] and the BCU Imaging Biobank at Bio Check Up Srl (Italy) [114]. The former facilitates excellent scientific research using biomaterials (biological samples, images, clinical and epidemiological data), the latter is a non-profit biorepository aimed at the collection and storage of diagnostic images, derived descriptors and clinical data to foster scientific advances in imaging and biomarkers discovery. The Brain Images of Normal Subjects (BRAINS) Imagebank is designed to provide detailed structural brain imaging data of healthy individuals across the human life-course. The image bank, hosted by the Brain Research Imaging Centre at Edinburgh University, Scotland, UK, is a searchable database of integrated data sets already collected as part of research studies which include healthy (or control) subjects [5].

Several European projects aiming to build data infrastructure containing radiological images that are adequately cross-linked to corresponding -omics and health datasets. The euCanSHare project [61] was the first project designed to link these infrastructures for secure and integrated storage of heterogeneous data samples (incl. imaging, -omics, bio-samples and health data), with pilot validation for cardiovascular personalized medicine. Some European projects are under development as part of SC1-DT-TDS-05-2020 H2020 call "AI for Health Imaging" [1]: PRIMAGE (Predictive in silico multiscale analytics to support childhood cancer personalized evaluation empowered by imaging biomarkers [213]), CHAIMELEON (Accelerating the lab to market transition of AI tools for cancer management [6]), EUCANIMAGE (A European Cancer Image Platform Linked to Biological and Health Data for Next-Generation Artificial Intelligence and Precision Medicine in Oncology [11]), INCISIVE (A multimodal AI-based toolbox and an interoperable health imaging repository for the empowerment of imaging analysis related to the diagnosis, prediction and follow-up of cancer [13]) and ProCAncer-I (An AI Platform integrating imaging

data and models, supporting precision care through prostate cancer's continuum [20]). All these projects are devoted to testing and developing AI tools and analytics focused on the prevention, prediction and treatment of the most common forms of cancer while providing solutions to securely share health images across Europe. Two other European calls devoted to the development of cutting-edge services for the enhancement of research infrastructures funded two big projects: ISIDOR-e [15] and CanServe [25]. The European funding initiatives underlined the importance being placed on testing and developing AI tools and analytics focused on preventing, predicting and treating the most common forms of cancer, while providing solutions to securely share knowledge of healthcare images across Europe.

2.1.1.2 Challenges

Given the above-mentioned evolution of biobanks, the associated challenges are also changing. In particular, the main challenges related to comprehensive biobanking can be sorted into three macro-groups related to standardization, reproducibility and integration. A detailed description of each category is provided below, with next sections describing potential solutions.

- **Standardization:** the first issue concerns the standardization of the formats for each domain (those vertically with respect to the Figure 2.2). Standardizing the formats for each domain is crucial for ensuring consistency and comparability of data across different sources. This includes not only the data itself, but also the methods used to acquire and store it. This can be challenging when dealing with domains that have not yet been fully standardized. However, standardizing the formats is necessary to ensure that the data can be easily shared, analyzed and compared, as well as made more suitable for creating predictive systems that support medical decision-making.

- **Reproducibility:** To guarantee the reproducibility of the step towards the digital content (numerical descriptor) for each domain, it is necessary to follow clear and consistent data management procedures. This is crucial for both research purposes and healthcare services, as it allows results comparison, validation, replication, and

dissemination. The basic assumption is that the user of the biobank should be able to expand the dataset starting from new biological samples or raw data (imaging or pathomics). Practically, a digital biobank should release, together with the raw and processed information, all the metadata useful for reproducing any procedure used to derive digital data from biological data.

- **Integration:** Today, an important challenge of the biobanks is the development of imaging technological tools required for radiomic/-pathomic analyses, integrating these features with genetic and clinical data in novel research approaches such as radiogenomics/radiopathomics. Features from different domains (last column of Figure 2.2) should be put together in a single format to facilitate bioinformatics, multiomics, and multiassay. In this context, the biobanks' contribution in integrating imaging and molecular data will represent an innovative approach to improve patients' clinical management supporting the creation of predictive systems that could impact decision-making.

Each of these challenges cannot be treated ignoring regulatory and bioethical issues. Although the biobanking and the sharing of extensive databases could favor innovation and research in healthcare, they are considered potentially critical because of the accessibility to sensitive data. Regarding the protection of the patient and human biological material, a milestone is represented by the Declaration of Helsinki guidelines (1964). This document sets ethical principles including the importance of protecting the dignity, autonomy, privacy of the participants of research projects. To face this issue for biobanks, the World Medical Association (WMA) published the Declaration of Taipei to provide guidelines on the collection, storage, and use of identifiable data and biological material beyond the individual care of patients [23]. This declaration is the first international guideline to provide ethical directions about the complex issues that arise with activities associated with human databases and biobanks. Furthermore, the GDPR 2016/679 represents one of the most complete and shared tools worldwide for the protection of sensitive data and defines the techniques for pseudonymisation and anonymisation of personal data [79]. Findable, Accessible, Interoperable and Reusable (FAIR) principles

point out a path to follow, suggesting that the maximization of the utility of clinical/research data is obtained if these data are traceable, accessible, interoperable, reproducible, and of good quality, allowing study findings to be imparted and shared in a clear and understandable way. FAIR principles are not to be intended as a standard, but simply as a definition of good data stewardship practices [324]. The term 'Findable' implies data can be found online, typically through indexing in search engines. 'Accessible' means data can be retrieved directly or via an approval process. 'Interoperable' imposes data to follow standards. Finally, 'Reusable' requires the context of the data generation (metadata) is documented so it can be compared to or integrated with other data sets. These principles were initially developed for the academic world but have become an indispensable part of clinical research. Following these principles requires an application of standards to the various aspects of data collection and sharing. In relation to this aspect, Holub P et al. proposed FAIR-HEALTH principles [141], including additional components such as quality aspects related to research reproducibility and meaningful reuse of the data [224, 98]; incentives to stimulate effective enrichment of data sets and biological material collections and their reuse on all levels [141, 34]; privacy-respecting approaches for working with the human material and data. Ultimately, to overcome these challenges, digital biobanks must ensure that data and information are standardized, reproducible and integrated, following regulatory and bioethical guidelines [302]. Additionally, adhering to ISO standards and SOPs, as well as following the FAIR principles, will contribute to the success of the digital biobank model. Various initiatives have been launched to tackle the highlighted issues.

2.1.1.3 Towards comprehensive biobanking

In this section, current biobanking initiatives will be critically analyzed considering the comprehensive biobanking perspective, identifying possible critical issues and solutions. Integration and digitalization efforts at various levels (standards, repositories, tools) will be examined.

2.1.1.3.1 Integration of standards The standardization of procedures, file format and vocabularies is the first step to integrate heterogeneous data and to guarantee the functional and semantic interoperability.

As will be emphasized in the following sections, each domain should have its own standard but in some specific domain, e.g., in radiomics, there is not yet a single standard universally approved (de facto) or imposed by the laws (de jure).

In fact, in the domain of radiomics, the lack of a universally approved standard poses several specific challenges to data integration efforts. The heterogeneity of imaging modalities, each with its unique file formats and data structures, complicates the standardization process. Diverse algorithms and software tools for feature extraction contribute to variations in feature definitions and extraction methodologies. Moreover, the absence of consensus on clinically relevant radiomic features hinders the development of standardized vocabularies. The interpretation of correlations between radiomic data and clinical outcomes varies among experts and institutions, adding to the complexity. The rapid evolution of imaging technologies and continuous advancements in analysis methods necessitate flexible standards that can adapt to technological progress. Interdisciplinary collaboration is crucial, requiring effective communication between radiologists, oncologists, data scientists, and other stakeholders. Additionally, addressing concerns related to data privacy, regulatory compliance, and the limited adoption of existing standards further underscores the need for a concerted effort to establish robust and widely accepted standards in the field of radiomics [312].

Therefore, a first significant effort is needed to bring together different standards to create a standard-based integration profile. This integration profile should aim to simultaneously describe clinical, imaging, biological, molecular and omics data, by defining the meaningful attributes to represent the data in each field in a standard way. This also requires an effort in the database design and in the establishment of how to organize and connect data so different from each other in a single structured repository. The ESR started a collaboration with BBMRI-ERIC in 2014, recognizing the importance of integrating imaging and "omics" data. Therefore, this challenge is being addressed in the several European projects mentioned in Section 2.1.1.1.3, which in fact aim to build data infrastructure containing radiological images cross-linked to corresponding -omics and health datasets. In particular, in the context of the PRIMAGE project, it has been proposed a first standard-based integration profile to link imaging

data to biological sample data, typically included in a traditional biobank [270, 213]. To build this model of interoperability among heterogeneous data, the existing formats and ontologies for image and data description, the Minimum Information About Biobank Data Sharing (MIABIS) and DICOM, were considered as standard of reference. In this DICOM-MIABIS model, MIABIS has been expanded to the imaging field to also include image collections, and the expansion has been realized by adding to the MIABIS core a module based on the DICOM standard metadata. Since, as said, it is challenging to find a well-accepted radiomic standard, a special attention shall be paid to the description of the radiomic features extraction and the biomarker validation, which are fundamental data that add value to an imaging biobank. But certainly, this DICOM-MIABIS integration profile represents a first effort and a starting point for standardization of imaging data and metadata representation for data sharing. In a recent update, MIABIS Core 3.0 has been developed with 32 attributes describing Biobanks, Collections, Research Resources and Networks according to a modular structure that makes it easier to adhere to and to extend the terminology. Additional aggregate-level components have been prepared for imaging (DICOM-MIABIS) and for SOPs.

2.1.1.3.2 Integrated repositories The critical issues related to the development of imaging biobanks often make it hard for a small research group to have its own biobank with a significant quantity of both radiomic and biological -omic data. A solution to this problem is the use of public databases containing multi-omic and imaging data with additional supporting data related to the images such as patient outcomes, treatment details, genomic, pathology and expert analyses [125].

Databases are available for different biomedical research fields with primary availability and development in the oncological and neurological fields. In the field of oncology, the US-based Cancer Imaging Archive (TCIA) [86] is a service that stores medical images of cancer patients in a large archive accessible for public download. DICOM is the primary file format used by TCIA for image storage, but TCIA does not enforce other standards for describing nonimage supporting data, such as treatment details and patient outcomes [54]. The connection of TCIA to the Cancer Genome Atlas (TCGA-TCIA) [86]represents the largest data repository

in cancer research containing several primary sites and a large amount of available data (over 20,000 primary cancer and matched normal samples crossing 33 cancer types) [336]. Of note, some of the available information is not compliant with standards [33]. Imaging Data Commons (IDC) is a repository of publicly available cancer imaging data (radiology collections from TCIA and subsequently digital pathology images from Human Tumor Atlas Network (HTAN)), often linked with other types of cancer data, and co-located with cloud-based computational resources and big data analysis tools provided by the Google Cloud Platform [12]. The NCI Cancer Research Data Commons (CRDC) is a cloud-based data science infrastructure that connects data sets with analytics tools to allow users to share, integrate, analyze, and visualize cancer research data to drive scientific discovery [18]. In the field of neurological and neurodegenerative as examples are available the National Institute of Mental Health (NIMH) Data Archive (NDA) [19] and the Laboratory of Neuro Imaging (LONI) [17]. Before the advent of open repositories, it was extremely difficult for an investigator to share and find datasets relevant to his research. The repositories allow researchers, engineers, educators to use their datasets collections to test and validate new hypotheses, to build new analysis tools and techniques, to show students interesting and specific use cases. In addition, a number of active research communities and collaborations have developed thanks to the sharing of specific multicentric collections. Some limitations include the lack of adequate descriptions of the collection, linkages to other databases, standard-compliant data formats, and the complete anonymization of metadata which leads to a loss of information fundamental to research. The main concern associated with using public databases is related to the risk of re-identification for individuals' sensitive data. A key aspect is the application of data curation procedures and of robust de-identification techniques. Anonymization may no longer be appropriate, especially if individual-level data is to be shared. Some repositories apply access restrictions, but the decisional procedure and the data access criteria must be transparent. A further barrier to data sharing is the insufficient attribution of credits and the (mistaken) authors' beliefs about ownership of data. Balancing the potential benefits of using public databases for research and healthcare advancements against the ethical and privacy concerns requires a delicate approach [167, 326].

2.1.1.3.3 Platforms for digital biobanking From the technological point of view, a comprehensive biobank should have an optimized software architecture for the massive extraction of quantitative data and its association with other variables [36]. The main functional requirements are: (i) integration with current health information systems (i.e. DICOM sources, PACS, electronic medical records), (ii) modular extensibility in different components (.i.e. medical image visualization, database searching engines, back-end, front-end), (iii) scalability allowing for the wake up process of new storage units or servers, (iv) easy accessibility for clinical users and collaborators, (v) inference of AI models and data mining. The following components can be found in data management systems for biomedical science. Historically, neuroimaging communities have been the most productive in developing platforms for the collection and management of DICOM diagnostic images. For example, the Extensible Neuroimaging Archive Toolkit (XNAT) is an open-source software suite developed by the Neuroinformatics Research Group of St. Louis, Missouri, to address and facilitate data management challenges in Neuroimaging studies (MRI, CT and PET). While XNAT supports mainly DICOM images and reports, it can, at least in principle, store data of different types [210]. XNAT relies on a three-tiered software architecture made of a PostgreSQL database back-end, a Java-based middleware tier usually deployed on an Apache Tomcat servlet container, and a web-based user interface. Specialized in the integration of neuroimaging data, the Collaborative Informatics and NeuroImaging Suites (COINS) is another platform that enables radiologists and researchers to easily manage questionnaires and neuroimaging data (MMRI, EEG, MEG, and genetic data) [183]. It was developed at the Mind Research Group (New Mexico, USA). COINS's main strengths are the adoption of a centralized infrastructure and the well structured taxonomy for data and data sharing. At the time of writing, COINS supports only the DICOM format, but other data types could be zipped and uploaded via the web interface. The metadata schemas of XNAT and COINS are both structured using XML. One of the main limitations of XNAT is that the creation of a new data type (e.g., clinical variable or assessment) requires the construction of a new XML document and other operations that demand manual changes by an administrator with good informatic skills. On the contrary COINS allows a greater level of exten-

sibility and the creation of customized clinical assessments to complement the neuroimaging scans. However, it does not support user-configurable fields for all the neuroimaging data types (only for MRI and MEG), and no explicit creation of new data types is available to integrate other data sources. Concerning the database, COINS stores all the metadata in an Entity-Attribute-Value (EAV) catalog, while XNAT adopts a mixed model using tables for the widely used data types along with an EAV representation for all the remaining metadata [147]. In terms of the scalability of the system, the EAV approach is less efficient in data retrieval and could affect the catalog performance. Both the repositories are equipped with a DICOM node to receive the imaging studies and a web portal for the users' access.

2.1.2 Comprehensive digital biobanking model

After the recognition of major challenges and criticisms in the clinical and research context, as well as of the integration and digitalization efforts, this section aims to define the key cases and scenarios of a comprehensive biobanking approach. We will account for the needs and requirements (related to standardization and harmonization) for the implementation of a CDB. A prerogative of the proposed approach will be "to invent as little as possible", thus including current "standards", when possible. On this premise, existing standards, procedures and initiatives will be introduced as pillars of the proposed comprehensive biobanking model, that will be discussed as a valuable tool for clinical decision-making in the field of precision medicine.

2.1.2.1 Use cases and scenarios

The identification of use cases is key to highlighting the main challenges and criticisms in the clinical and research context and will help to steer the needs and requirements for the implementation of a comprehensive biobanking approach. First, it should be emphasized that a comprehensive biobank user (namely a researcher/clinician) has two different ways to interface with the digital biobank:

- **"Data Catalog" mode**: a Data Catalog is a collection of metadata, combined with data management and search tools, that helps ana-

lysts and other data users to find the data that they need, serves as an inventory of available data, and supplies information to evaluate the suitability of data for intended uses. According to this interface modality, the user does not access the digital biobank data, but can design a research project with the support of information derived from the catalog. Only after the approval of the project by the Ethical Committee, the data will be transferred to the user according to the Material and Data Transfer Agreement (MDTA);

- **"Data Access" mode**: thanks to this modality, the user can preliminarily consult the catalog, then he/she can access biobank's collections with his/her account and thus perform the analyses directly on the platform. The details of which user has access to which data are typically managed on a project-by-project basis under the responsibility of the collection owners themselves.

It is worth noting that in the "Data Catalog" mode of interacting with a digital biobank, users are spared from immediate privacy concerns. Because this mode revolves around metadata and data management tools, researchers can design and assess research projects without accessing the actual biobank data. Since the focus is on the information derived from the catalog rather than the data itself, there is no immediate need to inform subjects about the specific research being conducted. This is in stark contrast to the "Data Access" mode, where privacy issues come to the fore. In this mode, researchers need not only to consult the catalog, but also access the biobank's collections for direct analysis. This requires informing patients about the use of their data, obtaining ethical approvals, and adhering to strict data access protocols, thus limiting the freedom to expose and analyze data without strict ethical considerations. Striking a balance between facilitating comprehensive research and ensuring patient privacy becomes imperative in the data access mode.

Concerning the possible use cases, these could be of two types:

- **Based on research questions**: a researcher/clinician needs a collection to address a specific clinical question about a pathology or population of interest (e.g., can some radiomic features predict survival in a particular disease?).

- **Based on data integration and reproducibility**: a researcher/-clinician needs external collection to augment his/her initial dataset, to reproduce results obtained by other researchers or to validate AI models.

Figure 2.3 shows both an example of a CDB that enables the implementation a study involving radiomic, pathomic, and genomic descriptors, and a use case where an external researcher or clinician that aims to explore a CDB based on specific criteria, for example to answer a research question. This figure is also representative of a use case where an external researcher aims to reproduce results obtained in another study. In this case all the above-mentioned challenges would come to the surface. To illustrate the outlined use cases, two real-world relevant scenarios can be considered, namely a multi-omic oncology study (e.g. radiogenomic or radiopathomic investigation), and, in view of the recently faced emergency, a study aiming at improving the accuracy and efficiency of COVID-19 diagnosis through AI-based segmentation. Concerning the multi-omic oncology study, a researcher may be interested in exploring the correlation between radiomic features extracted from radiological images and genomic markers [73] (or pathomic features [70]) associated with clinical/pathological outcomes in a specific cancer type and could leverage the CDB to access the required material (e.g. raw images, molecular data, image annotations) to obtain numerical descriptors and/or explore pre-extracted features to reproduce previously obtained results. Moreover, the second real-world example involves a clinician or researcher who aims to develop an AI-based segmentation model for COVID-19 lesions to accurately identify and delineate lung alterations in COVID-19 patients [31]. Also in this case, the researcher can leverage the CDB to access a wide set of raw chest CT images and corresponding masks from COVID-19 patients, as well as to augment his/her initial dataset with other data contained in the CDB, to reproduce results previously obtained by other researchers or to validate the developed AI models.

2.1.2.2 Standardization and harmonization initiatives

This section is focused on solutions to implement the comprehensive biobank approach, following the above defined requirements. The preroga-

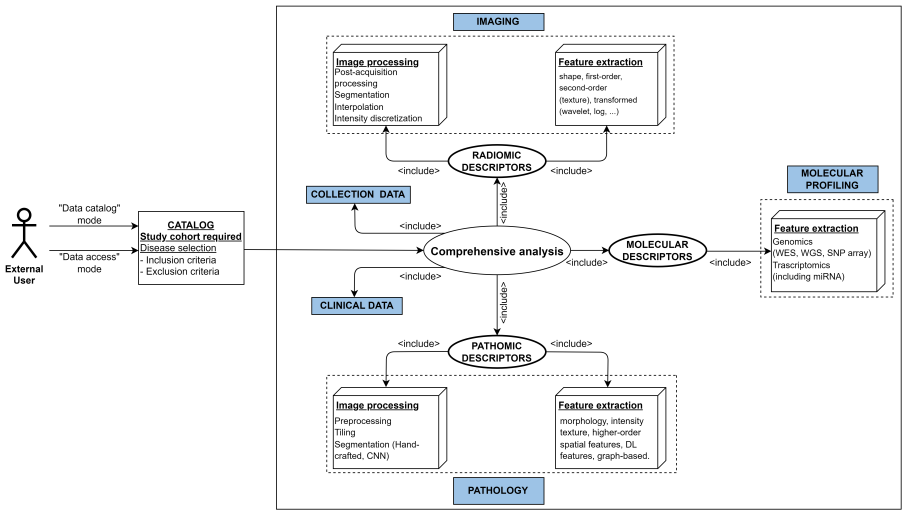


Figure 2.3. Use case diagram representing an external user (clinician/researcher) with the aim of performing a comprehensive analysis involving genomic, radiomic, and pathomic features of a specific tumor type or reproducing or integrating an already performed study. The figure also depicts the ways to interface with the digital biobank (e.g. "Data Catalog" or "Data Access" mode). CNN=Convolutional Neural Networks; DL=Deep- Learning; WES=Whole Exome Sequencing; WGS=Whole Genome Sequencing; SNP=Single Nucleotide Polymorphism; miRNA=MicroRNA.

tive will be to include standards, protocols, and standardization initiatives already existing and shared in common practice. A list of the current standards and standardization/harmonization initiatives related to all domains of Figure 2.2 and functional for the development of a comprehensive digital biobanking approach is reported in Table 2.1. Table 2.1 includes standards and initiatives selected from an online semi-systematic review. In addition to the well-known DICOM and Health Level Seven (HL7) standards, standardization initiatives concerning storage of biological sample, data acquisition (collection and clinical data, imaging, pathology and molecular profiling), data reporting (radiology and pathology report), data curation (clinical outcomes, annotations) and processing (including both procedures and feature extraction) were selected. Only standardization or harmonization initiatives/alliances/communities that had a web reference, were recently updated, and provided a complete picture of the formats, data models, and operating procedures to be followed were selected. Exclusion criteria were therefore the partial information on the website, the absence of recent publications, the absence of clear protocols and guidelines. Taken together, these initiatives represent a huge opportunity to converge towards the interoperability of digital biobanks with clinical data management systems used in common practice, promoting the collection and sharing of real-world data, with a notable impact on the data quality and volume. In addition, standards and standardization/harmonization will be analyzed and selected in light of well-known standards-related challenges, such as fragmentation of standards across institutions and countries or interoperability between different systems and platforms.

Table 2.1. Initiatives and standards functional to the development of a comprehensive biobank.

Acronym	Extended name	Domain	Date	Aim	Conformance Standard (Y/N)
MIABIS	Minimum Information About Biobank Data Sharing	Clinical	2012	Promoting a model of minimum information required to initiate collaborations between biobanks and enable the exchange of biological samples and data.	Yes (de facto)
HL7	Health Level 7	Electronic health information	1987	HL7 is a non-profit ANSI-accredited standards development organization that develops standards that provide for global health data interoperability.	Yes (consensus-based)
OMOP-CDM	Observational Medical Outcomes Partnership – Common Data Model	Health data	2008	Standardize the structure and content of observational data and to enable efficient analyses that can produce reliable evidence.	No
CDISC	Clinical Data Interchange Standards Consortium	Clinical	1997	Standards to support the acquisition and submission and archive of clinical research data.	No
DICOM (SR, SEG, WG 26, CP-1705, CP-1764)	Digital Imaging and Communications in Medicine	Medical Imaging, Digital Pathology, Radiomics	1993	Enables the transmission, storing, retrieval, processing, and displaying of medical imaging information in a standardized format.	Yes (de jure)
IBSI	Image Biomarker Standardization Initiative	Radiomics	2016	Standardizing the extraction of image biomarkers from acquired imaging for the purpose of high-throughput quantitative image analysis.	No
RQS	Radiomics Quality Score	Radiomics	2017	Checklist to ensure quality of radiomic studies.	No
OME	Open Microscopy Environment	Pathology	2012	Developing a data model for representation of image acquisition parameters, image experiment metadata, image analysis results.	No
Feature DB		Radiomics Patomics	2017	Building a comprehensive framework to support the generation, management and interrogation of large volumes of radiomic and pathomic feature sets.	No
MixS (MIGS, MIMS, MIMARKS, MISAG, MIMAG, MIUViG)	Minimum Information about (X) any Sequence (Genome, Metagenome, Marker gene, Single Amplified Genome, Metagenome-Assembled Genome, Uncultivated Virus Genome)	NGS	2005	Defining a set of core descriptors for genomes, metagenomes and the samples thereof.	Yes (de facto)
MIAME, MINSEQE	Minimum Information About a (Microarray Experiment, a high-throughput sequencing experiment)	NGS	2001	Accelerating and supporting the effective sharing and reproducibility of functional genomics data.	Yes (de facto)
QIBA EIBALL	Quantitative Imaging Biomarkers Alliance European Imaging Biomarker Alliance	Medical Imaging, Biomarkers	2007 2019	Standardizing methods to create biomarkers that meet a claimed performance (accurate and reproducible).	No

2.1.2.2.1 Clinical and regulatory data standardization initiatives A critical issue is related to the standardization of clinical data to favor the sharing of medical information. Biological researchers and biobankers, being the producers and often the end-users of such data, have a pivotal role in enabling biological data integration. In this context, one of the most promising approaches is adopted by the BBMRI-ERIC infrastructure. The concept of MIABIS was introduced in 2012 by the Sweden BBMRI to facilitate samples collection and data sharing. Subsequently, it was further updated in 2016, upgrading the components defining biobanks, collections, and studies on an aggregated level [123]. The integrated data can then be used for retrieving data in queries in a structured and organized form. The MIABIS Core version 2.0 is currently used in different biobank registers and catalogs, that is, in the BBMRI-ERIC Directory; the development and the improvement of MIABIS is currently coordinated by the Common Service IT operations of BBMRI-ERIC [109]. Another proposal of sharing of clinical data in digital format are the interoperability standards (V2.x, V3, CDA, FHIR) promoted by Health Level Seven International (HL7). HL7 is a not-for-profit, ANSI-accredited organization dedicated to providing a comprehensive framework and related standards for the exchange, integration, sharing and retrieval of electronic health information that supports clinical practice and the management, delivery and evaluation of health services. Another option is the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM). It is an open community data standard, designed to standardize the structure and content of observational data and to enable efficient analyses that can produce reliable evidence. The CDM is designed to include all observational health data elements (experiences of the patient receiving health care) that are relevant for research use cases to support the generation of reliable scientific evidence about disease natural history, healthcare delivery, effects of medical interventions, the identification of demographic information, health care interventions and outcomes. Finally, an important initiative is supported by the Clinical Data Interchange Standards Consortium (CDISC), establishing standards to support the acquisition and submission and archive of clinical research data. Taken together, these initiatives represent a huge opportunity to converge towards the interoperability of digital biobanks with clinical data management systems used in

common practice, promoting the collection and sharing of real world data, with a notable impact on the data quality and volume.

2.1.2.2.2 Imaging standardization initiatives The standardization of the medical imaging formats plays a crucial role in the effective use of the data and subsequent clinical decision-making. DICOM is the current de jure standard for the storage, retrieval and transmission of radiological and many other medical images, enabling the integration of multi-vendor medical imaging devices such as scanners, servers, workstations, printers, network hardware and PACS [33]. The DICOM file format contains required and optional metadata which describe the patient, exam details, and, in many cases, individual image technical details (e.g., rows, columns, modality, manufacturer). The DICOM file format contains mandatory and optional metadata describing the patient, examination details, and, in many cases, technical details of individual images (e.g., rows, columns, modality, manufacturer). The DICOM standard fully supports a series of key actions involved in the radiology workflow (de-identification, annotation, reporting), allowing to encapsulate in a single format much of the information necessary also for subsequent analytical phases. DICOM-Structured Reporting (SR) provides a versatile mechanism for communicating image-based measurements and supports both quantitative and qualitative evaluations using the TID 1500 template [230, 88]. A DICOM-Segmentation Object (SEG) is the standard way to encode segmentations defined as labeled image voxels [119]. For example, considering a typical radiomic workflow, the use of DICOM objects would allow an AI system to work with appropriately de-identified data, information related to the patient's clinical status (DICOM-SR) and information on the localization of the region of interest (DICOM-SEG) within a single DICOM folder that can be useful to calculate the radiomic descriptors, thus favoring the aggregation of suitable data to develop reliable systems for classification or prediction of clinical outcomes.

While DICOM is widely used and well established, the Neuroimaging Informatics Technology Initiative (NifTI) format is also gaining recognition as a de facto neuroimaging file format, but it also has several other advantages that make it a popular choice in various imaging applications. Unlike DICOM, which primarily contains technical details of the images

and is optimized for clinical use, NifTI provides a more straightforward and flexible way of storing and exchanging imaging data and its metadata, making it an attractive option for researchers and data scientists. One of its key benefits is its compatibility with a wide range of software platforms and applications, allowing for easy sharing of imaging data between different systems.

Harmonizing imaging data standards with other data types in the biobanking context poses significant challenges. The variety of data types collected by biobanks, including clinical, genomic, and imaging data, requires a comprehensive approach to standardization. Achieving interoperability between DICOM and NIfTI standards, as well as other formats, is complicated by the different storage systems and information models employed by different biobanks. Standardizing metadata across these disparate sources is crucial for seamless integration. Moreover, ethical and legal considerations, such as patient privacy and data sharing regulations, add layers of complexity to the harmonization process. Dealing with the large and complex nature of medical imaging data, along with ensuring both syntactic and semantic interoperability, further underscores the challenges. Ongoing updates to standards and the need for user training and adoption contribute to the multifaceted nature of harmonizing these standards in the biobanking landscape [117].

Another important aspect concerns standardization and harmonization of numerical descriptors associated with diagnostic imaging (e.g., radiomics), as well as of procedures for obtaining these numerical descriptors. The lack of shared reference standards concerning data storage, the missing agreement on analysis procedures, and the feature reliability and reproducibility limitations affect radiomics. However, several initiatives have been launched to address these issues. Quantitative Imaging Biomarkers Alliance (QIBA) and European Imaging Biomarkers Alliance (EIBALL) initiatives include collaborating to identify needs, barriers and solutions to the creation of quantitative biomarkers, and accelerating the development of hardware and software to obtain accurate and reproducible quantitative biomarkers. In addition, the Image Biomarker Standardization Initiative (IBSI) is an independent international collaboration dedicated to standardizing radiomic analysis. In particular, the IBSI aims to address many challenges in 4 different specific areas: (1) standard nomenclature and

common radiomic features, (2) radiomics image processing schemes, (3) data sets for validation and calibration, and (4) a set of reporting guidelines [342]. This group defined 174 radiomic features commonly used to quantify the morphologic characteristics and numerous others needed to define the quantitative information and tries to standardize the image processing steps of data conversion, post-acquisition processing, segmentation, interpolation, masking, and others. Such standardization is expected to make radiomics clinically useful and scalable for the integrated diagnosis service [175]. Limiting the radiomic analysis to the IBSI standardized features can facilitate radiomic features interchangeability across platforms [145]. Concerning procedures, a detailed reporting and documentation of radiomic studies is essential to develop this emerging field in terms of clinical translation and to improve the reproducibility of study outcomes. The Radiomics Quality Score (RQS) has been introduced to assess radiomic studies in terms of their compliance with best-practice procedures and to provide a reference guide for the drafting of manuscripts of radiomic studies [181].

2.1.2.2.3 Digital Pathology standardization initiatives Traditionally, pathologists assess and document features of traditional slides in diagnostic reports, which are then archived. With the introduction and advancement of digital pathology, the significance of slides has undergone a transformation, acquiring a "digital copy", thus allowing for an immediate possibility of reuse.

Digital pathology is a general term that refers to the process of digitizing histopathology, immunohistochemistry or cytology slides using whole-slide scanners, along with the interpretation, management, and analysis of these digitized whole-slide images (WSIs) using computational approaches (computational pathology) [28]. The digital pathology slides can be stored in a centralized repository, enabling remote access for manual review by a pathologist or automated evaluation by a data algorithm.

Computational pathology uses advanced computational methods, either hand-crafted (pathomics [134]) or deep-learning-based, to extract valuable information from high-resolution WSIs that can be correlated with phenotypic features in different types of malignancies in association with the traditional histopathologic evaluation performed by pathologists [97].

Despite the growing demand, digital pathology is currently still limited due to several aspects [315]. First, the introduction of digital pathology in clinical practice is highly dependent on the standardization of procedures and file formats. The process of transforming glass slides into WSIs involves a series of phases: (i) pre-analytical (tissue procurement to fixation, processing, cutting, etc.), (ii) analytical (stain selection, validation) and digital (scanning, evaluation of monitor resolution, number of colors and distribution image format), (iii) post-analytical (analysis of results, the reporting of data, and machine learning application and sharing) [52]. Although there are still no robust standardization criteria, several encouraging initiatives are proposed. The College of American Pathologists provided guidelines for the use of approaches involving digital pathology; moreover, other organizations such as the National Society for Histotechnology in the USA and the Royal College of Pathologists in Europe have initiated programs and recommendations for the implementation of digital pathology [236]. Based on the compelling need for data standardization and interoperability in digital pathology, there are ongoing efforts for the standardization of the representation and storage of pathology image data and analysis result [139, 87].

The DICOM Standard Committee WG-26 has put in a tremendous effort to support the use of DICOM in the pathology domain, and considerable progress has been made in incorporation of the information object model for pathology images, including WSIs. The use of DICOM for digital pathology images allows to achieve highly efficient pathology workflows and to easier manage WSIs together with images from other diagnostic domains [214, 161, 235, 154]. While the DICOM standard has been extended to support digital pathology, it has seen little adoption in pathology practice. At the time of this writing, no FDA-cleared digital pathology systems actually employ it natively [139]; however, at least one high-throughput WSI device (Leica Aperio GT 450 DX) outputs DICOM natively [3]. Generally, TIFF (Tagged Image File Format) or SVS (Aperio ScanScope Virtual Slide) file formats are preferred for various reasons. First, they can handle larger file sizes than DICOM thanks to lossy compression, which is important in digital pathology where WSIs can be several gigabytes in size. However, the compression potentially leads to degradation of image quality. Second, TIFF and SVS allow for faster access to images for review

and analysis, without the need for decompression or conversion, and are more commonly used in open-source projects in digital pathology. Instead, DICOM addresses primarily IT experts who have the necessary technical expertise to implement it. This disconnect has resulted in an apparent lack of prioritization of interoperability, and vendors lack a compelling return on investment for building DICOM turn-key solutions. On the other side, DICOM provides a rich set of metadata, including patient information, image acquisition parameters, and annotations, while SVS and TIFF lack this level of information. Moreover, DICOM is a widely accepted standard in medical imaging, allows for seamless interoperability between different systems and platforms and provides long-term archiving, ensuring that images will be accessible and usable in the future, while SVS and TIFF may not be as well-suited for this purpose. As these standards are refined and implemented, we expect that open source and commercial software products will adopt these formats to enable interoperability across different imaging and software systems.

Another important aspect concerns the standardization and harmonization of pathomics workflow. The challenges stem from several factors, including the heterogeneous nature of the data, the variability in image acquisition and processing, and the lack of consensus on the best methods for obtaining the numerical descriptors. One challenge is to ensure the comparability of results obtained from different imaging modalities, such as bright-field, fluorescence, and electron microscopy. Another challenge is to ensure the reproducibility of results, as the variability in image acquisition and processing can lead to different results even when the same image is analyzed multiple times. To address these, it is necessary to establish consensus-based standards and guidelines for image acquisition, processing, and analysis, as well as for the generation of numerical descriptors that ensure consistency and reproducibility.

Despite the previous considerations, the integration of digital pathology data in a digital biobank could represent a groundbreaking advancement, especially when compared to the more established digitalization in radiology. While radiology has been at the forefront of the digital revolution in medical imaging, the digitalization of pathology images is a more recent development that holds immense promise, especially in bridging the gap across multiple scales (e.g., molecular, microscopic, macroscopic) in

the study of diseases, particularly in oncology, where understanding the molecular and cellular intricacies of tumors is critical for accurate diagnosis, prognosis, and treatment decisions [313].

2.1.2.2.4 Next Generation Sequencing standardization initiatives The possibility of managing NGS data confers the opportunities to adopt a personalized approach to the patient. Despite the efforts of these international projects to encourage sharing in processing, analyses, and output of genomics data, there is not yet a single shared direction relating to the management of NGS data. Indeed, the actual proposed standardized procedures and data formats, as well as comprehensive quality management considerations, are not yet fully followed. There are many initiatives promoted and aimed at standardizing genomics data in the fields of (i) reporting standards initiatives, (ii) data analysis and quality metrics projects, (iii) file format, data analysis and quality control tools, and (iv) data integration initiatives. Concerning (i), the basic approach to better exchange and integration of data contributed by different laboratories using different sequencing technologies is the adoption of MIGS-MIMS (Minimum Information about a Metagenomic Sequence). MIGS represents a minimum information checklist that is aimed at standardizing the description of a genomic sequence maintained by the Genomic Standards Consortium; indeed, this organization has also developed an extension of MIGS to support metagenomic data sets called MIMS [296, 84]. For this aim, the Genomic Data Commons Data Portal requires one to provide a specific set of metadata, to contribute to the platform [126]. Other international initiatives such as MIAME (Minimum information about a microarray experiment), MINSEQE (Minimum information about a high-throughput sequencing experiment) are adopted to facilitate the workflow of genomic data standardization [75]. They are proposed by the Functional Genomics Data Society (FGED), defining a minimum set of metadata for high-throughput sequencing, to guarantee the quality, documentation, and reproducibility of experiments and sharing of data. Concerning (ii), the FDA – National Center for Toxicological Research has underlined the necessity of comparability between results obtained from different platforms. The MicroArray Quality Control Project (MAQC) is a project addressed to the reliability and reproducibility of cross-platform gene expression analy-

sis as well as the development of standards and quality guidelines [209, 190]. Other similar international projects focused on quality metrics efforts were the Critical Assessment of Microarray Data Analysis (CAMDA) and Normalization and Transformation Ontology (NTO).

About (iii), there are several data standards that have become *de facto*, meaning they are widely accepted and used without being officially sanctioned. These standards cover a range of topics including sequences, variants, and experiments. For sequences, the FASTQ format is widely used to store and exchange DNA and RNA sequence data along with their associated quality scores. This format is the starting point for most genomic analysis and has become a cornerstone of many genomic analysis pipelines, as the quality scores are critical for the assessment of the quality and reliability of the sequence data [83]. Another important data format in genomics is BAM (Binary Alignment Map) /SAM (Sequence Alignment Map), which is used for storing and sharing the results if aligning the sequences in FASTQ to a reference genome. BAM, the binary version of SAM, is a compact and efficient format for storing substantial amounts of alignment data, while SAM is a human-readable format. BAM/SAM files are commonly used for large-scale genomic analysis and can be used for tasks such as read visualization, quality control, and downstream analysis. There are several types of quantitative data that can be generated after processing BAM and SAM files, including sequence alignments, quality scores, coverage, structural variations, epigenetic analysis, genomic variants [261]. Concerning the latter, the VCF (Variant Call Format) format is widely used for storing and sharing information about genomic variants, such as single nucleotide polymorphisms and insertions/deletions. VCF files provide a standardized way of describing and comparing genomic variants across different samples and are widely used for genomic data exchange and storage [340].

It is not necessary to convert BAM/SAM files to VCF files in all cases, as BAM/SAM files contain information about the alignment of the sequences and VCF files contain information about genomic variants. Depending on the specific analysis tasks, either format may be more appropriate, or both formats may be used in conjunction. For example, if the goal is to perform variant calling, the BAM/SAM file would first be used to align the sequences and the resulting alignment would then be used as

input for a variant caller to generate a VCF file [340].

Regarding (iv), a lot of standard initiatives and efforts are described and available on the different datasets; they are continuously proliferating but unfortunately not necessarily in harmonizing ways. An important initiative, focusing mainly on the standardization of genomic data is carried out by Global Alliance for Genomics and Health (GAGH), an organization that could help to develop the interoperability to develop the great potential of genomic data [10]. To further accelerate the standardization process, several international organizations took part in the creation of physical standards for omics data. The US National Institute of Standards and Technology (NIST) focused on the standardization of sample preparation within the framework of truly diverse projects. An ongoing project is the Genome in a Bottle (GIAB) consortium [24] focused on adapting procedures established for whole-genome sequencing to the clinical environment [279]. Other initiatives are promoted by the Association of Biomolecular Resource Facilities (ABRF), a network focusing on standardization and optimization with the objective to develop guidelines; in detail, the ABRF-NGS group aims to identify the optimal methods and strategies for NGS projects, comparing the performances of different NGS platforms [116, 188]. It will be essential for biobanks to follow these international standardization initiatives to make the diverse types of stored data accessible, manageable, and reuse, promoting an upcoming application of NGS data in clinical practice.

It is worth noting that the implementation of genomic standards within the biobanking framework poses several practical challenges, particularly in the context of data sharing and privacy. This complexity stems from several remarkable features that make genomic data different from other health data, such as the direct relationship between genomic-associated information and prognosis, as well as the presence of significant commonality among blood relatives individuals. Genomic data, with its stability and identification potential, raises privacy concerns that challenge conventional health data privacy models. Advances in privacy technologies are improving genomic data sharing, but regulatory and ethical guidelines need further enhancement. Addressing these challenges is crucial for empowering individuals to actively contribute to scientific research, advancing genomic data sharing and benefiting medical research [68, 219].

2.1.2.3 Proposed approach

Based on some of the existing initiatives identified in the previous sections, a comprehensive biobanking approach that could fill the identified standardization, reproducibility and integration needs was proposed. The proposed approach is based on the use of standards and data formats already existing and shared in common practice (Table 2.1). Looking at Figure 2.4, the challenge is to develop an aggregate database model that is functional to the integration among multidisciplinary domains (vertical view) and the generation of numerical descriptors associated with each domain (horizontal view).

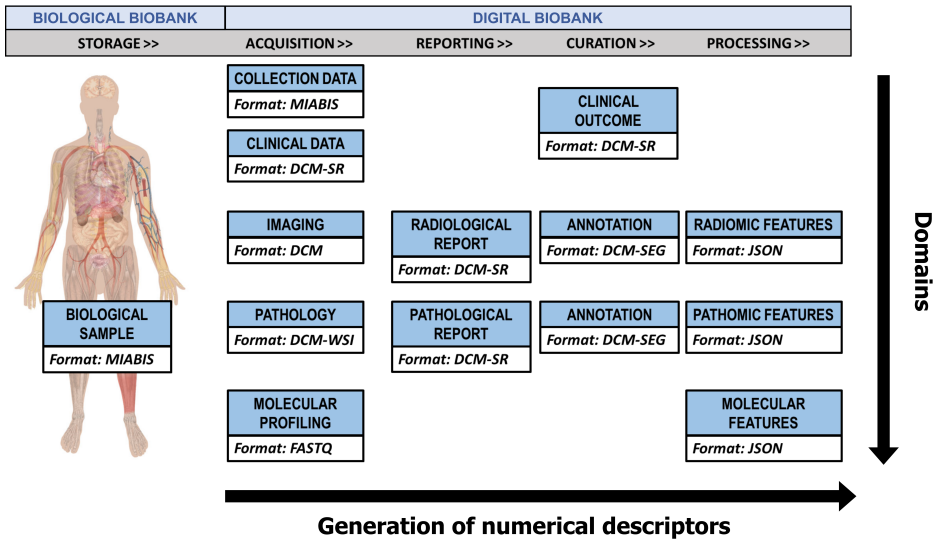


Figure 2.4. Proposed integrative approach accounting for standardization/harmonization of each diagnostic domain and integration among multidisciplinary domains, together with the harmonization/standardization concerning the generation of numerical descriptors associated with each single domain. The approach involved the use of MIABIS, DICOM and FASTQ as they are established standards in the common practice to describe raw and derived data from clinical imaging, pathology, and next-generation sequencing domains. JSON format was proposed to store and interchange domain-specific numerical descriptors. MIABIS=Minimum Information About BIobank data Sharing; DCM (DICOM)=Digital Imaging and Communications in Medicine; WSI=Whole Slide Imaging; SR=Structured Report; SEG=Segmentation; JSON=JavaScript Object Notation.

According to our proposed model (Figure 2.4), the storage and the acquisition of the biological samples will be managed using MIABIS standard. Instead, the key actions of imaging workflow will be supported by DICOM standard allowing to encapsulate in a single format much of the information necessary for subsequent analytical phases both for radiological and pathological domains [33]. Indeed, the proposed approach is based on the use of DICOM for the acquisition of both radiologic images and WSI images. On this line, DICOM-SR is proposed for the encoding of clinical data, radiology and pathology reports, as well as clinical out-

comes. Furthermore, all image annotations (both associated to radiology and pathology) will be encoded in DICOM format, by means of DICOM-SEG. These choices are also motivated by the existence of a theoretical model of DICOM-MIABIS integration [270]. In molecular domain, the use of FASTQ was proposed as the format for molecular profiling [89]. This choice is based on several advantages offered by the FASTQ format. FASTQ is a widely used and well-established format in the field of genomics and NGS data, and it provides reliable data storage, ease of use, and compatibility with existing NGS data analysis pipelines. Additionally, FASTQ format includes quality scores for each base, which can be used for error correction and quality control during data analysis. By using FASTQ for molecular profiling, we aim to ensure the integrity and accuracy of the molecular data stored in the biobank.

It should be considered that, while leveraging FASTQ for molecular profiling in a digital biobank brings numerous advantages, integrating this data with imaging and pathology data introduces may require developing custom interfaces or middleware to ensure seamless data interaction. This is because FASTQ files have a unique format tailored for molecular data, that is completely different from the imaging formats. In fact, FASTQ data typically operates at a granular level, dealing with individual nucleotides. Integrating this high-resolution molecular information with imaging data, which may be volumetric and multi-dimensional, poses challenges in scaling and correlating information accurately. In addition, integrating these results with imaging or pathology-derived features may introduce analytical complexities, requiring sophisticated computational approaches, robust infrastructures, potentially impacting storage costs and retrieval times, and expertise in both genomics and imaging informatics.

Concerning the data processing and the generation of numerical descriptors (namely radiomic, pathomic and molecular features features), we propose to use JSON (JavaScript Object Notation) format [16]. JSON is a lightweight data format that makes it easy to transfer and store huge amounts of data. This makes it ideal for use in a digital biobank. Moreover, JSON is formatted in a human-readable way, and it is supported by the most common programming languages and tools, allowing for easy and portable data management. JSON data can also be hierarchically structured and easily parsed by computer programs, making it easier to

organize and manage data in a digital biobank and to automate data processing and analysis tasks [16]. This is directly linked with the possibility of integrating JSON data with machine learning algorithms, allowing for the creation of complex and personalized prediction models that can handle large amounts of data [148]. DICOM is trialing the use of JSON for encoding the output of AI algorithms (e.g., risk prediction of skin disease) in DICOM-SR format. The goal of this trial is to harmonize with the machine learning community where JSON is the preferred format for algorithm output. Support for the JSON format has been added to the DICOM Standard in Part 18 as the DICOM JSON Model [21]. This model describes how different DICOM value representations can be encoded in JSON and allows for seamless integration of JSON data with DICOM objects.

Based on these considerations, JSON emerges as a powerful tool for structuring and organizing data within a digital biobank in several scenarios, thereby enhancing the efficiency, reproducibility, and compatibility of data processing and analysis tasks across different modalities and research objectives. For example, one can imagine a scenario where the biobank contains diverse datasets, including DICOM images from radiological studies, molecular data in FASTQ format, and associated clinical metadata. JSON can be employed to create a unified data structure that encapsulates information from these different modalities. The hierarchical nature of JSON allows for the representation of patient details, experimental parameters, and imaging annotations in a single, coherent format. This simplifies data retrieval and analysis by providing a standardized structure that machine learning algorithms can easily interpret and process.

In the context of radiomic analyses, JSON can be employed to store extracted features from DICOM images [311]. Each JSON object could represent a specific radiological study, encapsulating details such as image metadata, segmentation information, and a structured array of radiomic features. This format ensures that radiomic data are stored in a coherent and easily accessible manner, facilitating subsequent analyses, comparisons, and the application of machine learning algorithms.

The proposed approach also promoted the use of IBSI-compliant software for radiomic feature extraction, such as PyRadiomics [311], that can also be used starting from DICOM input images with the file name pointing

to a DICOM-SEG, thus automatically obtaining radiomic features without any intermediate steps. This allows for a reproducible feature extraction that can be achieved under real clinical conditions that usually involve DICOM objects. Of note, PyRadiomics supports the use of JSON as a format for storing quantitative data [311]. This compatibility allows for easy storage and analysis of radiomic features in a digital biobank and supports the use of JSON as a format for quantitative data in medical imaging and radiomics.

Overall, the promotion of IBSI-compliant software for radiomic feature extraction marks a forward-thinking approach that significantly strengthens the reproducibility and reliability of radiomic analyses. Adherence to the standards set by the IBSI ensures a consistent and uniform framework for defining and calculating radiomic features. This adherence eliminates variability in feature extraction methods across different software tools and platforms, providing a common language for the field. The standardized definitions not only enhance the transparency of radiomic analyses, but also enable seamless interoperability, allowing researchers to achieve comparable results and facilitating cross-study comparisons. The adoption of IBSI-compliant software within the research community contributes to a more standardized and reliable landscape in radiomics, addressing critical challenges in the pursuit of meaningful and reproducible insights from medical imaging data [121].

As these standards are refined and implemented, we expect that open source and commercial software products will adopt these formats as their default data models for image analysis results to enable interoperability across different imaging and software systems to facilitate easier development and integration of new data management capabilities.

Efforts should also be made to address the challenges related with the missing harmonization of ontologies between standards associated with multidisciplinary domains (the DICOM-MIABIS model constitutes a first example of effort towards this direction [270]), and the limited support for standards in existing healthcare data platforms, that could be addressed by integrating platforms or creating a platform with interconnected modules. The benefits of improved interoperability and data consistency make these efforts valuable for advancing the impact of biobanking in research and healthcare.

2.1.3 Discussion

Based on the huge amount of heterogeneous information belonging to different diagnostic domains, there is an urgent need to make this information available and suitable to promote scientific research and technological development [66]. The management of such cross-domain heterogeneity has always been an open challenge, especially within the intricate framework of precision medicine. The essence of precision medicine lies in tailoring treatments to the unique profiles of individual patients, which increases the complexity of handling diverse datasets. This challenge is particularly acute in oncology, where the lack of a universal therapeutic approach is acutely recognized. Tumor heterogeneity between patients and even within the same patient over time complicates the establishment of standardized treatments [149]. The effective organization of biomedical data in biobanking infrastructures is crucial to enable a comprehensive approach to clinical studies as well as the development of AI tools supporting decision-making in medicine. The main challenges are related to the (i) standardization, (ii) reproducibility and (iii) integration of data and procedures. In this study, we propose a comprehensive digital biobanking approach that could address the identified standardization and harmonization needs and serve as a valuable tool for clinical decision-making in the field of precision medicine. Of note, we harnessed digital biobanks as a source of multifactorial information containing standardized and curated imaging data along with clinical, molecular, and pathological data.

First, the concept of traditional and digital biobank was introduced reporting some of the most well-known biobanking research networks and projects, and the state-of-the-art of the harmonization initiatives in the field of biobanking (i.e., standards, repositories, platforms) was reviewed. Moving from traditional to digital biobanking presents several challenges, especially concerning infrastructure requirements and data security. Upgrading from traditional storage methods to digital systems requires significant and expensive infrastructure changes. Continuous monitoring, updates, and migration to newer storage technologies are essential to preserve data integrity. Sensitive information requires strong encryption methods, access controls, and continuous monitoring to prevent breaches or unauthorized access. Digital biobanking raises ethical concerns regarding the secondary use of stored biological samples and associated data. Addressing

these challenges requires collaboration among experts in biobanking, data management, cybersecurity, and regulatory compliance.

Furthermore, digital biobanks need to ensure that data and information are standardized, reproducible and integrated, following regulatory and bioethical guidelines and adhering to ISO standards and SOPs, as well as considering that the stewardship and management of scientific data needs to adhere to the FAIR principles [162]. This will contribute to the success of the digital biobank model to achieve its full potential as a scientific resource.

On this basis, an integrative standardization/harmonization approach that encompassed (wherever possible) the use of standards and procedures that are already used in common practice was proposed. From the point of view of the domains, we chose to focus on the most advanced diagnostic domains (primarily in the field of oncology), namely clinical imaging, pathology, and NGS. Both the integration of these domains, as well as the generation of numerical descriptors associated with each single domain through robust data curation and data processing procedures were considered. The proposed approach involved the use of MIABIS, DICOM, and FASTQ as they are established standards in the common practice to describe raw and derived data from the chosen domains. MIABIS was used to store and exchange data elements that describe the collection of biological samples [109]. DICOM was established for the acquisition, reporting, and curation (annotation and clinical outcomes definition) of both radiologic [7] and WSI images [154]. Although the DICOM standard is designed to incorporate in the metadata specific characteristics of the patient and the examination characteristics, much of the information needed for molecular imaging analysis is not included.

FASTQ was used for storing NGS-generated raw data [89]. It is worth noting that, although several types of genomic analyses (e.g., array comparative genomic hybridization [246] are used depending on the scientific question being asked and the resources available, we structured our proposal based on a typical NGS pipeline due to its massive use for large-scale, high-resolution ultra-high throughput, scalable, and fast genomic analyses [261].

Since the JSON format natively suits the hierarchical format of DICOM metadata, we proposed to use JSON format to store and interchange

domain-specific numerical descriptors. Furthermore, it is the preferred format for encoding the output of AI algorithms, is widely supported by major programming languages, and can be linked to formal ontologies [140].

To the best of our knowledge, this is the first study that foresees new models of standardization and integration among multidisciplinary domains, also proposing robust data curation and processing pipelines to obtain reproducible numerical descriptors associated with each diagnostic domain. In fact, dealing with standardized and harmonized procedures and data not only provides a conducive environment for advancing -omics studies (e.g., radiomics, pathomics, genomics), but also offers opportunities to explore the potential links between quantitative -omics data and clinical outcomes of patients with specific diseases [66, 198, 70]. In this regard, also Lu et al. [198] underscore in their review the significance of integrating various -omics modalities, including pathomics, radiomics, and genomics, to advance prognostic assays and delve into the potential links between quantitative -omics data and clinical outcomes in patients with specific diseases. Of note, they not only emphasize the importance of integration but also highlight the challenges, potential opportunities, and avenues for future works.

Another noteworthy study by Izzo et al.[148] focused on addressing the challenge of growing metadata heterogeneity in the biomedical field developing digital repositories with flexible and extensible data models, as in the case of modern integrated biobanks management. In particular, they proposed a novel flexible and extensible JSON-based data model to describe heterogeneous data in a generic biomedical scenario. They first describe how to incorporate the model inside the XTENS (eXTensible Environment for NeuroScience) digital repository to support heterogeneous data management in a generic biomedical science scenario. Then, they tested the model focusing on a specific use case of an integrated biobanking management, where different information (e.g. clinical, histopathological, genomic) could be queried, integrated, and shown in a structured view. The JSON metadata schema they proposed aimed at describing and integrating in a highly flexible but consistent format heterogeneous datasets and information in biomedical science, both for clinical and research support, with a focus on biobanking and multi-disciplinary biomedical research.

Overall, the implementation of the proposed comprehensive digital

biobanking approach offers a range of advantages for clinicians, researchers, and patients alike. Firstly, the digital biobank ensures a complete and centralized repository of study information for research purposes. By standardizing and integrating medical images, molecular profiles, and patient data, the biobank improves data management, ensures interoperability between different systems, and makes it easier to access, exchange, and analyze large amounts of data. Notably, the integration of clinical outcomes, radiological, and digital pathology annotations, as well as the use of radiomic, pathomic, and molecular features, allows for the development of predictive models for various diseases, improving diagnosis and treatment planning.

It is important to understand the compliance of the standards with FAIR principles; this would be fundamental to accomplish the requirements of clinical research in terms of data sharing and management. As highlighted in Aiello et al. for the DICOM standard, further efforts are needed by researchers, clinicians and companies to promote and facilitate the use of standards to increase the value of imaging data, according to FAIR [33]. In addition, implementation of the FAIR principles within the numerical descriptors' fields (e.g. radiomics) can facilitate its faster clinical translation [310].

Therefore, in the proposed CDB approach, adherence to FAIR principles is paramount for effective data management. This entails not only ensuring that all incorporated standards are FAIR-compliant but also implementing FAIR principles to address numerical descriptors, thereby maximizing the utility of clinical and research data, emphasizing traceability, accessibility, interoperability, reproducibility, reusability, and data quality. Importantly, the notion of "reusability" extends to both human and machine utilization. Consequently, the proposed approach prioritizes making data machine-readable to harness the full potential of modern technologies. Within the realm of numerical descriptors, the Radiomic Ontology project offers a Python library for FAIR radiomics analysis, serving as a valuable resource to facilitate the seamless transition of research efforts into clinical practice [275].

It is worth noting the importance of including numeric descriptors within digital biobanks not only for what concerns the improvement of diagnostic-molecular knowledge and the direct implications on decision

making, but also for what concerns the regulatory and economic points of view. At a regulatory level, the proposed approach promotes, and is directly applicable for, federated solutions, where the raw data can remain in the proprietary site, exposing only the derived numeric descriptors [316]. The federated approach not only adheres to regulatory requirements but also facilitates scalability by allowing each site to manage and control its raw data. This decentralized structure contributes to the sustainability of the model, as it aligns with data protection norms and supports long-term collaborative efforts.

From an economic point of view, the advantages of sharing digital information are undeniable, just consider the enormous difference in costs between the conservation of biological samples and digital data and the limitation of aliquots of the biological sample compared to the unlimited possibility of reproducing digital data [92]. However, as highlighted in this work, all these advantages require particular care in the management of digital data. This economic efficiency holds significant implications for the scalability and sustainability of the biobanking model. However, as highlighted in this work, all these advantages require particular care in the management of digital data.

The centralized storage of large amounts of patient data in a digital biobank provides researchers with access to a rich resource for medical research, enabling the development of new treatments for various diseases. By reducing the need for repeated tests and imaging, and by improving patient outcomes, the implementation of a digital biobank can also help to reduce healthcare costs in the long term.

Storing data in standardized formats and using structured reports helps to ensure the accuracy and consistency of the data, leading to improved data quality. Digital biobanks typically employ robust security measures, such as encryption and access controls, to protect sensitive patient data, ensuring privacy and confidentiality. They are designed to oversee enormous amounts of data and can be easily scaled to accommodate growth in the number of patients and types of data being collected. This results in time efficiency, as healthcare providers and researchers can save time and increase efficiency, leading to improved patient care and faster progress in medical research.

It is also crucial to emphasize that the effort towards the implemen-

tation of procedures aiming at standardization and harmonization of data associated with diagnostic imaging, histopathology and NGS is fundamental also to make these data usable by AI algorithms for predicting clinical outcomes.

Of note, the work by Kondylakis et al. [169] delves into the transformative potential of AI in the realm of medical imaging, emphasizing the necessity for extensive and harmonized datasets in AI development, particularly for cancer-related medical imaging. The collaborative efforts of five EU projects aim to create ethically compliant, quality-controlled, and GDPR-compliant big data infrastructures. These platforms seek to seamlessly integrate large-scale data with AI algorithms, establishing sustainable AI cloud-based systems dedicated to developing trustworthy and reliable models in cancer care. The study's keypoints include the challenges of data access, the development of a common data model, and the importance of a European Union meta-tool repository to streamline efforts and minimize duplication in a dynamic field such as medical imaging.

A completely different field of medical AI is connected to clinical reports, which clinicians typically tend to make like a narrative text. AI can help to structure or extract specific text parts from routine pathology reports for further scientific purposes [174]. All the reports could be better used for any scientific purpose if it would be easier to search for different disease entities, and this could increase the value of millions of biospecimens and images which are currently stored in biobanks.

However, although out of the scope of our work, it should be also mentioned that the pathway for regulatory approval is a key roadblock in the clinical adoption of AI-based prognostic and predictive tools. One of the principles for regulatory permission includes the necessary explanation of how the software works and their translation to clinical practice (AI-based techniques are perceived as being a black-box and lacking interpretability) [169, 58].

The trasformative impact of AI on biobanks encompasses streamlining processes, improving data accessibility, and synergizing with biobanking to revolutionize cancer research and enhance patient-centric healthcare strategies. AI-powered predictive modeling can also aid in the strategic prioritization of research initiatives within biobanks. By analyzing historical data and identifying patterns, AI algorithms can assist researchers in

identifying the most promising datasets for further investigation. This predictive capability not only optimizes research efforts but also contributes to resource efficiency by directing attention towards areas with higher potential for significant scientific advancements and contribution to a deeper understanding of diseases [124].

Among the issues emerging from our work, it should be considered that, besides the diversity of data, there is also a wide variety in the models built for homogenizing and storing them. This is the reason we opted for not including ontologies in our semi-systematic search for standardization initiatives. In fact, ontologies are usually developed for describing limited sets of data and cannot scale when other types of data need to be stored using the same model. More than this, various expert groups are performing extensions to ontologies that are not synchronized and compatible with each other, thus leading to several variations of the same ontology, which complicates the ontology selection and its reuse. In addition, ontologies have language-dependency [202].

Collaboration among expert groups to establish synchronized and compatible ontology extensions could enhance their usability and reduce variations. Alternatively, the exploration of domain-agnostic data representation frameworks may provide a solution to explore [157].

Moreover, we did not consider the domain of biomedical signals (EEG, ECG, ...), which, although not yet associated with -omics domains, it deserves attention since it can allow link physiological information to other diagnostic parameters [251]. Future research should thus extend its scope to incorporate biomedical signal domains, exploring standardized approaches for their integration into comprehensive data models.

Another topic of discussion emerging from our work concerns the type of data to be included in digital biobanks. Although the topic is open and debated, there is an example that could be followed, consisting of BCU Imaging Biobank [114]. This biobank is digital only and approved as a biobank by BBMRI. Therefore, although there is always a link to biological repositories in the definition of a digital biobank, we can also assume the absence of a biological sample. Along these lines, it can be said that the digital biobank would remain a biobank since it inherits from the traditional biobank all the procedures for certifying its content, including, for example, all issues related to the rights and ownership of

collections.

In addition, since the aim of the work was limited to frame possible functional solutions to the implementation of a digital biobank, trying to solve the issues arising from standardization, reproducibility, and integration, the development of an IT infrastructure for data sharing was out of the scope of the work. However, with the practical implementation of the approach at the infrastructure level, it becomes evident that a comprehensive requirement analysis will be essential to further refine and optimize the biobanking system. This analysis will involve identifying and documenting the specific functional and non-functional requirements necessary to ensure the effectiveness, reliability, and scalability of the digital biobank infrastructure. Given the complexity and interdisciplinary nature of biobanking, the requirement analysis will encompass various aspects, including but not limited to: *i*) defining standardized protocols for data acquisition, storage, and retrieval across different domains such as radiology, pathology, and molecular biology; *ii*) dentifying interoperability standards to facilitate seamless integration and exchange of data between different systems and platforms; *iii*) Addressing data security and privacy concerns to safeguard sensitive patient information and ensure compliance with regulatory requirements; *iv*) Establishing quality assurance measures to ensure the accuracy, reproducibility, and reliability of data generated and stored within the biobank; *v*) Assessing scalability requirements to accommodate future expansion and growth of the biobank, including considerations for increasing data volumes and user access. By conducting a thorough requirement analysis, the diverse needs and priorities of stakeholders, including researchers, clinicians, and data managers, could be systematically captured, and a robust roadmap for the continued advancement and optimization of the proposed approach digital biobank infrastructure.

In conclusion, this work has proposed a CDB model that could improve the management, standardization and sharing of data in compliance with ethical norms.In particular, examining the current state of the art, the need emerged to build, starting from the current reference standards, an integrative framework that can guarantee effective exploitation of the full potential of complete digital biobanks. This issue has been addressed by proposing a standardization model of numerical descriptors that can be derived from each specific diagnostic domain. Ultimately, this work shows

that with further standardization efforts it is possible to implement digital biobanks that represent the driving force to promote the development of data-driven and multi-domain tools that can facilitate the effective implementation of precision medicine.

Main contributions of the chapter

- Digital biobanks were identified as a multifactorial solution capable of containing curated and standardized data from the most advanced diagnostic domains in the field of oncology.
- An integrative standardization/harmonization approach aiming to incorporate existing standards and procedures already employed in common practice was proposed.
- With further standardization efforts it is possible to implement digital biobanks that represent the driving force to promote the development of data-driven and multi-domain tools that can facilitate the effective implementation of precision medicine.

Chapter 3

Supervised Analysis of Multiscale Data

This chapter is dedicated to the advancement of methodologies for extracting radiomic and pathomic features from radiological and histopathological images. It also encompasses the development of computational and statistical tools tailored for the analysis and integration of data across various diagnostic domains, including radiomics, pathomics, genomics, and clinical information. The overarching goal is to strengthen the clinical decision-making process with a multidimensional understanding of cancer, combining insights from diverse imaging scales and diagnostic realms.

In the intricate landscape of cancer diagnosis and treatment, understanding the interplay between data at different imaging scales is paramount. The initial subsection delves into the significance of comprehending the correlation between radiomic and pathomic features, offering a nuanced understanding that spans from microscopic intricacies to radiological resolutions. By correlating features from imaging modalities like DWI and DCE-MRI with pathomic features derived from digitized pathology images, the thesis aims to uncover synergies that can enhance the comprehension of tumor heterogeneity. This exploration not only enriches our diagnostic accuracy but also holds promising potential for predicting treatment responses and prognoses. Simultaneously, the chapter explores potential applications of pathomics in cancer diagnosis. Here, the focus is on studying how pathomic features can support clinical-decision. Pathomics, with

its wealth of information gleaned from histopathological images, offers a unique perspective into cellular structures, densities, and morphologies. Studying the integration of this information into the diagnostic workflow holds the promise of refining tumor characterization, ultimately aiding in more accurate diagnoses and potentially informing the development of CDSS. By bridging the gap between traditional pathology and emerging computational approaches, we aim to pave the way for a more comprehensive and personalized clinical practice in the field of oncology. The chapter dual focus on exploring correlations across imaging scales and studying potential applications of pathomics in CDSS reflects a holistic approach to cancer diagnosis and decision support, acknowledging the intricate interplay of diverse data sources in the pursuit of improved patient outcomes. Part of this chapter was published in [314, 70].

3.1 Exploration of correlation between data at different scales

In the complex world of cancer diagnosis and treatment, understanding how data from different imaging scales connect is crucial. This understanding is key to tackling the complexities of tumor diversity. A significant approach involves integrating and exploring the associations between radiomic and pathomic features, giving us a holistic view from tiny details to larger-scale images. The integration and correlation of data at different imaging scales can offer clinicians a more comprehensive understanding of tumors, helping with a detailed and accurate characterization. Since cancer shows its diversity at both the cellular and radiological levels, bridging these scales is essential for a full grasp of the disease. Connecting radiomic features from imaging methods like DWI and DCE-MRI with pathomic features from digitized pathology images provides a synergistic understanding, especially relevant in the broader context of cancer. Investigating correlations at diverse resolution scales provides richer insights into the biological aspects of tumors. Details captured by pathomic features, including cellular densities and structures, complement radiomic features, offering a more complete characterization of tumor diversity. Beyond improving diagnostic accuracy, the comprehensive knowledge of data at different scales arising from the exploration of their correlations has great potential in predict-

ing treatment responses, prognoses, and guiding personalized therapeutic strategies. Correlations between pathomics and radiomics, genomics allowed for establishing domain specific biological understanding of cancer morphology. Integration of pathomics with radiomics, genomics resulted in improved comprehensive signatures that were better associated with cancer sub-types and prognosticating treatment outcome. Exploring the correlation across imaging scales can be particularly pertinent in the context of Glioblastoma Multiforme (GBM), where the intricate interplay between the microscopic and radiological features is known to be highly complex. This subsection delves into the significance of exploring these correlations, with a focus on a study aiming at exploring the associations between radiomic features derived from DWI and DCE-MRI and pathomic features obtained from digitized pathology images in the context of GBM.

Glioblastoma multiforme, characterized by extensive intratumoral heterogeneity, poses a formidable challenge in clinical decision-making. To address this complexity, the study introduces a novel exploration, leveraging the power of radiomic and pathomic features. The study focuses on the correlation between radiomic features extracted from preoperative ADC maps and post-contrast T1 (T1C) images and pathomic features derived from H&E digitized pathology images, offering a unique perspective into the intricate interplay between microscopic and radiological tumor characteristics. The results of this study suggest that cross-scale associations may exist between digital pathology and ADC and T1C imaging. This can be useful not only to improve the knowledge concerning GBM intratumoral heterogeneity, but also to strengthen the role of radiomics approach and its validation in clinical practice as "virtual biopsy", introducing new insights for omics integration toward a personalized medicine approach.

This research aimed at identifying cross-scale associations between quanitative descriptors from radiology and pathology images of patients with GBM. The radiopathomic approach investigated in our preliminary study revealed interesting cross-scale relationships between radiomic features from functional magnetic resonance images and pathomic features from digital pathology images of patients with GBM The promises arising from the integration between data at different imaging scales would be surely the improvement of diagnostics and molecular knowledge about GBM, for diagnostic, prognostic and therapeutic purposes, and this would have direct

implications in clinical decision-making process. In particular, radiomics and pathomics could fill the need to assess tumor heterogeneity, which strongly characterize GBM. Moreover, the radiopathomic integration could be very useful for the validation of the radiomic approach in clinical practice as "virtual biopsy".

3.1.1 The relationship between radiomics and pathomics in Glioblastoma patients: Preliminary results from a cross-scale association study

3.1.1.1 Introduction

GBM is the most damaging tumor of the brain, characterized by an almost unavoidable propensity to relapse after rigorous treatment and carrying a fatal prognosis [44]. Despite advancement in surgical and medical therapies, the overall prognosis of GBM patients remains poor, with a median survival of 10-14 months [263]. One of the main reasons for the aggressive behavior and the poor outcomes of GBM is its intrinsic intratumor heterogeneity at both microscopic and radiological resolution scales, arising from the presence of clonal and subclonal differentiated tumor cell populations, glioma stem cells, and components of the tumor microenvironment, which affect multiple hallmark cellular functions in cancer. Recent studies have indicated that intratumor heterogeneity is partly responsible for the dismal outcome of GBM patients and this represents a significant challenge to the development of novel targeted therapies for GBM and evidence-based clinical decision-making [81, 284]. In this context, numerous quantitative approaches at different imaging scales have been taken to comprehensively characterize this disease [81, 170, 253]. On a radiological point of view, it is well-known that conventional MRI provides basic anatomic and morphological information for diagnosing brain tumors and lacks the capability to illustrate tumor heterogeneity and identifies the hot spot for biopsy and surgical resection [247]. DWI and dynamic contrast-enhanced DCE-MRI are two functional MRI techniques that are commonly employed in clinic for the assessment of GBM tumor characteristics [131]. T1-weighted imaging acquired following injection with a gadolinium contrast agent is used to identify regions where the active tumor has disrupted the blood-brain barrier, and contrast enhancement is used to define the

extent of the primary tumor region [111]. Hyperintense regions on fluid attenuated inversion recovery (FLAIR) images are thought to indicate a combination of tumor-related edema and infiltrative non-enhancing tumor. DWI can quantitatively and noninvasively reflect the random Brownian motion of water molecules within GBM tissues through quantitative apparent diffusion coefficient (ADC) maps. These maps identify areas of restricted diffusion that may indicate either hypercellular tumor or coagulative necrosis [85]. Although several studies suggested that ADC values can play a valuable role in GBM diagnosis, staging, assessment of response to treatment, and prognosis [74], the biophysical mechanisms underlying changes in ADC are not always fully understood [133, 152]. Radiomics is a new frontier of medicine based on extracting numerical descriptors from radiologic images that are imperceptible by the human eye and are potentially able to describe the intratumoral heterogeneity [182]. Moreover, concerning intratumoral heterogeneity, radiomics is often referred as a sort of «virtual biopsy», since it allows to enrich the traditional diagnostic radiologic workflow with more information not detectable by human eye and associated with processes not included in radiologic workflow [181]. Current radiomics studies in the glioma field have shown promising results in demonstrating correlations between MRI features and GBM differential diagnosis [71], molecular characteristics [191, 327], and prognoses [74, 159, 215]. However, correlating these MRI features with underlying histopathology remains challenging [30, 64]. An interesting approach to relate radiomic results to tumor pathologic findings relies on quantitative analysis of digitized histopathology images [198]. On a microscopic scale, the emerging and rapidly expanding field of pathomics aims to apply high-throughput image feature extraction techniques to interrogate the microscopic patterns in pathologic data, especially from hematoxylin-eosin–stained sections. Because of the close similarity of the approaches, the features from in vivo images may be compared with the features extracted from histopathological images, often benefiting from a clearer biological definition of the image patterns and hence a better understanding of the features [198, 134]. The founder hypothesis supporting the use of radiomics and pathomics in medical care is that data derived from images have a correlation with the underlying biological processes. More precisely, data derived from images would give additional information in relation to

the underlying biological processes compared to the visual interpretation of the image as a picture, which is the traditional way of interpreting images [266]. Therefore, as MRI features and digital pathology offer complementary sources of information about the tumor in vivo and in vitro, a natural question is whether radiomics and pathomics features might be connected. The promises arising from the integration between data at different imaging scales would surely be the improvement of diagnostics and molecular knowledge about GBM, for diagnostic, prognostic and therapeutic purposes, and this would have direct implications in clinical decision-making process [48]. Radiomics and pathomics could fill the need to assess tumor heterogeneity, which strongly characterize GBM. Moreover, the radiopathomic integration could be beneficial for validating the radiomic approach in clinical practice as "virtual biopsy" [181, 306]. Previously, it has been shown that both radiomic and pathomic image-based signatures can independently predict outcomes of interest in GBM [170, 253, 30]. Moreover, some studies on GBM and other cancer types support the hypothesis that combining radiomic and pathomic features will even further improve prognostication and enhance the understanding of the disease by means of predictive models [322, 91, 273, 252]. However, before making predictions it is important to understand if there are actually correlations between radiomic and pathomic characteristics. To date, few studies have explored the associations between MRI images of GBM patients (which interrogate tumor at macroscopic scale) and histopathologic images (which depict tumor at microscopic scale). Preliminary results reported an inverse correlation between ADC-based features and basic pathomic features such as tumor cellularity in GBM [288, 108]. Moreover, there have been recent efforts to identify cross-scale associations between radiology and pathology scales by investigating on more complicated radiomics and pathomic features in different cancer types (of which brain tumors) [64, 198, 38]. However, these findings were inconsistent mainly due to the small number of patients involved in the studies, and image-based tumor heterogeneity and radiomics features have not yet been exactly correlated with findings in histopathology. In light of the above, this works presents initial results aiming at determining if radiomics features extracted from preoperative ADC maps and post-contrast T1 (T1C) images are associated with histological features (pathomic features) arising from H&E digitized pathology

images of patients with GBM. In particular, the aim was to identify pathomic features from digitized histopathology that potentially reflect tissue composition basis of radiomic descriptors from MRI, towards improving the understanding of GBM heterogeneity.

3.1.1.2 Methods

3.1.1.2.1 Patients The study was conducted in accordance with the Declaration of Helsinki, and the study protocol was approved by the Ethics Committee of the Istituto Nazionale Tumori "Fondazione G. Pascale (protocol number 1/20). The subjects used for the study belong to the public database CPTAC-GBM (Clinical Proteomic Tumor Analysis Consortium-Glioblastoma Multiforme), accessed on October 2021 [95]. Radiology images, clinical data, digital histopathology slides, and associated quantified features (cellularity, necrosis, tumor nuclei, age, tumor weight) of samples included were downloaded from The Cancer Imaging Archive (TCIA) database [86]. The inclusion criteria were the following: availability of pretreatment T1C images and ADC maps, availability of corresponding digital pathology WSI, WSI slides with at least 70% of tumor nuclei and at most 20% necrosis. Moreover, patients were excluded in case of i) insufficient quality of MRI to perform imaging analysis and/or obtain measurements, ii) insufficient quality of WSI that did not meet the requirements for diagnosis (e.g., tissue folds, torn tissue) and iii) images with a positive value of Clinical Trial Time Point ID (corresponding to the number of days from the date the patient was initially diagnosed pathologically with the disease to the date of the scan). Finally, a total of 48 subjects were included in the study. Refer to Table 3.1 for clinical characteristics and outcomes of included patients. For further information on the employed dataset refer to B.1.1.

Table 3.1. Clinical and pathologic characteristics of the included patients. Abbreviations: BMI = Body Mass Index; Y = Yes; N = No; NR = not reported; OS = Overall Survival; mo = months.

Clinical and pathologic characteristics	**Value**
Age [mean ± SD]	62.3 ± 11.3
Sex [n (%)]	
Male	34 (70.8)

Female	14 (29.1)
BMI [mean ± SD]	62.3 ± 11.3
Risk factors [n (%)]	
Alcohol	
< 2 drinks per day (men) /< 1 (women)	15 (31.2)
> 2 drinks per day (men) /> 1 (women)	2 (4.2)
Consumed in the past	2 (4.2)
Lifelong non-drinker	16 (33.3)
NR	13 (27)
Tobacco	
Smoker < 15 years	2 (4.2)
Smoker > 15 years	6 (12.5)
Current smoker	8 (16.7)
Lifelong non-smoker	24 (50)
NR	8 (16.7)
Progression/Recurrence [n (%)]	
Y	19 (39.6)
N	16 (33.3)
NR	13 (27)
OS [n (%)]	
< 12 mo	22 (45.8)
> 12 mo	16 (33.3)
NR	10 (20.8)

3.1.1.2.2 MRI acquisition and processing MR examinations with contrast injection were performed on 1.5 and 3 Tesla (T) equipment (34 patients=1.5T; 12=3T and 2=the magnetic field was not reported). The acquisition protocol included: axial FLAIR sequence (TE (echo time)=71-155 ms; TR (repetition time)=7752-12000 ms; slice thickness=4-5 mm; acquisition matrix=128×256-512×192); axial DWI MR sequence with ADC (apparent diffusion coefficient) map (TE=27-149 ms; TR=1000-9601 ms; slice thickness=2-5 mm; acquisition matrix=120×180-512×192); axial post-contrast 3D/2D T1-weighted MR sequence (10=had contrast 2D acquisition; 19=with 2D sequence and 20=both 2D/3D acquisition) (TE=17-130 ms; TR=35-2140 ms; slice thickness=0.8-5 mm; acquisition matrix=252×250 - 384×256). The MRI scans were converted from DICOM format to NIfTI

format using dcm2niix software [189]. ADC maps were coregistered to T1C images by means of two-steps registration procedure both performed by using Elastix registration software [164]. Considering the multimodal registration strategy proposed by Leibfarth et al. [185], a first step a rigid registration was performed to obtain a rough alignment of the fixed and moving images. A two-level multiresolution approach was applied using a Gaussian smoothing without downsampling. A localized version of mutual information was considered as similarity measure and consisted in evaluating mutual information on multiple subregions. Specifically, the localization is obtained by constraining the sampling procedure to a cubic subregion of the image, randomly chosen in every iteration step from the fixed image domain [185, 166]. The standard gradient descent was applied for metric optimization [165]. The resulting transformation matrix was used to initialize the following deformable registration step. In particular, a two-level multiresolution approach using 3D Gaussian smoothing without downsampling was used together with a bending energy penalty term calculated to regularize the transformation. Finally, the similarity metric consisted in a combination of localized mutual information and bending energy penalty and the adaptive stochastic gradient descent optimizer was adopted for its minimization [165]. B-spline was used as interpolation method for the registration procedure. Volumes of interest (VOIs) were then manually delineated slice-by-slice by using ITK-SNAP (version 3.6.0, http://www.itksnap.org) on the T1C. VOIs consisted of T1C enhancing regions. Areas of intrinsic T1 hyperintensity representing hemorrhagic material were not included in T1C contour delineations. Necrotic/cystic regions and large vessels were excluded from all VOIs. The same 3D ROIs were applied to the registered ADC maps, and mean ADC values were extracted for each enhancing tumor VOI.

3.1.1.2.3 WSI acquisition and processing H&E-stained formalin-fixed paraffin-embedded (FFPE) tissues from surgical resection of primary tumor were used for pathological diagnosis. The slides were digitalized to SVS format at 20× magnification (resolution = 0.494 µm/px) [320]. All the WSI images were manually checked for artifacts, and the images free of all types of artifacts were chosen. WSIs of H&E staining slides without any preprocessing were imported in QuPath digital pathology software [51],

and regions covering the largest possible tissue area with viable tissue with vivid histopathologic characteristics and free of artifacts were delineated by an expert microscopist.

3.1.1.2.4 Radiomic features extraction Prior to radiomic features extraction, normalization was applied on T1C images intensities. Specifically, intensities were normalized by centering them at their respective mean value with the standard deviation of all gray values in the original image [107, 90, 311] 91 radiomics features were extracted from segmented VOIs on the enhanced region from T1C and ADC map by using the opensource Python package PyRadiomics (https://pyradiomics.readthedocs.io/en/Latest/). The extracted radiomics features were categorized into two groups: Firstorder features including 18 intensity statistics; 73 multi-dimensional texture features including 23 gray level co-occurrence matrix (GLCM), 16 gray level size zone matrix (GLSZM), 16 gray level run length matrix (GLRLM), 14 gray level dependence matrix (GLDM) and 5 neighboring gray tone difference matrix (NGTDM) features. TRefer to B.2 for complete list and further details on the extracted features. The computing algorithms can also be found at www.radiomics.io and the image biomarker standardization initiative (IBSI) presented a document to standardize the nomenclature and definition of radiomic features [342].

3.1.1.2.5 Pathomic features extraction

3.1.1.2.5.1 Detection measurements WSIs of H&E staining slides were imported in QuPath digital pathology software to carrying out cell and nuclear segmentation. By applying the cell detection function from the analysis module, the nuclear segmentation was performed to recognise objects through watershed cell detection based on segmentation parameters [207, 96], including morphology features of cell, nuclear, and cytoplasm. The setup parameter was set as hematoxylin OD for detection image, with pixel size of 0.5 µm. For nucleus parameters, the background radius and sigma were set as 8 and 1.5 µm, respectively. For intensity parameters, the threshold was set as 0.1 and the max background intensity was set as 2. Other parameters were set as their respective default values. The quality

of the automated cell detection was checked by an expert microscopist. Detection measurements were calculated for all cells using QuPath's add intensity features option (preferred pixel size: 0.5 µm, region: ROI, tile diameter: 25 µm, compute all features including Haralick features with 32 bins) and spatial analysis. Measurements included shape characteristics, Optical Density Sum (ODSum) Haralick texture features [137], and Delaunay triangulation measurements [240]. 65 features were measured for each detection on the candidate slide and exported into a tab delimited file using a QuPath script developed explicitly for this purpose. Refer to B.2 for complete list and further details on the extracted features. Detection measurements were aggregated across the case-level tiles by the mean of the values. For patients with more than one slide, the value of each feature was averaged across WSIs. Moreover, annotation measurements were also computed (WSI selected area, nuclei count, extracellular area, sum of cytoplasm, and extracellular area). Nuclear segmentation and feature extraction process was performed by means of Groovy scripts implemented on QuPath script editor. Exported measurements were imported into R version 3.4.2 for statistical analysis.

3.1.1.2.5.2 Cell density maps features Cell density maps associated with ROIs placed on WSI were calculated at different resolutions (50µm, 100µm, 150µm, 200µm). Specifically, the cellular density was estimated in each tile of dimension associated with the resolution by using the segmented nuclei and assigning a gray level to each of these tiles based on the number of nuclei estimated in each tile. For each resolution, this resulted in a spatial map of the cellular density of the digitized histopathology. 91 features were extracted from cell density maps of digital pathology images using PyRadiomics https://pyradiomics.readthedocs.io/en/latest/. Refer to B.2 for complete list and further details on the extracted features. For patients with more than one slide, the value of each feature was averaged across WSIs.

3.1.1.2.6 Radiopathomic analysis An integrative study design was defined and reported as radiopathomic workflow in Figure 3.1 to evaluate potential association between radiomic features and pathomic features for the included patients. After preliminary analyses, a deeper radiopathomic

analysis including correlation analysis and factor analysis was performed. The details of each analysis are reported in the next three paragraphs. Supplementary analyses were also performed to investigate radiopathomic associations including higher-order features obtained applying wavelet and local binary pattern (LBP) filters to the images (See [115]).

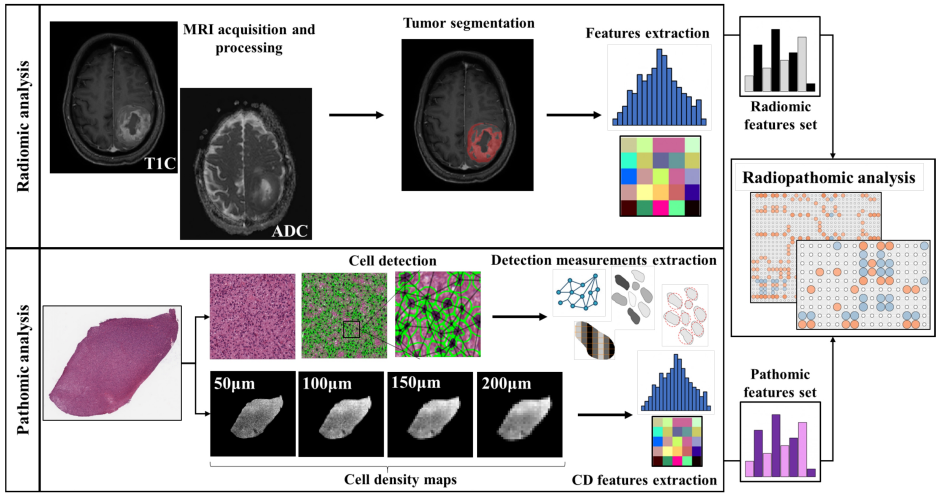


Figure 3.1. Workflow of the radiopathomic analysis implemented in the study. On the first row the radiomic analysis steps. On the second row the pathomic analysis steps.

3.1.1.2.6.1 Preliminary analysis Preliminary analysis aiming at investigating the relationship between ADC mean and basic histopathologic features related to cellular density and commonly associated with ADC meaning (nuclei count, area of the extracellular space and the sum of extracellular space, and cytoplasm) were performed [288, 122]. These features were normalized for the region considered in the WSI.

3.1.1.2.6.2 Correlation analysis Then, a deeper radiopathomic correlation analysis was conducted to explore more detailed associations between extracted pathomic and radiomic features detailed in the previous two paragraphs. Radiopathomic associations were performed considering two separate tasks for ADC and T1C images. A preliminary features selec-

tion was performed separately for radiomic features from ADC, radiomic features from T1C and pathomic features. Specifically, a correlation filter based on the absolute values of pairwise Spearman's correlation (ρ) coefficient was used to reduce feature redundancy. Threshold for ρ was set to 0.9. Briefly, if two features had $\rho > 0.9$, the function looks at the mean absolute correlation of each variable and the variable with the largest mean absolute correlation is removed. Spearman correlation analysis was also employed to determine the correlation of radiomics features with pathomic features. Specifically, we calculated ρ between ADC radiomic feature set selected after the correlation filter and the selected pathomic features surviving after the correlation filter step. The same analysis was performed considering T1C radiomic features. p values were adjusted for multiple correlations by using a false discovery rate (FDR) adjustment. An FDR-adjusted p-value (q-value) below 0.05 was considered statistically significant. Considering ρ in absolute value, the strength of correlation was described as very weak if ρ = 0.00-0.19, weak if ρ = 0.20-0.39, moderate if ρ = 0.40-0.59, strong if ρ = 0.60-0.79, and very strong if ρ = 0.80-1.00 [291]. Correlation analyses were complemented with Bayes Factors (BF) estimation. The Bayes Factor quantifies the evidence for or against the null hypothesis as the ratio of the likelihoods for the experimental and the null hypothesis [260]. It can be expressed as the logarithm of the ratio [151, 156], where negative numbers indicate that the null hypothesis is likely to be true, positive that it is false. By convention, absolute log Bayes factors greater than 0.5 are considered substantial evidence for or against, and absolute log-factors greater than 1 strong evidence [285, 319]. The statistical analysis was performed using R version 4.0.2.

3.1.1.2.6.3 Factor analysis Factor analysis was performed to project the radiomic and pathomic features onto a lower-dimensional latent-feature space that retaining most of the information contained in the whole feature set [242]. For each feature group (ADC, T1C and pathomic), the remaining correlation matrix after correlation filter was subjected to ridge-regularization and 5-fold cross-validation of the log-likelihood was used to determine the optimal value of the penalty-parameter. Feature normalization was performed using z-normalization to avoid the predominance of features with the largest scale in the analysis [136]. The regularized

feature-correlation matrix was used as input for projection by a maximum likelihood factor analysis procedure (factor analytic data compression step)[242]. In particular, a latent lower dimensional orthogonal meta-feature space representing the projection of the shared information in a feature set was obtained. The number of factors, corresponding to the dimension of the latent space, was determined by Guttman bounds [135]. The factor-solution was rotated to a simple orthogonal structure. After projection of the original variable-space onto the lower-dimensional factor-space, factor scores were obtained by regressing the latent features on the observed data by means of the obtained factor solution[242] Factor analysis was performed using the R FMradio (Factor Modelling for Radiomics Data) package (R version 4.0.2). The correlation between resulting factor scores associated with radiomics and pathomic factors were investigated and complemented with Bayes Factors estimation. An FDR-adjusted p-value (q-value) below 0.05 was considered statistically significant.

3.1.1.3 Results

3.1.1.3.1 Preliminary analysis Preliminary analysis of ADC value and basic histopathological features revealed weak but significant correlations between ADC and nuclei count (ρ = -0.317, p = 0.0286) and extracellular space (ρ = 0.3029, p = 0.0368). A positive but not significant correlation was found between ADC and the sum of extracellular space and cytoplasm (ρ = 0.2560, p = 0.0792).

3.1.1.3.2 Correlation analysis Concerning radiomic features, the correlation filter step reduced the feature set from 91 to 46 for ADC and from 91 to 53 for T1C. On the other hand, pathomic features were reduced from 429 to 232. Radiopathomic analysis between selected ADC radiomic features and pathomic features revealed 186 significant correlations (based on adjusted p-values after FDR correction), of which 31 negative correlations (-0.6909 $<\rho<$-0.4578, 22.9454 $<$BF$<$ 2.587×10^5) and 155 positive correlations (0.4556 $<\rho<$ 0.7395, 21.6563 $<$ BF $<6.461\times10^5$). Among features constituting the ADC-radiopathomic couples showing significant cross-scale associations, ADC radiomic features included 10 firstorder features and 23 texture features, while pathomic features included Mean Cytoplasm Eosin OD Min, 5 nuclear Haralick features and 22 cell-density

map features (of which 6 from 50µm resolution, 6 from 100µm resolution, 6 from 150µm resolution and 4 from 200µm resolution). Concerning ADC pipeline, most of the strong relationships (35/50) involved textural features from cell density maps, of which 10 correspond to associations with ADC firstorder features and the remaining 25 were with textural ADC features. The remaining strongest associations (15/50) involved intranuclear Haralick texture features (Haralick Angular Second Moment F0 and Information measure of correlation 2 F12) with ADC firstorder features (10 associations) and textural features (5 associations). Radiopathomic analysis between selected T1C radiomic features and pathomic features revealed 53 significant correlations (based on adjusted p-values after FDR correction), of which 24 negative correlations ($-0.6524 < \rho < -0.5078$, $22.9454 < BF < 2.587 \times 10^5$) and 29 positive correlations ($0.5064 < \rho < 0.6472$, $91.2984 < BF < 2.3688 \times 10^4$). Among features constituting the T1C-radiopathomic couples showing significant cross-scale associations, T1C radiomic features included 4 firstorder features and 11 texture features, while pathomic features included 2 nuclear Haralick features and 9 cell-density map features (of which 1 from 50µm resolution, 2 from 100µm resolution, 3 from 150µm resolution and 3 from 200µm resolution). Concerning T1C pipeline, similar finding were observed, with most of the strongest associations found between cell density map textural features and T1C firstorder (9/15) or textural features (3/15), and the remaining three associations involving the same two intranuclear Haralick texture features observed in the ADC results and two T1C textural features (glcm Joint Energy and ngtdm Coarseness). Similar findings, concerning both ADC and T1C radiopathomic tasks, were observed for significant moderate correlations (Appendix B, B.2). Figure 3.2 and 3.3 show the resulting correlation heatmaps, displaying Spearman's ρ between radiomic features (from ADC and T1C, respectively) and pathomic features. Strongest radiopathomic associations ($\rho \geq 0.6$) sorted by ρ strength are presented in Table 3.2 and 3.3.

Figure 3.2. Radiopathomic analysis between ADC radiomic features and pathomic features. Correlation matrix filtered from nonsignificant correlations (rows and columns with non significant values were deleted, while nonsignificant values surviving were set to zero). Abbreviations: CD = Cellular Density; ADC = Apparent Diffusion Coefficient; LALGLE = Large Area Low Gray Level Emphasis; IMOC = Information Measure Of Correlation; ASM = Angular Second Moment; LDLGLE = Large Dependence Low Gray Level Emphasis; LDHGLE = Large Dependence High Gray Level Emphasis; SDLGLE = Small Dependence Low Gray Level Emphasis; LALGLE = Large Area Low Gray Level Emphasis; GLNUN = Gray level non uniformity normalized; glcm = gray level co-occurrence matrix; gldm = Gray Level Dependence Matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix.

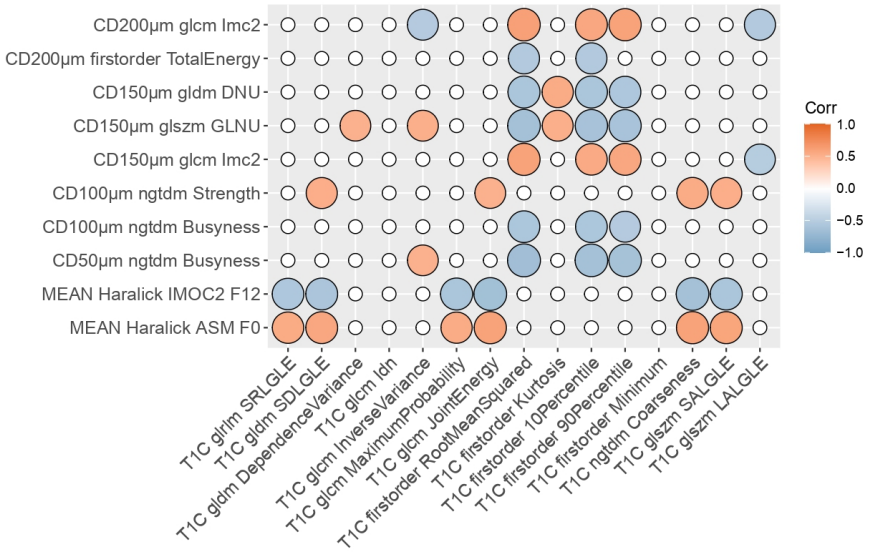


Figure 3.3. Radiopathomic analysis between T1C radiomic features and pathomic features. Correlation matrix filtered from nonsignificant correlations (rows and columns with non-significant values were deleted, while surviving nonsignificant values were set to zero). Abbreviations: DNU = Dependence non uniformity; GLNU = gray-level non-uniformity; IMOC = Information Measure Of Correlation; ASM = Angular Second Moment; SRLGLE = Short Run Low Gray Level Emphasis; SDLGLE = Small Dependence Low Gray Level Emphasis; T1C = post-contrast T1; SALGLE = Small Area Low Gray Level Emphasis; LALGLE = Large Area Low Gray Level Emphasis; glcm = gray level co-occurrence matrix; gldm = Gray Level Dependence Matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix.

Table 3.2: Summary of the highest-correlated (ρ>0.6) radiomic-pathomic features, with radiomic features extracted from ADC. Abbreviations: ADC = Apparent Diffusion Coefficient; CD = Cellular Density; FDR = False Discovery Rate; BF = Bayes Factor; LDLGLE = Large Dependence Low Gray Level Emphasis; LALGLE = Large Area Low Gray Level Emphasis; GLNUN = Gray level non uniformity normalized; LDLGLE = Large Dependence Low Gray Level Emphasis; ASM = Angular Second Moment; IMOC = Information Measure Of Correlation; glcm = gray level co-occurrence matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix.

Radiomic feature Name (ADC)	Pathomic feature name	ρ	FDR q	BF
ngtdm Contrast	CD200µm ngtdm Strength	0.739	1.48×10^{-5}	6.46×10^{6}
firstorder Minimum	CD200µm ngtdm Strength	0.729	1.48×10^{-5}	3.15×10^{6}
firstorder 10Percentile	CD200µm ngtdm Strength	0.725	1.48×10^{-5}	2.36×10^{6}

Continued on next page

Table 3.2: Summary of the highest-correlated (ρ>0.6) radiomic-pathomic features, with radiomic features extracted from ADC. Abbreviations: ADC = Apparent Diffusion Coefficient; CD = Cellular Density; FDR = False Discovery Rate; BF = Bayes Factor; LDLGLE = Large Dependence Low Gray Level Emphasis; LALGLE = Large Area Low Gray Level Emphasis; GLNUN = Gray level non uniformity normalized; LDLGLE = Large Dependence Low Gray Level Emphasis; ASM = Angular Second Moment; IMOC = Information Measure Of Correlation; glcm = gray level co-occurrence matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix. (Continued)

firstorder Median	CD200µm ngtdm Strength	0.698	4.12×10^{-5}	4.06×10^{5}
ngtdm Contrast	MEAN Haralick ASM F0	0.697	4.12×10^{-5}	3.74×10^{5}
firstorder Minimum	MEAN Haralick ASM F0	0.694	4.12×10^{-5}	3.16×10^{5}
gldm LDLGLE	CD100µm glrlm RV	0.693	4.12×10^{-5}	3.00×10^{5}
glszm LALGLE	CD100µm glrlm RV	0.693	4.12×10^{-5}	2.93×10^{5}
firstorder Mean	CD200µm ngtdm Strength	0.692	4.12×10^{-5}	2.75×10^{5}
ngtdm Contrast	MEAN Haralick IMOC2 F12	-0.691	4.12×10^{-5}	2.59×10^{5}
ngtdm Busyness	CD100µm glrlm RV	0.69	4.12×10^{-5}	2.42×10^{5}
glszm ZoneVariance	CD100µm glrlm RV	0.687	4.12×10^{-5}	2.07×10^{5}
glcm DifferenceVariance	MEAN Haralick ASM F0	0.687	4.12×10^{-5}	2.05×10^{5}
firstorder 10Percentile	MEAN Haralick ASM F0	0.687	4.12×10^{-5}	2.01×10^{5}
glcm DifferenceVariance	CD200µm ngtdm Strength	0.679	6.01×10^{-5}	1.31×10^{5}
firstorder Median	MEAN Haralick ASM F0	0.676	6.67×10^{-5}	1.09×10^{5}
glcm Imc2	CD100µm glrlm RV	-0.675	6.67×10^{-5}	1.05×10^{5}
firstorder Mean	MEAN Haralick ASM F0	0.672	7.40×10^{-5}	9.00×10^{4}
firstorder 90Percentile	CD200µm ngtdm Strength	0.66	1.44×10^{-4}	4.52×10^{4}
firstorder 90Percentile	MEAN Haralick ASM F0	0.656	1.67×10^{-4}	3.71×10^{4}
gldmDependenceVariance	CD100µm glrlm RV	0.655	1.67×10^{-4}	3.55×10^{4}
firstorder Minimum	MEAN Haralick IMOC2 F12	-0.654	1.67×10^{-4}	3.27×10^{4}
glszm LALGLE	CD150µm glszm LALGLE	0.653	1.67×10^{-4}	3.17×10^{4}
gldm LDLGLE	CD150µm glszm LALGLE	0.653	1.67×10^{-4}	3.14×10^{4}
ngtdm Busyness	CD150µm glszm LALGLE	0.648	2.05×10^{-4}	2.48×10^{4}
glszm ZoneVariance	CD150µm glszm LALGLE	0.644	2.44×10^{-4}	2.03×10^{4}
glcm DifferenceVariance	MEAN Haralick IMOC2 F12	-0.641	2.72×10^{-4}	1.76×10^{4}
glrlm RunVariance	CD100µm glrlm RV	0.64	2.78×10^{-4}	1.67×10^{4}
firstorder 10Percentile	CD100µm ngtdm Strength	0.634	3.58×10^{-4}	1.27×10^{4}
glcm Imc2	CD150µm glszm LALGLE	-0.63	4.39×10^{-4}	1.01×10^{4}
firstorder Minimum	CD100µm ngtdm Strength	0.628	4.46×10^{-4}	9.50×10^{3}
firstorder 10Percentile	MEAN Haralick IMOC2 F12	-0.628	4.46×10^{-4}	9.29×10^{3}
glszm GLNUN	CD100µm glrlm RV	0.627	4.46×10^{-4}	9.10×10^{3}
glszm GLNUN	CD150µm glszm LALGLE	0.622	5.71×10^{-4}	6.98×10^{3}
ngtdm Contrast	CD100µm ngtdm Strength	0.621	5.87×10^{-4}	6.62×10^{3}
glcm ClusterProminence	MEAN Haralick ASM F0	0.614	7.70×10^{-4}	4.98×10^{3}
firstorder Median	CD100µm ngtdm Strength	0.613	7.74×10^{-4}	4.83×10^{3}
ngtdm Complexity	CD50µm glcm Contrast	0.61	8.56×10^{-4}	4.21×10^{3}
gldmDependenceVariance	CD150µm glszm LALGLE	0.61	8.56×10^{-4}	4.11×10^{3}
firstorder TotalEnergy	CD50µm glcm Contrast	0.61	8.56×10^{-4}	4.07×10^{3}
glcm InverseVariance	CD100µm glrlm RV	0.609	8.64×10^{-4}	3.94×10^{3}
firstorder Mean	CD100µm ngtdm Strength	0.607	9.13×10^{-4}	3.66×10^{3}
firstorder Maximum	MEAN Haralick ASM F0	0.606	9.37×10^{-4}	3.42×10^{3}
firstorder Median	MEAN Haralick IMOC2 F12	-0.606	9.37×10^{-4}	3.41×10^{3}

Continued on next page

Table 3.2: Summary of the highest-correlated (ρ>0.6) radiomic-pathomic features, with radiomic features extracted from ADC. Abbreviations: ADC = Apparent Diffusion Coefficient; CD = Cellular Density; FDR = False Discovery Rate; BF = Bayes Factor; LDLGLE = Large Dependence Low Gray Level Emphasis; LALGLE = Large Area Low Gray Level Emphasis; GLNUN = Gray level non uniformity normalized; LDLGLE = Large Dependence Low Gray Level Emphasis; ASM = Angular Second Moment; IMOC = Information Measure Of Correlation; glcm = gray level co-occurrence matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix. (Continued)

gldm LDLGLE	CD100µm glrlm LRE	0.604	9.43×10^{-4}	3.23×10^{3}
firstorder Mean	MEAN Haralick IMOC2 F12	-0.604	9.43×10^{-4}	3.22×10^{3}
glszm LALGLE	CD100µm glrlm LRE	0.604	9.43×10^{-4}	3.19×10^{3}
ngtdm Busyness	CD100µm glrlm LRE	0.602	9.95×10^{-4}	2.97×10^{3}
glrlm RV	CD150µm glszm LALGLE	0.601	1.02×10^{-3}	2.84×10^{3}
glszm ZoneVariance	CD100µm glrlm LRE	0.6	1.04×10^{-3}	2.73×10^{3}

Table 3.3: Summary of the highest-correlated (ρ>0.6) radiomic-pathomic features, with radiomic features extracted from T1C. Abbreviations: T1C = post-contrast T1; CD = Cellular Density; FDR = False Discovery Rate; BF = Bayes Factor; SDLGLE = Small Dependence Low Gray Level Emphasis; IMOC = Information Measure Of Correlation; GLNU = Grey Level Non Uniformity; ASM = Angular Second Moment; glcm = gray level co-occurrence matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix.

Radiomic feature Name (T1C)	**Pathomic feature name**	**ρ**	**FDR q**	**BF**
firstorder RootMeanSquared	CD50µm ngtdm Bµsyness	-0.652	2.48×10^{-3}	3.09×10^{4}
glcm JointEnergy	CD200µm ngtdm Strength	0.647	2.48×10^{-3}	2.37×10^{4}
ngtdm Coarseness	CD200µm ngtdm Strength	0.644	2.48×10^{-3}	2.05×10^{4}
glcm JointEnergy	MEAN Haralick IMOC2 F12	-0.636	2.48×10^{-3}	1.39×10^{4}
firstorder RootMeanSquared	CD150µm glszm GLNU	-0.632	2.48×10^{-3}	1.12×10^{4}
firstorder 10Percentile	CD50µm ngtdm Bµsyness	-0.631	2.48×10^{-3}	1.05×10^{4}
ngtdm Coarseness	MEAN Haralick IMOC2 F12	-0.629	2.48×10^{-3}	9.88×10^{3}
firstorder RootMeanSquared	CD200µm glcm Imc2	0.621	3.15×10^{-3}	6.73×10^{3}
firstorder 90Percentile	CD50µm ngtdm Bµsyness	-0.619	3.15×10^{-3}	6.18×10^{3}
firstorder 10Percentile	CD150µm glszm GLNU	-0.61	4.26×10^{-3}	4.20×10^{3}
ngtdm Coarseness	MEAN Haralick ASM F0	0.604	4.27×10^{-3}	3.24×10^{3}
firstorder 90Percentile	CD150µm glszm GLNU	-0.603	4.27×10^{-3}	3.09×10^{3}
firstorder 90Percentile	CD200µm glcm Imc2	0.602	4.27×10^{-3}	2.98×10^{3}
firstorder RootMeanSquared	CD150µm glcm Imc2	0.602	4.27×10^{-3}	2.91×10^{3}
gldm SDLGLE	CD200µm ngtdm Strength	0.601	4.27×10^{-3}	2.86×10^{3}

3.1.1.3.3 Factor analysis The factor analytic data compression of the regularized correlation matrix resulted in 24, 32, and 27 latent factors, respectively for ADC, T1C and pathomic. These retained 76%, 89% and 81% of the covariation between the original 46 (for ADC), 53 (for T1C) and 232 (for pathomic) features. Loadings measuring the association between features and factors for ADC, DCE and pathomic feature analysis were reported in Appendix B (B.2). Having estabilished a compact representation of ADC, T1C and pathomic features in terms of factors, correlation analysis between factors scores revealed significant correlations based on

adjusted p-values after FDR correction. To facilitate the reading of factor analysis results, the term "Factor" was abbreviated to "F". Concerning ADC pipeline, 5 significant correlations were found, of which 3 negative correlations (ρ = -0.54, BF = 2.99×10^2 between ADC F5 and pathomic F11; ρ = -0.53, BF = 2.55×10^2 between ADC F4 and pathomic F11; ρ = -0.52, BF = 1.65×10^2 between ADC F1 and pathomic F5) and 2 positive correlations (ρ = 0.51, BF = 1.24×10^2 between ADC F1 and pathomic F17; ρ = 0.48, BF = 36.3 between ADC F2 and pathomic F14). Among factors constituting the ADC-radiopathomic couples showing significant cross-scale associations, ADC F1 consisted mainly of firstorder features (mean absolute deviation, maximum, minimum, 10^{th} and 90^{th} percentile, median and interquartile range) and textural glcm features (difference average and cluster shade, tendency and prominence); ADC F2 consisted mainly of glcm dependence variance and large dependence low gray level emphasis, Informational Measure of Correlation from glcm and three glszm features (zone variance, large area low gray level emphasis, gray level non-uniformity normalized); ADC F4 consisted mainly of textural features associated with entropy (gldm dependence entropy, glcm difference entropy and glszm zone entropy); ADC F5 consisted mainly of ngtdm complexity and firstorder total energy. Considering pathomic factors, F11 was mainly related to texture strength from ngtdm of cell density maps at different resolutions (100µm, 150µm, 200µm); F14 consisted mainly of morphological nuclear features (circularity and eccentricity); F17 is mainly related to dependence variance and non-uniformity normalized from gldm of 150µm cell density maps. Concerning T1C pipeline, 2 significant and negative correlations were found, in particular between T1C F2 and pathomic F11 (ρ = -0.63, BF = 9.14×10^3) and between T1C F5 and pathomic F8 (ρ = -0.53, BF = 2.52×10^2). Among factors constituting the T1C-radiopathomic couples showing significant cross-scale associations, T1C F2 is mainly related to short run emphasis, small dependence and small area low gray level from glrlm, gldm and glszm, respectively, and maximum probability and joint energy from glcm; T1C F5 is mainly composed of firstorder features (10^{th} percentile, root mean squared, minimum). Considering pathomic factors, F11 was described above since it also appeared in ADC-radiopathomic associations, while F8 consisted mainly of busyness from ngtdm of cell density maps at different resolutions (50µm, 100µm) and informational

measure of correlation from glcm of cell density maps at 150µm and 200µm resolutions.

3.1.1.4 Discussion

The possibility to integrate data from different imaging scales in a common framework opens the improvement of diagnostics and molecular knowledge about GBM and its heterogeneity and provides an advantageous context to validate radiomic approach as a clinical "virtual biopsy" tool [181, 198]. Multiparametric MRI and digital pathology images from biopsy samples are currently acquired as standard clinical practice for GBM and provide information that are essential to making correct diagnoses, appropriate patient management and treatment decisions. However, the isolation of radiology and pathology workflows, and consequently of the analysis of quantitative data arising from radiology (radiomics) and digital pathology (pathomics) makes it hard to harness the potential arising from the integration of quantitative data at different imaging scales [266, 283]. In this preliminary study, we evaluated cross-scale associations between radiomic and pathomic descriptors in patients with GBM, quantitatively correlating imaging features extracted by functional MRI images with histological features extracted from H&E digitized slices from surgical resections. In this study, we focused on commonly acquired functional MR sequences, including T1C and ADC from DWI, as they are both part of the routine examination for patients with GBM and are functional modalities, meaning they can provide functional information about GBM and are able to detect tumor volume and physiological changes beyond the lesions shown on conventional morphological MRI [131]. Firstly, we investigated the relationship between ADC mean value and basic histopathologic features related to cellular density, namely nuclei count, area of the extracellular space and the sum of the area of the latter with that of the cytoplasm, that commonly associated with ADC meaning [288, 122]. Indeed, in the field of glial tumors, quantitative assessment of ADC with DWI has mainly been considered as imaging biomarker for its estimation of cellularity based on its inverse relation with water diffusivity in the extracellular compartment (the higher the tumor grade, the lower the mean tumor ADC values) [104]. Our results supported this assumption, showing a significant negative correlation between ADC and nuclei count, and a positive correlation between

ADC and extracellular space and the sum of extracellular space and cytoplasm. These findings were in line with results from the meta-analysis by Surov et al. [288], which found a strong inverse correlation between ADC and cellularity in glioma. An inverse correlation between ADC and cellularity in GBM was also found by Eidel et al. [108] who adopted an approach including trajectory analysis and automatic nuclei counting for the analysis. However, these results should be carefully interpreted mainly due to the extremely small populations investigated in the studies. Of note, we found that the value of correlation with ADC mean decreased when considering the sum of extracellular space and cytoplasm area than extracellular space area alone, and this could be associated with the bias introduced by water in cytoplasm [108]. Preliminary analyses were not performed for T1C since, differently from ADC map that corresponds to a quantitative measurement that is supposed to reflect water diffusivity and thus to be affected by tissue proprieties (e.g., cellularity, cell size, nuclear size, necrosis, extracellular space) directly detectable from nuclear-based features in H&E slides, the relationship between information provided from T1C and nuclear-based features from H&E is not trivial. Further research is needed to investigate on the association between radiomics features from T1C and pathomic features directly associable to this kind of images, such as those associated to microvessel density and quantitative perfusion maps [307]. After preliminary analysis for ADC, a deeper radiopathomic analysis was conducted to explore more detailed radiopathomic associations. In particular, considering two separate tasks for ADC and T1C images, we first investigated correlations between radiomic and pathomic features. It should be highlighted that we opted for analysing the extracted radiomic and pathomic features filtered with a correlation filter to make the analyses more controllable and manageable and eliminate redundancies between the features. Furthermore, although we have not built predictive models, it should be considered that the elimination of highly correlated features is part of the main remedies against overfitting. On the other hand, we used a high threshold ($\rho = 0.90$) to lose as little information as possible after features drop out. In addition, this value was recommended as threshold by Peeters et al. [242] for redundancy filtering to perform factor analysis. Significant cross-scale associations were identified between pathomics features and ADC radiomic features, with correlation strength ranging from 0.45 to

0.74 in absolute value. Significant but fewer and with a lower upper bound ρ values were also found concerning the association between pathomics and radiomics features from T1C, with correlation strength ranging from 0.5 to 0.65 in absolute value. These results were corroborated by very large values of BF (ranging from 21.7 to 6.46×10^5 for ADC task and from 22.9 to 2.58×10^5 for T1C task) that gave strong evidence that the observed data supported the alternative hypothesis. The radiopathomic analysis was enforced by a factor analysis aiming at establishing a compact representation of radiomic and pathomic features in terms of factors that retained most of the information contained in the full data set, using an approach dealing with both the high-dimensionality as well as the collinearity burden of feature sets [242]. The correlation between resulting radiomic and pathomic factors revealed five significant cross-scale associations between pathomics and ADC factors (with correlation strength ranging from 0.51 to 0.54 in absolute value) and only two significant ρ values were found concerning the association between pathomics and T1C radiomics features (ρ = -0.63 and ρ = -0.53). Also results of correlation analysis between factors were corroborated by very large values of BF (ranging from 36.3 to 9.14×10^3) that gave strong evidence that the observed data supported the alternative hypothesis. While it is difficult to demonstrate a causal relationship between radiomics and pathomic features, we can hypothesize about the underlying connections between them. As highlighted, the number of significant radiopathomic associations was almost four times higher when considering ADC than T1C features and more than half higher when considering ADC than T1C factors. This could be related to the quantitative information provided from ADC maps that is supposed to be more influenced by cellularity-associated properties extracted from the H&E images with respect to T1C. However, it is interesting to highlight the significant relationships identified between features associated with cell density and T1C radiomic features since this could provide info on the complementarity of diffusion and perfusion that are usually considered independent characteristics associated with the aggressiveness of the lesions and are independently assessed [248]. Moreover, these results are in line with those from a recent study by Bobholz et al. [63] who found that multiparametric MR imaging intensity values, of which ADC and T1C, were associated with tumor cellularity. However, it should be highlighted that tumor cellu-

larity was quantified on postmortem data, and any other pathomic feature than cellularity was evaluated in their study. Another interesting point to highlight was that most significant radiopathomics associations involved textural features extracted from cell density maps, both considering ADC and T1C tasks. Concerning ADC correlation analysis, almost 1/3 correspond to associations with first order features and the remaining 2/3 were with textural features. Concerning T1C correlation analysis, almost 4/5 correspond to associations with firstorder features and the remaining 1/3 were with textural features. A low number of significant associations involved two intranuclear Haralick texture features (Haralick Angular Second Moment F0 and Information measure of correlation 2 F12) with ADC firstorder features (10 associations) and textural features (5 associations) and two T1C textural features (glcm Joint Energy and ngtdm Coarseness). These results were in line with those obtained in the factor analysis, in which all but one pathomic factors associated with ADC and T1C radiomics were mainly composed by features from cell density maps. The remaining one was mainly associated with nuclear circularity and eccentricity, and this could support the hypothesis that a link exists between specific aspects of tissue heterogeneity and parameters from diffusion MRI [292]. The prevalence of correlations involving features and factors mainly associated with features from cell-density maps highlights that the cell-density map-based feature extraction technique gave rise to novel interesting potential markers that could reflect macroscopic properties from radiologic images. Moreover, this could be intuitively related to the intrinsic meaning of these features that, arising from cell-density maps, are associated with a higher scale level than cellular/intranuclear features and are supposed to be closer to the macro-scale of radiomic features. In addition, it should be also highlighted that, among these significant results involving features from cell density maps, most of them involved features at lower resolutions (100µm, 150µm and 200 µm per pixel), that are supposed to be closer to the macro-scale of radiomic features with respect to those from 50µm resolution maps. On the other side, any of delaunay triangulation features, that describe inter-nuclear characteristics, showed significant correlations with radiomics. Given the promising results involving features and factors associated with original images, we considered it appropriate to perform supplementary analyses by extracting the first and second order features

following application of wavelet and local binary pattern (LBP) filters to the images. Supplementary analysis results followed the trend of results obtained in the main analyses, with more significant results involving ADC than T1C pipeline and most of correlations involving features and factors mainly associated with features from cell-density maps. Of note, several wavelet and LBP features appeared in significant results both in correlation and factor analysis. Of note, we opted for investigating these two higher-order feature groups due to the wavelet and LBP features ability to decipher textural information from different scales [305, 60] and their power for texture analysis and classification [245]. Moreover, these features have also been used for analyzing histopathological images [308, 222]. To our knowledge, this is the first study aiming at investigating radiopathomic associations between radiomic features extracted from ADC maps and T1C images and handcrafted pathomic features arising both from cell segmentations and cell density maps at different resolutions. Only a few works have similarly correlated radiology and histopathology information, and efforts have been rather limited to establishing qualitative clinical correlation. In the context of brain tumors, Rathore et al. [252] aimed to assess the power of radiomic and pathomic features, both in comparison and in combination, for the prediction of survival in GBM patients from TCIA. They found that performances of machine learning models based on the combination of radiomic and pathomic features lead to better results than those based on radiomics and pathomic features taken separately. However, they have conducted any analysis to assess relationships between radiomic and pathomic features. Moreover, radiomic features from ADC were not evaluated. Bobholz et al. [64] examined the localized relationship between MR-based radiomic features from T1, T1C, FLAIR and ADC and their corresponding histomics features in patients with brain tumors. They found several significant radiopathomic associations ($\rho > 0.2$) and suggested that radiomic features were able to capture underlying histopathology. However, a direct comparison with our results was not possible due to differences in extracted features, statistical analyses, and study setting (e.g., pathomic features extracted from autopsy samples, non-specificity on GBM patients). In the context of other cancer types, Shao et al. [273] found that the combination of radiomics and pathomics features was helpful in terms of pretreatment prediction of pathological response in patients with rectal cancer. Interest-

ingly, they also investigated on the correlation between extracted radiomic and pathomic features before the construction of the predictive signatures. Alvarez-Jimenez et al. [38] identified significant cross-scale associations between CT radiomic features and pathomic features that independently showed discriminative power for differentiating between non-small cell lung cancer subtypes. It is worth noting that, similarly to what we did in our study, they examined pathomic features extracted from cell density maps at different resolutions. Of note, Lu et al. [198] performed a brief review on fusion of pathomics, radiomics and genomics, also providing an overview of research works combining radiomic and pathomics in the specific field of cancer prognosis. Despite some promising results emerging in our preliminary study, there are still many limitations to overcome. First, despite utilizing a large public database, our final cohort size was still limited. However, we ensured our analysis was conducted as rigorously as possible. We attempted to account for these limited numbers by appropriately adjusting the FDR threshold. Another important limitation affecting our work concerns the missing information on the exact localization and orientation of the surgical sample. This certainly constitute a bias related to the lack of MRI-histology correspondence, but at the same time it could be an advantage as regards the generalizability of the approach developed in a clinical context since the dataset investigated could be representative of real-world data acquired in clinical practice. However, further studies performed on radiomics and pathomics data arising from colocalized MRI-pathology images are needed, although care should be taken concerning issues that can introduce bias when studying correlations between imaging and histological data. In particular, differences in orientation between the MRI scan planes and the surgical sample can determine a significant mismatch, the tissue deformation occurring when the histological sample is placed outside its anatomical background can determine important and locally nonlinear alignment inconsistencies, and the different spatial resolution of the two methods (1–5mm for MR and 3–5um, for histology) does not allow accurate co-localization [48]. However, we attempted to account for the missing MRI-histology co-localization bias as best as possible by proceeding to accurately avoid necrotic/cystic zones and large vessels in both macroscopic VOIs on MRI and microscopic segmentations on WSI images, in order to allow for multiscale quantification providing

complementary information contributing to the understanding of tumor characteristics [300]. In addition, we could not investigate the association between pathomic features and features deriving from quantitative parameters associated with perfusion (e.g. Ktrans, Kep, Ve, CBV) due to the non-availability of DCE/DSC images for almost all included patients [339]. Furthermore, as previously mentioned, the features associated with cell density and intranuclear ones are not directly associable (regarding the functional meaning) to those of T1C. It would have been interesting to conduct preliminary analyses on microvessel cell density (MVD) to evaluate the associations between more advanced features associated with MVD segmentations and radiomic features from T1C [307, 331]. Moreover, we have not carried out survival analyses since the information on treatments that the patients have carried out were missing. Of note, we opted for evaluating a limited number of features in this study in order to be able to directly translate descriptors from one scale to the next as done in our pathomic-radiomic analysis experiments. An essential criterion for having examined these hand-crafted features is that they are easier to interpret with respect to features obtained from neural networks [298, 237]. However, our framework may be used to study associations between an expanded suite of radiomic and pathomic measurements, as well as extended to characterizing other cancers. Another critical limitation affecting our study concern the well-known lack of shared reference standards concerning data storage, the missing agreement on analysis procedures, and the feature reliability and reproducibility limitations affecting both radiomics and pathomics [304]. In particular, the existing lack of standardization in terms of image acquisition, processes, segmentation methods, and radiomics/pathomics analysis tools, could lead to discrepancies in feature measurements that are not due to underlying biological variations [234, 332]. For example, differences in scanners and image acquisition parameters between considered patients may have affected radiomics results [80]. In particular, it should be highlighted that MR images were acquired at different magnetic field strengths, and this represented a confounding factor, mainly for T1 and T2 images [43], and exacerbated the potential for batch effects. Batch effects, arising from technical variations such as differences in image acquisition parameters or equipment, can introduce systematic biases in the data and confound the interpretation of results. These effects are particu-

larly relevant in multi-center studies or when data is collected over an extended period, as variations in imaging protocols or equipment calibration can significantly impact feature measurements [80]. While the stability of ADC radiomics features was found to be unaffected by differences in magnetic field strength (matching the field-independent nature of ADC) [241], efforts were made to mitigate potential batch effects. For instance, normalization techniques were applied to T1C raw images to account for the varying intensity ranges of MRI data and improve the robustness of radiomics features, as indicated by the IBSI guidelines [342]. Moreover, we proceeded to report in detail all steps of radiomic and pathomic workflow performed in our study since it is essential to develop this emerging field in terms of clinical translation and to improve the reproducibility of study outcomes [181, 303]. Although our findings require careful interpretation due to the limitations mentioned above, the implemented radiopathomic approach revealed interesting cross-scale relationships between radiology and pathology in patients with GBM and may represent a starting point for future research on GBM. Our results strengthen the role of radiomics approach and its validation in clinical practice as "virtual biopsy", introducing new insights for omics integration toward a personalized medicine approach. Further prospective and retrospective studies involving larger groups of patients are essential to validate obtained results, perform in-depth analyses and extend this approach to other cancer types.

3.2 Towards clinical practice: Leveraging Pathomics to support Cancer Diagnosis

This section focuses on the power of pathomics in bridging the gap between research and clinical practice, highlighting two distinct lines of research that harness pathomic approaches for clinical applications in the field of oncology. The focus will be on pathomic features since they can provide additional information concerning the underlying biological processes compared to the WSI visual interpretation, also potentially providing a tangible and interpretable representation of patterns and structures within pathological images, aiding clinicians and researchers in understanding the underlying characteristics of tissues. Moreover, pathomic analyses often require less data than DL methods while still capturing relevant information.

The pivotal role of pathomics in shaping the future of cancer diagnosis and treatment will be emphasized. In exploring this field, two distinct studies were spotlighted to exemplify the potential of pathomic approaches to bridge the gap between research endeavors and tangible clinical applications in oncology. Each study represents a unique exploration, leveraging pathomic features to address for specific challenges in breast cancer and prostate cancer diagnosis, offering a glimpse into the transformative possibilities that such approaches bring to the forefront of clinical practice.

The first study tackles the challenge of quantifying tumor-infiltrating lymphocytes (TILs) in breast cancer (BC) since traditional methods reliant on visual assessments of histopathological WSIs often result in non-standardized and time-consuming processes. This study pioneers a pathomic approach, employing high-throughput image feature extraction techniques to analyze microscopic patterns in BC WSIs. The proposed approach introduces a potential breakthrough, offering a more standardized and efficient assessment of TILs in BC—a crucial factor in decision-making support for BC treatment strategies.

The second study addresses the complexities of Gleason grading in prostate cancer (PCa) given due to the well-known subjectivity and labor-intensive nature of the traditional manual assignment of Gleason Scores (GS). In response to these challenges, this study introduces a multiscale, pathomic-based approach for PCa Gleason Score classification from H&E-stained digital pathology images. Leveraging pathomic features extracted at different magnification scales, the proposed approach sets the stage for a more efficient and consistent Gleason grading process. The study hints at a potential shift in the clinical landscape, envisioning a future where technology enhances the precision and reliability of PCa diagnosis. Through these investigations, the aim of this section is to navigate the path from research innovation to real-world impact, envisioning a landscape where pathomics becomes an integral tool for enhancing diagnostic precision and treatment strategies in cancer care.

3.2.1 A pathomic approach for tumor-infiltrating lymphocytes classification on breast cancer digital pathology images

3.2.1.1 Introduction

The treatment of BC is largely determined by the biology of the tumor. It is becoming more evident that the patient's immunity can be an important indicator of what treatment is needed and how their own immune system can significantly contribute to their chances of long-term survival [264]. In recent years, the role of the tumor microenvironment (TME) has received increasing attention in the immuno-oncology scientific community, focusing on the interaction between tumor cells and the host immune system[187]. TILs, namely lymphocytes and plasma cells that have invaded the tumor tissue, are proving to be an important biomarker in cancer patients as they can play a role in killing tumor cells, particularly in some types of BC. In recent years, several studies have shown the predictive and prognostic value of visually scored TILs in triple-negative BC (TNBC) and human epidermal growth factor receptor 2 (HER2+) BC, making TILs a powerful clinical biomarker [110, 178]. Therefore, accurate identification and measurement of TILs can support less aggressive target treatments, particularly immunotherapy, reducing the need for more invasive options like chemotherapy. Currently, the degree of TILs infiltration is assessed by simple visual evaluation of H&E-stained WSIs of tumor sections, following recommendations from the International TILs Working Group (TILs-WG) [265, 42]. According to TILs-WG guidelines, the boundaries of the invasive tumor must be identified, and only TILs inside those boundaries were evaluated. In particular, the TILs-WG suggests only considering stromal TILs (sTILs), which are defined as mononuclear hosts immune cells (predominantly lymphocytes) present within the tumor margin and located within the stroma among carcinoma cells without directly contacting or infiltrating tumor cell nests (Figure 3.4(A-C)).

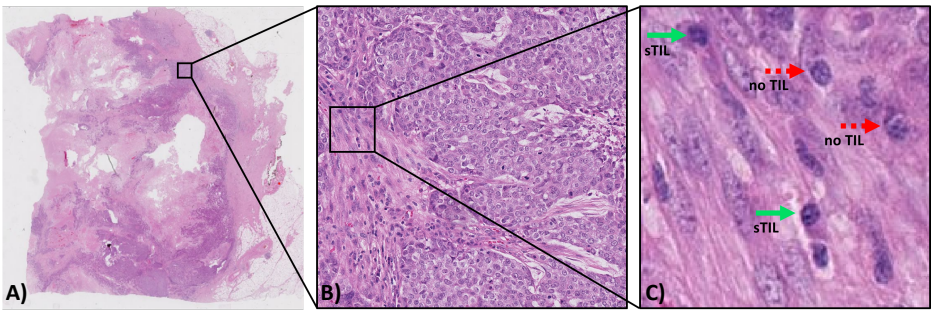


Figure 3.4. Identifying stromal Tumor Infiltrating Lymphocyte (TIL) regions in a H&E-stained WSI of breast cancer. (A) H&E-stained WSI of breast cancer. (B) Example of a region of tissue. (C) Example of stromal TILs (green solid arrow) and cells that are no TILs (red dotted arrow). The image shows how it can be challenging to visually differentiate TILs from other cells.

This is because sTILs are more prevalent, more variable in amount and have been demonstrated to be more reproducibly assessed when compared to intratumoral TILS (iTILs), defined as lymphocytes within nests of carcinoma characterized by cell-to-cell contact with no presence of stroma among them [172]. sTILs are then reported as a percentage, which refers to the percentage of stromal area occupied by mononuclear inflammatory cells over the total within-tumor stromal area [172]. Specifically, the TIL score is manually estimated as the percentage of total tumor-associated stromal area occupied by TILs, where areas of necrosis, ductal and lobular carcinoma in situ (DCIS/LCIS), and normal breast tissue are excluded[265]. However, despite these successful standardization efforts and the rising amounts of evidence concerning the association of TILs with favorable prognosis [249], manual quantification of TILs has several drawbacks: is time-consuming, limited in precision, requires extensive training by pathologists, inter-observer variability will likely remain an issue given the difficulty in quantitatively evaluating histological properties, and lacks the ability to evaluate more complex properties such as TIL distribution patterns or features associated with tissue phenotype. Therefore, this lack of standardized, objective, and effective TILs quantification strategies make TILs marker still not considered robust enough for being used in clinical practice [265]. In several TIL studies, pathologists count by eye the quantity of lymphocytes within a limited area, an obviously very error-prone

and time-consuming task [110]. To favor the reporting of TILs in BC it is therefore necessary to provide the pathologists with automated image analysis tools. The availability of a computerized image analysis scheme for automated quantification of TILs will enable the development of an inexpensive image-based system for predicting disease survival and outcome. In particular, computational image analysis approaches are crucial for improving TILs quantification accuracy, saving time, enabling the analysis of more complex spatial patterns, and providing standardized metrics for validation by expert pathologists and regulatory agencies [163, 26, 280].

Over the last decade, the rapidly increasing technical advancements in the field of digital pathology enabled the development of increasingly advanced computerized image analysis approaches for automating manual tasks such as TILs quantification and addressing their associated limitations. Different automated TILs quantification approaches have been proposed, ranging from those based on classical image processing methods, where cell boundaries are identified and the resulting objects classified according to certain properties, to deep learning (DL)-based approaches that directly classify cells without the need for explicit segmentation [163, 93, 199, 184]. Although the most common strategies proposed to quantify TILs are those exploiting DL supervised techniques, achieving high quality results with this group of methods requires large scale dataset for training with precisely annotated masks accurately representing the cell borders (in case of semantic segmentation approaches such as the U-Net [259]) or the exact locations (in case of object-detection approaches such as Fast R-CNN [128] and YOLO [255]) of TILs in training images. Moreover, DL-based approaches are mostly neither interpretable nor easily explainable [132, 262], and fail to aid pathologists with novel insights on the tissue since they lack in providing a structured representation of the input image data.

The emerging and rapidly expanding field of pathomics aims to apply high-throughput image feature extraction techniques to interrogate the microscopic patterns in pathologic data, especially from H&E-stained sections[134]. The founding hypothesis in support of the use of pathomics in medical care is that quantitative features derived from digital pathology images could give additional information in relation to the underlying biological processes compared to the visual interpretation of the image as a

picture, which is the traditional way of interpreting images. In this context, features extracted from pathomic analysis could be a promising tool for automatically identifying TILs in BC H&E images since they may contain information helping to discover novel biological associations that could be useful for TILs detection [134]. Moreover, they provide more easily interpretable and explainable results than those obtained using DL-based approaches for TILs detection [132, 287, 295]. Only a few previous studies aimed at detecting TILs using quantitative features from H&E images of different cancer types[93, 53, 39], while most studies aimed at developing DL-based computational TIL assessment methods broadly following some or all steps of visual TILs-WG guidelines [287, 295, 41, 267]. However, automated TILs quantification methods are still far from reaching clinical practice for many reasons (e.g., requirement of more rigorous design, greater sample size, more standardized survival analyses, high demand for powerful infrastructures for training data etc.) and studies contributing to achieving a quantitative, automatic, interpretable, and as simple as possible (both considering the implementation and the hardware requirements) are needed. In light of this, this study aims to develop an approach based on a comprehensive set of hand-crafted pathomic features to automatically classify TILs in H&E-stained WSI of BC sections. It is worth remarking that pathomic features are explicit, therefore easy to interpret, as opposed to the DL approach. Several predictive models built using these features will be evaluated to study how different classifiers influence the results.

3.2.1.2 Materials and Methods

In this section, data used in this study as well as the preprocessing methods and the algorithm for nuclei segmentation are introduced. Moreover, the extraction of pathomic features, feature selection and classifier modeling are described in detail. Finally, the metrics used for the evaluation are provided. The overall methodological workflow of this study is shown in Figure 3.5(A-D).

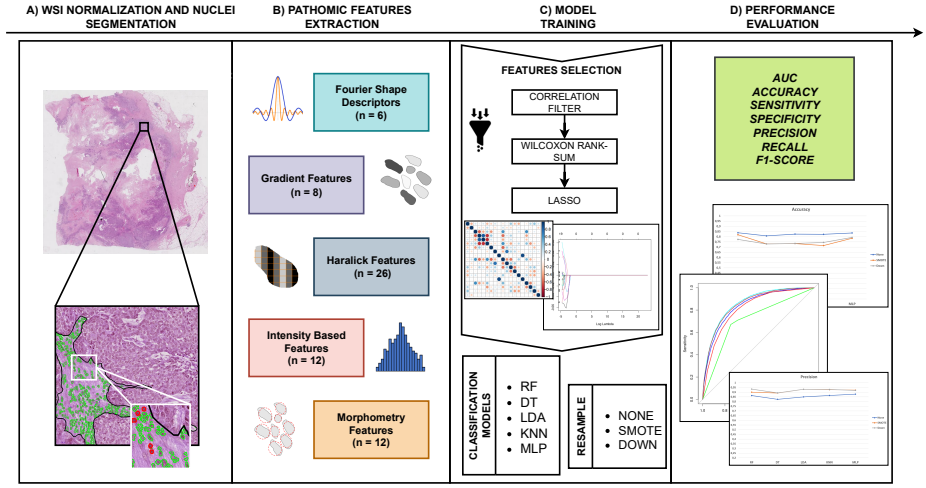


Figure 3.5. Workflow of the approach developed in the study. Starting from H&E images from HER2+ and TN breast cancer, the approach starts with the segmentation of nuclei within tumor-associated stroma (A). Pathomic features were then extracted from cell segmentation within tumor-associated stroma (B). After the feature selection step, five classification models were trained (C) and their performance evaluated (D).

3.2.1.2.1 Dataset A public dataset of diagnosed TNBC and HER2+ BC patients was used https://tiger.grand-challenge.org/. In particular, for the development of the pipeline the 'WSIROIS' dataset was employed. This dataset consists of 195 WSIs of core-needle biopsies and surgical resections of BC patients with annotated tissue compartments and TILs, combining cases from three sources:

- TCGA: TNBC cases from TCGA-BRCA [86] archive (n = 151). The annotations provided for this dataset were generated by adapting the publicly available BCSS (https://bcsegmentation.grand-challenge.org/) and NuCLS (https://nucls.grand-challenge. org/, [39])datasets.
- RUMC: 26 cases of TNBC and HER2+ cases from Radboud University Medical Center (Netherlands), with annotations by a panel of board-certified breast pathologists
- JB: 18 cases of TNBC and HER2+ cases from Jules Bordet Institute

(Belgium). Annotations for these were made by a panel of board-certified breast pathologists.

All data were provided at 20× magnification (resolution of 0.5 micron-per-pixel) for processing. All images have ROI-level annotation masks for use in processing. Concerning annotations corresponding to tissue compartments, each image had ROI-level annotations of the following tissue classes: invasive tumor, tumor-associated stroma, in-situ tumor, healthy glands, necrosis not in-situ, inflamed stroma and other. ROIs of tissue compartments were manually annotated using polygons. TILs were manually annotated using point annotations and then squared bounding boxes of 8x8 µm were constructed centered on the point annotation [290]. Five breast pathologists independently performed tissue and TILs annotations and any disagreement was solved by consensus. The classes 'tumor-associated stroma' and 'inflamed stroma' were merged to form a 'stroma' class. Then, all other classes were set to zero. In this way, masks only consisted of background and stroma, and TILs. ROIs without stromal tissue annotation were excluded. For further information on the employed dataset refer to B.1.2.

3.2.1.2.2 Preprocessing Stain normalization was performed in the preprocessing step to each image using the method proposed by Vahadane et al. [309]. For each image, this method involves a stain separation by sparse non-negative matrix factorization, in which the image is decomposed into sparse and non-negative stain density maps, followed by a structure-preserving color normalization step aimed at altering only the image color appearances while preserving the structure [309]. Nuclear segmentation was performed on H&E-stained slides to recognize objects through a watershed-based cell detection method. Of note, nuclear segmentation is a crucial but challenging task in pathology image analysis, and the search for a robust, practically usable nucleus segmentation module is still ongoing [268, 150, 195]. Various different approaches have been proposed, ranging from relatively simple thresholding techniques (e.g. Otsu [232], watershed [207]) to more advanced methods, such as, for instance, those based on deep learning [268]. In this study, a watershed-based method was used for its documented advantages, namely simplicity, speed and the ease with which parameters can be adjusted and fine-tuned. This

method applies a set of mathematical operations (fast radial symmetry transform and regional minima) at different scales to identify candidate locations for nuclei [207] and is part of the most widely used nuclear segmentation methods among pathologists due to its implementation within the most popular digital pathology tools (e.g. Cell Profiler [216], Qupath [267], ImageJ [271], OpenCV [11]). The cell detection function from the analysis module of QuPath was used to perform watershed-based nuclei segmentation within tumor-associated stroma [207, 96, 317, 51]. The setup parameter was set as hematoxylin OD for the detection image, with a pixel size of 0.5 µm. Nucleus parameters were set as follows: background radius: 8 µm; median filter radius: 0 µm; σ: 1.5 µm; minimum cell area: 10 μm^2; maximum cell area: 400 μm^2 [27]. For intensity parameters, the threshold was set to 0.1, and the maximum background intensity was set to 2. An expert microscopist verified the accuracy of the automatic cell detection.

3.2.1.2.3 Feature extraction and selection 71 pathomic features were extracted using the open-source HistomicsTK package https://github.com/DigitalSlideArchive/HistomicsTK/ from each segmented nucleus. The extracted radiomics features were categorized into five groups: 6 Fourier Shape Descriptors (FSD) features, 8 gradient features, 26 Haralick features, 12 intensity-based features, 19 morphometry features. The computing algorithms can be found at https://digitalslidearchive.github.io/HistomicsTK/histomicstk.features.html. Features selection was performed using a correlation filter based on the absolute values of pairwise Spearman's correlation (ρ) coefficient to reduce feature redundancy and prevent model overfitting (To reduce the dimension of features and solve the overfitting problem). The threshold for ρ was set to 0.9. Briefly, if two features had $\rho > 0.9$, the function looks at the mean absolute correlation of each variable, and the variable with the largest mean absolute correlation is removed. A further feature reduction through a univariate analysis using a nonparametric Wilcoxon rank-sum test was performed to investigate their statistical significance with respect to the outcome (TIL/noTIL). The significantly different features ($p < 0.05$) were then selected. The optimal feature set from the remaining features was selected using the least absolute shrinkage and selection operator (LASSO) algorithm method, which was used to identify the most outcome-related

features. The LASSO is a regularization technique used to minimize the number of non-zero elements and make the solution unique. In the LASSO algorithm, the shrinkage parameter lambda was identified when the misclassification error was smallest in 5-fold cross-validation [297, 256]. The surviving features were used to build pathomic-based prediction models. Feature selection procedures were implemented in R studio software, version 4.0.2 (downloadable from http://www.R-project.org).

3.2.1.2.4 Data Processing To avoid the predominance of features with the largest scale in the analysis, z normalization was performed on the training set prior model training and applied to the test set. According to z normalization, each feature was normalized as (x-x')/s where x, x', and s are the feature, the mean, and the standard deviation, respectively.

3.2.1.2.5 Model Building and Performance Evaluation Before data processing, data were randomly split into training (70%) and testing sets (30%) but stratified to ensure that the ratio of positive samples to negative samples was the same between the training and test set. Five different models were used to evaluate the discrimination power of pathomic features: Linear Discriminant Analysis (LDA), K-Nearest Neighbors (KNN), Decision Tree (DT), Random Forest (RF), Multi-layer Perceptron (MLP). Five common classifiers from different families were chosen to see which of them provides the best classification performance. K-Fold cross-validation (CV) with K = 5 was applied to the training data set for model selection (hyperparameter tuning). Concerning model parameters, the number of neighbors k in k-NN was selected from [5,7,9]; the number of trees to grow and the number of variables randomly sampled as candidates at each split was set to $ntree = 500$ and $mtry = \sqrt{numberofpredictors}$, respectively; the confidence factor for DT was chosen from cp $\in [0.1, 0.5]$; the number of hidden layers in MLP was selected from hl $\in [1, 5]$ no tuning parameters were needed for LDA. Additionally, to overcome the imbalance in the distribution of the TILs condition, models were also trained using both downsampling and a synthetic minority oversampling technique (SMOTE) procedure on the training dataset to achieve balanced classes. The downsampling technique is a resampling method that decreases the size of the majority class to be the same or closer to the minority class

size by just taking out a random sample. On the other hand, SMOTE is a hybrid resampling method that downsamples the majority class, and artificially generates new examples of the minority class using the nearest neighbors of these cases. Furthermore, the majority class examples are also under-sampled, leading to a more balanced dataset [82]. Receiver operating characteristic (ROC) curve analysis was applied to evaluate the predictive performance of the classification models (in all three settings: not resampled, downsampled, and SMOTE-sampled training set), and the area under the curve (AUC) was calculated. Accuracy (Eq. (3.1)), sensitivity or recall (Eq. (3.1)), specificity (Eq. (3.3)), precision (Eq. (3.4)), and F1-score (Eq. (3.5)) were employed to evaluate the diagnostic performance of resulting models in the hold-out test set. The associated formulas are listed below:

$$Accuracy = \frac{(TP + TN)}{TP + TN + FP + FN} \tag{3.1}$$

$$Sensitivity = \frac{(TP)}{TP + FN} \tag{3.2}$$

$$Specificity = \frac{(TN)}{TN + FP} \tag{3.3}$$

$$Precision = \frac{(TP)}{TP + FP} \tag{3.4}$$

$$F1 = \frac{(precision * recall)}{precision + recall} \tag{3.5}$$

where TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative. The AUC-ROCs were statistically compared between different classifiers and between different resampled strategies using DeLong test with Bonferroni correction. P values less than 0.05 were considered significant. Models building and evaluation were implemented using Caret package in R studio software, version 4.0.2, downloadable from http://www.R-project.org.

3.2.1.3 Results

A total of 1037 ROIs with tissues compartments and TILs annotated were considered. Nuclei segmentation resulted in 92141 nuclei, of which 20111 TILs. Selection of pathomic features returned 21 features, of which 2 from FSD group, 2 from gradient group, 9 from Haralick group, 3 from morphometry group and 5 from intensity group (Figure 3.6). Table 3.4 and 3.5 reported the predictions of the five built machine learning models. Moreover, Figure 3.7 shows the ROC curves representing performances of each classifier on the test set, both in the original setting, in the downsampled setting and the SMOTE-sampled setting. The performance of RF were the best according to AUC and the DeLong test, with or without applying resampling techniques, both in the training ($AUCROC_{RF}$ = 0.857-0.858) and the test set ($AUCROC_{RF}$ = 0.855-0.857). DeLong test revealed that all except 9 AUC comparisons were significant. It can be observed that the use of resample technique did not significantly affect MLP and RF prediction performances. In addition, performances of KNN with downsampling were not significantly different from those of KNN with smote resampling. The same behavior was observed for LDA. In addition, classification performances of KNN with downsampling were not significantly different from those of LDA without resampling. Refer to figure 3.8 for an overview of DeLong test results. Of note, although the sensitivity of models with SMOTE or down resampling in training and validation sets declined, the specificity improved greatly. However, RF produced the best classification performances, suggesting that these models could potentially be employed to discriminate TILs from noTILs. Comparing the AUCROC for SMOTE, downsampling, and the original data, applying resampling does not improve performances (Refer to Table 3.5).

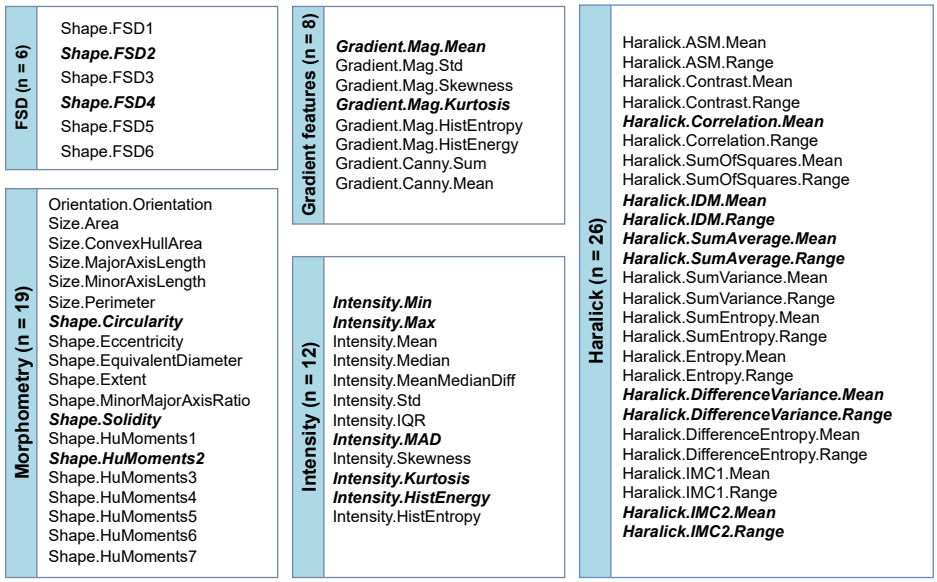


Figure 3.6. Selected pathomic features (n = 21) after the three steps of feature selection. Surviving features were reported in bold. Abbreviations: FSD = Fourier Shape Descriptor; MAD = Median Absolute Deviation; IDM = Inverse Difference Moment; IMC = Information Measure of Correlation.

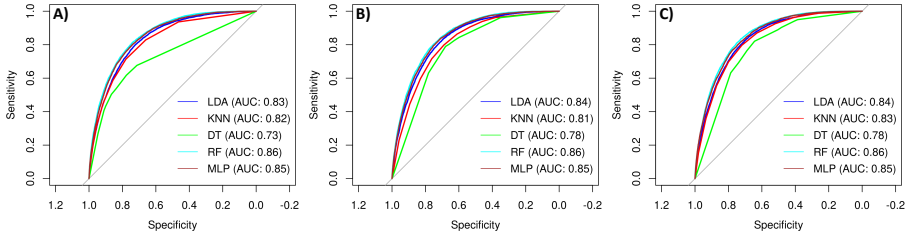


Figure 3.7. Receiver Operating Characteristic (ROC) curves of prediction models without resampling (A), with SMOTE-resampling (B) and with down-sampling (C). Abbreviations: LDA = Linear Discriminant Analysis; KNN = K-Nearest Neighbour; DT = Decision Tree; RF = Random Forest; MLP = Multi-layer Perceptron.

Table 3.4. Average prediction performance of different pathomic models for classifying TILs from noTILs in the training set with 5-fold CV. Bold indicates best performing models. Abbreviations: AUCROC = Area Under the Receiver Operating Characteristic curve; SMOTE = synthetic minority oversampling technique; LDA = Linear Discriminant Analysis; KNN = K-Nearest Neighbour; DT = Decision Tree; RF = Random Forest; MLP = Multi-layer Perceptron.

Models	Resample	AUCROC	Accuracy	Sensitivity (Recall)	Specificity	Precision	F1-score
RF	**None**	**0.8576 ± 0.0029**	**0.8306 ± 0.0024**	**0.9398 ± 0.0025**	**0.4386 ± 0.0036**	**0.8572 ± 0.0009**	**0.8966 ± 0.0015**
	SMOTE	**0.8541 ± 0.0027**	**0.8162 ± 0.0034**	**0.8744 ± 0.0033**	**0.6077 ± 0.0083**	**0.8888 ± 0.0022**	**0.8815 ± 0.0023**
	Down	**0.8572 ± 0.003**	**0.7617 ± 0.0025**	**0.7486 ± 0.0029**	**0.8086 ± 0.0073**	**0.9335 ± 0.0023**	**0.8309 ± 0.0019**
DT	None	0.7383 ± 0.0046	0.8085 ± 0.0018	0.9413 ± 0.0133	0.3322 ± 0.0556	0.8351 ± 0.0096	0.8849 ± 0.0006
	SMOTE	0.7712 ± 0.0165	0.7162 ± 0.01	0.7013 ± 0.0216	0.7699 ± 0.032	0.9164 ± 0.008	0.7943 ± 0.0106
	Down	0.7702 ± 0.021	0.7012 ± 0.0113	0.6763 ± 0.0203	0.7905 ± 0.027	0.9207 ± 0.0075	0.7796 ± 0.0115
LDA	None	0.8365 ± 0.0019	0.813 ± 0.0016	0.9467 ± 0.003	0.3332 ± 0.0069	0.8359 ± 0.0011	0.8878 ± 0.0011
	SMOTE	0.8395 ± 0.0024	0.725 ± 0.003	0.6949 ± 0.005	0.833 ± 0.0069	0.9372 ± 0.0021	0.798 ± 0.0029
	Down	0.8392 ± 0.0024	0.7237 ± 0.0026	0.693 ± 0.0037	0.8335 ± 0.0065	0.9372 ± 0.0022	0.7968 ± 0.0023
KNN	None	0.8232 ± 0.0036	0.817 ± 0.0022	0.9176 ± 0.002	0.4563 ± 0.0067	0.8582 ± 0.0015	0.8869 ± 0.0014
	SMOTE	0.8151 ± 0.0025	0.7159 ± 0.0039	0.6905 ± 0.0048	0.8069 ± 0.0079	0.9277 ± 0.0027	0.7917 ± 0.0033
	Down	0.8335 ± 0.0032	0.7385 ± 0.0056	0.7208 ± 0.0061	0.8017 ± 0.0044	0.9288 ± 0.002	0.8117 ± 0.0046
MLP	None	0.8524 ± 0.0035	0.8253 ± 0.0034	0.9279 ± 0.0093	0.4569 ± 0.0313	0.8598 ± 0.0059	0.8925 ± 0.0024
	SMOTE	0.8518 ± 0.0024	0.7628 ± 0.0254	0.7568 ± 0.0461	0.7844 ± 0.0492	0.9271 ± 0.0117	0.8324 ± 0.0235
	Down	0.8534 ± 0.0029	0.7559 ± 0.0196	0.743 ± 0.0344	0.8023 ± 0.0341	0.9313 ± 0.0085	0.826 ± 0.0186

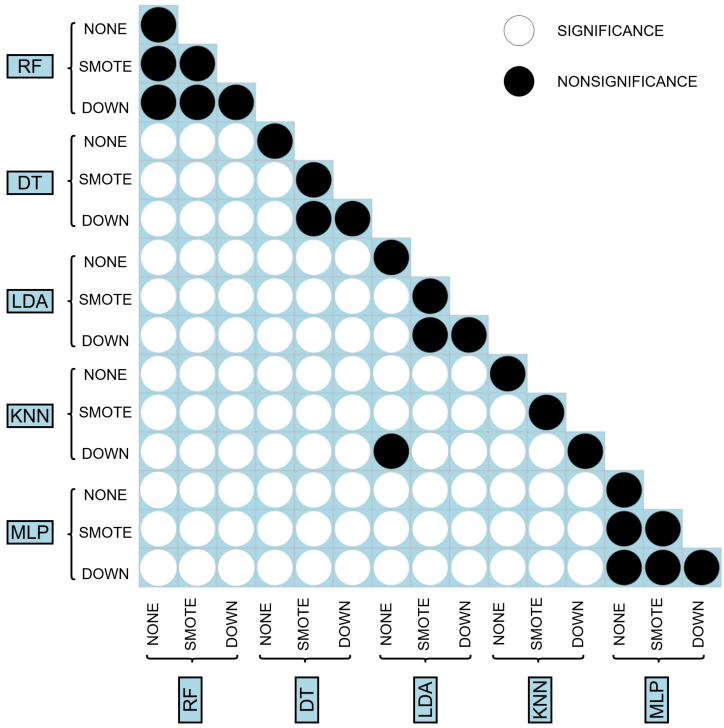


Figure 3.8. Comparison of AUCROCs between different classification models and different resampling techniques. Abbreviations: SMOTE = synthetic minority oversampling technique; LDA = Linear Discriminant Analysis; KNN = K-Nearest Neighbour; DT = Decision Tree; RF = Random Forest; MLP = Multi-layer Perceptron.

Table 3.5. Average prediction performance of different pathomic models for classifying TILs from noTILs in the test set. Bold indicates best performing models. Abbreviations: AUCROC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; SMOTE = synthetic minority oversampling technique; LDA = Linear Discriminant Analysis; KNN = K-Nearest Neighbour; DT = Decision Tree; RF = Random Forest; MLP = Multilayer Perceptron

Models	Classifier	AUCROC (95% CI)	Accuracy (95% CI)	Sensitivity (Recall)	Specificity	Precision	F1-score
RF	**None**	**0.8567 (0.8518, 0.8567)**	**0.8296 (0.8251, 0.834)**	**0.9395**	**0.4373**	**0.8563**	**0.8960**
	SMOTE	**0.8552 (0.8503, 0.8552)**	**0.8158 (0.8112, 0.8204)**	**0.8733**	**0.6107**	**0.8890**	**0.8811**
	Down	**0.8569 (0.8521, 0.8569)**	**0.7617 (0.7567, 0.7667)**	**0.7475**	**0.8125**	**0.9344**	**0.8305**
DT	None	0.8182 (0.8124, 0.8182)	0.8003 (0.7955, 0.805)	0.9107	0.4062	0.8455	0.8769
	SMOTE	0.7751 (0.7693, 0.7751)	0.7037 (0.6983, 0.7091)	0.6796	0.7896	0.9202	0.7818
	Down	0.8284 (0.823, 0.8284)	0.6825 (0.677, 0.688)	0.6438	0.8206	0.9276	0.7601
LDA	None	0.834 (0.8286, 0.834)	0.8134 (0.8087, 0.818)	0.9467	0.3376	0.8361	0.8880
	SMOTE	0.8372 (0.8319, 0.8372)	0.7225 (0.7172, 0.7278)	0.6906	0.8365	0.9378	0.7954
	Down	0.8367 (0.8314, 0.8367)	0.722 (0.7166, 0.7272)	0.6903	0.8352	0.9373	0.7950
KNN	None	0.8182 (0.8124, 0.8182)	0.8136 (0.809, 0.8182)	0.9159	0.4487	0.8557	0.8848
	SMOTE	0.8097 (0.8039, 0.8097)	0.7125 (0.7072, 0.7179)	0.6883	0.7991	0.9244	0.7891
	Down	0.8284 (0.823, 0.8284)	0.7362 (0.7309, 0.7413)	0.7189	0.7977	0.9269	0.8098
MLP	None	0.8491 (0.8439, 0.8491)	0.823 (0.8185, 0.8275)	0.9476	0.3782	0.8447	0.8932
	SMOTE	0.8481 (0.8431, 0.8481)	0.7662 (0.7612, 0.7712)	0.7632	0.7768	0.9243	0.7632
	Down	0.8503 (0.8453, 0.8503)	0.7476 (0.7424, 0.7527)	0.7290	0.8140	0.9333	0.8186

3.2.1.4 Discussion

Growing evidence has revealed that TILs are an important prognostic positive biomarker in BC patients [110]. Despite standardization efforts, the quantification of TILs is still based on a visual assessment, thus being time-consuming, labor-intensive, requiring large experience and expertise, prone to the inter-pathologist discrepancy, and associated with a mere qualitative evaluation of histological properties [265]. Therefore, the TILs marker is still not considered robust enough to be used in clinical practice, mainly because there are no standardized, objective, and effective TILs quantification strategies. Despite efforts from many researchers, the pathologist does not regularly assess the TIL infiltration due to a lack of automatic image analysis tools. Automatic detection of TILs using computational image analysis approaches could contribute to developing objective and precise mechanisms to measure the infiltration grade of BC, with a look towards the possibility of reaching a standardized TILs assessment that pathologists could then use in clinical practice for decision-making and treatment planning. Most current literature on automated TILs quantification relies on DL-based methods, which could be highly effective for the task, but due to their not immediate interpretability, fail to aid pathologists with novel insights on the tissue. Given the need for studies contributing to achieving a quantitative, automatic, interpretable, and as simple as possible TILs evaluation, a pathomic-based approach for classifying TILs on BC digital pathology images was proposed, exhibiting several types of interpretable features for automatic TILs classification. Results showed that some specific pathomic features extracted from nuclei in tumor-associated stromal regions of H&E images of BC could classify if nuclei were TILs or not. In particular, several machine learning models with selected pathomic features that could classify sTILs with high prediction performances after a feature selection process were developed. Considering the prediction performances of ML models, RF revealed the highest predictive value among different classifiers, with AUCs reaching 86%. Downsampling and SMOTE sub-sampling techniques were applied to tackle the unbalanced data problem and showed improvement in specificity. Machine learning algorithms tend to sacrifice the minority group for an unbalanced dataset to achieve higher accuracy [57]. Although both two subsampling methods improved specificity, SMOTE was able to enhance

AUC in a balanced way: while maintaining high sensitivity, it was also able to increase specificity significantly. In addition, SMOTE has been shown to be robust to the variation of unbalanced ratios with various classifiers [82]. Previous studies aimed at detecting TILs using visual features from H&E images of different cancer types [53, 39, 267]. Basavanhally et al. [53] developed an approach combining a region-growing algorithm and Markov random fields to identify lymphocytes in HER2+ BC images. Kuo et al. [177], in a study aiming at lymphocytes infiltration for BC patients ante-lymphadenectomy, obtained AUCs up to 0.87 using a SIFT algorithm for classifying lymphocytes from non-lymphocytes. However, these studies involved smaller datasets and used different segmentation approaches and features with respect to this study. Moreover, they were published before the development of TILs-WG guidelines. Corredor et al. [93] performed a study aiming at detecting TILs on lung cancer H&E images using shape, texture, and color features. They also automatically detected nuclei with a watershed approach and obtained 89%, 83% and 86% in precision, recall and F1 score for TILs detection. Of note, these studies developed the classification approaches training SVM classifiers. In this study, for exploring the behaviour of machine learning models that were suitable in case of training data much larger than the number of features, thus not exploring models that perform better with a low amount of training data and large features [62]. Moreover, differently from the above-mentioned studies, this work was focused on sTILs, according to the recommendations of TILs-WG [265]. Other studies aimed at developing DL-based computational TIL assessment methods broadly following some or all steps of visual TILs-WG guidelines [287, 295, 39]. Lu et al. [199] developed a U-Net model for lymphocyte detection showing 0.9536, 0.901, 0.9266 of recall, F1-score and precision, and strong association with immune response of outcomes. Le et al. [184], in a study aiming at exploring the correlation between tils distribution and invasive tumor through CNNs, obtained AUCs up to 0.950 for TILs. Sun et al. [287] developed a deep learning-based tool for an automatic til score assessment and achieved a F1-score of 0.856 for nuclei (also including TILs) classification. Similar approaches were also developed for detecting TILs in other cancer types [280, 47, 328, 272]. Although the promising results achieved by DL-based approaches, a hand-crafted pathomics approach was preferred rather than a fully automated DL one as

this could give better control over the initial data analysis and subsequent ML model construction and a better understanding of the whole pipeline. In fact, although the majority of DL-based methods have strong predictive performance and are comparable to the traditional hand-crafted approaches, these methods have limited ability to explain the deep features with neither a set of diagnostic plans nor an insight into the results. In particular, it was considered appropriate and constructive to explore hand-crafted pathomic features since pathomic and in general digital pathology is a recently explored field and, given the additional information related to the underlying biological processes that pathomics could provide, this study may add value in this direction. In this regard, while it might be difficult to demonstrate a causal relationship between pathomic features and underlying biological processes, it is possible to hypothesize about their underlying connections. It is worth noting that among features that passed the feature selection step and thus constitute ML models, there are features associated with shape (e.g. circularity, FSD) and intensity (mean, median absolute deviation and energy), and this could be related to the well-known properties of TILs of having a regular shape, clearer margins and higher peak intensity than other cells. Moreover, the contribution of features associated with gradient (mean, kurtosis) and Haralick features could be attributable to the more homogeneous enhancement and different textural distribution pattern of TILs than other cells [94]. Of note, Haralick features from nuclear were shown to be prognostic among different cancer types [334]. Despite some promising results emerging in this preliminary study, there are still many limitations which have to be acknowledged. The first concerns the unbalanced nature of the dataset, which has been tried to account for by applying SMOTE and downsampling techniques that have demonstrated their value in the setting of medical imaging ML analysis [82]. Moreover, although the dataset was very large given the high number of nuclei used for training and test models, the number of corresponding WSIs is from a limited number of patients. On a positive note, the study has the advantage of involving patients from different institutions. Another issue concerns the prerequisite of pixel level annotation for individual nuclei segmentation within tumor-associated stroma, thus implying that any segmentation errors can have a direct impact on the final performance. Although the choice of an automatic and well-established

approach guarantees robustness and reproducibility, the lack of a segmentation ground truth does not allow to quantitatively evaluate and control the segmentation error (e.g. through Dice index or F1-score [341, 180]) and also to test more advanced segmentation methods (e.g. training neural networks), also by making comparisons (as done in previous studies) with other existing nuclear segmentation algorithms implemented in different software/studies [269, 176, 171]. It must be emphasized, however, that the problem of segmentation is inherent in hand-crafted approaches that aim to extract quantitative descriptors from regions of interest (e.g. radiomics, pathomics), and embraces all the advantages and disadvantages associated with the different segmentation techniques (manual, semi-automatic, automatic) [29]. It is worth noting that the implemented segmentation has the advantage of being fully automatic, and so not prone to inter- and intra-observer variability, replicable and not time-consuming for the pathologist [138]. As previously highlighted, a limited number of features was explored in this study to be able to interpret them directly and give a physical meaning as reliable as possible. An essential criterion for having examined these features is that they are easy to interpret, as opposed to features obtained from neural networks [132, 287, 295]. However, this framework could be used in future studies evaluating an expanded suite of pathomic measurements, as well as extended to characterizing TILs in other cancer types. For example, it might have been interesting to explore nuclear architectural information and local cell cluster graph-based measurements that have been reported to be of prognostic value in non-small cell lung cancer [94]. However, this is related to another issue concerning the missing information on both diagnostic and prognostic outcomes associated with patients (e.g., it is only known that the patients had HER2+ or TNBC immunophenotype, but the distribution was not specified; moreover, no prognostic outcomes were provided), which prevented us from performing diagnostic/prognostic predictions using the TIL assessment obtained using the developed classification method. Moreover, also based on previous studies tried to assess TILs using immunohistochemistry (IHC) images [49, 335, 333], it could be interesting to investigate pathomic features from IHC to add value to the developed approach. While TIL scoring in conventional H&E sections has the advantage that no additional stains are required, lymphocyte subtyping into cytotoxic, helper, regulatory, and other T cells not possible

in H&E may become clinically relevant in the future. Therefore, in addition to TIL detection through morphological (H&E) features, subtyping of TILs through IHC could be relevant [163]. Another critical limitation affecting the study concerns the well-known lack of shared reference standards concerning data storage, the missing agreement on analysis procedures, and the feature reliability and reproducibility limitations affecting pathomics [186, 100]. In particular, the existing lack of standardization in image acquisition, preprocessing, segmentation methods, and pathomics analysis tools, could lead to discrepancies in feature measurements that are not due to underlying biological variations [134, 226]. However, all steps of pathomic workflow performed in this study were reported in detail since it is essential to develop this emerging field in terms of clinical translation and to improve the reproducibility of study outcomes [324]. It is worth noting that the proposed approach has followed a fragment of the TIL-WG guideline, by leveraging on annotations of tumor-associated stroma in which sTILs were annotated. In particular, having worked on the tumor-associated stroma annotations masked some intrinsic pitfalls associated with the identification of stromal regions (e.g. the exclusion of areas of necrosis, ductal and lobular carcinoma in situ (DCIS/LCIS), and normal breast tissue) [172]. Studies trying to address all the concepts of stromal and intratumoral TILs and accounting for confounding morphologies specific to different tumor sites, subtypes, and histologic patterns as envisioned by the TIL-WG are required [295]. In addition, it should also be highlighted that a general limitation of an automated approach for TIL evaluation concern whether its performance should be measured i) as the concordance between manual and automated sTIL score (namely the agreement between automatic TIL score and TIL score visually evaluated by pathologist or ii) according to the clinical outcome of the patient (e.g., the ability to predict survival or response to treatment, or a mix of both [184, 287]. Although findings of this study require careful interpretation due to the limitations mentioned above, the proposed hand-crafted pathomic approach revealed interesting results that may represent a starting point for future research aiming at developing a quantitative, standardized, interpretable, automatic and as simple as possible method for TIL assessment. In the future, the generalizability of the proposed method should be validated on more datasets with larger WSI, incorporating it

into a more complex pipeline that includes all the steps of the TIL-WG guidelines, also exploring the impact of different intermediate processing steps (e.g. nuclear segmentation, stain normalization) on the TILs classification performance.

3.2.1.5 Conclusions

The proposed pathomic approach accurately classified sTILs on H&E images of HER2+ and TNBC. This study brought to attention several types of quantitative features that are potentially easier to interpret and linked to underlying biological processes that can assist pathologists in improving diagnosis by increasing their understanding of the pathology. Compared with the state-of-the-art DL approaches, the proposed method is therefore an important step towards quantitative, reproducible, interpretable, and rater-independent TILs quantification analysis in BC histopathology and is also easily applicable to new datasets and domains.

3.2.2 An automated pathomic-based approach for the estimation of prostate cancer grade on digital pathology images

3.2.2.1 Introduction

As highlighted in the previous sections, AI is rapidly reshaping the landscape of medical pathology, particularly in the oncologic field, where the computational analysis of digital pathology images, known as pathomics, is garnering substantial interest for a wide range of applications. Pathomics converts digitized WSIs into mineable datasets based on AI algorithms and links these extracted and quantified pathological features to clinically related indicators. This innovative approach has the potential to facilitate tasks such as cancer classification with notable advantages, including improved workflow, efficiency, reproducibility, reduced costs, and enhanced training opportunities. In pathomic approaches, a computer is "trained" using a data set of sample images of tumors that have been annotated and/or classified by pathologists. Machine learning-based techniques extract a set of predefined hand-crafted features from WSIs patches (smaller images in which WSIs are often divided due to their considerable

size) or regions of interest, and the computer uses the classification information to develop its own pattern-recognition criteria with which to perform the classification task. The computed features are then fed into a standard classifier such as support vector machines, random forests, or Bayesian classifiers to perform the classification. Another class of methods that have become popular in recent years are based on Deep Learning, an advanced subfield of machine learning characterized by the use of artificial neural networks featuring multiple, hidden layers. These methods are not limited to the hand-crafted features and avoid the feature engineering step by performing the feature learning and classification in a single framework. Given sufficient training data and proper training, these should be able to discover better discriminative visual patterns and lead to superior results, but in the medical domain, especially in digital pathology, these are challenged by a scarcity of labeled training data. The accurate identification of PCa on a WSI and its grading represents one of the most important open challenges in the field of digital pathology. This because PCa, other than being one of the most common cancers worldwide, accounting for a large proportion of all cancer-related deaths [59], is also characterized by an extremely high heterogeneity, both within and between individuals. The Gleason grade is the well-established histological parameter used to assess the aggressiveness of PCa based on gland structures identified in biopsy or radical prostatectomy specimens and, along with other clinical variables, is commonly used to stratify the risk for managing PCa patients. In particular, PCa is graded according to a Gleason pattern scale and assigned a score corresponding to the Gleason grades of the two most predominant patterns, which are then used to classify patients into five grade groups that predict prognosis [112, 113, 193]. In more datail, the Gleason grading system classifies histological patterns of prostate tumors into five levels, ranging from least aggressive (grade 1) to most aggressive (grade 5), based on gland structures observed in the tumor biopsy. Typically, tumors exhibit two patterns, and the original Gleason score is determined by adding the two most prevalent patterns, resulting in scores from 2 to 10. However, the current application of the Gleason grading system has undergone changes from its original version. Presently, pathologists primarily assign Gleason scores of 6-10 based on the two dominant tumor patterns, as assigning scores of 2-5 has shown poor reproducibility and weak correlation

with radical prostatectomy grades. For instance, patients with a Gleason score of 6 or below are typically considered low risk, those with a score of 7 fall into the intermediate-risk category, and patients with a Gleason score of 8 or higher are classified as high risk. Figure 3.9 displays examples of prostatic tumor regions with different Gleason patterns and benign prostatic tissues. As observed in Figure 1, the textural appearance of prostate cancers varies between grades due to abnormal changes in gland structures.

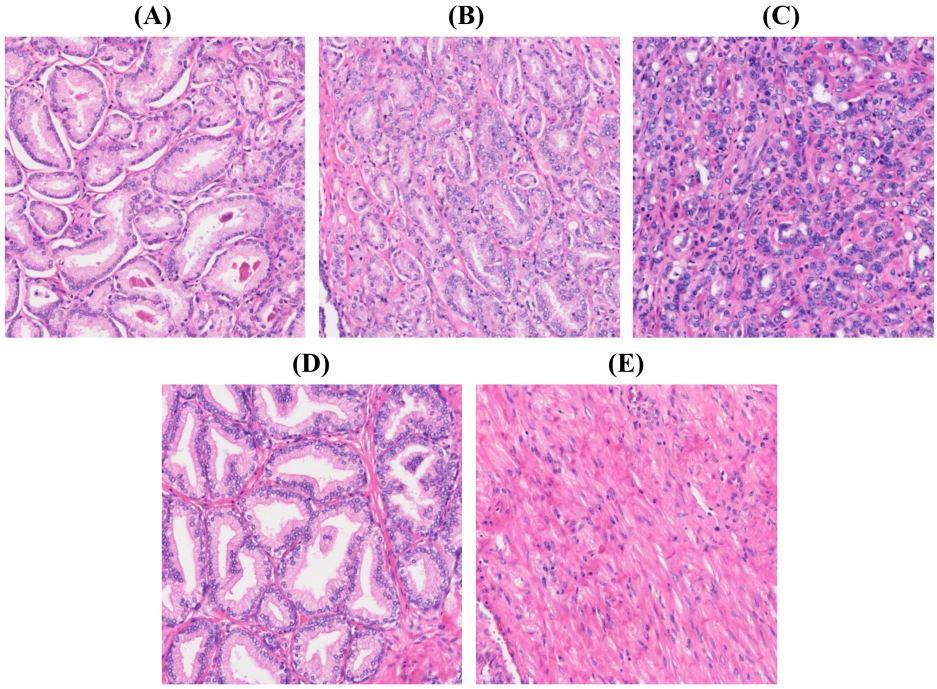


Figure 3.9. Examples of prostatic cancerous and benign tissue regions: (A) well-formed glands (G3); (B) poorly formed glands (G4); (C) poorly formed glands and single cells (G5); (D) Normal tissue; (E) Stroma.

Gleason grading of tumor biopsies remains one of the most powerful prognostic predictors in prostate cancer and its accuracy is crucial to determine the appropriate treatment. However, this process is time-consuming, labor-intensive and prone to inter- and intra-observer variability due to the extreme PCa biological heterogeneity and its characteristic variable

molecular alterations, thus posing risks of misdiagnosis and over-diagnosis [277]. Therefore, the ability to automatically diagnose PCa and identify

Gleason patterns to assign Gleason scores from diagnostic pathology slides would have significant impacts on clinical decision making, and is required to improve the efficiency, accuracy, and consistency of Gleason grading. Recent advances in histology scanning technology, Artificial Intelligence (AI) and computer vision offer great opportunities to improve the PCa diagnosis and the fidelity of Gleason grading. Several published studies on digital pathology have addressed the automatic diagnosis and grading of PCa [329, 143]. Among the early studies in automated PCa grading, gland-nuclei-based approaches sought to leverage the distinctive characteristics of glandular structures and nuclei within tissue samples [228]. However, these methods encountered notable challenges when confronted with the task of effectively classifying regions of high-grade carcinoma where some of the tissue structures (e.g. as glands) cannot be differentiated. In contrast, texture-based have the ability to operate at the tissue level, alleviating the need for intricate cellular or glandular segmentations, thus simplifying the computational process and allowing to broaden the applicability of texture-based methods across diverse prostate cancer presentations. Textural features can translate intricate textural patterns within WSIs into meaningful insights, representing a paradigm of interpretability with direct implications into the decision-making process [221]. In recent years, the advent of DL has heralded a new era in PCa computational pathology. Harnessing the power of neural networks, deep learning methods exhibit an unparalleled capacity to discern complex patterns within PCa pathology images [299]. Despite their potency, these approaches grapple with a critical trade-off associated with the lack of interpretability. The intricate layers of abstraction achieved by deep learning models often render them as 'black boxes,' hindering our ability to understand the underlying rationales for specific classifications. Moreover, their voracious appetite for data, particularly large annotated datasets, poses a logistical challenge, limiting their accessibility and applicability in settings where such extensive datasets may be unavailable [250]. Subsequent breakthroughs in the applicability of DL models toward reproducible prostate cancer AI diagnostic tools started in 2019 with the introduction of models trained and tested on very large cohorts [78, 225, 76]. One of these approaches [78]

has also led to the first FDA-approved AI-based pathology product for in vitro diagnostics. Although all these efforts have been made in improving the utility of digital pathology within the realm of PCa, several advancements are still required before massive scaling and adoption of these models should occur in the clinical realm. In navigating the complex ground of PCa diagnosis and grading, it becomes evident that each approach brings its unique strengths and limitations to the fore. The quest for a comprehensive and reliable solution requires a deep understanding of the interplay between these methodologies and a keen exploration of innovative pathways that synergize their advantages while mitigating their respective shortcomings. Of note, the inability to interpret the extracted signatures and the predictions made by the model is considered to be one of the major issues that limit the acceptance of AI models in medicine. In particular, although many pathomic models have been developed and achieved good results, pathologists are often hesitant to adopt such models due to the unclear internal mechanisms or unexplained features [330]. In other words, interpretability largely hinders the general application of the pathomic model. In this study, the focus has been placed on addressing well-established issues such as PCa diagnosis and grading, which are actively being tackled by the scientific community. PCa classification tasks are being utilized to explore pathomic features and their classification performance. In particular, we explored multiscale, multi-channel pathomic-based classification tasks for PCa diagnosis and Gleason Score classification from H&E-stained digital pathology images. The study design encompasses not only a thorough evaluation of the efficacy of the developed models in distinguishing malignant from non-malignant tissue and classifying different Gleason patterns, but aims to detect which are the most discriminant features for automated diagnosis and gleason grade estimation. By delving into an extensive range of features, the aim is to aspire to offer novel pathomic insights into PCa diagnosis and grading. In particular, unlike conventional methods that predominantly operate within a specific magnification range, multiple magnification scales will be explored, capturing the intricate details of prostate cancer pathology at both 20×, 10× and 5× magnifications, aligning with the commonly employed diagnostic practices of pathologists. Texture-based pathomic features were opted for exploration since they can provide a tangible and interpretable

representation of patterns and structures within pathological images, aiding clinicians and researchers in understanding the underlying characteristics of tissues. Moreover, texture-based analyses often require less data than DL methods while still capturing relevant information. Unlike the majority of studies that are traditionally limited to specific feature subsets such as wavelet, color, or local binary patterns (LBP) [243, 60, 293], this study explores an expansive feature space, integrating multiple channels alongside the commonly used grayscale-converted RGB images. This expansive feature exploration broadens the diagnostic spectrum, capturing a more comprehensive understanding of PCa heterogeneity. Through rigorous evaluation of predictive models, as well as a deep analysis of variable importance across models and resolutions, an attempt will be made to unravel the intricate dynamics between pathomic features and PCa clinical and histopathological outcomes, aiming to provide insights on PCa pathomics.

3.2.2.2 Materials and Methods

In this section, data used in this study, as well as the preprocessing, feature extraction and selection methods are introduced. Moreover, the classifier modeling and metrics used for the evaluation are provided. The overall methodologic workflow is shown in Figure 3.10

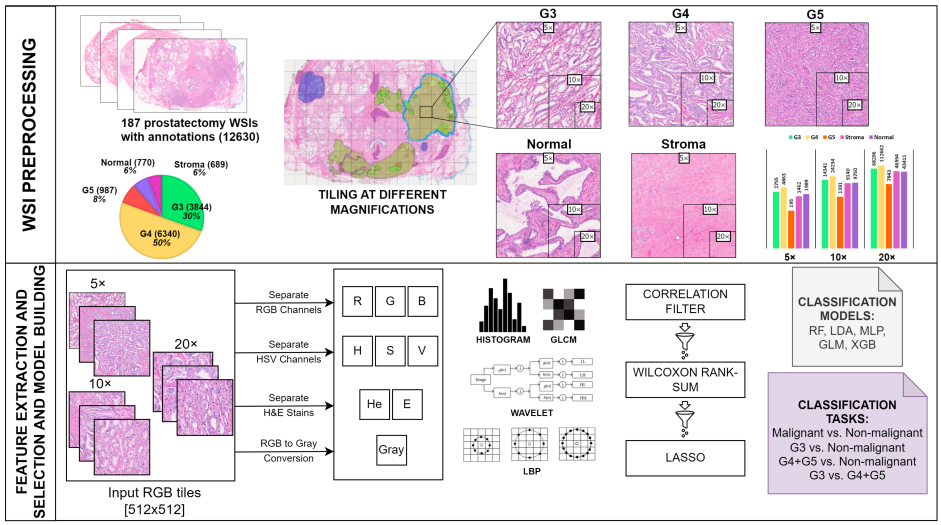


Figure 3.10. Workflow of the pathomic approach implemented in the study. Starting from H&E prostatectomy WSIs, the approach starts with the WSI tiling at three different magnification scales (5×, 10× and 20×). Pathomic features were then extracted from each tile and selected to train classification models.

3.2.2.2.1 Dataset A public dataset of H&E-stained whole slide image dataset of prostatectomy specimens was used (https://aggc22.grand-challenge.org/AGGC22/). The dataset examined in the study consists of 187 H&E-stained WSIs dataset of prostatectomy with annotations performed by experienced pathologists. The WSIs were acquired with brightfield imaging at 0.5µm × 0.5µm per pixel resolution (20× magnification). Class labels of the annotations include Gleason pattern 3 (G3), Gleason pattern 4 (G4), Gleason pattern 5 (G5), normal prostetic tissue, and stroma tissue. For further information on the employed dataset refer to B.1.3.

3.2.2.2.2 Preprocessing Because of different storage time, procedures and stains of biopsy slides, the digitized histological images generally have different color appearances. To suppress the influence of staining variations on textural feature computations, a color deconvolution based method [203] was applied to normalize dermis pixels. For each image, we first adaptively determine the H&E stain vectors by using a singu-

lar value decomposition method [203]. The color deconvolution is then performed to normalize the image into standard color appearance based on pre-determined H&E stain vectors. The feature extraction step implemented in this study is based on patches that are cropped from the digitized slide images. As suggested in Doyle et al. [103], histological features may be meaningful at different resolutions. Thus, we process image patches at different magnification scales. On the one hand, the larger the patches are, and the lower the scale is, the less localized are the features. On the other hand, smaller patches in higher scales may not contain enough information to allow extraction of descriptive features. Also, the smaller the sampling rate of the patches is, the more repetitive the data set is, and the more computation time is required. Thus, the patch size, scales and sampling rate were selected such that most histologic structures (e.g., glands) are captured, while maintaining enough data samples for a fine representation of the annotated labels. Three magnification levels, 5×, 20× and 20×, were investigated, and a patch size of 512×512 pixels for tiling without overlap (thus, the scales correspond to 1024×1024µm^2, 512×512µm^2 and 256×256µm^2, respectively). The patches are cropped from the native 20× resolution of the core image, and lower scales (10× and 5×) are downsampled into patches of size 512×512 pixels (3.11). The tiling algorithm that developed selects tiles from the images while obeying the following rules: at least 90% overlap with tissue annotations and tile overlap of 0%. ROIs that were too small to extract 5 tiles from were excluded.

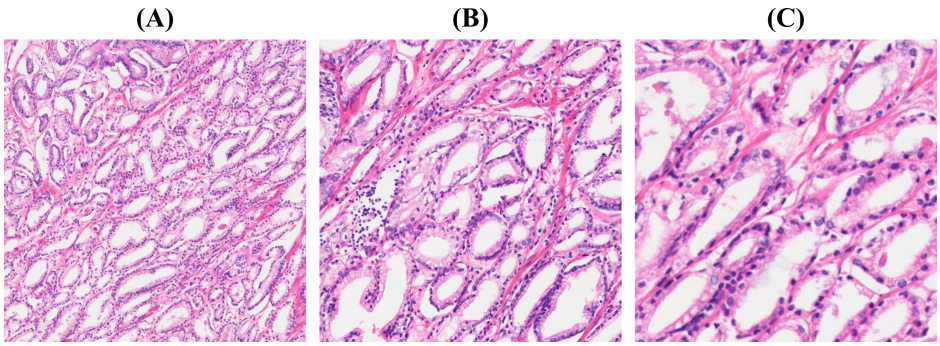


Figure 3.11. Example of 512×512 tiles of a Gleason pattern 3 mask at 5× (A), 10× (B), 20× (C) magnification at 5×, 10×, 20× magnification.

3.2.2.2.3 Feature extraction To describe the characteristics of histological images we investigated four sets of descriptors: histogram, texture features computed in terms of co-occurrence (GLCM), wavelet and LBP. All these features were extracted from the provided images (that are native red/green/blue (RGB) images) converted to grayscale and also from the RGB and HSV color spaces, as well as for H and E stains separated applying color deconvolution. Tiles were processed using a custom, in-house MATLAB function to extract pathological features. The gray level histogram of a given image was used to construct a histogram feature sets containing 18 features: Minimum, Maximum, Mean, Median, Standard-Deviation, Variance, Skewness, Kurtosis, Tenth Percentile, Ninetieth Percentile, Energy, Entropy, Mean Absolute Deviation, Robust Mean Absolute Deviation, Range, Interquartile Range, Root Mean Squared, Uniformity. The second feature set was based on GLCM features. In particular, we calculated 14 Haralick features according to Haralick et al. [137]: Angular Second Moment, Energy, Contrast, Correlation, Variance Inverse Different Moment, Sum Average, SumVariance, Sum Entropy, Entropy Difference Variance, Difference Entropy, Information Measure Of Correlation I, Information Measure Of Correlation II, Maximal Correlation Coefficient. The third feature set consists of LBP uniform features, that were obtained using MATLAB's extractLBPFeatures function. LBP consists of computing the distribution of binary patterns in the circular neighbourhood of each pixel. The neighbourhood is characterized by a radius R and a number of neighbours P. The principle is to threshold neighbouring pixels compared to the central pixel. The value 1 is assigned to each of the P neighbours if the current pixel intensity is superior or equal to the central pixel intensity. Otherwise, the value 0 is assigned. Thus, for each pixel, a binary pattern is obtained from the neighbourhood. The number of neighbours P was set as 8, and 59 features were obtained. The fourth feature set included wavelet features. In particular, a discrete wavelet transform (DWT) was performed on each pathology image using the Haar wavelet function. During the two-dimensional wavelet transform process, a digital image is decomposed into four images (hereafter called LL, HL, LH, and HH). Here, 'H' stands for high-pass filter, 'L' stands for low-pass filter, and 'HL' means a high-pass filter corresponding along the row direction of the input image and a low-pass filter following along the column direction.

The same nomenclature applies to the LH, HH, and LL images. Then, the histogram and GLCM features were calculated to extract texture features from the wavelet-transformed images, thus obtaining 128 wavelet features for each image. Overall, 219 features were extracted from 9 images, namely the rgb-to-grayscale conversion image, Red (R), Green (G), Blue (B), Hue (Hu), Saturation (S), Value (V), H stain (H), E stain (E), for a total of 1971 features for each tile, regardless of the associated resolution (5×, 10×, 20×).

3.2.2.2.4 Feature selection Features selection was performed using a correlation filter based on the absolute values of pairwise Spearman's correlation (ρ) coefficient to reduce feature redundancy (also considering the inter-channel correlations [278]) and prevent model overfitting. The threshold for ρ was set to 0.9. Briefly, if two features had $\rho > 0.9$, the function looks at the mean absolute correlation of each variable, and the variable with the largest mean absolute correlation is removed. A further feature reduction through a univariate analysis using a nonparametric Wilcoxon rank-sum test was performed to investigate their statistical significance with respect to the binary outcome, depending on the investigated classification task. The significantly different features ($p < 0.05$) were then selected. The optimal feature set from the remaining features was selected using the least absolute shrinkage and selection operator (LASSO) algorithm method, which was used to identify the most outcome-related features. The LASSO is a regularization technique used to minimize the number of non-zero elements and make the solution unique. In the LASSO algorithm, the shrinkage parameter lambda was identified when the misclassification error was smallest in 5-fold cross-validation. The surviving features were used to build pathomic-based prediction models. Feature selection procedures were implemented in R studio software, version 4.0.2 (downloadable from http://www.R-project.org).

3.2.2.2.5 Data processing To avoid the predominance of features with the largest scale in the analysis, z normalization was performed on the training set prior model training and applied to the test set. According to z normalization, each feature was normalized as (x-x')/s where x, x', and s are the feature, the mean, and the standard deviation, respectively.

For binary classification, we further grouped the G3, G4, and G5 tissue as "Malignant", whereas normal and stroma tissue comprise the "Non-malignant" group. Then, the following four classification tasks were investigated: Malignant vs. Non-malignant, G3 vs. Non-malignant, G4+G5 vs. Non-malignant, G3 vs. G4+G5, for each magnification, for a total of 12 classification tasks.

3.2.2.2.6 Model building and performance evaluation Before data processing, data were randomly split into training (80%) and testing sets (20%) but stratified to ensure that the ratio of positive samples to negative samples was the same between the training and test set. This strategy maintained the proportional distribution of classes (positive or negative) in both the training and testing sets, ensuring an accurate representation of the various sample categories. Additionally, in order to preserve the individuality of each patient in the training and testing data, an additional step was taken for balancing. While each patient may contribute multiple tiles to the overall dataset, it was ensured that patient IDs did not appear simultaneously in both the training and testing sets. This approach was implemented to prevent any potential loss of model generalization due to the simultaneous presence of the same patients in both sets, thus ensuring a fair and balanced representation of data diversity in the model training and evaluation process. Five different models were used to evaluate the discrimination power of pathomic features: Gradient Boosting Machine (GBM), Linear Discriminant Analysis (LDA), Random Forest (RF), Multi-layer Perceptron (MLP), eXtreme Gradient Boosting (XGB). The Random forests were constructed using the ranger R package [325]. K-Fold cross-validation (CV) with $K = 5$ was applied to the training data set for model selection (hyperparameter tuning). Additionally, to overcome the imbalance in the distribution of the positive vs negative condition, models were also trained using downsampling, namely a resampling method that decreases the size of the majority class to be the same or closer to the minority class size by just taking out a random sample. ROC curve analysis was applied to evaluate the predictive performance of the classification models, both without resampled and downsampled. The AUC-ROC was calculated and accuracy, sensitivity or recall, specificity, precision, F1-score, balanced accuracy, and G-score were employed to eval-

uate the diagnostic performance of resulting models in the hold-out test set. The AUC-ROCs were statistically compared between different classifiers and between different resampled strategies using DeLong test with Bonferroni correction. P values less than 0.05 were considered significant. Models building and evaluation were implemented using Caret package in R studio software, version 4.0.2, downloadable from http://www.R-project.org.

3.2.2.3 Results

The total number of tiles according to tissue compartments and magnification scales are shown in Table 3.6.

Table 3.6. Number of tiles according to tissue compartments and magnification scale. Abbreviation: G3 = Gleason pattern 3; G4 = Gleason Pattern 4; G5 = Gleason pattern 5.

	Magnification		
	5×	**10×**	**20×**
G3	2755	14341	68296
G4	4863	24234	112662
G5	195	1391	7843
Normal	1462	9140	46994
Stroma	1989	9750	43411

Results of the selection of pathomic features according to the classification tasks are shown in Appendix B (B.1-B.12). In summary, for the classification task between Malignant and Non-malignant tiles, selection of pathomic features returned 53, 41, and 38 features, respectively for 5×, 10× and 20× magnification. For the classification task between G3 and Non-malignant tiles, selection of pathomic features returned 58, 51, and 47 features, respectively for 5×, 10× and 20× magnification. 57, 44 and 44 features were selected for the classification task between G4 + G5 and Non-malignant tiles, respectively for 5×, 10× and 20× magnification. Finally, the feature selection process returned 59, 53, and 51 pathomic features for the classification task between low grade and high grade, respectively

for 5×, 10× and 20× magnification. Tables from 3.7 to 3.18 reported the predictions of the five built classification models, both in the original setting, as well as in the downsampled setting, respectively for 5×, 10× and 20× magnification. Moreover, Figures 3.12,3.13, and 3.14 show the ROC curves representing performances of each classifier on the test set, both in the original setting and in the downsampled setting, at each magnification.

Table 3.7: Average prediction performance of different pathomic models for the "Malignant vs Non-malignant" classification task in the test set at 5× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting.

Mdl	AUC (95%CI)	Accuracy (95%CI)	Sens	Spec	Prec	Recall	F1	BAcc	GSc
GBM (d)	0.967 (0.961,0.973)	0.905 (0.894,0.916)	0.869	0.921	0.824	0.869	0.846	0.895	0.895
GBM (n)	0.967 (0.961,0.974)	0.916 (0.905,0.926)	0.799	0.966	0.909	0.799	0.85	0.882	0.878
LDA (d)	0.97 (0.963,0.976)	0.92 (0.909,0.93)	0.901	0.928	0.842	0.901	0.87	0.914	0.914
LDA (n)	0.971 (0.964,0.977)	0.92 (0.909,0.93)	0.833	0.957	0.892	0.833	0.862	0.895	0.893
MLP (d)	0.979 (0.974,0.984)	0.939 (0.929,0.948)	0.917	0.948	0.883	0.917	0.899	0.932	0.932
MLP (n)	0.978 (0.972,0.984)	0.943 (0.933,0.951)	0.899	0.961	0.908	0.899	0.904	0.93	0.93
RF (d)	0.965 (0.959,0.972)	0.894 (0.882,0.906)	0.866	0.906	0.798	0.866	0.831	0.886	0.886
RF (n)	0.965 (0.959,0.971)	0.907 (0.896,0.918)	0.79	0.957	0.887	0.79	0.836	0.874	0.87
XGB (d)	0.972 (0.966,0.978)	0.92 (0.909,0.93)	0.888	0.933	0.85	0.888	0.869	0.911	0.91
XGB (n)	0.977 (0.972,0.983)	0.938 (0.928,0.947)	0.864	0.97	0.924	0.864	0.893	0.917	0.915

Table 3.8: Average prediction performance of different pathomic models for the "Malignant vs Non-malignant" classification task in the test set at 10× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting.

Mdl	AUC (95%CI)	Accuracy (95%CI)	Sens	Spec	Prec	Recall	F1	BAcc	GSc
GBM (d)	0.999 (0.998,0.999)	0.987 (0.985,0.989)	0.983	0.99	0.982	0.983	0.982	0.986	0.986
GBM (n)	0.998 (0.998,0.999)	0.985 (0.982,0.987)	0.968	0.994	0.989	0.968	0.978	0.981	0.981
LDA (d)	0.992 (0.99,0.994)	0.967 (0.963,0.969)	0.927	0.989	0.979	0.927	0.952	0.958	0.957
LDA (n)	0.992 (0.99,0.994)	0.96 (0.956,0.963)	0.899	0.994	0.989	0.899	0.942	0.947	0.945
MLP (d)	0.998 (0.998,0.999)	0.989 (0.987,0.991)	0.984	0.992	0.985	0.984	0.985	0.988	0.988
MLP (n)	0.999 (0.998,0.999)	0.99 (0.988,0.991)	0.978	0.996	0.993	0.978	0.985	0.987	0.987
RF (d)	0.999 (0.999,0.999)	0.988 (0.986,0.989)	0.984	0.99	0.982	0.984	0.983	0.987	0.987

Continued on next page

Table 3.8: Average prediction performance of different pathomic models for the "Malignant vs Non-malignant" classification task in the test set at 10× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting. (Continued)

RF (n)	0.999 (0.999,0.999)	0.987 (0.985,0.989)	0.972	0.995	0.991	0.972	0.982	0.984	0.983
XGB (d)	0.999 (0.999,0.999)	0.99 (0.988,0.992)	0.988	0.992	0.985	0.988	0.986	0.99	0.99
XGB (n)	0.999 (0.999,0.999)	0.991 (0.99,0.993)	0.986	0.994	0.989	0.986	0.988	0.99	0.99

Table 3.9: Average prediction performance of different pathomic models for the "Malignant vs Non-malignant" classification task in the test set at 20× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting.

Mdl	**AUC (95%CI)**	**Accuracy (95%CI)**	**Sens**	**Spec**	**Prec**	**Recall**	**F1**	**BAcc**	**GSc**
GBM (d)	0.995 (0.995,0.996)	0.973 (0.972,0.974)	0.97	0.974	0.945	0.97	0.957	0.972	0.972
GBM (n)	0.996 (0.995,0.996)	0.976 (0.975,0.978)	0.959	0.985	0.966	0.959	0.962	0.972	0.971
LDA (d)	0.986 (0.985,0.987)	0.951 (0.949,0.953)	0.92	0.965	0.923	0.92	0.922	0.943	0.942
LDA (n)	0.988 (0.987,0.989)	0.95 (0.948,0.952)	0.885	0.98	0.954	0.885	0.918	0.932	0.931
MLP (d)	0.997 (0.996,0.997)	0.974 (0.972,0.975)	0.974	0.974	0.944	0.974	0.959	0.974	0.974
MLP (n)	0.996 (0.996,0.997)	0.976 (0.975,0.978)	0.973	0.978	0.953	0.973	0.963	0.975	0.975
RF (d)	0.997 (0.997,0.997)	0.972 (0.971,0.974)	0.969	0.974	0.944	0.969	0.956	0.971	0.971
RF (n)	0.997 (0.997,0.998)	0.978 (0.977,0.979)	0.963	0.985	0.968	0.963	0.965	0.974	0.974
XGB (d)	0.998 (0.998,0.998)	0.982 (0.981,0.983)	0.981	0.983	0.963	0.981	0.972	0.982	0.982
XGB (n)	0.999 (0.998,0.999)	0.986 (0.985,0.987)	0.979	0.989	0.976	0.979	0.978	0.984	0.984

Table 3.10: Average prediction performance of different pathomic models for the "Gleason 3 vs Non-malignant" classification task in the test set at 5× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting.

Mdl	**AUC (95%CI)**	**Accuracy (95%CI)**	**Sens**	**Spec**	**Prec**	**Recall**	**F1**	**BAcc**	**GSc**
GBM (d)	0.951 (0.94,0.962)	0.883 (0.865,0.901)	0.931	0.798	0.892	0.931	0.911	0.865	0.862
GBM (n)	0.944 (0.932,0.956)	0.87 (0.85,0.888)	0.919	0.782	0.883	0.919	0.901	0.851	0.848
LDA (d)	0.948 (0.935,0.961)	0.898 (0.88,0.914)	0.936	0.831	0.908	0.936	0.922	0.883	0.882
LDA (n)	0.947 (0.934,0.96)	0.896 (0.878,0.912)	0.939	0.82	0.903	0.939	0.921	0.879	0.877
MLP (d)	0.957 (0.944,0.97)	0.917 (0.9,0.931)	0.961	0.837	0.914	0.961	0.937	0.899	0.897
MLP (n)	0.958 (0.947,0.969)	0.897 (0.879,0.913)	0.939	0.822	0.904	0.939	0.921	0.88	0.878

Continued on next page

Table 3.10: Average prediction performance of different pathomic models for the "Gleason 3 vs Non-malignant" classification task in the test set at 5× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting. (Continued)

RF (d)	0.935 (0.922,0.947)	0.857 (0.836,0.876)	0.92	0.743	0.865	0.92	0.892	0.832	0.827
RF (n)	0.937 (0.925,0.949)	0.857 (0.836,0.876)	0.926	0.732	0.861	0.926	0.892	0.829	0.823
XGB (d)	0.958 (0.948,0.968)	0.89 (0.871,0.906)	0.933	0.813	0.899	0.933	0.916	0.873	0.871
XGB (n)	0.96 (0.95,0.97)	0.893 (0.875,0.909)	0.942	0.804	0.896	0.942	0.919	0.873	0.871

Table 3.11: Average prediction performance of different pathomic models for the "Gleason 3 vs Non-malignant" classification task in the test set at 10× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting.

Mdl	**AUC (95%CI)**	**Accuracy (95%CI)**	**Sens**	**Spec**	**Prec**	**Recall**	**F1**	**BAcc**	**GSc**
GBM (d)	1 (0.999,1)	0.989 (0.986,0.991)	0.984	0.996	0.998	0.984	0.991	0.99	0.99
GBM (n)	0.999 (0.998,0.999)	0.99 (0.988,0.992)	0.988	0.994	0.996	0.988	0.992	0.991	0.991
LDA (d)	0.992 (0.989,0.994)	0.963 (0.959,0.967)	0.946	0.99	0.993	0.946	0.969	0.968	0.968
LDA (n)	0.991 (0.989,0.994)	0.964 (0.96,0.968)	0.948	0.988	0.992	0.948	0.97	0.968	0.968
MLP (d)	0.998 (0.997,0.999)	0.986 (0.984,0.989)	0.98	0.996	0.997	0.98	0.989	0.988	0.988
MLP (n)	0.999 (0.998,0.999)	0.988 (0.986,0.99)	0.985	0.994	0.996	0.985	0.99	0.989	0.989
RF (d)	0.999 (0.998,0.999)	0.988 (0.985,0.99)	0.985	0.992	0.995	0.985	0.99	0.988	0.988
RF (n)	0.999 (0.999,0.999)	0.989 (0.987,0.991)	0.988	0.991	0.994	0.988	0.991	0.99	0.99
XGB (d)	0.999 (0.999,0.999)	0.992 (0.99,0.994)	0.99	0.996	0.997	0.99	0.994	0.993	0.993
XGB (n)	0.999 (0.999,0.999)	0.993 (0.991,0.995)	0.993	0.994	0.996	0.993	0.995	0.994	0.994

Table 3.12: Average prediction performance of different pathomic models for the "Gleason 3 vs Non-malignant" classification task in the test set at 20× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting.

Mdl	**AUC (95%CI)**	**Accuracy (95%CI)**	**Sens**	**Spec**	**Prec**	**Recall**	**F1**	**BAcc**	**GSc**
GBM (d)	0.999 (0.998,0.999)	0.977 (0.976,0.979)	0.969	0.993	0.996	0.969	0.982	0.981	0.981
GBM (n)	0.999 (0.999,0.999)	0.98 (0.979,0.982)	0.973	0.994	0.997	0.973	0.985	0.983	0.983
LDA (d)	0.997 (0.996,0.998)	0.966 (0.964,0.968)	0.948	0.997	0.998	0.948	0.973	0.973	0.972
LDA (n)	0.997 (0.996,0.997)	0.97 (0.968,0.972)	0.955	0.996	0.998	0.955	0.976	0.975	0.975
MLP (d)	0.999 (0.998,0.999)	0.991 (0.99,0.992)	0.989	0.995	0.997	0.989	0.993	0.992	0.992

Continued on next page

Table 3.12: Average prediction performance of different pathomic models for the "Gleason 3 vs Non-malignant" classification task in the test set at 20× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting. (Continued)

MLP (n)	0.999 (0.998,0.999)	0.994 (0.993,0.995)	0.993	0.997	0.998	0.993	0.996	0.995	0.995
RF (d)	0.999 (0.999,0.999)	0.979 (0.978,0.981)	0.969	0.998	0.999	0.969	0.984	0.984	0.983
RF (n)	0.999 (0.999,0.999)	0.981 (0.979,0.982)	0.972	0.997	0.998	0.972	0.985	0.984	0.984
XGB (d)	0.999 (0.999,0.999)	0.987 (0.985,0.988)	0.981	0.996	0.998	0.981	0.989	0.989	0.989
XGB (n)	0.999 (0.999,0.999)	0.986 (0.985,0.988)	0.981	0.996	0.998	0.981	0.989	0.988	0.988

Table 3.13: Average prediction performance of different pathomic models for the "Gleason 4 + Gleason 5 vs Gleason 3" classification task in the test set at 5× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting.

Mdl	**AUC (95%CI)**	**Accuracy (95%CI)**	**Sens**	**Spec**	**Prec**	**Recall**	**F1**	**BAcc**	**GSc**
GBM (d)	0.646 (0.62,0.672)	0.625 (0.603,0.647)	0.55	0.659	0.422	0.55	0.478	0.605	0.602
GBM (n)	0.648 (0.622,0.674)	0.654 (0.633,0.676)	0.37	0.783	0.435	0.37	0.4	0.576	0.538
LDA (d)	0.732 (0.708,0.756)	0.683 (0.662,0.704)	0.638	0.703	0.493	0.638	0.556	0.671	0.67
LDA (n)	0.726 (0.701,0.75)	0.711 (0.69,0.731)	0.538	0.789	0.536	0.538	0.537	0.664	0.652
MLP (d)	0.729 (0.703,0.754)	0.689 (0.668,0.71)	0.61	0.725	0.501	0.61	0.55	0.667	0.665
MLP (n)	0.714 (0.689,0.739)	0.689 (0.667,0.709)	0.55	0.751	0.5	0.55	0.524	0.651	0.643
RF (d)	0.626 (0.598,0.653)	0.62 (0.598,0.641)	0.547	0.653	0.416	0.547	0.472	0.6	0.597
RF (n)	0.63 (0.603,0.658)	0.672 (0.65,0.692)	0.293	0.843	0.457	0.293	0.357	0.568	0.497
XGB (d)	0.662 (0.636,0.688)	0.639 (0.617,0.66)	0.557	0.676	0.437	0.557	0.49	0.616	0.613
XGB (n)	0.715 (0.69,0.739)	0.706 (0.685,0.726)	0.5	0.799	0.529	0.5	0.514	0.649	0.632

Table 3.14: Average prediction performance of different pathomic models for the "Gleason 4 + Gleason 5 vs Gleason 3" classification task in the test set at 10× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting.

Mdl	**AUC (95%CI)**	**Accuracy (95%CI)**	**Sens**	**Spec**	**Prec**	**Recall**	**F1**	**BAcc**	**GSc**
GBM (d)	0.696 (0.684,0.708)	0.646 (0.635,0.657)	0.681	0.618	0.589	0.681	0.632	0.65	0.649
GBM (n)	0.69 (0.678,0.702)	0.654 (0.643,0.664)	0.46	0.81	0.66	0.46	0.542	0.635	0.61
LDA (d)	0.723 (0.712,0.734)	0.664 (0.653,0.674)	0.696	0.638	0.607	0.696	0.649	0.667	0.666
LDA (n)	0.718 (0.707,0.73)	0.676 (0.666,0.687)	0.503	0.816	0.688	0.503	0.581	0.659	0.641

Continued on next page

Table 3.14: Average prediction performance of different pathomic models for the "Gleason 4 + Gleason 5 vs Gleason 3" classification task in the test set at 10× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting. (Continued)

MLP (d)	0.698 (0.686,0.709)	0.641 (0.631,0.652)	0.724	0.575	0.578	0.724	0.643	0.649	0.645
MLP (n)	0.734 (0.723,0.745)	0.689 (0.678,0.699)	0.575	0.78	0.678	0.575	0.622	0.678	0.67
RF (d)	0.681 (0.67,0.693)	0.641 (0.63,0.652)	0.642	0.64	0.59	0.642	0.615	0.641	0.641
RF (n)	0.686 (0.675,0.698)	0.631 (0.62,0.641)	0.317	0.883	0.686	0.317	0.434	0.6	0.529
XGB (d)	0.717 (0.705,0.728)	0.657 (0.646,0.667)	0.683	0.636	0.602	0.683	0.64	0.659	0.659
XGB (n)	0.712 (0.701,0.723)	0.665 (0.655,0.676)	0.533	0.772	0.653	0.533	0.587	0.652	0.641

Table 3.15: Average prediction performance of different pathomic models for the "Gleason 4 + Gleason 5 vs Gleason 3" classification task in the test set at 20× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting.

Mdl	**AUC (95%CI)**	**Accuracy (95%CI)**	**Sens**	**Spec**	**Prec**	**Recall**	**F1**	**BAcc**	**GSc**
GBM (d)	0.688 (0.683,0.694)	0.611 (0.606,0.616)	0.711	0.557	0.468	0.711	0.564	0.634	0.629
GBM (n)	0.684 (0.679,0.69)	0.664 (0.659,0.669)	0.44	0.786	0.53	0.44	0.481	0.613	0.588
LDA (d)	0.711 (0.706,0.717)	0.621 (0.616,0.626)	0.715	0.569	0.476	0.715	0.572	0.642	0.638
LDA (n)	0.708 (0.703,0.714)	0.687 (0.682,0.692)	0.487	0.797	0.567	0.487	0.524	0.642	0.623
MLP (d)	0.711 (0.706,0.717)	0.621 (0.616,0.625)	0.742	0.554	0.477	0.742	0.58	0.648	0.641
MLP (n)	0.696 (0.69,0.701)	0.657 (0.652,0.662)	0.488	0.75	0.516	0.488	0.502	0.619	0.605
RF (d)	0.672 (0.666,0.677)	0.614 (0.609,0.619)	0.664	0.588	0.468	0.664	0.549	0.626	0.624
RF (n)	0.665 (0.66,0.671)	0.66 (0.655,0.665)	0.334	0.838	0.53	0.334	0.41	0.586	0.529
XGB (d)	0.681 (0.676,0.687)	0.618 (0.613,0.622)	0.678	0.584	0.472	0.678	0.556	0.631	0.629
XGB (n)	0.674 (0.668,0.679)	0.647 (0.642,0.652)	0.484	0.736	0.501	0.484	0.493	0.61	0.597

Table 3.16: Average prediction performance of different pathomic models for the "Gleason 4 + Gleason 5 vs Non–malignant" classification task in the test set at 5× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting.

Mdl	**AUC (95%CI)**	**Accuracy (95%CI)**	**Sens**	**Spec**	**Prec**	**Recall**	**F1**	**BAcc**	**GSc**
GBM (d)	0.986 (0.981,0.991)	0.903 (0.885,0.919)	0.863	0.972	0.982	0.863	0.918	0.918	0.916
GBM (n)	0.985 (0.98,0.99)	0.876 (0.857,0.894)	0.817	0.979	0.985	0.817	0.893	0.898	0.894
LDA (d)	0.991 (0.986,0.995)	0.938 (0.923,0.95)	0.913	0.981	0.988	0.913	0.949	0.947	0.946

Continued on next page

Table 3.16: Average prediction performance of different pathomic models for the "Gleason 4 + Gleason 5 vs Non–malignant" classification task in the test set at 5× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting. (Continued)

LDA (n)	0.991 (0.986,0.995)	0.915 (0.899,0.93)	0.874	0.987	0.992	0.874	0.929	0.93	0.929
MLP (d)	0.985 (0.98,0.99)	0.933 (0.918,0.946)	0.907	0.979	0.987	0.907	0.945	0.943	0.942
MLP (n)	0.99 (0.986,0.994)	0.925 (0.91,0.939)	0.89	0.987	0.992	0.89	0.938	0.938	0.937
RF (d)	0.984 (0.978,0.989)	0.907 (0.89,0.922)	0.871	0.968	0.979	0.871	0.922	0.92	0.918
RF (n)	0.984 (0.979,0.989)	0.883 (0.865,0.9)	0.827	0.981	0.987	0.827	0.9	0.904	0.901
XGB (d)	0.989 (0.985,0.993)	0.925 (0.909,0.938)	0.891	0.983	0.989	0.891	0.937	0.937	0.936
XGB (n)	0.99 (0.986,0.994)	0.908 (0.89,0.923)	0.861	0.987	0.992	0.861	0.922	0.924	0.922

Table 3.17: Average prediction performance of different pathomic models for the "Gleason 4 + Gleason 5 vs Non–malignant" classification task in the test set at 10× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting.

Mdl	**AUC (95%CI)**	**Accuracy (95%CI)**	**Sens**	**Spec**	**Prec**	**Recall**	**F1**	**BAcc**	**GSc**
GBM (d)	0.998 (0.998,0.999)	0.984 (0.981,0.986)	0.981	0.987	0.985	0.981	0.983	0.984	0.984
GBM (n)	0.998 (0.998,0.999)	0.983 (0.98,0.985)	0.976	0.989	0.988	0.976	0.982	0.982	0.982
LDA (d)	0.991 (0.989,0.992)	0.96 (0.956,0.964)	0.939	0.98	0.977	0.939	0.958	0.959	0.959
LDA (n)	0.992 (0.99,0.993)	0.959 (0.955,0.963)	0.93	0.986	0.984	0.93	0.956	0.958	0.958
MLP (d)	0.999 (0.998,0.999)	0.99 (0.988,0.992)	0.986	0.994	0.994	0.986	0.99	0.99	0.99
MLP (n)	0.998 (0.997,0.998)	0.987 (0.984,0.989)	0.98	0.993	0.992	0.98	0.986	0.986	0.986
RF (d)	0.999 (0.999,0.999)	0.984 (0.982,0.987)	0.982	0.987	0.986	0.982	0.984	0.984	0.984
RF (n)	0.999 (0.999,0.999)	0.985 (0.983,0.987)	0.978	0.992	0.991	0.978	0.984	0.985	0.985
XGB (d)	0.999 (0.999,0.999)	0.987 (0.985,0.989)	0.983	0.991	0.991	0.983	0.987	0.987	0.987
XGB (n)	0.999 (0.999,0.999)	0.991 (0.989,0.993)	0.988	0.994	0.993	0.988	0.991	0.991	0.991

Table 3.18: Average prediction performance of different pathomic models for the "Gleason 4 + Gleason 5 vs Non–malignant" classification task in the test set at 20× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting.

Mdl	**AUC (95%CI)**	**Accuracy (95%CI)**	**Sens**	**Spec**	**Prec**	**Recall**	**F1**	**BAcc**	**GSc**
GBM (d)	0.995 (0.994,0.996)	0.975 (0.973,0.977)	0.964	0.984	0.981	0.964	0.972	0.974	0.974
GBM (n)	0.996 (0.995,0.996)	0.974 (0.972,0.975)	0.957	0.987	0.985	0.957	0.971	0.972	0.972

Continued on next page

Table 3.18: Average prediction performance of different pathomic models for the "Gleason 4 + Gleason 5 vs Non–malignant" classification task in the test set at 20× magnification level. Abbreviations: AUC = Area Under the Receiver Operating Characteristic curve; CI = Confidence Interval; Sens = Sensitivity; Spec = Specificity; Prec = Precision; BAcc = Balanced Accuracy; GSc = G-Score; GBM = Gradient Boosting Machine; LDA = Linear Discriminant Analysis; MLP = Multi-layer Perceptron; RF = Random Forest; XGB = eXtreme Gradient Boosting. (Continued)

LDA (d)	0.989 (0.988,0.99)	0.955 (0.953,0.957)	0.928	0.977	0.971	0.928	0.949	0.952	0.952
LDA (n)	0.989 (0.988,0.99)	0.951 (0.949,0.953)	0.914	0.981	0.976	0.914	0.944	0.948	0.947
MLP (d)	0.997 (0.997,0.998)	0.978 (0.976,0.979)	0.968	0.987	0.984	0.968	0.976	0.977	0.977
MLP (n)	0.998 (0.997,0.998)	0.984 (0.983,0.985)	0.975	0.992	0.991	0.975	0.983	0.984	0.983
RF (d)	0.997 (0.997,0.997)	0.979 (0.978,0.981)	0.967	0.989	0.987	0.967	0.977	0.978	0.978
RF (n)	0.997 (0.997,0.997)	0.979 (0.978,0.981)	0.963	0.993	0.991	0.963	0.977	0.978	0.978
XGB (d)	0.993 (0.992,0.994)	0.984 (0.982,0.985)	0.973	0.993	0.991	0.973	0.982	0.983	0.983
XGB (n)	0.992 (0.991,0.993)	0.984 (0.982,0.985)	0.971	0.994	0.993	0.971	0.982	0.983	0.983

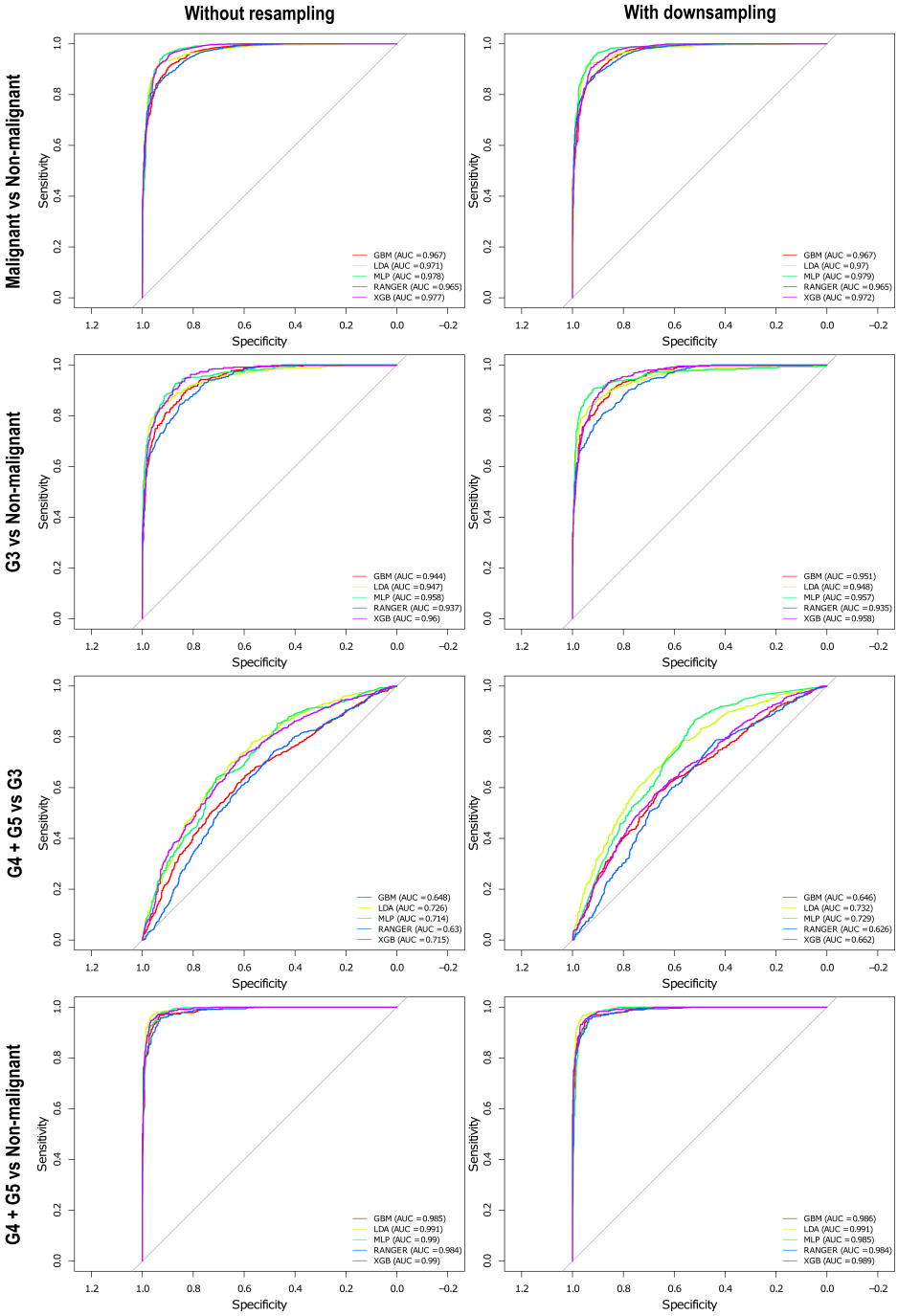


Figure 3.12. Receiver Operating Characteristic (ROC) curves of prediction models without resampling (first column) and with downsampling (second column for each classification task (shown on the rows), at 5× magnification.

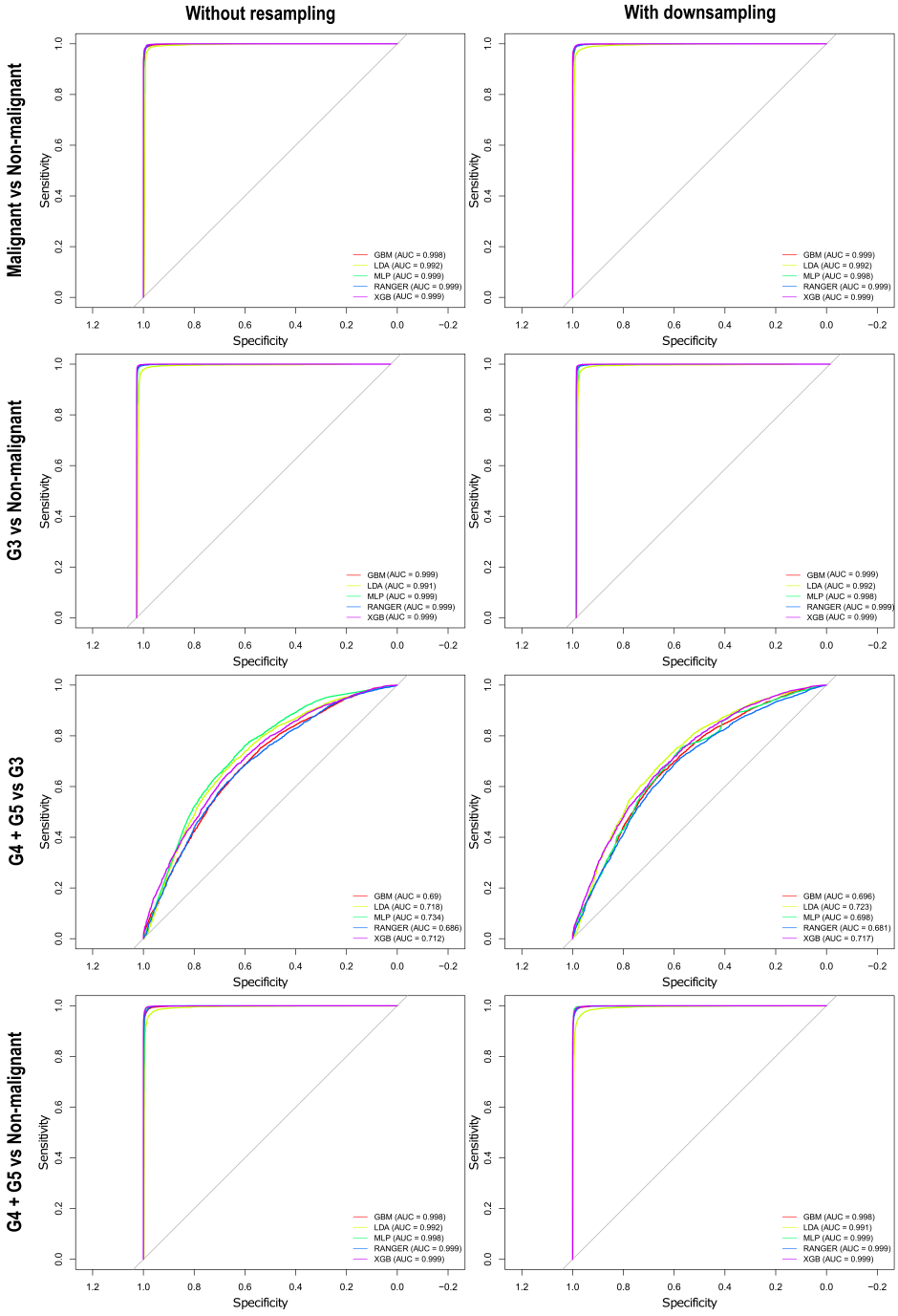


Figure 3.13. Receiver Operating Characteristic (ROC) curves of prediction models without resampling (first column) and with downsampling (second column for each classification task (shown on the rows), at 10× magnification.

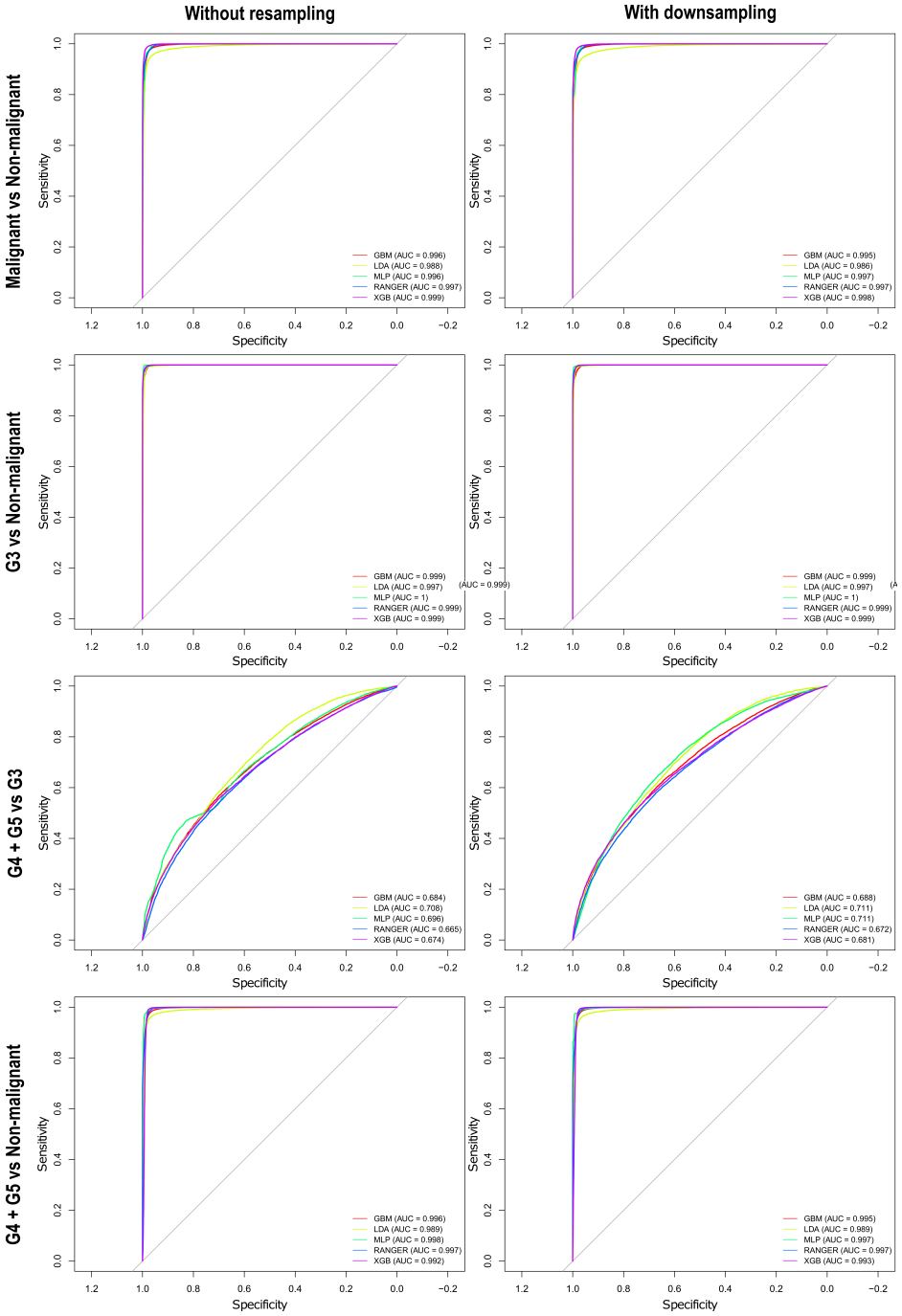


Figure 3.14. Receiver Operating Characteristic (ROC) curves of prediction models without resampling (first column) and with downsampling (second column for each classification task (shown on the rows), at 20× magnification.

3.2.2.4 Discussion

The study presents a multiscale, multi-channel, pathomic-based approach for estimating PCa grade on digital pathology images, addressing critical challenges in the current Gleason grading system. In particular, the existing manual PCa grading process suffers from reproducibility issues, subjectivity, and time constraints, thus necessitating technology-driven solutions that could contribute to develop accurate methods to assess PCa grade. Automatic PCa grading using computational image analysis approaches could contribute in this direction, with a look towards the possibility to support and improve clinical decision making. In this context, the aim was to develop a quantitative, automatic, interpretable, and as simple as possible Gleason Grading approach. Although most recent literature on automatic Gleason Grading is based on DL approaches, the intention of the work was to use a pathomic based approach to exploit the potential immediate interpretability of hand-crafted features, that can aid pathologists with novel insights on the examined tissue. This choice was motivated not only by the interpretability of hand-crafted features but also by the consideration that DL-based methods often require extensive datasets to reach accurate performance. This strategic decision aligns with our goal of developing a method that not only enhances interpretability but also proves effective in scenarios where obtaining large datasets may be impractical or challenging. Of note, the use of texture-based methods at tissue level was preferred due to their ability to alleviating the need for intricate cellular or glandular segmentations. This characteristic not only simplifies the computational process but also broadens the applicability of texture-based methods across diverse prostate cancer presentations. The investigated tissues involved stroma, normal tissue, and Gleason patterns 3, 4, and 5. Tiling was performed at multiple magnification scales (5×, 10×, and 20×), extracting an extensive set of pathomic features from each tile. The investigation of first-order, GLCM, LBP, and wavelet features stems from their extensive use in literature for a wide range of applications and their potential to capture crucial aspects of prostate cancer pathology. Histogram-based features are first-order statistics describing the distribution of intensity values in an image. They measure the degree of dispersion of the grey values, the presence/absence of outliers and other properties which reflect the overall structure of the texture. Local texture

descriptors such as GLCM and LBP, as second-order statistics, consider the joint variability of grey levels in pairs or groups of pixels. Widely employed in various applications, these descriptors prove effective in providing detailed insights. GLCM, quantifying spatial relationships, provides insights into glandular structures. LBP, capturing local texture variations, reveals detailed information, and wavelet features ensure a comprehensive understanding by addressing the multifaceted nature of prostate cancer pathology. The main rationale for using wavelets features is their ability to decipher textural information from different scales, which has been shown to significantly improve image classification [258, 60, 158]. The incorporation of explicit texture-based pathomic features prioritizes interpretability, demystifying the decision-making process for clinicians. Unlike the black-box nature associated with deep learning models, the presented approach emphasizes transparency, aligning with clinical interpretability requirements and empowering practitioners with actionable insights. The decision to utilize multiple resolutions aligns with the diverse diagnostic practices of pathologists. Choosing 5×, 10×, and 20× magnifications is based on a thoughtful consideration of what can be perceived at each level. This aligns with routine practices, ensuring a holistic understanding of tissue architecture without overwhelming complexity [35]. WSIs are characterized by a pyramidal structure consisting of the same image stored at different spatial resolutions. Most of the studies in digital pathology (either ML- or DL-based) use one fixed magnification level to perform the analysis. These models do not take advantage of the multi-scale nature of this type of data. Different magnification levels are usually required to recognize different features at a macroscopic scale such as the organ to which the image belongs and at a microscopic scale such as tumor-related information. Pathologists generally conduct their analysis under a microscope combining different scales: they start looking at the tissue at low magnification levels for macroscopic features and then they zoom in into the region of interest to examine the microscopic features at high magnification. For example, at middle magnification levels (such as 5–10×) it is possible to distinguish between glands, while at the highest ones (such as 20–40×) it is possible to better resolve cells [35]. The pathologists' approach highlights the importance of combining macroscopic and microscopic information obtained from different scales. Of note, some few ap-

proaches to integrate multiple magnification levels have recently been explored [78, 99]. This departure from focusing on specific scales enriches the diagnostic landscape, providing enhanced granularity crucial for nuanced Gleason grading. Unlike traditional studies that often rely on singular or paired features, this preliminary work embraces a diverse feature space, exploring eight channels alongside grayscale-converted RGB images. This comprehensive feature exploration can capture a spectrum of information essential for nuanced Gleason grading. The deliberate choice of incorporating color features aligns with the perceived characteristics of the human eye since color channels provide a richer representation of visual information, acknowledging and leveraging the importance of color information in pathology assessment. At present, grayscale texture analysis technology has witnessed significant advancements, and many gray-level texture descriptors have been developed and successfully employed in numerous domains for image classification. However, texture classification is performed on grayscale images, thus discarding color information, which is an important cue for visual perception. Therefore, the exclusion of color information can be limiting in capturing the full complexity of pathological features, also because the pathological assessment is done based on color images. However, it should be considered that the use of color features encounters challenges arising from the correlation among various color channels [278]. The color channels in images tend to exhibit high levels of correlation, meaning that values in one channel are good predictors of the values in other channels. However, the correlation filter used in the feature selection process was used to mitigate the limitations associated with the dependence and correlation among different color channels, ensuring our approach remains innovative and clinically robust. Obtained preliminary results showed that some specific pathomic features could classify prostatic tissue compartments. In particular, several machine learning models with selected pathomic features able to classify prostatic tissues with high prediction performances after a feature selection process were developed. The presented results demonstrate very high classification performances across the various classification tasks investigated. As expected, G3, G4, and G5 could be easily distinguished from benign tissues. In fact, very high AUC values were reached for PCa diagnosis tasks (e.g. Malignant vs Non-Malignant, G3 vs Non-Malignant, and G4 + G5 vs Non-Malignant classifi-

cation tasks). Lower AUC values for tasks involving the low- vs high-grade classification, with maximum AUC values between 0.72 and 0.73 reached for MLP and LDA models, both for 5×, 10× and 20× magnification scale. These results were in line with those obtained by Bhattacharjee et al. [60] who examined wavelet-based and color features to train MLP model to perform different classification tasks . Their results also revealed lower classification performance for the low-grade vs high-grade classification task with respect to classification tasks involving benign tiles versus malignant tiles. Similar results were found by Alexandratou et al. [37], who compared the performance of 16 supervised machine learning algorithms for prostate cancer diagnosis and Gleason grading, also addressing the classification tasks tumour vs. non-tumour, low vs. high grade and finding and accuracy of 97.9% for diagnosis (tumour-non-tumour) and a lower accuracy of 80.8% for low-high grade discrimination. Kim et al. performed texture analysis using GLCM method and carried out classification using SVM and KNN models for multiple tasks. The highest accuracy of 90% was achieved for benign vs grade 4 and 5 [2]. Hongming et al. [329] also presented an automatic approach for PCa grading based on LBP features descriptors integrated together across WSIs, thus obtaining a patient-level representation. Nevertheless, these results were not directly comparable with these this studies as their WSI preprocessing, feature extraction methods and features types were different. Moreover, the change in prediction performances depends on the features being extracted as well as the type of classification methods being used for detecting prostate carcinomas in histological sections. The strengths of the study lie in its holistic and explicit approach, encompassing multiple magnification scales and channels for feature extraction. The transparent emphasis on interpretability addresses a critical concern in AI-based pathology grading. The inclusion of an extensive feature space enhances the versatility of the proposed methodology. Looking ahead, there are still several intriguing future perspectives to explore. First, understanding the fusion of information across scales by simulating the practices of pathologists who scroll through the microscope at varying resolutions, ranging from low (e.g., 5×) to high (e.g., 20×), could provide valuable insights. Additionally, it could be interesting to investigate how the individual groups of features perform in PCa grading [158]. Of note, a limited number of features was evaluated in this study in order

to be able to interpret descriptors. An essential criterion for having examined these hand-crafted features is that they are easier to interpret with respect to features obtained from neural networks. However, the proposed framework may be used to study associations between an expanded suite of radiomic and pathomic measurements, as well as extended to characterizing other cancers. Furthermore, an intriguing avenue for future exploration involves comparing the developed approach with deep learning methods. While the study already allows for a comparison of results with similar tasks in the existing literature, a particularly interesting dimension would be to employ the same settings for a comprehensive hand-crafted versus deep learning-based features comparison. This comparative analysis has the potential to contribute not only to the ongoing refinement of our approach but also to the broader understanding of the optimal methodologies for PCa Gleason Score classification. It is noteworthy to emphasize that the primary objective was focused on investigating hand-crafted features. While the prospect of comparing the obtained results with DL methods is indeed compelling, it is essential to clarify that our current study primarily aims to delve into the intricacies of hand-crafted features. It was intend to reserve the exploration and comparison with DL-based features for future research endeavors. By prioritizing the examination of hand-crafted features, the aim was to establish a solid foundation and comprehensive understanding of their effectiveness in the context of PCa Gleason Score classification. This deliberate choice allows to thoroughly explore the nuances and intricacies of the features before extending our investigation to the realm of DL-based approaches in subsequent studies. This strategic approach aligns with the goal of providing a detailed and nuanced analysis of PCa Gleason Score classification, laying the groundwork for more comprehensive and insightful future comparisons and eventually integration with DL techniques [142, 301]. In conclusion, the preliminary results introduced a promising multiscale, pathomic-based approach for PCa Gleason Score classification, offering a potential paradigm shift in Gleason grading. The study opens avenues for future research, emphasizing the need for ongoing refinement, validation on diverse datasets, and collaboration with pathologists for real-world applicability. Moreover, by increasing the dimensionality of the dataset, deep learning approaches can be reconsidered. However, the proposed technology-driven solution holds significant promise

in revolutionizing Gleason grading, contributing to more informed clinical decision-making, and ultimately improving patient outcomes in prostate cancer diagnosis.

Main contributions of the chapter

- The correlation between radiomic and pathomic features in Glioblastoma multiforme was preliminarly explored, revealing significant cross-scale relationships and advancing tumor heterogeneity understanding for diagnostic, prognostic, and therapeutic considerations.
- Pathomics applications in cancer diagnosis were investigated, in particular:
  - A pathomic approach to quantifying tumor-infiltrating lymphocytes in breast cancer was developed, addressing traditional visual assessments' limitations. Results indicated potential for TIL classification using specific pathomic features, introducing a standardized, interpretable, automatic, and simple method crucial for breast cancer decision-making.
  - A multiscale, multi-channel, pathomic-based approach for Gleason Score classification in PCa was proposed. Texture-based pathomic features distinguished malignant from non-malignant prostatic tissue and classified different Gleason patterns, enhancing PCa diagnosis precision and reliability.

Chapter 4

Conclusions

In the era of precision medicine, integrating information across multiple scales is crucial for anticipating and steering cancer diagnosis and prognosis. Modern diagnostic techniques and computational analysis methods are generating a huge amount of heterogeneous data, yet their integration remains limited. CDSS utilizing Machine Learning, and data analytics have immense potential to aggregate, structure, and understand these data to support clinical decision. The presented PhD thesis focuses on two key objectives to advance precision medicine in oncology. The first objective concerns the integration and organization of heterogeneous data from the most advanced diagnostic domains for cancer management, namely diagnostic imaging, histopathology, and genetics. Building upon these foundational considerations, the second objective of the thesis involves the supervised computational analysis of biomedical images at various scales. This has been achieved through the implementation of processing pipelines designed for the extraction of radiomics and pathomics features.

4.1 Main Findings

The key findings obtained in this thesis are listed below:

- Given the urgency of organizing complex biomedical data belonging to different diagnostic domains, digital biobanks were identified as a multifactorial solution capable of containing curated and standardized data from the most advanced diagnostic domains in the

field of oncology (clinical imaging, pathology, and NGS). An integrative standardization/harmonization approach was proposed, aiming to incorporate existing standards and procedures already employed in common practice.

- The correlation between radiomic and pathomic features was preliminarly explored in the context of Glioblastoma Multiforme. Findings revealed significant cross-scale relationships, offering a better understanding of tumor heterogeneity and impacting diagnostic, prognostic, and therapeutic considerations.

- The applications of pathomics in cancer diagnosis were explored, highlighting its potential to support clinical decision-making. In particular:

  - A pathomic approach to quantifying tumor-infiltrating lymphocytes in breast cancer was developed to overcome the limitations of the traditional visual assessments. Results showed that some selected specific pathomic features extracted from nuclei in tumor-associated stromal regions of H&E images of breast cancer could classify TILs. The proposed approach introduces a quantitative, standardized, interpretable, automatic and simple method for TIL assessment, crucial for decision-making support in breast cancer treatment strategies.

  - To address the complexities of Gleason grading in PCa, a multiscale, multi-channel, pathomic-based approach for Gleason Score classification was proposed. Texture-based pathomic features distinguishing malignant from non-malignant prostatic tissue and classifying between different Gleason patterns. This approach offered an efficient and consistent Gleason grading process, potentially enhancing the precision and reliability of PCa diagnosis.

4.2 Summary and Discussion of Results

Concerning the first objective, Chapter 2 emphasized the urgent need to make the huge amount of heterogeneous information belonging to differ-

ent diagnostic domains available and suitable to promote scientific research and technological development. In particular, this increasing volume of heterogeneous and complex biomedical data, driven by advancements in acquisition and computational methods, led us to consider developing a comprehensive approach for organizing and making this information accessible for scientific research and technological advancement, with direct implication for support decision making. Digital biobanks were framed as a strategic multifactorial solution capable of containing curated and standardized imaging data, along with clinical, molecular, and pathologic data. This because the effective organization of biomedical data in biobanking infrastructures is crucial to enable a comprehensive approach to clinical studies, as well as the development of AI tools supporting decision making in medicine. We firstly highlighted that digital biobanks must ensure that data and information are standardized, reproducible and integrated, following regulatory and bioethical guidelines and adhering to ISO standards and SOPs, as well as considering that the stewardship and management of scientific data needs to adhere to the FAIR principles. This will contribute to the success of the digital biobank model to achieve its full potential as scientific resource. Based on this foundation, an integrative standardization/harmonization approach was proposed, aiming to incorporate existing standards and procedures already employed in common practice. We focused on the most advanced diagnostic domains in the field of oncology, namely clinical imaging, pathology, and NGS. We considered both the integration of these domains, as well as the generation of numerical descriptors associated with each single domain through robust data curation and data processing procedures. Our approach involved the use of MIABIS, DICOM and FASTQ as they are established standards in the common practice to describe raw and derived data from the chosen domains. MIABIS was used to storage and exchange of data elements that describe the collection of biological samples. DICOM was established for the acquisition, reporting, and curation (annotation and clinical outcomes definition) of both radiologic and WSI images. Since the JSON format natively suits the hierarchical format of DICOM metadata, we proposed to use JSON format to store and interchange domain-specific numerical descriptors. Furthermore, it is the preferred format for encoding the output of AI algorithms, is widely supported by major programming languages,

and can be linked to formal ontologies. This study provides new models of standardization and integration among multidisciplinary domains, also incorporating robust data curation and processing pipelines to obtain reproducible numerical descriptors associated with each diagnostic domain. Standardized procedures and data not only foster advancements in -omics studies (e.g., radiomics, pathomics, genomics) but also open avenues to explore connections between quantitative -omics data and clinical outcomes in specific diseases. The aim of the work was limited to framing potential functional solutions for implementing a digital biobank, addressing challenges in standardization, reproducibility, and integration. The development of an IT infrastructure for data sharing is reserved for future research. The proposed comprehensive digital biobank model could improve the management, standardization and sharing of data in adherence with ethical norms. This work serves as an encouragement for the scientific community to develop tools supporting effective multi-domain integration in biobanking.

Findings and considerations raised in Chapter 2 on integration of data from different diagnostic domains, as well as the generation of numerical descriptors associated with each single domain, paved the way for investigations outlined in Chapter 3. Chapter 3 unfolds the development of computational and statistical tools designed for analyzing and integrating data across different scales of biomedical images. The first part of the chapter emphasizes the significance of understanding the correlation between radiomic and pathomic features to enhance the comprehension of tumor heterogeneity at different scales. Cancer, manifesting its diversity at both cellular and radiological levels, requires bridging these scales for a full understanding of the disease. The information achieved through the exploration of associations between radiomic and pathomic features provides a holistic view of cancer from microscopic details to larger-scale images and can offer clinicians a more comprehensive understanding of tumors, contributing to a detailed and accurate characterization. The correlation across imaging scales in the context of Glioblastoma Multiforme, where the intricate interplay between the microscopic and radiological features is known to be highly complex, was explored. The aim was to identify cross-scale associations between quantitative descriptors from radiology and pathology images of patients with glioblastoma. The

radiopathomic approach investigated, although preliminary, revealed interesting cross-scale relationships between radiomic features from functional magnetic resonance images and pathomic features from digital pathology images of patients with GBM. This, in turn, directly impacts the clinical decision-making process, covering diagnostic, prognostic, and therapeutic considerations. In summary, this study on Glioblastoma Multiforme, although an initial example paving the way for future radiopathomic studies, underscores the crucial role of integrating radiomics and pathomics for addressing the need to evaluate tumor heterogeneity. This combined approach not only enhances our understanding of the complex interplay between microscopic and radiological features but also provides a valuable way to validate the practical use of radiomics. "Virtual biopsy" captures the essence of this concept. The idea of a "virtual biopsy" marks a significant shift in diagnostic methods. Unlike traditional invasive biopsies, it relies on advanced analyses of radiological images. Bringing together radiopathomic data in this approach offers a comprehensive view of the tumor's structure and composition, giving crucial insights without the need for invasive procedures. This progressive method represents a significant step forward in both confirming and applying the information obtained through radiomics, offering a detailed yet non-invasive perspective for informed clinical decision-making.

Additionally, in Chapter 3 the applications of pathomics in cancer diagnosis were explored, highlighting its potential to support clinical decision-making and contribute to CDSS. The integration of pathomic information into the diagnostic workflow is seen as a promising avenue for refining tumor characterization and improving the accuracy of diagnoses. Two distinct studies were presented in this part to exemplify the power of pathomics in addressing specific challenges in breast cancer and prostate cancer diagnosis, that are both the second most common cancers in men and women respectively. This thesis section marked a pivotal shift towards practical clinical applications. The first study pioneers a pathomic approach to quantifying tumor-infiltrating lymphocytes in breast cancer, overcoming the limitations of traditional visual assessments. Results showed that some selected specific pathomic features extracted from nuclei in tumor-associated stromal regions of H&E images of breast cancer could classify if nuclei were TILs or not. In particular, several machine

learning models with selected pathomic features that could classify sTILs with high prediction performances after a feature selection process were developed. The proposed approach introduces a standardized and efficient assessment of TILs, crucial for decision-making support in BC treatment strategies. The second study addresses the complexities of Gleason grading in PCa, introducing a multiscale, pathomic-based approach for Gleason Score classification. This approach, leveraging pathomic features at different magnification scales, offers a more efficient and consistent Gleason grading process, potentially enhancing the precision and reliability of PCa diagnosis. Together, these investigations navigate the path from research innovation to real-world impact, envisioning a landscape where pathomics becomes an integral tool for enhancing diagnostic precision and treatment strategies in cancer care. The results of these studies underscore the transformative potential of pathomics in bringing about positive changes in clinical practice, positioning it as a key player in the future of oncology. The dual focus of Chapter 3 on exploring correlations across imaging scales and studying potential applications of pathomics in CDSS reflects a holistic approach to cancer diagnosis and decision support. By acknowledging the intricate interplay of diverse data sources, the research aims to contribute to improved patient outcomes in oncology.

4.3 Methodological Choices

This section delves into the methodological decisions underpinning the research.

It is worth drawing attention to the decision to choice to rely on handcrafted features, that is grounded in the pragmatic consideration of data availability and interpretability.

One primary consideration that guided the choice of feature extraction methods is the dimensionality of the available data. Deep learning methods, particularly CNNs, has ushered in a paradigm shift in computational image analysis, particularly in the domains of radiomics and pathomics. Although these methods are becoming increasingly popular because of their remarkable classification performance, they inherently require a large and diverse datasets samples representing real-world data to achieve the optimal performances [192, 201]. Moreover, although their advantage

of automated feature extraction and superior performance in capturing complex patterns, they may pose challenges in interpretability, requiring mechanisms for explaining model decisions. Meanwhile, handcrafted features have been implemented for decades and still serve as a powerful tool when combined with machine learning classifiers. Supervised learning algorithms, i.e., learning a mapping from input data to output (labels) from a set of training examples, are widely used in medical image data analysis [102]. Traditional supervised learning algorithms, such as RF, SVM and KNN, require prespecified hand-crafted features, while CNNs learn image features from the inputs to classify labels. However, CNNs typically require large numbers of training examples, which can be difficult to obtain in the medical imaging space, due to confidentiality constraints, financial limitations and time required for expert annotations [254]. Thus, traditional methods using hand-crafted features remain useful in dealing with limited samples of medical imaging data. Therefore, the decision to start with hand-crafted approach in 3 stems primarily from the pragmatic recognition of data limitations, encompassing both the aspect of pathological image data and annotation availability. In particular, the number of patients involved in the three studies investigated was relatively small, namely 48 for the study investigating radiopathomic correlations in patients with GBM [70], 195 in the study aiming at classifying TILs in BC [314], and 187 in that on Gleason Score classification for PCa patients. The limited sample size is directly linked with the mentioned challenges associated with the digitization of WSIs, a crucial prerequisite for the effective application of deep learning methodologies. Unlike radiological images, for which the transition to digital formats has been widespread and transformative, WSI technology in pathology is viewed more as an augmentation rather than a substitution for conventional practices, with WSI integrated alongside traditional microscopy rather than completely replacing it. Pathologists still routinely examine glass slides under a microscope, and the adoption of WSI is often supplemental. This coexistence stems from the intricacies and nuances of pathological analysis, where the fine details of cellular structures and tissue morphology are crucial for accurate diagnoses. Moreover, pathology images require a more complex acquisition process, may be subject to resource and cost limitations, and poses logistical and technical challenges. Therefore, while WSI technology

offers advantages in terms of remote access, sharing, and archiving, it has not fully replaced traditional microscopy due to the necessity of retaining the tactile and detailed aspects of glass slide examination [46]. This situation presents a practical barrier to the use of deep learning approaches that rely on significant volumes of digitized data, especially in the field of digital pathology.

Another aspect to be considered is that, although the limited number of patients included in the studies, a substantial volume of items were analyzed in the studies. For instance, the study on TILs classifications involves around 90,000 segmented cells, and the prostate study deals also with tens of thousands of mosaic tiles for classifications at 20× magnification. Machine learning and deep learning pipelines tailored for histopathological data often necessitate the partitioning of WSIs into multiple patches or tiles to enhance the original training data and accommodate GPU memory constraints. This partitioning strategy is crucial for handling large images, as a single WSI can be divided into numerous tiles, each serving as a distinct training input. These tiles are randomly selected and subjected to verification processes to ensure the preservation of sufficient tissue information. Additionally, data augmentation techniques, such as random rotation, flipping, or affine transformation, are applied to these tiles to mitigate the risk of overfitting. However, this approach comes with its challenges, notably the increased likelihood of data leakage [77]. The multiple subimages originating from the same histological specimen, despite being processed independently, are still intrinsically linked to the source. This interdependence raises concerns about the potential for models to inadvertently learn from the same specimen across different tiles, compromising the generalizability and reliability of the trained model. Therefore, while the augmented dataset addresses the challenge of limited patient numbers, the risk of data leakage necessitates careful consideration. The balance between enhancing dataset size through augmentation and maintaining the integrity of the training process is crucial for ensuring the robustness and validity of deep learning models applied to histopathological data [77]. Since data leakage affect both hand-crafted approaches and DL-approah, the advantage of explainability in traditional machine learning, including the use of hand-crafted features, becomes especially pertinent when addressing challenges related to data leakage. In traditional ma-

chine learning models, the manual or transparent selection of features provides a clearer understanding of which features are driving model predictions. This transparency facilitates the identification of potential sources of data leakage and aids in crafting effective mitigation strategies. The incorporation of hand-crafted features enhances interpretability, enabling a more detailed examination of the features contributing to the model's decisions. This heightened interpretability becomes an asset in pinpointing whether manually selected features inadvertently contribute to unwanted influences on the model, shedding light on potential areas of data leakage. The clarity offered by hand-crafted features empowers researchers to navigate the model's decision-making process more effectively. Conversely, the black-box nature of deep learning models poses challenges in terms of interpretability. The intricate relationships learned by hidden layers in deep neural networks make it challenging to comprehend why a specific decision was made, making the detection and management of data leakage more complex.

Another critical issue related to data limitations is the annotation availability. Given the large datasets required to obtain high DL performance, also a large number of annotations are required. The availability of high-quality annotations is pivotal for training accurate and robust deep learning models. In the realm of medical imaging, especially digital pathology, obtaining annotations involves meticulous outlining and labeling of structures, cells, or regions of interest within images. The sheer volume of annotations needed to achieve high performance in deep learning exacerbates the challenge, often necessitating substantial resources and time [118, 220]. The scarcity of annotations is a significant obstacle that demands attention alongside the challenge of obtaining sufficient raw data, regardless of the medical imaging scale. Both challenges are intricately interconnected and represent critical hurdles in the realm of medical imaging. To address this, approaches like transfer learning, active learning, and semi-supervised learning emerge as potential strategies [318, 153]. These techniques can leverage knowledge from pre-existing annotated datasets, focus on acquiring annotations for the most informative examples, and utilize both annotated and unlabeled data, respectively. It's crucial to note that while these approaches can partially mitigate the annotation challenge, they are not comprehensive solutions. Annotations form the bedrock of learning

and, notably, model validation. In the absence of ample annotated examples, these techniques may enhance the utilization of available annotations but cannot replace the need for sufficient annotated data. In the specific studies investigated this issue was highlighted. For example, when investigating on TILs classification task, the issue concerning the prerequisite of pixel level annotation for individual nuclei segmentation within tumor-associated stroma emerged, thus implying that any segmentation errors can have a direct impact on the final performance. Although the choice of an automatic and well-established approach guarantees robustness and reproducibility, the lack of a segmentation ground truth does not allow to quantitatively evaluate and control the segmentation error (e.g. through Dice index or F1-score) and also to test more advanced segmentation methods (e.g. training neural networks), also by making comparisons with other existing nuclear segmentation algorithms implemented in different software/studies. In the investigation of Gleason Score classification, the dataset comprises 187 H&E-stained WSIs of prostatectomy specimens, each meticulously annotated by experienced pathologists. The challenge of obtaining comprehensive annotations in medical imaging, as highlighted earlier, is particularly pronounced in intricate tasks like Gleason Score classification. In this context, leveraging texture-based methods becomes pivotal. The unique advantage of these methods lies in their capability to classify high-grade prostatic tumors, characterized by minimal glandular differentiation. Importantly, these methods operate at the tissue level, alleviating the need for intricate cellular or glandular segmentations, which are often labor-intensive and prone to annotation challenges. This characteristic not only simplifies the computational process but also aligns with the broader challenges discussed earlier, emphasizing the importance of innovative approaches to address annotation scarcity and avoiding extremely difficult or time-consuming annotation tasks. The theoretical foundation for this ambition lies in the potential to construct a robust model that does not heavily rely on complex segmentations, such as those involving glands or nuclei, which can be time-consuming compared to annotations focusing on tissue zones. Unlike the study on TILs classifications, where the segmentation of individual cells is fundamental, the Gleason Score classification study envisions a more streamlined approach. The intention is to develop a model that, once well-trained on a wealth of annotations re-

flecting real-world data, could transition to a whole-slide approach for new data. In this envisioned scenario, the need for segmentations diminishes, as the model could efficiently process WSIs directly. These WSIs would be partitioned into tiles, allowing the model to identify and categorize regions without the need for elaborate cellular or glandular segmentations. It underscores a forward-looking strategy, acknowledging the potential evolution of the model from a training phase heavily reliant on annotations to a deployment phase where the focus shifts towards the efficiency and applicability of whole-slide analysis in real-world scenarios.

Another pivotal aspect in the decision-making process is the explainability of features. Hand-crafted features, being explicitly designed and defined, offer a level of interpretability that deep learning models often lack. In the context of medical diagnostics, where transparency and interpretability are paramount, the use of hand-crafted features allows for a more comprehensible understanding of the features driving the analysis. Handcrafted features, defined by transparent mathematical formulas, offer a level of interpretability distinct from deep learning models. While the complexity of deep features may hinder straightforward interpretation, handcrafted features provide clear insights into the specific information they represent. However, it's essential to recognize that interpreting the nuances of handcrafted features may still require domain expertise. The radiomic and pathomic features extracted manually hold clinical significance, and their interpretability can directly contribute to the clinical decision-making process. The transparency of these features is invaluable, providing clinicians with insights into the characteristics of oncological pathologies that influence the derived numerical descriptors.

Highlights

- ***Data availability and Interpretability:*** The decision to rely on hand-crafted features was driven by pragmatic considerations regarding data availability and interpretability. This choice acknowledges the challenges associated with deep learning methods, particularly in the context of limited sample sizes and annotation availability.
- ***Dimensionality of available data:*** The choice of feature extraction methods is influenced by the dimensionality of the available data. While deep learning methods offer remarkable performance, they necessitate large and diverse datasets, posing challenges in data acquisition, particularly in medical imaging.
- ***Potential power of handcrafted features:*** Hand-crafted features, despite being manually specified, remain a powerful tool when combined with machine learning classifiers, especially in scenarios with limited sample sizes or annotation availability, as commonly encountered in medical imaging datasets.
- ***Challenges of deep learning approaches:*** Although deep learning methods excel in automated feature extraction and capturing complex patterns, their black-box nature can hinder interpretability, complicating model explanations and management of data leakage.
- ***Pragmatic Constraints and Data Leakage Risk:*** Practical challenges, such as limited sample sizes and the complexities of WSI technology in pathology, present barriers to the widespread adoption of deep learning approaches that rely on extensive digitized data. While data augmentation techniques enhance dataset size, they also introduce the risk of data leakage, underscoring the importance of maintaining the integrity of the training process.
- ***Annotation Scarcity:*** The scarcity of annotations in medical imaging datasets poses a significant challenge, necessitating innovative approaches such as transfer learning, active learning, and semi-supervised learning to leverage available data effectively.
- ***Clinical Significance of Features:*** The interpretability of hand-crafted features holds clinical significance, providing clinicians with valuable insights into oncological pathologies and contributing directly to the clinical decision-making process.

4.4 Future Directions

As the research progresses, there emerges an opportunity to delve into the realm of deep learning. The ongoing work on Gleason score classification, as mentioned, involves a more substantial dataset. This provides a logical segue into exploring the potential benefits of deep learning, considering the availability of data that aligns more closely with the requirements of deep neural networks. The identified need for larger, more representative datasets acknowledges the evolution of the research strategy. With a growing repository of diverse and voluminous data, the exploration of deep learning becomes a logical progression, ensuring that the models trained can generalize well to real-world scenarios. In conclusion, the decision to employ hand-crafted features in the initial stages of the research is a judicious one, considering data limitations and the imperative for interpretability in medical contexts. As the research advances and datasets expand, the inclusion of deep learning methodologies and the exploration of deep features in comparison to handcrafted descriptors becomes a strategic next step, paving the way for enhanced predictive modeling and nuanced insights into complex oncological data. This adaptive approach ensures a thoughtful integration of both traditional and cutting-edge techniques, aligning the research with the evolving landscape of data availability and computational methodologies in the field of oncological data analysis. As the thesis progresses into future work, the discussion on the transition to deep learning methodologies becomes even more pertinent. The acknowledgment of data limitations and the imperative for interpretability has steered the initial focus towards hand-crafted features. However, with the growing datasets, particularly in the ongoing Gleason score classification study, there arises a strategic opportunity to delve into the capabilities of deep learning, paving the way for more nuanced insights and predictive modeling in oncological data analysis. In essence, the decision to utilize hand-crafted features in Chapter 3 is part of a strategic narrative that encompasses the interpretability, clinical relevance, and real-world impact of the presented research [192]. As the journey continues, the integration of deep learning methodologies into the evolving research paradigm positions the work on a trajectory that combines the best of traditional and cutting-edge approaches, ultimately contributing to improved patient outcomes in

oncology.

4.5 Concluding Remarks

Overall, the thesis presents a comprehensive framework that addresses the challenges posed by the vast amount of heterogeneous biomedical data. The proposed integrative approach and the development of computational tools pave the way for a better understanding of cancer, supporting the transition towards personalized and precise clinical decision-making in oncology. Concerning the first objective, the effort made was to address issues concerning data organization and integration. The proposed approach provides a framework that overcomes challenges in standardizing and harmonizing data across diagnostic imaging, histopathology, and genetics. This initiative aligns with the broader goal of collecting numerical descriptors from patients with oncological diseases and organizing this wealth of information within a digital biobanking infrastructure. The biobanking system is designed to curate and standardize imaging data alongside clinical, molecular, and pathologic data, facilitating a comprehensive approach to cancer disease. Simultaneously, the second objective delved into supervised analysis at different imaging scales, with efforts addressing cross-scale integration challenges and explore the potential of these integrated datasets in supporting clinical decision-making. A concerted effort was made to adhere to standardized procedures, exemplified by the use of IBSI for radiomics; however, it is noted that there is still progress to be made in achieving comparable standards for pathomics. The studies conducted within this framework underscored deficiencies in data and procedure standardization and harmonization, which were diligently addressed in preceding chapters. In summary, the thesis strives to contribute to the fields of data science and medical informatics by proposing methodologies for data integration and organization, as well as advancing the supervised analysis of multiscale data. The ultimate goal is to enhance our understanding of oncological diseases and provide valuable support for clinical decision-making through the integration of data at different imaging scales, also considering genomics, and clinical data.

Appendix A

Preliminaries and basic concepts

This chapter introduces briefly the main domains related to this dissertation by clarifying some basic concepts.

A.1 Oncology

A.1.1 Defining a tumor

A tumor is an abnormal mass of tissue within the body, primarily arising from uncontrolled cell growth caused by mutated or damaged DNA (deoxyribonucleic acid). This DNA damage disrupts the normal cell cycle, inhibiting programmed cell death and promoting unregulated replication, resulting in the formation of a tumor mass. Broadly, tumors can be categorized as malignant or benign. Malignant tumors are cancerous, invading new areas of the body, while benign tumors are non-cancerous and pose no risk of spreading. Tumor metastasis, the spread of cancer cells to other parts of the body, is a major contributor to cancer-related deaths. Accurate characterization of tumor types is crucial for effective patient care.

A.1.2 Diagnosis and Staging

Cancer diagnosis involves a comprehensive evaluation of disease progression in a patient through a combination of clinical and pathological

observations. Typically, this process includes medical imaging and biopsy procedures to confirm the cancer type. Cancer staging aims to categorize patients based on similar tumor characteristics obtained from these procedures, along with their prognostic outlook or chance of survival. The widely adopted clinical staging system is Tumor Node Metastasis (TNM), developed by the American Joint Committee on Cancer (AJCC) in collaboration with the Union for International Cancer Control (UICC). In summary, the T component assesses primary tumor aspects such as size and contiguous growth into surrounding tissue, the N component defines the extent of cancer in regional lymph nodes, and the M component indicates the absence or presence of distant metastases. TNM characterization is specific to each cancer site, with values rising (e.g., T0-T4, N0-N3, M0-M1) corresponding to increased cancer extent and severity. As part of staging, cancer receives a histologic grade through a biopsy, evaluating tumor cell differentiation on a scale of 1-4. Grade 1 indicates well-differentiated cells resembling the surrounding tissue, while grades 3-4 denote poorly or undifferentiated cells, making it difficult to identify the biopsy's origin. The combination of T, N, and M components assigns patients to an overall prognostic group identified with Roman numerals (I-IV), further subdivided with characters (e.g., A-C). Higher stage groups and TNM stages correlate with poorer prognosis and increased disease severity. Accurate staging is crucial for treatment planning, as different stages warrant distinct treatment approaches, including invasive (e.g., surgical), non-invasive (e.g., radiotherapy), or combined strategies.

A.1.3 Conventional Treatments

There are many treatment options available for cancer management,each with varying applications and success rates contingent upon the current extent of the disease. Treatments are administered with either curative or palliative intent, and clinicians endeavor to choose the most suitable techniques based on various patient factors, such as staging and age, drawing insights from prior clinical outcomes. Before a treatment methodology gains widespread adoption, its efficacy is typically evaluated through clinical trials, wherein different treatments are compared across distinct arms of the trial—an ongoing process in the exploration of novel treatment opportunities. Frequently, treatments involve carefully managed rounds at

multiple time-points, and combinations may be utilized. For instance, neoadjuvant therapy precedes the primary treatment modality, as seen in neoadjuvant radiotherapy administered before surgery. Conversely, adjuvant therapy follows the primary treatment, as exemplified by adjuvant chemotherapy scheduled after primary radiotherapy to reduce the risk of relapse. Terminology is often integrated, such as chemoradiotherapy, signifying the concurrent delivery of radiotherapy and chemotherapy within a treatment plan. Traditional treatments are typically administered based on how an "average" patient with similar symptoms would respond. However, cancer has proven to be a dynamic and heterogeneous disease, with each patient representing a unique combination of genetic and environmental factors. This diversity becomes evident in clinical settings, as patients with the same cancer type and staging often exhibit varied responses to identical treatments. Over the past few decades, this realization has sparked a movement away from a one-size-fits-all treatment paradigm, ushering in a new era of tailored strategies that leverage extensive data from each individual. This data-driven approach forms the foundation of the National Health Service (NHS) initiative to enhance patient outcomes, known as Personalized Medicine, also commonly referred to as Precision Medicine.

A.1.4 Personalized medicine in Oncology

The hope of personalized or precision medicine in oncology is to leverage various personal data from a patient, primarily focusing on tumor biology, to carefully choose the most effective treatment strategy. Precision medicine suggests that by identifying and clinically quantifying the critical biological characteristics at each stage of a disease, targeting and eliminating these crucial steps could be a successful treatment approach. Equally important, this method has the potential to pinpoint treatments that may offer limited benefit. The goal is to administer the optimal therapy from the available options initially, steering clear of "trial and error" treatment approaches and more effectively managing undesirable and unnecessary treatment side effects. It's crucial to note that utilizing tumor data to inform clinical decisions is not a new concept; tissue biopsy and blood analysis have been common data points for decades in attempts to enhance the diagnosis and treatment of patient cancers. The recent ad-

vancements in high-throughput omics analysis hold the promise of providing further insights and introducing new therapeutic strategies. However, Precision Medicine faces a significant challenge posed by one of the limitations of omics strategies, which rely on single tumor biopsies: intratumoral heterogeneity, a topic explored in the following section.

A.1.5 Tumor Heterogeneity

Heterogeneity and homogeneity represent opposite aspects of the same observation: a description of either high dissimilarity or high similarity, respectively. The presence of diverse cell types within a tumor, known as tumor heterogeneity, poses a significant challenge in clinical treatment. Typically, cancerous tumors become more diverse as they grow, exhibiting both differences between patients with the same type of cancer and variations within the subgroups of cells within a single tumor mass. Variations within the subpopulations of cells arise from their distinct characteristics. For instance, differences in the genetic capabilities of subpopulations to metabolize or create vascular structures can result in necrotic regions—areas with dead cells—within a tumor. As previously mentioned, this thesis explores the potential identification of these differences in medical imaging as a useful quantitative measure. Intra-tumor heterogeneity is considered a primary mechanism for tumors adapting to targeted therapies, following Darwinian principles of selection that foster drug resistance. In highly heterogeneous tumors, therapies directed at a specific mutation identified from a single biopsy may only succeed in creating an environment where cancer cells lacking that mutation can thrive. Consequently, the entire tumor survives and becomes resistant to the treatment. Growing evidence suggests that patients with high intra-tumor heterogeneity experience a diminished response to treatment and inferior outcomes.

A.2 Machine learning

Machine learning is a field that focuses on the construction of algorithms to make predictions based on training data. A machine learning task aims to identify (to learn) a function f that maps the input domain X (of data) onto output domain Y (of possible predictions). Functions f are

chosen from different function classes, dependent on the type of learning algorithm that is being used. Mitchell defines "learning" as follows: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E". The performance measure P tells the researchers quantitatively how well a certain machine learning algorithm is performing. For a classification task, the accuracy of the system is usually chosen as the performance measure, where accuracy is defined as the proportion for which the system correctly produces the output. Experience E that machine learning algorithms undergo are datasets. These datasets contain a set of examples that are used to train and test these algorithms. Machine learning can be accomplished using either a supervised or an unsupervised approach. In supervised learning, the system receives a dataset with different example parameter values and decisions/classification, from which it infers a mathematical function, which automatically maps an input signal to an output signal. In supervised learning, the aim of the model is to estimate a hypothesis function h that approximates the real and true function f, given a dataset of N training samples. The values of the training set may be of any value, and can range from images, binary values to regression values. The performance of the supervised algorithm can be evaluated by applying the hypothesis function h on a test data set, which only contains the x values. Where the hypothesis is considered to generalise good if it is able to correctly predict a large portion of y values. In supervised learning there are two primary types of problems, classification and regression. A problem is defined as a classification problem if the value of the output y belongs to a class. This thesis will focus on binary classification, where the output y either has the value 1 (true) or 0 (false). A regression problem is when the model aims to predict y as either a discrete or continuous number. The supervised learning method was used in all the studies mentioned in this dissertation. Unsupervised learning, on the other hand, means that the system acts and observes the consequences of its actions, without referring to any predefined type cases other than those previously observed. This is pure 'learning by doing' or trial-and-error. In unsupervised learning the aim is similar to that of a supervised model. For the unsupervised case the model also aims to estimate a hypothesis function h. In unsupervised learning the

model aims to learn the pattern of the input data x0, x1, ...,xN, without receiving feedback of what prediction is correct. Unsupervised algorithms are most commonly used for clustering, association and dimensionality reduction problems. However, unsupervised algorithms can also be useful for classification problems, as these algorithms are able to discover the hidden structure of data. Compared to supervised learning, unsupervised methods typically perform poorly in the beginning, when they are untuned, but as they tune themselves, performance increases. It can be argued that using unsupervised learning, a classifying system should be able to set up hypotheses that no human can figure out, due to their complexity. To evaluate classifier performance given by a machine learning scheme, either a special testing dataset or a cross validation technique may be employed. A test dataset contains pre-classified data different to those in the training dataset, and is used only for evaluation, not for training. If data are scarce, it is sensible to use cross-validation in order not to waste any data, which could be useful to enhance classifier performance; all data are used both for training the classifier and for testing its performance. More data do not necessarily mean better classifier performance. Even though the classifier becomes better on the training set it could actually perform worse on the testing data. This is due to the overfitting of the classifier transfer function, so that it fits too tightly to the training data and the border between classes is jagged rather than smooth, unlike how it usually should be. Although a large number of machine learning classifiers have been developed and are currently available, some of the most commonly used machine learning algorithms or classifiers were used in this thesis. The basic learning concept and characteristics of these machine learning classifiers are described briefly in the next subparagraphs.

A.2.1 Random Forest

Random forest is a common type of supervised ML algorithm first proposed by Leo Breiman in the paper Random Forests. Randomforest is a tree-based ensemble method consisting of multiple decision tree classifiers. A decision tree make sdecisions by posing questions about each feature contained in the dataset. Each question is connected to a node in the tree, and every internal node is connected to a child node, where the child node represents an answer. Each question thereby forms a tree like struc-

ture, for binary classification every answer is either a true or a false. In random forest, each decision tree is used to make predictions about the input dataset, the final decision is made by aggregating the result from all trees through voting. It is desired that each decision tree is uncorrelated since this provides better and more consistent predictions and avoids overfitting. The random forest algorithm does this by utilising bootstrap aggregation(bagging) and feature randomness. In bootstrap aggregation, each decision tree is created using a random selection from the training dataset with out replacement. The algorithm behind random forest is successful for most prediction tasks since decision tree is uncorrelated which allow for errors to occur without affecting the outcome. In addition, random forest is widely used since it has a fast convergence rate and has a high prediction accuracy.

A.2.2 Support Vector Machines

Support Vector Machines (SVMs) are one of the most popular classification algorithms to perform supervised classification over the dataset. The aim of SVM is to find the most optimal classification function that differentiates between units of classes in training data. With a linearly separable dataset, the most optimal classification function can be decided by constructing a hyperplane which maximizes the margin between two datasets and thus creates the largest possible distance between datasets. The idea behind SVMs is that by finding the maximum margin, which represents the most optimal hyperplane and has the best generalization ability that it can reach. This results in the best classification performance for both the training data as well as future data. Thus, a Support Vector Machine is strictly not a machine, but a simple and powerful algorithm.

A.2.3 K Nearest Neighbor

k-Nearest-Neighbor classification, or kNN, is another popular machine learning algorithm that localizes a group of k objects in a training case that has the closest proximity or similarity to the test object, and then assigns a label derived from the prevalence of a class in the closest proximity. Three important components are needed for this algorithm: a group of labeled objects; a proximity metric; and the number k of nearest neighbors. A

popular proximity metric that is used for kNN classification is "Euclidian Distance". Other metrics exist for defining the distance between instances. of a dataset. Examples include the Minkowsky, Camberra or Chebychev metrics although often weighing strategies are used that alter the voting influence for more accurate results.

A.2.4 Decision Tree

Decision Tree is a powerful and highly adaptable machine learning method. The method is based on multiple decisions made in a tree-like structure, as the name suggests. Starting with a main problem of dividing a set of data points into two or more classes and correctly doing so by going through a series of judgements or decisions. The decisions in question are made from testing the different features in the data set. Each decision in the process either leads to the final conclusion or another test with the previous decision kept as a condition. The Decision Tree is generated recursively, following the divide-and-conquer method. Starting from a root node, the Decision Tree is further made up of internal nodes and leaf nodes, where the leaf nodes correspond to the decision results and all other nodes correspond to the feature testing. The recursion is stopped if all samples in the current node belongs to the same class, if the current feature set is empty or if the current node is emptied of samples. An important part of the Decision Tree algorithm is selecting the splitting procedure. Reviewing Information Gain, the Gain Ratio or the Gini Index are all frequently used for this purpose. Information Gain is based on measuring the purity of each node via evaluation of the information entropy, where lower entropy means higher purity of the node. The purity of a node is a measurement of how many samples in the node belongs to a single class. Taking the varying number of samples in different nodes into account, the Information Gain of splitting the data can be evaluated for each feature, where the highest Information Gain indicates which feature to use as the splitting feature. A downside of the Information Gain procedure is that it proves biased to features with many possible values. This problem is handled by instead using the Gain Ratio which also accounts for the intrinsic value of each feature which is related to the number of possible values. However, the Gain Ratio is instead biased towards the features with few possible values. A third splitting procedure, the Gini index, uses the Gini value which

represents the probability that two randomly selected samples from the data set belongs to different classes.

A.2.5 Multilayer Perceptron

The MLP is a type of feed-forward Neural Network commonly used in binary classification, recognition, approximation and problems that are not linearly separable. The MLP consists of three types of layers; an input layer, hidden layers and an output layer, as shown in Figure 3.1. The input layer receives the signal to be processed and passes it to the rest of the network. The number of hidden layers is arbitrary to the task of the network, and these layers perform the computations critical to the function of the classification or regression. The output layer performs the classification and prediction based on the output of the hidden layers. As an MLP is a type of feed–forward network the data is passed from the input layer to the output layer in a forward direction. However, the neurons are trained with the back propagation learning algorithm. The back propagation algorithm allows for the network to repeatedly adjust the weights of the connections between the neurons, the goal of the algorithm is to minimise the difference between the networks output vector and the desired output vector. Thus, back propagation allows for the network to learn hidden patterns in the input or output.

A.2.6 Linear Discriminant Analysis

The Linear Discriminant Analysis (LDA), is a linear classification and dimensionality reduction technique suitable for both binary and multi-class classification problems. The main idea of LDA is to project data from a higher dimensional feature space to a lower dimensional space to separate classes and minimize the variability within the classes. LDA is based on the statistical properties variance or covariance and the mean values of the data of interest for each class. Besides being highly relevant for comparison with other linear algorithms, such as Logistic Regression, LDA also address some weak points and limitations of other algorithms. LDA is more suited for the multi-class classification problems in comparison with other linear algorithms. Furthermore, an algorithm such as Logistic Regression can become unstable when classes are well separated or when too few examples

are used to estimate the parameters. These instabilities are also considered when using LDA.

A.2.7 Gradient Boosting Machines

Gradient boosting came from the AdaBoost that weak learner can be modified to become stronger, decision trees with a single split. Weak learner is defined as a classifier that is slightly correlated to true classification, but slightly better than random probability. According to Friedman, each step consists weighted minimization and recomputation of classifier and weight. Generalization of AdaBoost then becomes gradient boosting. According to Friedman's gradient boosting algorithm, on each iteration, the gradient descent is first computed in order to fit a new base learner function. Once the best gradient descent step size is found, the function estimation is updated. Here is the formula of gradient descent:
$\theta_{n+1} = \theta_n - \eta L'(\theta_n)$
where η is the step size or learning rate, and $L'(\theta_n)$ is the gradient. The new model is fitted using information of errors from the previous model. The resulting predictors are the combination of all models. GBMs build trees at the same time, and new trees help to correct errors made by previous trees. Typically, number of trees, depth of trees and learning rate are the most important parameters. The recommendation is to choose smaller trees rather than larger ones. Other regularization method includes stochastic gradient boosting which is the boosting with sub sampling per split, and regularized gradient boosting which is the boosting with L1 or L2 regularization. At each iteration, the data are sampled into without replacement and the tree is built only using the sample of data. Compared to random forest, GBM tends to have overfitting, but at the same time, if the hyperparameter setting is correct, GBM can have strong prediction to the data.

A.2.8 EXtreme Gradient Boosting

Extreme Gradient Boosting (XGBoost) is a sophisticated machine learning algorithm celebrated for its remarkable performance across various supervised learning domains. Founded on the principles of gradient boosting, XGBoost introduces several innovations aimed at enhancing both com-

putational efficiency and predictive accuracy. At its essence, XGBoost operates through iterative optimization of a base learner, typically a decision tree, to progressively refine the predictive model's efficacy. The algorithm initiates with an initial approximation of the target function, with subsequent iterations focusing on minimizing the ensemble's errors by incorporating new decision trees trained to rectify the residual errors of preceding ones. A pivotal advancement of XGBoost lies in its adoption of a more efficient and scalable algorithm for tree construction. By leveraging gradient-based optimization techniques, XGBoost facilitates expedited convergence and mitigates computational complexity, rendering it particularly adept for handling extensive datasets and high-dimensional feature spaces. Furthermore, XGBoost integrates regularization strategies like L1 and L2 regularization, alongside tree pruning and column subsampling, to forestall overfitting and amplify model generalization. These regularization mechanisms effectively manage model complexity, thereby enhancing performance on unseen data instances. Moreover, XGBoost showcases versatility in accommodating diverse data types and objective functions. Whether tasked with classification, regression, ranking, or other supervised learning objectives, XGBoost offers customizable functionalities to adapt the model's behavior to specific task requirements. XGBoost learns by building an ensemble of decision trees iteratively, minimizing an objective function composed of the training loss and regularization terms. It leverages gradient descent to find the optimal trees, employing first and second-order derivatives of the loss function. XGBoost utilizes hyperparameters such as maximum depth, regularization parameters, and subsampling strategies for features and instances to prevent overfitting and improve computational efficiency. It is worth noting that subsampling, in particular, introduces randomness and diversity to the model, reducing the chances of overfitting and speeding up the training process by processing fewer data points during each iteration.

Appendix B

Additional Materials

B.1 Datasets

B.1.1 The Clinical Proteomic Tumor Analysis Consortium Glioblastoma Multiforme Collection (CPTAC-GBM)

This subsection describes data employed in the subsection 3.1.1. The information presented in the paragraph is sourced from https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=30671232#30671232cd1f5cd3cd8441738857b6174cea1d35. The investigated collection includes subjects from the National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium Glioblastoma Multiforme (CPTAC-GBM) cohort. CPTAC is a national effort to accelerate the understanding of the molecular basis of cancer through the application of large-scale proteome and genome analysis, or proteogenomics. Radiology and pathology images from CPTAC patients are being collected and made publicly available by The Cancer Imaging Archive to enable researchers to investigate cancer phenotypes which may correlate to corresponding proteomic, genomic and clinical data. Radiology imaging is collected from standard of care imaging performed on patients immediately before the pathological diagnosis, and from follow-up scans where available. Pathology imaging is collected as part of the CPTAC qualification workflow. Table B.1 reports the detailed description of the Radiology and Pathology images. For the development of the study described in subsection 3.1.1, proteomic and ge-

Table B.1. Characteristics of Clinical Proteomic Tumor Analysis Consortium Glioblastoma Multiforme (CPTAC-GBM) Radiology and Pathology images

	Radiology Image Statistics	**Pathology Image Statistics**
Modalities	CT, MR	Pathology
Number of Participants	66	178
Number of Studies	148	N/A
Number of Series	1,752	N/A
Number of Images	156,259	462
Images Size (GB)	39.4	86

nomic data were not considered since the study was imaging-oriented. 48 patients were selected according to the following inclusion criteria: availability of pretreatment T1C images and ADC maps, availability of corresponding digital pathology WSI, WSI slides with at least 70% of tumor nuclei and at most 20% necrosis. Moreover, patients were excluded in case of i) insufficient quality of MRI to perform imaging analysis and/or obtain measurements, ii) insufficient quality of WSI that did not meet the requirements for diagnosis (e.g., tissue folds, torn tissue) and iii) images with a positive value of Clinical Trial Time Point ID (corresponding to the number of days from the date the patient was initially diagnosed pathologically with the disease to the date of the scan).

B.1.2 Tumor Infiltrating lymphocytes in breast cancER (TIGER)

This subsection describes data employed in the subsection 3.2.1. The information presented in the paragraph is sourced from https://tiger.grand-challenge.org/Home/. The dataset was released by the TIGER Challenge that is the first challenge on fully automated assessment of tumor-infiltrating lymphocytes (TILs) in H&E breast cancer slides. It was organized by the Diagnostic Image Analysis Group (DIAG) of the Radboud University Medical Center (Radboudumc) in Nijmegen (The Netherlands), in close collaboration with the International Immuno-Oncology Biomarker

working Group (www.tilsinbreastcancer.org). The challenge was ideated to evaluate new computer algorithms for the automated assessment of TILs in HER2 positive (HER2+) and Triple Negative (TNBC) breast cancer histopathology slides. The organizers of the TIGER challenge publicly released training data containing digital pathology images of HER2+ and TNBC breast cancer whole-slide images, together with manual annotations. Training data comes from multiple sources. A subset of Her2+ and TNBC cases was provided by the Radboud University Medical Center (RUMC) (Nijmegen, Netherlands). A second subset of HER2+ and TNBC cases was provided by the Jules Bordet Institut (JB) (Bruxelles, Belgium). A third subset of TNBC cases only was derived from the TCGA-BRCA archive obtained from the Genomic Data Commons Data Portal. Training data were released in the format of three datasets, which they called WSIROIS (used in this thesis), WSIBULK,and WSITILS, that are described in the following subsections. All data, both at WSI and at ROI level, is released at a spacing (pixel size) of approximately 0.5 um/px.

B.1.2.1 WSIROIS: Whole-slide images with manual annotations in regions of interest

This set was used in this thesis. In this set, the TIGER Challenge organizers released n=195 whole-slide images of breast cancer, both (core-needle) biopsies, and surgical resections, with regions of interest selected and manually annotated. This dataset contains images and annotations from multiple sources:

- **TCGA: n=151** WSIs of TNBC cases from the TGCA-BRCA archive. Annotations are extracted and adapted from the publicly available BCSS [40]and NuCLS [39] datasets.
- **RUMC: n=26** WSIs of TNBC and HER2+ cases from Radboud University Medical Center (Netherlands). Annotations were made by a panel of board-certified breast pathologists.
- **JB: n=18 WSIs** of TNBC and HER2+ cases from Jules Bordet Institute (Belgium). Annotations were made by a panel of board-certified breast pathologists.

In each WSI, ROIs are manually annotated with both polygons indicating different tissue compartments, and with point annotations indicating lymphocytes and plasma cells. In each ROI, the following regions are annotated, with the corresponding labels:

- **invasive tumor**: this class contains regions of the invasive tumor, including several morphological subtypes, such as invasive ductal carcinoma and invasive lobular carcinoma;
- **tumor-associated stroma**: this class contains regions of stroma (i.e., connective tissue) that are associated with the tumor. This means stromal regions contained within the main bulk of the tumor and in its close surrounding; in some cases, the tumor-associated stroma might resemble the "healthy" stroma, typically found outside of the tumor bulk;
- **in-situ tumor**: this class contains regions of in-situ malignant lesions, such as ductal carcinoma in situ (DCIS) or lobular carcinoma in situ (LCIS).
- **healthy glands**: this class contains regions of glands with healthy epithelial cells;
- **necrosis not in-situ**: this class contains regions of necrotic tissue that are not part of in-situ tumor; for example, ductal carcinoma in situ (DCIS) often presents a typical necrotic pattern, which can be considered as part of the lesion itself, such a necrotic region is not annotated as "necrosis" but as "in-situ tumor";
- **inflamed stroma**: this class contains tumor-associated stroma that has a high density of lymphocytes (i.e., it is "inflamed"). When it comes to assessing the TILs, inflamed stroma and tumor-associated stroma can be considered together, but were annotated separately to take into account for differences in their visual patterns;
- **rest**: this class contains regions of several tissue compartments that are not specifically annotated in the other categories; examples are healthy stroma, erythrocytes, adipose tissue, skin, nipple, etc.

Additionally, most ROIs contain annotations of **lymphocytes and plasma cells** in the form of bounding boxes. Cells were annotated using point annotations and then squared bounding boxes of 8x8 microns were constructed centered on the point annotation.

B.1.2.2 WSIBULK: Whole-slide images with coarse manual annotation of the tumor bulk

A set of n=93 WSIs of both biopsies and surgical resections of TNBC and HER2+ breast cancer tissue from RUMC and JB was also released. For each WSI, coarse annotations of the "tumor bulk" were provided, specifically including a coarse manual annotation of one or more regions in the slide that contain invasive tumor cells. The organizers made sure that all cancer cells belonging to the invasive part of the tumor are confined within the manually annotated regions and that all tissue outside of these annotations does not contain tumor cells belonging to the invasive part of the tumor.

B.1.2.3 WSITILS: Whole-slide images with visual estimation of the TILs at slide level

An additional set of n=82 WSIs of both biopsies and surgical resections of TNBC and HER2+ breast cancer tissue from RUMC and JB where the visual assessment of the TILs has been done at WSI level was provided. In this dataset, no manual annotations are provided, but a list of TIL values, one per slide, was released. TILs in this set of cases have been assessed by a board-certified breast pathologist, following the recommendation of the TILs working group [265]. Several cases contain comments indicating potential pitfalls in either visual TILs scoring or machine-based TILs scoring, some of which were addressed in [172].

For the development of the pipeline described in subsection 3.2.1, the 'WSIROIS' dataset was employed. This dataset consists of 195 WSIs of core-needle biopsies and surgical resections of BC patients with annotated tissue compartments and TILs. The classes 'tumor-associated stroma' and 'inflamed stroma' were merged to form a 'stroma' class. Then, all other classes were set to zero. In this way, masks only consisted of background

and stroma, and TILs. ROIs without stromal tissue annotation were excluded. There are plans to use the remaining part of the dataset released by the organizers of the TIGER challenge to develop future studies incorporating it into a more complex pipeline that includes all the steps of the TIL-WG guidelines (e.g. background removal, identification of invasive tumor and stroma tissue), also exploring the impact of different intermediate processing steps (e.g. nuclear segmentation, stain normalization) on the TILs classification performance.

B.1.3 Automated Gleason Grading Challenge 2022 (AGGC22)

This subsection describes data employed in the subsection 3.2.2. The information presented in the paragraph is sourced from https://aggc22.grand-challenge.org/. The dataset was released by the organizers of the Automated Gleason Grading Challenge (AGGC) 2022. This challenge was organized with the aim to help drive forward research and innovation for Automated Gleason Grading in computational pathology . The challenge required researchers to develop algorithms that identify distinct Gleason Patterns within the H&E-stained whole slide image dataset. The published dataset included H&E-stained whole slide images of prostatectomy and biopsy specimens with annotations performed by experienced pathologists. Additionally, a set of images scanned by multiple scanners was also provided with the aim to assess the algorithm performance of handling variations caused by image digitalization. The dataset for this challenge is obtained from National University Hospital, Singapore. Below are listed and detailed the released data cohorts:

- **Subset 1**: Whole mount images scanned by a Akoya Biosciences scanner, including 150 cases, of which 105 cases for training set and 45 for test set.
- **Subset 2**: Biopsy images scanned by a Akoya Biosciences scanner, including 53 cases, of which 37 for training set and 16 for test set.
- **Subset 3**: Whole mount images scanned by multiple scanners. The training set was composed of WSIs scanned by multiple scanners (Akoya Biosciences, Olympus, Zeiss, Leica, KFBio, Philips). Each scanner scanned 26 cases except 25 for Olympus and 15 for Zeiss. In

total, 144 cases were released. Same procedures were performed to obtain the test set, for which each scanner scanned 12 cases except 7 for Zeiss. In total, 67 cases were released.

In training set, each image comes with a set of binary masks of annotations ("Gleason Pattern3", "Gleason Pattern4", "Gleason Pattern5", "Normal", "Stroma") performed by pathologists. The number of binary masks varies from case to case and maximum is 5. The size of the pixel is 0.5 µm/pixel (20x) for images and annotations. All H&E images are converted from its original formats to " .tiff" using MATLAB. Manual annotations are performed on images scanned by Akoya Biosciences only. For subset 3, images scanned by different scanners are not aligned with ones scanned by Akoya Biosciences. Therefore, we performed the image registration to transform the original masks to match with images scanned by other scanners. Considering that prostatectomy specimens are much larger and thus contain more information than biopsies, the annotations in prostatectomy WSIs were used in this thesis to train and test the models developed in 3.2.2. The 187 radical prostatectomy WSIs scanned by a Akoya Biosciences (train and test of Subset 1 + train and test of subset 3) scanner were selected. Preprocessing steps described in the 3.2.2 were applied to these 187 slides. There are plans to use the remaining part of the dataset released by the organizers of the AGGC22 challenge to develop future studies to evaluate the models developed both on WSI images derived from biopsies obtained from the same scanner (Subset 2) and on images obtained from different scanners. In addition, there are plans to study the reproducibility of features extracted from images of the same slides obtained from different scanners (Subset 3).

B.2 The relationship between radiomics and pathomics in Glioblastoma patients: Preliminary results from a cross-association study

B.2.1 Extracted radiomic and pathomic features

This subsection shows the complete list of radiomic and pathomic features extracted in the paragraphs 3.1.1.2.4 and 3.1.1.2.5.

B.2.1.1 Radiomic features

Below are listed the extracted radiomic features from ADC and T1+C images: First-order features including 18 intensity statistics; 73 multi-dimensional texture features including 23 Gray Level Co-occurrence Matrix (GLCM), 16 Gray Level Size Zone Matrix (GLSZM), 16 Gray Level Run Length Matrix (GLRLM), 14 Gray Level Dependence Matrix (GLDM) and 5 Neighboring Gray Tone Difference Matrix (NGTDM) Features. All the mathematical formulations of the computed features can be found at https://pyradiomics.readthedocs.io/en/latest/features.html.

First-order features (n = 18)

- firstorder InterquartileRange
- firstorder Median
- firstorder Uniformity
- firstorder Range
- firstorder Entropy
- firstorder Minimum
- firstorder Kurtosis
- firstorder RootMeanSquared
- firstorder 90Percentile
- firstorder Energy
- firstorder Skewness
- firstorder Mean
- firstorder Variance
- firstorder RobustMeanAbsoluteDeviation
- firstorder MeanAbsoluteDeviation
- firstorder Maximum
- firstorder 10Percentile

Second-order texture features (n = 73)

- glrlm RunLengthNonUniformity
- glrlm GrayLevelNonUniformityNormalized
- glrlm LongRunLowGrayLevelEmphasis
- glrlm HighGrayLevelRunEmphasis
- glrlm RunVariance
- glrlm RunLengthNonUniformityNormalized

- glrlm ShortRunEmphasis
- glrlm LongRunEmphasis
- glrlm LongRunHighGrayLevelEmphasis
- glrlm GrayLevelNonUniformity
- glrlm ShortRunHighGrayLevelEmphasis
- glrlm ShortRunLowGrayLevelEmphasis
- glrlm RunPercentage
- glrlm GrayLevelVariance
- glrlm RunEntropy
- glrlm LowGrayLevelRunEmphasis
- ngtdm Coarseness
- ngtdm Complexity
- ngtdm Busyness
- ngtdm Strength
- ngtdm Contrast
- glcm JointEntropy
- glcm ClusterShade
- glcm ClusterProminence
- glcm Autocorrelation
- glcm SumSquares
- glcm SumEntropy
- glcm DifferenceVariance
- glcm Idmn
- glcm Imc2
- glcm Imc1
- glcm DifferenceAverage
- glcm JointAverage
- glcm InverseVariance
- glcm DifferenceEntropy
- glcm Idm
- glcm Idn
- glcm Correlation
- glcm MCC
- glcm ClusterTendency
- glcm Contrast
- glcm Id

- glcm MaximumProbability
- glcm JointEnergy
- gldm SmallDependenceHighGrayLevelEmphasis
- gldm LargeDependenceLowGrayLevelEmphasis
- gldm DependenceNonUniformity
- gldm LargeDependenceHighGrayLevelEmphasis
- gldm DependenceNonUniformityNormalized
- gldm LowGrayLevelEmphasis
- gldm LargeDependenceEmphasis
- gldm SmallDependenceEmphasis
- gldm DependenceEntropy
- gldm DependenceVariance
- gldm SmallDependenceLowGrayLevelEmphasis
- gldm GrayLevelNonUniformity
- gldm HighGrayLevelEmphasis
- gldm GrayLevelVariance
- glszm GrayLevelNonUniformityNormalized
- glszm SmallAreaHighGrayLevelEmphasis
- glszm GrayLevelVariance
- glszm LargeAreaEmphasis
- glszm SizeZoneNonUniformity
- glszm LargeAreaLowGrayLevelEmphasis
- glszm HighGrayLevelZoneEmphasis
- glszm ZonePercentage
- glszm ZoneVariance
- glszm GrayLevelNonUniformity
- glszm LargeAreaHighGrayLevelEmphasis
- glszm SizeZoneNonUniformityNormalized
- glszm SmallAreaEmphasis
- glszm ZoneEntropy
- glszm SmallAreaLowGrayLevelEmphasis
- glszm LowGrayLevelZoneEmphasis

B.2.1.2 Pathomic features

Below are listed the extracted pathomic features from detection measurements (shape characteristics, intranuclear Haralick texture features,

Delaunay triangulation) and cell-density maps features including 18 firstorder intensity statistics and 73 multi-dimensional texture features including 23 Gray Level Co-occurrence Matrix (GLCM), 16 Gray Level Size Zone Matrix (GLSZM), 16 Gray Level Run Length Matrix (GLRLM), 14 Gray Level Dependence Matrix (GLDM) and 5 Neighboring Gray Tone Difference Matrix (NGTDM) Features. All the mathematical formulations of the computed features can be found at https://qupath.readthedocs.io/en/stable/docs/intro/index.html and https://pyradiomics.readthedocs.io/en/latest/features.html

Detection measurements

- MEAN Nucleus Area
- MEAN Nucleus Perimeter
- MEAN Nucleus Circularity
- MEAN Nucleus MaxCaliper
- MEAN Nucleus MinCaliper
- MEAN Nucleus Eccentricity
- MEAN Nucleus HematoxylinODMean
- MEAN Nucleus HematoxylinODSum
- MEAN Nucleus HematoxylinODStdDev
- MEAN Nucleus HematoxylinODMax
- MEAN Nucleus HematoxylinODMin
- MEAN Nucleus HematoxylinODRange
- MEAN Nucleus EosinODMean
- MEAN Nucleus EosinODSum
- MEAN Nucleus EosinODStdDev
- MEAN Nucleus EosinODMax
- MEAN Nucleus EosinODMin
- MEAN Nucleus EosinODRange
- MEAN Cell Area
- MEAN Cell Perimeter
- MEAN Cell Circularity
- MEAN Cell MaxCaliper
- MEAN Cell MinCaliper
- MEAN Cell Eccentricity
- MEAN Cell HematoxylinODMean
- MEAN Cell HematoxylinODStdDev

- MEAN Cell HematoxylinODMax
- MEAN Cell HematoxylinODMin
- MEAN Cell EosinODMean
- MEAN Cell EosinODStdDev
- MEAN Cell EosinODMax
- MEAN Cell EosinODMin
- MEAN Cytoplasm HematoxylinODMean
- MEAN Cytoplasm HematoxylinODStdDev
- MEAN Cytoplasm HematoxylinODMax
- MEAN Cytoplasm HematoxylinODMin
- MEAN Cytoplasm EosinODMean
- MEAN Cytoplasm EosinODStdDev
- MEAN Cytoplasm EosinODMax
- MEAN Cytoplasm EosinODMin
- MEAN Nucleus CellAreaRatio
- MEAN Delaunay MeanDistance
- MEAN Delaunay MedianDistance
- MEAN Delaunay MaxDistance
- MEAN Delaunay MinDistance
- MEAN Delaunay MeanTriangleArea
- MEAN Delaunay MaxTriangleArea
- MEAN ODSum Mean
- MEAN ODSum Std dev
- MEAN ODSum Min
- MEAN ODSum Max
- MEAN ODSum Median
- MEAN ODSum HaralickAngularSecondMoment F0
- MEAN ODSum HaralickContrast F1
- MEAN ODSum HaralickCorrelation F2
- MEAN ODSum HaralickSumOfSquares F3
- MEAN ODSum HaralickInverseDifferenceMomen
- MEAN ODSum HaralickSumAverage F5
- MEAN ODSum HaralickSumVariance F6
- MEAN ODSum HaralickSumEntropy F7
- MEAN ODSum HaralickEntropy F8
- MEAN ODSum HaralickDifferenceVariance F9

- MEAN ODSum HaralickDifferenceEntropy F10
- MEAN ODSum HaralickInformationMeasureOfCorrelation 1 F11
- MEAN ODSum HaralickInformationMeasureOfCorrelation 2 F12

Cell-density map-based (first-order features)

- firstorder InterquartileRange
- firstorder Median
- firstorder Uniformity
- firstorder Range
- firstorder Entropy
- firstorder Minimum
- firstorder Kurtosis
- firstorder RootMeanSquared
- firstorder 90Percentile
- firstorder Energy
- firstorder Skewness
- firstorder Mean
- firstorder Variance
- firstorder RobustMeanAbsoluteDeviation
- firstorder MeanAbsoluteDeviation
- firstorder Maximum
- firstorder 10Percentile

Cell-density map-based second-order texture features (n = 73)

- glrlm RunLengthNonUniformity
- glrlm GrayLevelNonUniformityNormalized
- glrlm LongRunLowGrayLevelEmphasis
- glrlm HighGrayLevelRunEmphasis
- glrlm RunVariance
- glrlm RunLengthNonUniformityNormalized
- glrlm ShortRunEmphasis
- glrlm LongRunEmphasis
- glrlm LongRunHighGrayLevelEmphasis
- glrlm GrayLevelNonUniformity
- glrlm ShortRunHighGrayLevelEmphasis
- glrlm ShortRunLowGrayLevelEmphasis

- glrlm RunPercentage
- glrlm GrayLevelVariance
- glrlm RunEntropy
- glrlm LowGrayLevelRunEmphasis
- ngtdm Coarseness
- ngtdm Complexity
- ngtdm Busyness
- ngtdm Strength
- ngtdm Contrast
- glcm JointEntropy
- glcm ClusterShade
- glcm ClusterProminence
- glcm Autocorrelation
- glcm SumSquares
- glcm SumEntropy
- glcm DifferenceVariance
- glcm Idmn
- glcm Imc2
- glcm Imc1
- glcm DifferenceAverage
- glcm JointAverage
- glcm InverseVariance
- glcm DifferenceEntropy
- glcm Idm
- glcm Idn
- glcm Correlation
- glcm MCC
- glcm ClusterTendency
- glcm Contrast
- glcm Id
- glcm MaximumProbability
- glcm JointEnergy
- gldm SmallDependenceHighGrayLevelEmphasis
- gldm LargeDependenceLowGrayLevelEmphasis
- gldm DependenceNonUniformity
- gldm LargeDependenceHighGrayLevelEmphasis

- gldm DependenceNonUniformityNormalized
- gldm LowGrayLevelEmphasis
- gldm LargeDependenceEmphasis
- gldm SmallDependenceEmphasis
- gldm DependenceEntropy
- gldm DependenceVariance
- gldm SmallDependenceLowGrayLevelEmphasis
- gldm GrayLevelNonUniformity
- gldm HighGrayLevelEmphasis
- gldm GrayLevelVariance
- glszm GrayLevelNonUniformityNormalized
- glszm SmallAreaHighGrayLevelEmphasis
- glszm GrayLevelVariance
- glszm LargeAreaEmphasis
- glszm SizeZoneNonUniformity
- glszm LargeAreaLowGrayLevelEmphasis
- glszm HighGrayLevelZoneEmphasis
- glszm ZonePercentage
- glszm ZoneVariance
- glszm GrayLevelNonUniformity
- glszm LargeAreaHighGrayLevelEmphasis
- glszm SizeZoneNonUniformityNormalized
- glszm SmallAreaEmphasis
- glszm ZoneEntropy
- glszm SmallAreaLowGrayLevelEmphasis
- glszm LowGrayLevelZoneEmphasis

B.2.2 Summary of the significantly moderately correlated radiomic-pathomic features

Table B.2: Summary of the significantly moderately correlated radiomic-pathomic features, with radiomic features extracted from ADC. Abbreviations: CD = Cellular Density; ADC = Apparent Diffusion Coefficient; LALGLE = Large Area Low Gray Level Emphasis; IMOC = Information Measure Of Correlation; ASM = Angular Second Moment; LDLGLE = Large Dependence Low Gray Level Emphasis; LDHGLE = Large Dependence High Gray Level Emphasis; SDLGLE = Small Dependence Low Gray Level Emphasis; LALGLE = Large Area Low Gray Level Emphasis; GLNUN = Gray level non uniformity normalized; glcm = gray level co-occurrence matrix; gldm = Gray Level Dependence Matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix.

Radiomic feature Name (ADC)	Pathomic Feature name	ρ	FDR q	BF

Continued on next page

Table B.2: Summary of the significantly moderately correlated radiomic-pathomic features, with radiomic features extracted from ADC. Abbreviations: CD = Cellular Density; ADC = Apparent Diffusion Coefficient; LALGLE = Large Area Low Gray Level Emphasis; IMOC = Information Measure Of Correlation; ASM = Angular Second Moment; LDLGLE = Large Dependence Low Gray Level Emphasis; LDHGLE = Large Dependence High Gray Level Emphasis; SDLGLE = Small Dependence Low Gray Level Emphasis; LALGLE = Large Area Low Gray Level Emphasis; GLNUN = Gray level non uniformity normalized; glcm = gray level co-occurrence matrix; gldm = Gray Level Dependence Matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix. (Continued)

glcm ClusterShade	MEAN Haralick ASM F0	0.589	1.66×10^{-3}	1.73×10^{3}
glcm Imc2	CD100µm glrlm LRE	-0.589	1.67×10^{-3}	1.68×10^{3}
firstorder Maximum	CD200µm ngtdm Strength	0.588	1.71×10^{-3}	1.62×10^{3}
firstorder 90Percentile	MEAN Haralick IMOC F12	-0.586	1.83×10^{-3}	1.49×10^{3}
glcm DifferenceAverage	MEAN Haralick ASM F0	0.585	1.84×10^{-3}	1.46×10^{3}
glcm DifferenceVariance	CD100µm ngtdm Strength	0.584	1.88×10^{-3}	1.40×10^{3}
glcm InverseVariance	CD150µm glszm LALGLE	0.582	1.99×10^{-3}	1.31×10^{3}
firstorder 90Percentile	CD100µm ngtdm Strength	0.579	2.26×10^{-3}	1.14×10^{3}
gldm DependenceVariance	CD100µm glrlm LRE	0.578	2.34×10^{-3}	1.09×10^{3}
gldm SDLGLE	MEAN Haralick IMOC F12	-0.577	2.42×10^{-3}	1.04×10^{3}
glcm ClusterProminence	CD200µm ngtdm Strength	0.573	2.72×10^{-3}	9.15×10^{2}
ngtdm Complexity	CD150µm glcm Contrast	0.573	2.76×10^{-3}	8.89×10^{2}
ngtdm Complexity	CD50µm glcm DV	0.572	2.82×10^{-3}	8.56×10^{2}
gldm SDLGLE	CD200µm ngtdm Strength	0.57	2.91×10^{-3}	8.19×10^{2}
firstorder TotalEnergy	CD50µm glcm DV	0.564	3.67×10^{-3}	6.50×10^{2}
glrlm RunVariance	CD100µm glrlm LRE	0.563	3.73×10^{-3}	6.30×10^{2}
firstorder 10Percentile	CD150µm ngtdm Strength	0.558	4.55×10^{-3}	5.15×10^{2}
glcm ClusterProminence	CD50µm glcm DV	0.546	7.04×10^{-3}	3.37×10^{2}
ngtdm Strength	MEAN Haralick ASM F0	0.543	7.74×10^{-3}	3.01×10^{2}
firstorder Median	CD150µm ngtdm Strength	0.543	7.74×10^{-3}	3.00×10^{2}
firstorder Minimum	CD150µm ngtdm Strength	0.542	7.88×10^{-3}	2.92×10^{2}
glszm GLNUN	CD100µm glrlm LRE	0.541	7.93×10^{-3}	2.83×10^{2}
firstorder TotalEnergy	CD150µm glcm Contrast	0.541	7.93×10^{-3}	2.82×10^{2}
glcm InverseVariance	CD100µm glrlm LRE	0.539	8.34×10^{-3}	2.66×10^{2}
ngtdm Complexity	CD50µm ngtdm Complexity	0.539	8.45×10^{-3}	2.59×10^{2}
firstorder Mean	CD150µm ngtdm Strength	0.535	9.40×10^{-3}	2.32×10^{2}
glcm ClusterTendency	CD50µm ngtdm Complexity	0.531	1.07×10^{-2}	2.02×10^{2}
glcm ClusterProminence	CD50µm glcm Contrast	0.531	1.07×10^{-2}	2.00×10^{2}
firstorder Range	CD50µm ngtdm Complexity	0.53	1.09×10^{-2}	1.95×10^{2}
firstorder TotalEnergy	CD50µm glcm DA	0.53	1.09×10^{-2}	1.92×10^{2}
firstorder Maximum	MEAN Haralick IMOC2 F12	-0.525	1.27×10^{-2}	1.65×10^{2}
glcm DifferenceAverage	CD200µm ngtdm Strength	0.524	1.28×10^{-2}	1.61×10^{2}
glcm ClusterTendency	CD50µm glcm Contrast	0.524	1.30×10^{-2}	1.57×10^{2}
glcm ClusterTendency	CD50µm glcm DV	0.523	1.32×10^{-2}	1.53×10^{2}
glcm ClusterProminence	MEAN Haralick IMOC2 F12	-0.522	1.35×10^{-2}	1.48×10^{2}
glcm ClusterTendency	MEAN Haralick ASM F0	0.521	1.37×10^{-2}	1.45×10^{2}
glcm ClusterShade	CD200µm ngtdm Strength	0.519	1.44×10^{-2}	1.37×10^{2}
ngtdm Contrast	MEAN HaralickDV F9	0.519	1.46×10^{-2}	1.34×10^{2}
glcm ClusterProminence	CD100µm ngtdm Strength	0.518	1.46×10^{-2}	1.32×10^{2}
ngtdm Strength	CD50µm ngtdm Complexity	0.518	1.48×10^{-2}	1.29×10^{2}
ngtdm Complexity	CD50µm glcm DA	0.517	1.49×10^{-2}	1.28×10^{2}

Continued on next page

Table B.2: Summary of the significantly moderately correlated radiomic-pathomic features, with radiomic features extracted from ADC. Abbreviations: CD = Cellular Density; ADC = Apparent Diffusion Coefficient; LALGLE = Large Area Low Gray Level Emphasis; IMOC = Information Measure Of Correlation; ASM = Angular Second Moment; LDLGLE = Large Dependence Low Gray Level Emphasis; LDHGLE = Large Dependence High Gray Level Emphasis; SDLGLE = Small Dependence Low Gray Level Emphasis; LALGLE = Large Area Low Gray Level Emphasis; GLNUN = Gray level non uniformity normalized; glcm = gray level co-occurrence matrix; gldm = Gray Level Dependence Matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix. (Continued)

firstorder Maximum	CD100µm ngtdm Strength	0.516	1.52×10^{-2}	1.23×10^{2}
glcm MaximumProbability	CD150µm glszm LALGLE	0.516	1.52×10^{-2}	1.22×10^{2}
glcm MaximumProbability	CD100µm glrlm RunVariance	0.515	1.53×10^{-2}	1.20×10^{2}
firstorder MAD	CD50µm ngtdm Complexity	0.515	1.53×10^{-2}	1.19×10^{2}
firstorder IR	CD50µm ngtdm Complexity	0.515	1.53×10^{-2}	1.18×10^{2}
glcm ClusterShade	MEAN Haralick IMOC2 F12	-0.514	1.57×10^{-2}	1.14×10^{2}
ngtdm Coarseness	CD200µm glcm JointEnergy	0.512	1.62×10^{-2}	1.10×10^{2}
firstorder Range	CD50µm glcm Contrast	0.511	1.66×10^{-2}	1.07×10^{2}
ngtdm Complexity	CD100µm glcm Contrast	0.51	1.70×10^{-2}	1.03×10^{2}
firstorder TotalEnergy	CD100µm glcm Contrast	0.51	1.70×10^{-2}	1.02×10^{2}
firstorder Minimum	MEAN HaralickDV F9	0.51	1.71×10^{-2}	1.01×10^{2}
glcm DifferenceVariance	MEAN HaralickDV F9	0.509	1.73×10^{-2}	9.89×10^{1}
glcm DifferenceAverage	MEAN Haralick IMOC2 F12	-0.508	1.73×10^{-2}	9.72×10^{1}
gldm SDLGLE	MEAN Haralick ASM F0	0.508	1.73×10^{-2}	9.71×10^{1}
firstorder MAD	MEAN Haralick ASM F0	0.508	1.75×10^{-2}	9.50×10^{1}
glcm ClusterShade	CD50µm glcm DV	0.507	1.76×10^{-2}	9.30×10^{1}
firstorder 90Percentile	CD150µm ngtdm Strength	0.507	1.76×10^{-2}	9.28×10^{1}
ngtdm Contrast	CD150µm ngtdm Strength	0.507	1.77×10^{-2}	9.17×10^{1}
ngtdm Complexity	CD150µm glcm DA	0.502	2.02×10^{-2}	8.05×10^{1}
firstorder Maximum	CD50µm glcm DV	0.501	2.06×10^{-2}	7.84×10^{1}
glszm ZoneVariance	CD200µm glrlm RunVariance	0.5	2.13×10^{-2}	7.54×10^{1}
firstorder 10Percentile	MEAN HaralickDV F9	0.5	2.13×10^{-2}	7.47×10^{1}
ngtdm Contrast	CD200µm glcm ClusterShade	0.498	2.23×10^{-2}	7.09×10^{1}
gldm LDLGLE	CD200µm glrlm RunVariance	0.497	2.25×10^{-2}	6.94×10^{1}
firstorder Minimum	MEAN HaralickEntropy F8	-0.497	2.25×10^{-2}	6.90×10^{1}
firstorder 10Percentile	MEAN HaralickEntropy F8	-0.497	2.25×10^{-2}	6.84×10^{1}
glszm LALGLE	CD200µm glrlm RunVariance	0.497	2.25×10^{-2}	6.81×10^{1}
firstorder Median	MEAN HaralickEntropy F8	-0.497	2.25×10^{-2}	6.76×10^{1}
ngtdm Busyness	CD200µm glrlm RunVariance	0.496	2.29×10^{-2}	6.59×10^{1}
glcm ClusterShade	CD50µm ngtdm Complexity	0.495	2.31×10^{-2}	6.50×10^{1}
glcm ClusterShade	CD50µm glcm Contrast	0.494	2.38×10^{-2}	6.28×10^{1}
glcm DifferenceVariance	CD150µm ngtdm Strength	0.494	2.38×10^{-2}	6.22×10^{1}
glcm ClusterProminence	CD150µm glcm DV	0.492	2.49×10^{-2}	5.92×10^{1}
firstorder Mean	MEAN HaralickEntropy F8	-0.492	2.50×10^{-2}	5.85×10^{1}
glcm DifferenceVariance	MEAN HaralickEntropy F8	-0.491	2.55×10^{-2}	5.71×10^{1}
glcm ClusterProminence	CD150µm glcm Contrast	0.49	2.56×10^{-2}	5.64×10^{1}
firstorder Median	MEAN HaralickDV F9	0.49	2.57×10^{-2}	5.59×10^{1}
firstorder TotalEnergy	CD150µm glcm DA	0.489	2.64×10^{-2}	5.40×10^{1}
ngtdm Coarseness	CD150µm glcm JointEnergy	0.488	2.67×10^{-2}	5.31×10^{1}
firstorder Mean	MEAN HaralickDV F9	0.488	2.69×10^{-2}	5.25×10^{1}
firstorder IR	CD50µm glcm Contrast	0.487	2.71×10^{-2}	5.12×10^{1}

Continued on next page

Table B.2: Summary of the significantly moderately correlated radiomic-pathomic features, with radiomic features extracted from ADC. Abbreviations: CD = Cellular Density; ADC = Apparent Diffusion Coefficient; LALGLE = Large Area Low Gray Level Emphasis; IMOC = Information Measure Of Correlation; ASM = Angular Second Moment; LDLGLE = Large Dependence Low Gray Level Emphasis; LDHGLE = Large Dependence High Gray Level Emphasis; SDLGLE = Small Dependence Low Gray Level Emphasis; LALGLE = Large Area Low Gray Level Emphasis; GLNUN = Gray level non uniformity normalized; glcm = gray level co-occurrence matrix; gldm = Gray Level Dependence Matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix. (Continued)

gldm DependenceVariance	CD200µm glrlm RunVariance	0.487	2.71×10^{-2}	5.10×10^{1}
firstorder Maximum	CD50µm glcm Contrast	0.487	2.71×10^{-2}	5.10×10^{1}
glcm ClusterProminence	MEAN HaralickEntropy F8	-0.486	2.73×10^{-2}	5.03×10^{1}
glcm DifferenceVariance	CD200µm glcm ClusterShade	0.486	2.76×10^{-2}	4.95×10^{1}
firstorder Range	CD50µm glcm DV	0.485	2.80×10^{-2}	4.84×10^{1}
firstorder IR	CD50µm glcm DV	0.484	2.83×10^{-2}	4.76×10^{1}
ngtdm Complexity	CD100µm glcm DV	0.484	2.85×10^{-2}	4.70×10^{1}
firstorder TotalEnergy	CD50µm ngtdm Complexity	0.482	3.03×10^{-2}	4.42×10^{1}
glcm ClusterProminence	CD150µm ngtdm Strength	0.481	3.08×10^{-2}	4.29×10^{1}
glcm Imc2	CD200µm glrlm RunVariance	-0.481	3.08×10^{-2}	4.29×10^{1}
firstorder IR	MEAN Haralick ASM F0	0.48	3.11×10^{-2}	4.22×10^{1}
firstorder MAD	CD50µm glcm DV	0.48	3.12×10^{-2}	4.19×10^{1}
firstorder MAD	CD50µm glcm Contrast	0.479	3.16×10^{-2}	4.11×10^{1}
firstorder 90Percentile	MEAN HaralickEntropy F8	-0.478	3.25×10^{-2}	3.98×10^{1}
firstorder 90Percentile	CD50µm glcm DV	0.475	3.54×10^{-2}	3.66×10^{1}
firstorder 90Percentile	MEAN HaralickDV F9	0.475	3.55×10^{-2}	3.62×10^{1}
firstorder Median	CD50µm glcm DV	0.474	3.60×10^{-2}	3.56×10^{1}
firstorder TotalEnergy	CD100µm glcm DV	0.474	3.62×10^{-2}	3.51×10^{1}
gldm LDHGLE	CD50µm glcm Contrast	0.473	3.65×10^{-2}	3.47×10^{1}
firstorder Mean	CD50µm glcm DV	0.472	3.81×10^{-2}	3.31×10^{1}
firstorder 10Percentile	CD200µm glcm ClusterShade	0.47	3.94×10^{-2}	3.19×10^{1}
ngtdm Contrast	MEAN HaralickEntropy F8	-0.47	3.97×10^{-2}	3.15×10^{1}
ngtdm Complexity	CD150µm glcm DV	0.468	4.09×10^{-2}	3.05×10^{1}
glcm DifferenceAverage	CD50µm glcm DV	0.468	4.12×10^{-2}	3.01×10^{1}
ngtdm Strength	MEAN Haralick IMOC2 F12	-0.467	4.15×10^{-2}	2.95×10^{1}
glcm ClusterProminence	CD50µm ngtdm Complexity	0.467	4.15×10^{-2}	2.95×10^{1}
firstorder 10Percentile	CD50µm ngtdm Strength	0.467	4.15×10^{-2}	2.93×10^{1}
glcm ClusterShade	CD100µm ngtdm Strength	0.467	4.15×10^{-2}	2.92×10^{1}
gldm LDHGLE	CD200µm ngtdm Strength	0.466	4.24×10^{-2}	2.85×10^{1}
firstorder Mean	CD200µm glcm ClusterShade	0.465	4.35×10^{-2}	2.75×10^{1}
glcm DifferenceAverage	CD50µm ngtdm Complexity	0.465	4.35×10^{-2}	2.75×10^{1}
glcm ClusterProminence	CD100µm glcm DV	0.464	4.35×10^{-2}	2.73×10^{1}
firstorder Range	CD50µm glcm DA	0.463	4.44×10^{-2}	2.65×10^{1}
glrlm RunVariance	CD200µm glrlm RunVariance	0.463	4.44×10^{-2}	2.65×10^{1}
gldm LDHGLE	MEAN Haralick ASM F0	0.463	4.44×10^{-2}	2.63×10^{1}
firstorder 90Percentile	CD200µm glcm ClusterShade	0.463	4.44×10^{-2}	2.62×10^{1}
firstorder Median	CD200µm glcm ClusterShade	0.463	4.45×10^{-2}	2.60×10^{1}
firstorder TotalEnergy	CD100µm glcm DA	0.462	4.56×10^{-2}	2.53×10^{1}
glcm DifferenceAverage	CD100µm ngtdm Strength	0.461	4.67×10^{-2}	2.46×10^{1}
ngtdm Coarseness	CD50µm glcm Idmn	-0.459	4.78×10^{-2}	2.39×10^{1}
glcm JointAverage	CD50µm glcm Contrast	0.459	4.78×10^{-2}	2.36×10^{1}

Continued on next page

Table B.2: Summary of the significantly moderately correlated radiomic-pathomic features, with radiomic features extracted from ADC. Abbreviations: CD = Cellular Density; ADC = Apparent Diffusion Coefficient; LALGLE = Large Area Low Gray Level Emphasis; IMOC = Information Measure Of Correlation; ASM = Angular Second Moment; LDLGLE = Large Dependence Low Gray Level Emphasis; LDHGLE = Large Dependence High Gray Level Emphasis; SDLGLE = Small Dependence Low Gray Level Emphasis; LALGLE = Large Area Low Gray Level Emphasis; GLNUN = Gray level non uniformity normalized; glcm = gray level co-occurrence matrix; gldm = Gray Level Dependence Matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix. (Continued)

ngtdm Complexity	MEAN Cytoplasm EosinODMin	0.459	4.78×10^{-2}	2.36×10^{1}
firstorder 10Percentile	MEAN HaralickSumEntropy F7	-0.458	4.78×10^{-2}	2.33×10^{1}
glcm ClusterShade	MEAN HaralickEntropy F8	-0.458	4.78×10^{-2}	2.33×10^{1}
glcm MaximumProbability	MEAN Haralick IMOC2 F12	-0.458	4.78×10^{-2}	2.33×10^{1}
firstorder Median	MEAN HaralickSumEntropy F7	-0.458	4.78×10^{-2}	2.32×10^{1}
glcm SumEntropy	MEAN Haralick IMOC2 F12	0.458	4.78×10^{-2}	2.31×10^{1}
firstorder Minimum	MEAN HaralickSumEntropy F7	-0.458	4.79×10^{-2}	2.29×10^{1}
firstorder Maximum	CD150µm ngtdm Strength	0.457	4.81×10^{-2}	2.27×10^{1}
firstorder Minimum	CD50µm ngtdm Strength	0.457	4.81×10^{-2}	2.27×10^{1}
glcm DifferenceAverage	CD50µm glcm Contrast	0.457	4.85×10^{-2}	2.24×10^{1}
glcm ClusterProminence	MEAN HaralickDV F9	0.456	4.94×10^{-2}	2.17×10^{1}
firstorder Maximum	CD50µm ngtdm Complexity	0.456	4.94×10^{-2}	2.17×10^{1}
glcm ClusterShade	MEAN HaralickDV F9	0.456	4.94×10^{-2}	2.17×10^{1}

Table B.3: Summary of the significantly moderately correlated radiomic-pathomic features, with radiomic features extracted from T1C. Abbreviations: DNU = Dependence non uniformity; GLNU = gray-level non-uniformity; IMOC = Information Measure Of Correlation; ASM = Angular Second Moment; SRLGLE = Short Run Low Gray Level Emphasis; SDLGLE = Small Dependence Low Gray Level Emphasis; T1C = post–contrast T1; SALGLE = Small Area Low Gray Level Emphasis; LALGLE = Large Area Low Gray Level Emphasis; glcm = gray level co-occurrence matrix; gldm = Gray Level Dependence Matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix.

Radiomic feature Name (T1C)	**Pathomic feature name**	**ρ**	**FDR q**	**BF**
glszm SALGLE	CD200µm ngtdm Strength	0.599	4.44×10^{-3}	2.59×10^{3}
glcm JointEnergy	MEAN Haralick ASM F0	0.597	4.66×10^{-3}	2.33×10^{3}
firstorder RootMeanSquared	CD150µm gldm DNU	-0.594	4.91×10^{-3}	2.11×10^{3}
glszm SALGLE	MEAN Haralick IMOC2 F12	-0.591	5.27×10^{-3}	1.86×10^{3}
glcm MaximumProbability	MEAN Haralick IMOC2 F12	-0.59	5.27×10^{-3}	1.78×10^{3}
glszm SALGLE	MEAN Haralick ASM F0	0.586	5.63×10^{-3}	1.54×10^{3}
firstorder 10Percentile	CD150µm gldm DNU	-0.586	5.63×10^{-3}	1.53×10^{3}
firstorder 90Percentile	CD150µm glcm Imc2	0.578	7.49×10^{-3}	1.11×10^{3}
firstorder 10Percentile	CD200µm glcm Imc2	0.577	7.49×10^{-3}	1.05×10^{3}
gldm SDLGLE	MEAN Haralick IMOC2 F12	-0.576	7.49×10^{-3}	1.03×10^{3}
gldm SDLGLE	MEAN Haralick ASM F0	0.576	7.49×10^{-3}	9.98×10^{2}
glcm MaximumProbability	CD200µm ngtdm Strength	0.572	8.45×10^{-3}	8.59×10^{2}
glrlm SRLGLE	CD200µm ngtdm Strength	0.57	8.52×10^{-3}	8.15×10^{2}
firstorder 10Percentile	CD100µm ngtdm Busyness	-0.57	8.52×10^{-3}	7.97×10^{2}
firstorder RootMeanSquared	CD100µm ngtdm Busyness	-0.569	8.56×10^{-3}	7.69×10^{2}
firstorder 10Percentile	CD150µm glcm Imc2	0.562	1.01×10^{-2}	6.00×10^{2}
firstorder 90Percentile	CD150µm gldm DNU	-0.562	1.01×10^{-2}	5.96×10^{2}
glrlm SRLGLE	MEAN Haralick IMOC2 F12	-0.561	1.01×10^{-2}	5.78×10^{2}

Continued on next page

Table B.3: Summary of the significantly moderately correlated radiomic-pathomic features, with radiomic features extracted from T1C. Abbreviations: DNU = Dependence non uniformity; GLNU = gray-level non-uniformity; IMOC = Information Measure Of Correlation; ASM = Angular Second Moment; SRLGLE = Short Run Low Gray Level Emphasis; SDLGLE = Small Dependence Low Gray Level Emphasis; T1C = post–contrast T1; SALGLE = Small Area Low Gray Level Emphasis; LALGLE = Large Area Low Gray Level Emphasis; glcm = gray level co-occurrence matrix; gldm = Gray Level Dependence Matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix. (Continued)

firstorder 10Percentile	CD200µm ngtdm Strength	0.561	1.01×10^{-2}	5.78×10^{2}
firstorder Minimum	CD200µm ngtdm Strength	0.561	1.01×10^{-2}	5.70×10^{2}
glrlm SRLGLE	MEAN Haralick ASM F0	0.554	1.26×10^{-2}	4.49×10^{2}
glcm MaximumProbability	MEAN Haralick ASM F0	0.545	1.76×10^{-2}	3.20×10^{2}
firstorder Kurtosis	CD150µm gldm DNU	0.539	2.09×10^{-2}	2.66×10^{2}
ngtdm Coarseness	CD100µm ngtdm Strength	0.537	2.21×10^{-2}	2.42×10^{2}
glcm Idn	CD200µm ngtdm Strength	-0.536	2.21×10^{-2}	2.40×10^{2}
firstorder RootMeanSquared	CD200µm firstorder TE	-0.533	2.40×10^{-2}	2.14×10^{2}
firstorder 10Percentile	CD200µm firstorder TE	-0.533	2.40×10^{-2}	2.13×10^{2}
glszm SALGLE	CD100µm ngtdm Strength	0.531	2.50×10^{-2}	2.00×10^{2}
gldm SDLGLE	CD100µm ngtdm Strength	0.529	2.62×10^{-2}	1.87×10^{2}
glszm LALGLE	CD200µm glcm Imc2	-0.525	2.91×10^{-2}	1.67×10^{2}
firstorder 90Percentile	CD100µm ngtdm Busyness	-0.523	3.09×10^{-2}	1.54×10^{2}
glcm InverseVariance	CD150µm glszm GLNU	0.522	3.10×10^{-2}	1.48×10^{2}
glcm InverseVariance	CD200µm glcm Imc2	-0.522	3.10×10^{-2}	1.48×10^{2}
firstorder Kurtosis	CD150µm glszm GLNU	0.516	3.64×10^{-2}	1.25×10^{2}
glcm JointEnergy	CD100µm ngtdm Strength	0.514	3.87×10^{-2}	1.16×10^{2}
glcm InverseVariance	CD50µm ngtdm Busyness	0.51	4.32×10^{-2}	1.03×10^{2}
glszm LALGLE	CD150µm glcm Imc2	-0.508	4.59×10^{-2}	9.51×10^{1}
gldm DependenceVariance	CD150µm glszm GLNU	0.506	4.70×10^{-2}	9.13×10^{1}

B.2.3 Factor analysis

Table B.4: ADC factor loadings for the five ADC factors. A darker grey in the column cell correspond to a higher loadings measuring the association between features and factors. Abbreviations: ADC = Apparent Diffusion Coefficient; LALGLE = Large Area Low Gray Level Emphasis; LDLGLE = Large Dependence Low Gray Level Emphasis; LDHGLE = Large Dependence High Gray Level Emphasis; SDLGLE = Small Dependence Low Gray Level Emphasis; LALGLE = Large Area Low Gray Level Emphasis; GLNUN = Gray level non uniformity normalized; SZNU = Size Zone Non-Uniformity; glcm = gray level co-occurrence matrix; gldm = Gray Level Dependence Matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix; F = Factor.

Radiomic feature name (ADC)	***ADC_F1***	***ADC_F2***	***ADC_F3***	***ADC_F4***	***ADC_F5***
firstorder MAD	0.857	-0.106	-0.114	0.241	-0.016
firstorder Maximum	0.902	-0.088	-0.029	-0.004	0.002
firstorder Minimum	0.814	-0.03	-0.137	-0.309	-0.068
firstorder Kurtosis	-0.027	-0.006	0.48	-0.09	0.057
firstorder 10Percentile	0.86	-0.057	-0.065	-0.251	-0.051
firstorder IR	0.836	-0.112	-0.144	0.286	0.013
firstorder Median	0.876	-0.068	-0.076	-0.184	-0.027
firstorder 90Percentile	0.892	-0.076	-0.085	-0.102	-0.046
firstorder Skewness	0.121	0.09	0.243	0.141	-0.116
firstorder Range	0.82	-0.13	0.083	0.307	0.073
firstorder Mean	0.881	-0.069	-0.076	-0.17	-0.035
firstorder TotalEnergy	0.554	-0.042	0.194	0.331	0.491

Continued on next page

Table B.4: ADC factor loadings for the five ADC factors. A darker grey in the column cell correspond to a higher loadings measuring the association between features and factors. Abbreviations: ADC = Apparent Diffusion Coefficient; LALGLE = Large Area Low Gray Level Emphasis; LDLGLE = Large Dependence Low Gray Level Emphasis; LDHGLE = Large Dependence High Gray Level Emphasis; SDLGLE = Small Dependence Low Gray Level Emphasis; LALGLE = Large Area Low Gray Level Emphasis; GLNUN = Gray level non uniformity normalized; SZNU = Size Zone Non-Uniformity; glcm = gray level co-occurrence matrix; gldm = Gray Level Dependence Matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix; F = Factor. (Continued)

gldm LDLGLE	-0.002	0.898	-0.099	-0.012	-0.034
gldm DepEntropy	-0.008	-0.217	0.49	0.721	0.068
gldm GLNU	-0.148	0.637	0.559	-0.135	-0.024
gldm LDHGLE	0.814	-0.066	0.089	0.131	0.103
gldm DV	-0.16	0.867	0.11	-0.079	-0.006
gldm SDLGLE	0.501	0.202	-0.497	-0.424	-0.16
ngtdm Busyness	-0.043	0.899	-0.054	-0.033	-0.027
ngtdm Strength	0.845	-0.068	-0.184	0.161	0.015
ngtdm Coarseness	-0.263	-0.027	-0.677	-0.292	0.157
ngtdm Complexity	0.588	-0.022	-0.034	0.393	0.559
ngtdm Contrast	0.777	-0.007	-0.087	-0.369	-0.273
glrlm RLNU	-0.038	-0.025	0.835	0.178	-0.058
glrlm RunVariance	-0.206	0.853	0.107	-0.135	0.025
glcm Imc2	0.041	-0.901	-0.021	0.028	0.017
glcm DiffVar	0.873	-0.019	-0.109	-0.184	-0.113
glcm InvVar	-0.324	0.783	0.031	-0.269	0.083
glcm Imc1	-0.382	0.346	0.678	-0.236	0.03
glcm JointEntropy	-0.198	-0.2	0.774	0.355	0.029
glcm ClusterShade	0.874	-0.014	-0.074	0.039	0.138
glcm Correlation	-0.402	0.056	0.327	0.297	0.089
glcm DiffEntropy	0.136	-0.451	0.198	0.729	0.068
glcm SumEntropy	-0.084	-0.338	0.491	0.65	0.066
glcm ClusterTendency	0.859	-0.038	-0.104	0.192	0.165
glcm JointAverage	0.71	-0.157	0.112	0.326	0.088
glcm MaxProbability	0.25	0.629	-0.453	-0.344	-0.081
glcm ClusterProminence	0.878	-0.009	-0.089	-0.03	0.2
glcm Idn	-0.241	-0.016	0.757	0.106	0.078
glcm DiffAverage	0.882	-0.092	-0.145	0.095	-0.049
glcm Idmn	-0.387	-0.01	0.71	0.128	0.161
glszm ZoneVariance	-0.061	0.905	-0.006	-0.014	-0.019
glszm ZoneEntropy	0.038	-0.302	0.413	0.74	0.067
glszm LALGLE	-0.006	0.899	-0.092	-0.005	-0.033
glszm GLNUN	-0.046	0.806	-0.161	-0.364	0.027
glszm SZNU	0.015	-0.116	0.774	0.294	-0.048

Table B.5: T1C factor loadings for the 8 T1C factors. A darker grey in the column cell correspond to a higher loading measuring the association between features and factors. Abbreviations: T1C = post-contrast T1; F = Factor; glcm = gray level co-occurrence matrix; gldm = Gray Level Dependence Matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix.

Radiomic feature name (T1C)	***T1C_F1***	***T1C_F2***	***T1C_F3***	***T1C_F4***	***T1C_F5***	***T1C_F6***	***T1C_F7***	***T1C_F8***
glrlm RunPercentage	0.432	-0.104	-0.241	-0.747	0.158	-0.166	-0.251	-0.069
glrlm GLNU	-0.25	0.226	0.746	0.481	-0.093	0.071	-0.06	0.056
glrlm RLNU	0.044	0.29	0.897	0.204	-0.045	0.004	-0.004	0.053
glrlm RunEntropy	0.776	0.522	0.079	-0.166	0.015	-0.148	-0.01	-0.009
glrlm SRLGLE	-0.171	-0.856	-0.291	0.026	0.061	-0.148	0.11	-0.184

Continued on next page

Table B.5: T1C factor loadings for the 8 T1C factors. A darker grey in the column cell correspond to a higher loading measuring the association between features and factors. Abbreviations: T1C = post-contrast T1; F = Factor; glcm = gray level co-occurrence matrix; gldm = Gray Level Dependence Matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix. (Continued)

gldm SmallDLGLE	-0.102	-0.87	-0.287	-0.07	0.073	-0.17	0.028	-0.13
gldm DependenceEntropy	0.327	0.704	0.416	0.275	-0.094	0.002	0.262	0.041
gldm SmallDHGLE	0.825	0.19	0.157	-0.239	0.145	0.073	-0.085	0.264
gldm DNU	0.278	0.289	0.875	-0.088	-0.004	0.044	-0.045	0.011
gldm LargeDLGLE	-0.322	-0.356	-0.138	0.673	-0.134	0.045	0.269	-0.268
gldm LargeDHGLE	0.285	0.354	0.544	0.185	0.028	-0.087	0.156	0.588
gldm DependenceVariance	-0.385	0.087	0.183	0.781	-0.154	0.208	0.215	0.055
glcm SumSquares	0.936	0.156	-0.09	-0.118	0.049	-0.045	0.046	-0.009
glcm JointAverage	0.675	0.336	0.296	-0.254	0.122	-0.067	-0.079	0.43
glcm Idn	-0.044	0.569	0.454	0.248	-0.203	0.434	0.206	0.149
glcm Imc2	0.402	-0.166	-0.673	-0.376	0.077	-0.044	-0.027	-0.071
glcm ClusterShade	0.551	0.03	0.033	-0.056	-0.031	0.497	-0.04	-0.044
glcm Correlation	0.081	0.495	0.173	0.336	-0.202	-0.005	0.456	0.021
glcm Imc1	-0.251	0.479	0.684	0.213	-0.055	0.113	0.257	0.061
glcm ClusterTendency	0.917	0.172	-0.053	-0.078	0.034	-0.061	0.152	0
glcm InverseVariance	-0.599	-0.04	0.147	0.57	-0.182	0.275	0.32	0.041
glcm SumEntropy	0.649	0.655	0.191	-0.122	-0.018	-0.103	0.13	-0.033
glcm ClusterProminence	0.902	0.063	0.022	0.009	0.066	0.155	0.163	0.039
glcm MaximumProbability	-0.299	-0.858	-0.253	0.136	0.025	0.047	-0.022	0.067
glcm DifferenceVariance	0.87	0.114	-0.118	-0.17	0.035	0.083	-0.166	-0.02
glcm JointEnergy	-0.207	-0.912	-0.181	0.06	0.076	0.005	-0.025	0.055
glcm DifferenceEntropy	0.745	0.37	-0.035	-0.391	0.105	-0.147	-0.193	-0.039
glcm JointEntropy	0.385	0.657	0.542	-0.058	0.001	-0.045	0.179	0.024
firstorder RMS	0.428	-0.048	0.021	-0.282	0.797	-0.19	-0.045	0.091
firstorder Kurtosis	-0.128	0.091	0.148	0.131	-0.161	0.793	0.013	0.131
firstorder Maximum	0.757	0.163	0.143	-0.282	0.419	0.169	-0.08	-0.017
firstorder 10Percentile	0.111	-0.151	0.081	-0.234	0.902	-0.105	-0.065	0.11
firstorder Skewness	0.186	0.16	-0.098	0.12	-0.231	0.689	0.015	-0.212
firstorder Range	0.784	0.331	0.247	-0.224	0.01	0.257	-0.087	0.106
firstorder IR	0.84	0.235	-0.163	-0.201	0.022	-0.289	0.059	-0.018
firstorder 90Percentile	0.61	0.024	-0.018	-0.288	0.663	-0.205	-0.044	0.056
firstorder Minimum	-0.117	-0.354	-0.223	-0.093	0.799	-0.188	0.022	-0.253
firstorder TotalEnergy	0.169	0.299	0.855	0.101	0.201	-0.03	0.069	0.11
ngtdm Complexity	0.879	0.137	0.142	-0.122	0.07	0.212	-0.03	0.044
ngtdm Strength	0.551	-0.223	-0.598	-0.167	-0.055	0.014	-0.171	0.04
ngtdm Coarseness	-0.095	-0.899	-0.282	-0.06	0.075	-0.115	0.002	0.046
ngtdm Busyness	-0.445	0.195	0.539	0.535	-0.088	-0.012	0.019	-0.231
ngtdm Contrast	0.555	-0.193	-0.511	-0.228	0.092	-0.314	-0.262	-0.069
glszm ZoneVariance	-0.173	0.057	0.141	0.889	-0.153	0.017	-0.17	0.007
glszm SZNU	0.323	0.28	0.845	-0.175	0.007	0.049	-0.045	0.002
glszm SALGLE	-0.134	-0.879	-0.258	-0.022	0.05	-0.152	0.071	-0.116
glszm ZoneEntropy	0.407	0.715	0.397	0.218	-0.088	-0.01	0.176	0.053
glszm LALGLE	-0.252	-0.181	-0.008	0.852	-0.172	0.046	-0.003	-0.122
glszm GLNUN	-0.691	-0.566	-0.008	0.237	-0.009	0.214	0.133	0.022
glszm LAHGLE	0.012	0.215	0.397	0.759	-0.065	-0.063	-0.104	0.3

Continued on next page

Table B.5: T1C factor loadings for the 8 T1C factors. A darker grey in the column cell correspond to a higher loading measuring the association between features and factors. Abbreviations: T1C = post-contrast T1; F = Factor; glcm = gray level co-occurrence matrix; gldm = Gray Level Dependence Matrix; glszm = Gray Level Size Zone Matrix; ngtdm = Neighbouring Gray Tone Difference Matrix; glrlm = Gray Level Run Length Matrix. (Continued)

glszm GLV	0.939	0.178	-0.052	-0.11	0.065	-0.016	0.021	0.032
glszm GLNU	-0.149	0.276	0.899	0.1	-0.066	0.066	-0.024	0.031
glszm SZNUN	0.507	-0.177	-0.325	-0.552	0.17	-0.095	-0.404	-0.116

B.3 An automated pathomic-based approach for the estimation of prostate cancer grade on digital pathology images

Table B.6. Selected features for the "Malignant vs Non-Malignant" classification task at 20× magnification level.

Malignant vs Non-Malignant
R firstorder Maximum
R firstorder Median
R wav LL firstorder StandardDeviation
R wav HH firstorder IR
R wav HH glcm Correlation
R wav HH glcm IMCII
R lbpU histfunc Var58
G wav HH firstorder Minimum
G wav HH firstorder Maximum
G wav HH firstorder RootMeanSquared
B firstorder NinetiethPercentile
B wav LL firstorder IR
B lbpU histfunc Var50
B lbpU histfunc Var52
H wav HL glcm IMCII
E wav LL firstorder Skewness
E wav LH glcm IMCII
E wav HL glcm IMCII
E lbpU histfunc Var54
E lbpU histfunc Var56
E lbpU histfunc Var58

E lbpU histfunc Var59
Hu firstorder NinetiethPercentile
Hu glcm AngularSecondMoment Energy
Hu glcm Correlation
Hu glcm Variance glcm
Hu wav HH glcm SumVariance
Hu lbpU histfunc Var2
Hu lbpU histfunc Var5
S firstorder NinetiethPercentile
S wav LL firstorder Skewness
S wav HH firstorder Kurtosis
S lbpU histfunc Var52
S lbpU histfunc Var54
S lbpU histfunc Var56
V wav LL firstorder Maximum
V wav HL firstorder Range
V lbpU histfunc Var34

Table B.7. Selected features for the "Gleason 3 vs Non-Malignant" classification task at 20× magnification level. G3 = Gleason 3.

G3 vs Non-Malignant
R firstorder Maximum
R firstorder Median
R wav LL firstorder StandardDeviation
R wav HH firstorder IR
R wav HL glcm MCC
R wav HH glcm IMCII
R lbpU histfunc Var19
R lbpU histfunc Var25
R lbpU histfunc Var58
G wav LL firstorder Range
G wav HH firstorder RootMeanSquared
G wav HL glcm SumEntropy

G lbpU histfunc Var3
G lbpU histfunc Var9
B firstorder Maximum
B glcm MCC
B wav LL firstorder Maximum
B wav LL firstorder IR
B lbpU histfunc Var9
B lbpU histfunc Var50
B lbpU histfunc Var52
H glcm IMCII
H wav HL glcm IMCII
E wav LL firstorder Skewness
E wav LH glcm IMCII
E wav HL glcm IMCII
E lbpU histfunc Var54
E lbpU histfunc Var56
E lbpU histfunc Var59
Hu firstorder NinetiethPercentile
Hu glcm AngularSecondMoment Energy
Hu glcm Correlation
Hu glcm Variance glcm
Hu glcm IMCI
Hu wav HH glcm SumVariance
Hu lbpU histfunc Var5
Hu lbpU histfunc Var40
S firstorder NinetiethPercentile
S wav LL glcm MCC
S wav HH glcm DifferenceVariance
S lbpU histfunc Var54
S lbpU histfunc Var56
V firstorder Maximum
V wav LL firstorder Maximum
V wav HL firstorder Maximum
V wav HL firstorder Range

V lbpU histfunc Var34

Table B.8. Selected features for the "Gleason 4 + Gleason 5 vs Gleason 3" at 20× magnification level. G3 = Gleason 3; G4 = Gleason 4; G5 = Gleason 5.

G4+G5 vs G3
Gray lbpU histfunc Var52
Gray lbpU histfunc Var56
R wav HL firstorder Minimum
R wav HL firstorder Maximum
R wav HL glcm DifferenceEntropy
R wav HL glcm MCC
R lbpU histfunc Var36
G wav HL firstorder Maximum
G lbpU histfunc Var7
G lbpU histfunc Var9
G lbpU histfunc Var52
B firstorder NinetiethPercentile
B wav LL firstorder TenthPercentile
B wav LH firstorder Minimum
B wav LH firstorder Maximum
B wav HL firstorder Minimum
B wav HH firstorder Variance
H glcm AngularSecondMoment Energy
H wav LL firstorder Median
H wav LL firstorder Skewness
H wav LL firstorder IR
H wav LL glcm MCC
H wav LH glcm MCC
H wav HL glcm DifferenceVariance
H wav HL glcm IMCII
E wav LL firstorder Skewness
E wav LL glcm IMCII

E wav LH glcm Variance glcm
E wav HL glcm IMCII
E lbpU histfunc Var50
Hu firstorder Range
Hu glcm SumVariance
Hu glcm DifferenceEntropy
Hu wav LH firstorder Minimum
Hu wav LH firstorder Maximum
Hu wav HL firstorder Minimum
Hu wav HH firstorder Minimum
Hu wav HH firstorder Maximum
Hu wav HH firstorder Kurtosis
Hu lbpU histfunc Var3
Hu lbpU histfunc Var10
Hu lbpU histfunc Var34
S wav LL firstorder Uniformity
S wav LL glcm SumEntropy
S wav LL glcm DifferenceVariance
S wav LH glcm IMCII
S wav HL glcm DifferenceVariance
S wav HL glcm IMCII
S lbpU histfunc Var30
V wav LL firstorder Variance
V wav HH firstorder Variance

Table B.9. Selected features for the "Gleason 4 + Gleason 5 vs Non-Malignant" classification task at 20× magnification level. G4 = Gleason 4; G5 = Gleason 5.

G4+G5 vs Non-Malignant
R firstorder Maximum
R firstorder Median
R glcm Correlation
R wav HH firstorder IR

R wav HH glcm IMCII
R lbpU histfunc Var19
R lbpU histfunc Var25
R lbpU histfunc Var58
G wav HH firstorder Minimum
G wav HH firstorder Maximum
G lbpU histfunc Var9
B glcm MCC
B wav LL firstorder TenthPercentile
B wav LL firstorder IR
B wav HL firstorder IR
B wav HL firstorder RootMeanSquared
B lbpU histfunc Var9
B lbpU histfunc Var16
H wav LL firstorder IR
H wav LH firstorder Uniformity
H wav LH glcm IMCII
H wav HL glcm IMCII
E wav LL firstorder Maximum
E wav LL firstorder Skewness
E wav LH glcm IMCII
E wav HL glcm IMCII
E lbpU histfunc Var54
E lbpU histfunc Var56
E lbpU histfunc Var59
Hu firstorder NinetiethPercentile
Hu firstorder Entropy
Hu glcm AngularSecondMoment Energy
Hu glcm Correlation
Hu glcm Variance glcm
Hu lbpU histfunc Var5
S firstorder NinetiethPercentile
S wav LL firstorder Range
S wav LL firstorder Uniformity

S wav LH glcm IMCII
S lbpU histfunc Var54
S lbpU histfunc Var56
V wav LL firstorder Maximum
V wav LL firstorder Variance
V wav HL firstorder Range

Table B.10. Selected features for the "Malignant vs Non-Malignant" classification task at 10× magnification level.

Malignant vs Non-Malignant
Gray firstorder Kurtosis
Gray wav LH firstorder Minimum
Gray wav HL firstorder Minimum
Gray wav HL firstorder Maximum
R firstorder Median
R wav LL firstorder StandardDeviation
R wav HL glcm SumVariance
R wav HH glcm IMCI
G wav HH firstorder Range
B glcm MCC
B wav LL firstorder InterquartileRange
B wav HL glcm SumEntropy
B wav HH glcm SumEntropy
H firstorder Median
H wav LH glcm SumVariance
H wav HL glcm DifferenceVariance
H wav HL glcm IMCII
H wav HH glcm DifferenceVariance
E wav LL firstorder Maximum
E wav LH glcm Variance glcm
E wav LH glcm IMCII
E lbpU histfunc Var56
Hu firstorder Median

Hu firstorder NinetiethPercentile
Hu glcm Correlation
Hu glcm Variance glcm
Hu glcm Entropy glcm
Hu wav HH glcm SumVariance
Hu lbpU histfunc Var1
Hu lbpU histfunc Var24
Hu lbpU histfunc Var34
S firstorder NinetiethPercentile
S wav HL firstorder Kurtosis
S wav HH firstorder Variance
S wav HH firstorder Kurtosis
S wav LH glcm IMCII
S wav HL glcm MCC
V glcm IMCI
V wav HH firstorder Minimum
V wav HH firstorder Maximum
V lbpU histfunc Var26

Table B.11. Selected features for the "Gleason 3 vs Non-Malignant" classification task at 10× magnification level. G3 = Gleason 3.

G3 vs Non-Malignant
Gray wav LH firstorder Minimum
Gray wav HL firstorder Minimum
Gray wav HL firstorder Maximum
Gray wav HH firstorder Kurtosis
R firstorder Median
R wav LL firstorder StandardDeviation
R wav LH firstorder InterquartileRange
R wav LH glcm SumVariance
R wav HH glcm IMCI
G firstorder Range
G wav HH firstorder Skewness

G lbpU histfunc Var58
B firstorder InterquartileRange
B glcm MCC
B wav HL glcm Correlation
B wav HL glcm SumEntropy
B wav HH glcm SumEntropy
H glcm IMCI
H wav LL firstorder Minimum
H wav LH glcm SumVariance
H wav HL glcm DifferenceVariance
H wav HL glcm IMCII
H wav HH glcm DifferenceVariance
E wav LL firstorder Maximum
E wav LL glcm DifferenceVariance
E wav LH glcm Variance glcm
E wav LH glcm SumVariance
E wav LH glcm IMCII
E wav HH glcm DifferenceVariance
E lbpU histfunc Var50
Hu firstorder Median
Hu firstorder NinetiethPercentile
Hu glcm AngularSecondMoment Energy
Hu glcm Correlation
Hu glcm Variance glcm
Hu wav HH glcm SumVariance
Hu lbpU histfunc Var7
Hu lbpU histfunc Var9
Hu lbpU histfunc Var18
Hu lbpU histfunc Var24
S firstorder NinetiethPercentile
S wav HH firstorder Variance
S wav HH firstorder Kurtosis
S wav HH firstorder Range
S wav LH glcm IMCII

V glcm IMCI
V wav HH firstorder Minimum
V wav HH firstorder Maximum
V wav LH glcm MCC
V lbpU histfunc Var1
V lbpU histfunc Var26

Table B.12. Selected features for the "Gleason 4 + Gleason 5 vs Gleason 3" classification task at 10× magnification level. G3 = Gleason 3; G4 = Gleason 4; G5 = Gleason 5.

G4+G5 vs G3
Gray wav LH firstorder Skewness
R wav LH firstorder Minimum
R wav HL firstorder Minimum
R wav HL firstorder Maximum
R wav LH glcm SumEntropy
R wav HL glcm MCC
R wav HH glcm IMCI
G wav HL firstorder Minimum
G wav HL firstorder Maximum
G lbpU histfunc Var3
G lbpU histfunc Var9
G lbpU histfunc Var22
B firstorder Entropy
B glcm Variance glcm
B wav LL firstorder TenthPercentile
B wav HL glcm MCC
B lbpU histfunc Var56
H firstorder Median
H glcm MCC
H wav LL firstorder InterquartileRange
H wav HL firstorder Minimum
H wav LH glcm SumVariance

H wav HL glcm SumVariance
H wav HL glcm DifferenceVariance
E firstorder StandardDeviation
E firstorder NinetiethPercentile
E wav LL firstorder Maximum
E wav LL firstorder Entropy
E wav LL glcm MCC
E wav LH glcm MCC
E lbpU histfunc Var26
E lbpU histfunc Var27
E lbpU histfunc Var42
Hu glcm Variance glcm
Hu glcm Entropy glcm
Hu wav LH firstorder Minimum
Hu wav HL firstorder Maximum
Hu lbpU histfunc Var34
Hu lbpU histfunc Var38
Hu lbpU histfunc Var44
Hu lbpU histfunc Var50
Hu lbpU histfunc Var52
S wav LL firstorder NinetiethPercentile
S wav LH firstorder Maximum
S wav HL firstorder Minimum
S wav LH glcm Variance glcm
S wav HL glcm DifferenceVariance
V wav LH firstorder Skewness
V wav LL glcm MCC
V wav HH glcm InverseDifferentMoment
V lbpU histfunc Var34
V lbpU histfunc Var36
V lbpU histfunc Var41

Table B.13. Selected features for the "Gleason 4 + Gleason 5 vs Non-malignant" classification task at 10× magnification level.

G4+G5 vs Non-Malignant
Gray wav LH firstorder Minimum
Gray wav HL firstorder Minimum
Gray wav HL firstorder Maximum
R firstorder Median
R wav LL firstorder StandardDeviation
R wav HL glcm SumVariance
G wav HH firstorder Skewness
G wav HL glcm IMCI
G lbpU histfunc Var58
B firstorder NinetiethPercentile
B firstorder InterquartileRange
B glcm MCC
H wav LL firstorder Minimum
H wav LH glcm SumVariance
H wav HL glcm DifferenceVariance
H wav HL glcm IMCII
H wav HH glcm DifferenceVariance
E firstorder MeanAbsoluteDeviation
E wav LL firstorder Maximum
E wav LL glcm MCC
E wav LH glcm IMCII
E wav LH glcm MCC
E wav HH glcm DifferenceVariance
E lbpU histfunc Var56
E lbpU histfunc Var58
Hu firstorder Median
Hu firstorder NinetiethPercentile
Hu glcm AngularSecondMoment Energy
Hu glcm Correlation
Hu wav HH glcm SumVariance
Hu lbpU histfunc Var34
S firstorder NinetiethPercentile

S wav LL firstorder Skewness
S wav HL firstorder Kurtosis
S wav HH firstorder Variance
S wav HH firstorder Kurtosis
S wav LH glcm IMCII
S wav HL glcm DifferenceVariance
V glcm IMCI
V wav LL firstorder Maximum
V wav HH firstorder Minimum
V wav HH firstorder Maximum
V wav LH glcm IMCII
V lbpU histfunc Var58

Table B.14. Selected features for the "Malignant vs Non-malignant" classification task at 5× magnification level.

Malignant vs Non-Malignant
Gray glcm MCC
Gray wav LL firstorder Range
Gray wav LH firstorder Minimum
Gray wav LH firstorder Maximum
R firstorder Median
R firstorder Skewness
R wav LL firstorder StandardDeviation
R wav LH firstorder Minimum
R wav HL firstorder Minimum
R wav HH firstorder Minimum
R wav HH glcm SumVariance
G wav HH firstorder Range
G wav LH glcm DifferenceEntropy
G lbpU histfunc Var58
B glcm SumEntropy
B glcm MCC
B wav LL firstorder TenthPercentile

B lbpU histfunc Var50
B lbpU histfunc Var52
H wav LL glcm IMCI
H wav HH glcm Correlation
E wav LL firstorder Skewness
E wav LL firstorder Entropy
E wav LL glcm DifferenceVariance
E wav LH glcm Variance glcm
E wav HL glcm DifferenceVariance
E wav HH glcm DifferenceVariance
E lbpU histfunc Var1
Hu firstorder Median
Hu firstorder NinetiethPercentile
Hu firstorder Uniformity
Hu glcm AngularSecondMoment Energy
Hu glcm MCC
Hu wav LL firstorder Minimum
Hu wav LL firstorder Maximum
Hu wav LH glcm IMCII
Hu wav HL glcm IMCII
Hu wav HH glcm IMCII
S firstorder NinetiethPercentile
S glcm AngularSecondMoment Energy
S wav LL firstorder Minimum
S wav LH firstorder Minimum
S wav LH firstorder Skewness
S wav HH firstorder Kurtosis
S wav HL glcm DifferenceVariance
V firstorder Maximum
V wav LL firstorder Maximum
V wav LL firstorder Uniformity
V wav LH glcm Variance glcm
V wav LH glcm IMCII
V lbpU histfunc Var37

V lbpU histfunc Var54
V lbpU histfunc Var56

Table B.15. Selected features for the "Gleason 3 vs Non-malignant" classification task at 5× magnification level. G3 = Gleason 3.

G3 vs Non-Malignant
Gray glcm MCC
Gray wav LH firstorder Maximum
Gray wav HL firstorder Minimum
Gray lbpU histfunc Var54
R firstorder Maximum
R wav LL firstorder Minimum
R wav LH firstorder Minimum
R wav LH firstorder Maximum
R wav HL firstorder Minimum
R wav HL firstorder Maximum
R wav HH firstorder Minimum
R wav LH glcm SumVariance
G wav LL firstorder Maximum
G wav HH firstorder Minimum
G wav HH firstorder Range
G lbpU histfunc Var58
B firstorder Kurtosis
B glcm MCC
B wav LL firstorder Median
B wav LL firstorder TenthPercentile
B wav LL firstorder Range
B wav HL firstorder Minimum
B wav HL firstorder Maximum
B wav HH firstorder Maximum
B wav HH firstorder Skewness
B lbpU histfunc Var50
H wav LL glcm IMCI

H wav LL glcm MCC
H wav HL glcm DifferenceVariance
H wav HH glcm Correlation
E wav HH firstorder Skewness
E wav LL glcm DifferenceVariance
E wav HL glcm DifferenceVariance
E wav HH glcm DifferenceVariance
E lbpU histfunc Var1
Hu firstorder NinetiethPercentile
Hu glcm AngularSecondMoment Energy
Hu glcm Variance glcm
Hu glcm MCC
Hu wav LL firstorder Minimum
Hu wav LH glcm AngularSecondMoment Energy
Hu wav LH glcm IMCII
Hu wav LH glcm MCC
Hu wav HL glcm IMCII
Hu lbpU histfunc Var32
Hu lbpU histfunc Var36
Hu lbpU histfunc Var41
S firstorder NinetiethPercentile
S glcm SumEntropy
S wav LL firstorder Minimum
S wav LH firstorder Minimum
S wav LH glcm IMCII
V wav LL firstorder Maximum
V wav LH glcm Variance glcm
V wav LH glcm IMCII
V lbpU histfunc Var22
V lbpU histfunc Var37
V lbpU histfunc Var54

Table B.16. Selected features for "Gleason 4 + Gleason 5 vs Gleason 3" classification task at 5× magnification level. G3 = Gleason 3; G4 = Gleason 4; G5 = Gleason 5.

G4+G5 vs G3
Gray wav HH glcm AngularSecondMoment Energy
R wav LH firstorder Maximum
R wav HH firstorder Range
R wav HL glcm IMCI
R lbpU histfunc Var59
G wav LL firstorder Range
G wav LH firstorder Maximum
G wav LH firstorder Range
G wav HL firstorder Maximum
G lbpU histfunc Var42
G lbpU histfunc Var45
B firstorder Maximum
B firstorder Median
B firstorder Range
B glcm AngularSecondMoment Energy
B glcm MCC
H wav LL firstorder MeanAbsoluteDeviation
H wav LL glcm AngularSecondMoment Energy
H wav LH glcm Correlation
H wav LH glcm IMCII
H wav HL glcm SumVariance
H wav HH glcm IMCI
E firstorder Range
E wav LL firstorder Maximum
E wav LL firstorder Range
E wav HH firstorder Minimum
E wav LH glcm IMCII
E wav LH glcm MCC
E lbpU histfunc Var25
E lbpU histfunc Var52
Hu firstorder TenthPercentile

Hu glcm IMCII
Hu wav LL firstorder Maximum
Hu wav HL firstorder Kurtosis
Hu wav LH glcm MCC
Hu wav HL glcm SumVariance
Hu wav HH glcm SumVariance
Hu wav HH glcm MCC
Hu lbpU histfunc Var5
Hu lbpU histfunc Var18
Hu lbpU histfunc Var23
Hu lbpU histfunc Var28
Hu lbpU histfunc Var34
S wav LL firstorder Skewness
S wav LL firstorder NinetiethPercentile
S wav LL firstorder Entropy
S wav LH firstorder Range
S wav HL firstorder Minimum
S wav HL firstorder Range
S wav HH firstorder Kurtosis
S wav LL glcm SumEntropy
S wav LH glcm Variance glcm
S lbpU histfunc Var24
S lbpU histfunc Var54
V firstorder Maximum
V wav LL firstorder Minimum
V wav LL firstorder Variance
V wav LH firstorder Skewness
V wav HH glcm DifferenceVariance

Table B.17. Selected features for the "Gleason 4 + Gleason 5 vs Non-malignant" at 5× magnification level. G4 = Gleason 4; G5 = Gleason 5.

G4+G5 vs Non-Malignant
Gray wav LL firstorder Range

Gray wav LH firstorder Minimum
Gray wav LH firstorder Maximum
Gray wav LH firstorder Skewness
Gray wav HL firstorder Maximum
R firstorder Maximum
R firstorder Median
R wav LH firstorder Minimum
R wav HL firstorder Minimum
R wav HH firstorder Minimum
R wav LH glcm SumVariance
R wav HH glcm SumVariance
G wav LH glcm DifferenceEntropy
G lbpU histfunc Var43
G lbpU histfunc Var44
G lbpU histfunc Var58
B glcm MCC
B wav LL firstorder TenthPercentile
B wav HH firstorder Range
B lbpU histfunc Var50
B lbpU histfunc Var52
H wav LL glcm IMCI
H wav LH glcm Correlation
H wav HH glcm Correlation
H lbpU histfunc Var59
E wav LL glcm DifferenceVariance
E wav LH glcm SumVariance
E wav HL glcm SumVariance
E wav HL glcm DifferenceVariance
E lbpU histfunc Var1
Hu firstorder Median
Hu firstorder NinetiethPercentile
Hu firstorder Uniformity
Hu glcm AngularSecondMoment Energy
Hu glcm SumVariance

Hu glcm MCC
Hu wav LH firstorder NinetiethPercentile
Hu wav LH glcm IMCII
Hu wav HL glcm SumVariance
Hu wav HL glcm IMCII
Hu wav HH glcm IMCII
Hu lbpU histfunc Var36
S firstorder NinetiethPercentile
S glcm SumEntropy
S wav LL firstorder Minimum
S wav LL firstorder Skewness
S wav LH firstorder Minimum
S wav LH glcm IMCII
S wav HL glcm DifferenceVariance
V firstorder Maximum
V firstorder Variance
V wav LL firstorder Maximum
V wav LL firstorder Uniformity
V wav LH glcm Variance glcm
V lbpU histfunc Var35
V lbpU histfunc Var36
V lbpU histfunc Var54

Bibliography

[1] AI4Medimaging - AI for better workflow, better outcomes in medical imaging answering unmet medical needs.

[2] Analysis of Texture Features and Classifications for the Accurate Diagnosis of Prostate Cancer -Journal of Korea Multimedia Society | Korea Science.

[3] Aperio GT 450 DX - Automated, High Capacity Digital Pathology Slide Scanner.

[4] BBMRI-ERIC and National Nodes publish 2020 Annual Report.

[5] Brains.

[6] Chaimeleon – Accelerating the lab to market transition of AI tools for cancer management.

[7] DICOM.

[8] Erasmus MC: Patiëntenzorg.

[9] EuroBioBank – EuroBioBank website.

[10] GA4GH.

[11] Home.

[12] Imaging Data Commons | CRDC.

[13] Incisive Project - Home.

[14] ISBER.

[15] ISIDORe project: services for infectious disease outbreak research.

[16] JSON.

[17] Laboratory of NeuroImaging.

[18] NCI Cancer Research Data Commons | CBIIT.

[19] NIMH Data Archive - NDA.

[20] ProCAncer-I | An AI Platform integrating imaging data and models, supporting precision care through prostate cancer's continuum.

[21] PS3.18.

[22] UK Biobank - UK Biobank.

[23] WMA - The World Medical Association-Home.

[24] Genome in a Bottle. *NIST*, August 2012. Last Modified: 2023-11-29T13:41-05:00.

[25] canSERV - Cutting Edge Cancer Research Services Across Europe, January 2023.

[26] Shahira Abousamra, Rajarsi Gupta, Le Hou, Rebecca Batiste, Tianhao Zhao, Anand Shankar, Arvind Rao, Chao Chen, Dimitris Samaras, Tahsin Kurc, and Joel Saltz. Deep Learning-Based Mapping of Tumor Infiltrating Lymphocytes in Whole Slide Images of 23 Types of Cancer. *Frontiers in Oncology*, 11, 2022.

[27] Balazs Acs, Fahad Shabbir Ahmed, Swati Gupta, Pok Fai Wong, Robyn D. Gartrell, Jaya Sarin Pradhan, Emanuelle M. Rizk, Bonnie Gould Rothberg, Yvonne M. Saenger, and David L. Rimm. An open source automated tumor infiltrating lymphocyte algorithm for prognosis in melanoma. *Nature Communications*, 10(1):5440, November 2019.

[28] Famke Aeffner, Mark D. Zarella, Nathan Buchbinder, Marilyn M. Bui, Matthew R. Goodman, Douglas J. Hartman, Giovanni M. Lujan, Mariam A. Molani, Anil V. Parwani, Kate Lillard, Oliver C. Turner, Venkata N. P. Vemuri, Ana G. Yuil-Valdes, and Douglas Bowman. Introduction to Digital Image Analysis in Whole-slide Imaging: A White Paper from the Digital Pathology Association. *Journal of Pathology Informatics*, 10:9, 2019.

[29] Parnian Afshar, Arash Mohammadi, Konstantinos N. Plataniotis, Anastasia Oikonomou, and Habib Benali. From Handcrafted to Deep-Learning-Based Cancer Radiomics: Challenges and Opportunities. *IEEE Signal Processing Magazine*, 36(4):132–160, July 2019.

[30] Kiran Aftab, Faiqa Binte Aamir, Saad Mallick, Fatima Mubarak, Whitney B. Pope, Tom Mikkelsen, Jack P. Rock, and Syed Ather Enam. Radiomics for precision medicine in glioblastoma. *Journal of Neuro-Oncology*, 156(2):217–231, January 2022.

[31] Marco Aiello, Dario Baldi, Giuseppina Esposito, Marika Valentino, Marco Randon, Marco Salvatore, and Carlo Cavaliere. Evaluation of AI-Based Segmentation Tools for COVID-19 Lung Lesions on Conventional and Ultra-low Dose CT Scans. *Dose-Response: A Publication of International Hormesis Society*, 20(1):15593258221082896, 2022.

[32] Marco Aiello, Carlo Cavaliere, Antonio D'Albore, and Marco Salvatore. The Challenges of Diagnostic Imaging in the Era of Big Data. *Journal of Clinical Medicine*, 8(3):316, March 2019.

[33] Marco Aiello, Giuseppina Esposito, Giulio Pagliari, Pasquale Borrelli, Valentina Brancato, and Marco Salvatore. How does DICOM support big data management? Investigating its use in medical imaging community. *Insights into Imaging*, 12(1):164, December 2021.

[34] Lina Al-Ebbini, Omar F. Khabour, Karem H. Alzoubi, and Almuthanna K. Alkaraki. Biomedical Data Sharing Among Researchers: A Study from Jordan. *Journal of Multidisciplinary Healthcare*, 13:1669–1676, 2020.

[35] Shaimaa Al-Janabi, André Huisman, and Paul J Van Diest. Digital pathology: current status and future perspectives. *Histopathology*, 61(1):1–9, 2012. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2559.2011.03814.x.

[36] Ángel Alberich-Bayarri, Rafael Hernández-Navarro, Enrique Ruiz-Martínez, Fabio García-Castro, David García-Juan, and Luis Martí-Bonmatí. Development of imaging biomarkers and generation of big data. *La radiologia medica*, 122(6):444–448, June 2017.

[37] Eleni Alexandratou, Vassilis Atlamazoglou, Trias Thireou, George Agrogiannis, Dimitrios Togas, Nikolaos Kavantzas, Efstratios Patsouris, and Dido Yova. Evaluation of machine learning techniques for prostate cancer diagnosis and Gleason grading. *International Journal of Computational Intelligence in Bioinformatics and Systems Biology*, 1(3):297–315, January 2010. Publisher: Inderscience Publishers.

[38] Charlems Alvarez-Jimenez, Alvaro A. Sandino, Prateek Prasanna, Amit Gupta, Satish E. Viswanath, and Eduardo Romero. Identifying Cross-Scale Associations between Radiomic and Pathomic Signatures of Non-Small Cell Lung Cancer Subtypes: Preliminary Results. *Cancers*, 12(12):3663, December 2020.

[39] Mohamed Amgad, Lamees A. Atteya, Hagar Hussein, Kareem Hosny Mohammed, Ehab Hafiz, Maha A. T. Elsebaie, Ahmed M. Alhusseiny, Mohamed Atef AlMoslemany, Abdelmagid M. Elmatboly, Philip A. Pappalardo, Rokia Adel Sakr, Pooya Mobadersany, Ahmad Rachid, Anas M.

Saad, Ahmad M. Alkashash, Inas A. Ruhban, Anas Alrefai, Nada M. Elgazar, Ali Abdulkarim, Abo-Alela Farag, Amira Etman, Ahmed G. Elsaeed, Yahya Alagha, Yomna A. Amer, Ahmed M. Raslan, Menatalla K. Nadim, Mai A. T. Elsebaie, Ahmed Ayad, Liza E. Hanna, Ahmed Gadallah, Mohamed Elkady, Bradley Drumheller, David Jaye, David Manthey, David A. Gutman, Habiba Elfandy, and Lee A. D. Cooper. NuCLS: A scalable crowdsourcing, deep learning approach and dataset for nucleus classification, localization and segmentation. *GigaScience*, 11:giac037, May 2022.

[40] Mohamed Amgad, Habiba Elfandy, Hagar Hussein, Lamees A. Atteya, Mai A. T. Elsebaie, Lamia S. Abo Elnasr, Rokia A. Sakr, Hazem S. E. Salem, Ahmed F. Ismail, Anas M. Saad, Joumana Ahmed, Maha A. T. Elsebaie, Mustafijur Rahman, Inas A. Ruhban, Nada M. Elgazar, Yahya Alagha, Mohamed H. Osman, Ahmed M. Alhusseiny, Mariam M. Khalaf, Abo-Alela F. Younes, Ali Abdulkarim, Duaa M. Younes, Ahmed M. Gadallah, Ahmad M. Elkashash, Salma Y. Fala, Basma M. Zaki, Jonathan Beezley, Deepak R. Chittajallu, David Manthey, David A. Gutman, and Lee A. D. Cooper. Structured crowdsourcing enables convolutional segmentation of histology images. *Bioinformatics (Oxford, England)*, 35(18):3461–3467, September 2019.

[41] Mohamed Amgad, Roberto Salgado, and Lee A. D. Cooper. MuTILs: explainable, multiresolution computational scoring of Tumor-Infiltrating Lymphocytes in breast carcinomas using clinical guidelines, January 2022.

[42] Mohamed Amgad, Elisabeth Specht Stovgaard, Eva Balslev, Jeppe Thagaard, Weijie Chen, Sarah Dudgeon, Ashish Sharma, Jennifer K. Kerner, Carsten Denkert, Yinyin Yuan, Khalid AbdulJabbar, Stephan Wienert, Peter Savas, Leonie Voorwerk, Andrew H. Beck, Anant Madabhushi, Johan Hartman, Manu M. Sebastian, Hugo M. Horlings, Jan Hudeček, Francesco Ciompi, David A. Moore, Rajendra Singh, Elvire Roblin, Marcelo Luiz Balancin, Marie-Christine Mathieu, Jochen K. Lennerz, Pawan Kirtani, I.-Chun Chen, Jeremy P. Braybrooke, Giancarlo Pruneri, Sandra Demaria, Sylvia Adams, Stuart J. Schnitt, Sunil R. Lakhani, Federico Rojo, Laura Comerma, Sunil S. Badve, Mehrnoush Khojasteh, W. Fraser Symmans, Christos Sotiriou, Paula Gonzalez-Ericsson, Katherine L. Pogue-Geile, Rim S. Kim, David L. Rimm, Giuseppe Viale, Stephen M. Hewitt, John M. S. Bartlett, Frédérique Penault-Llorca, Shom Goel, Huang-Chun Lien, Sibylle Loibl, Zuzana Kos, Sherene Loi, Matthew G. Hanna, Stefan Michiels, Marleen Kok, Torsten O. Nielsen, Alexander J. Lazar, Zsuzsanna Bago-Horvath, Loes F. S. Kooreman, Jeroen A. W. M. van der Laak, Joel Saltz, Brandon D. Gallas, Uday Kurkure, Michael Barnes, Roberto Salgado,

and Lee A. D. Cooper. Report on computational assessment of Tumor Infiltrating Lymphocytes from the International Immuno-Oncology Biomarker Working Group. *npj Breast Cancer*, 6(1):1–13, May 2020.

[43] Samy Ammari, Stephanie Pitre-Champagnat, Laurent Dercle, Emilie Chouzenoux, Salma Moalla, Sylvain Reuze, Hugues Talbot, Tite Mokoyoko, Joya Hadchiti, Sebastien Diffetocq, Andreas Volk, Mickeal El Haik, Sara Lakiss, Corinne Balleyguier, Nathalie Lassau, and Francois Bidault. Influence of Magnetic Field Strength on Magnetic Resonance Imaging Radiomics Features in Brain Imaging, an In Vitro and In Vivo Study. *Frontiers in Oncology*, 10:541663, 2020.

[44] Komal Anjum, Bibi Ibtesam Shagufta, Syed Qamar Abbas, Seema Patel, Ishrat Khan, Sayed Asmat Ali Shah, Najeeb Akhter, and Syed Shams ul Hassan. Current status and future therapeutic perspectives of glioblastoma multiforme (GBM) therapy: A review. *Biomedicine & Pharmacotherapy*, 92:681–689, August 2017.

[45] Laura Annaratone, Giuseppe De Palma, Giuseppina Bonizzi, Anna Sapino, Gerardo Botti, Enrico Berrino, Chiara Mannelli, Pamela Arcella, Simona Di Martino, Agostino Steffan, Maria Grazia Daidone, Vincenzo Canzonieri, Barbara Parodi, Angelo Virgilio Paradiso, Massimo Barberis, Caterina Marchiò, and Alleanza Contro il Cancro (ACC) Pathology and Biobanking Working Group. Basic principles of biobanking: from biological samples to precision medicine for patients. *Virchows Archiv: An International Journal of Pathology*, 479(2):233–246, August 2021.

[46] Nehal M. Atallah, Michael S. Toss, Clare Verrill, Manuel Salto-Tellez, David Snead, and Emad A. Rakha. Potential quality pitfalls of digitalized whole slide image of breast pathology in routine practice. *Modern Pathology*, 35(7):903–910, July 2022. Number: 7 Publisher: Nature Publishing Group.

[47] Thazin Nwe Aung, Saba Shafi, James S. Wilmott, Saeed Nourmohammadi, Ioannis Vathiotis, Niki Gavrielatou, Aileen Fernandez, Vesal Yaghoobi, Tobias Sinnberg, Teresa Amaral, Kristian Ikenberg, Kiarash Khosrotehrani, Iman Osman, Balazs Acs, Yalai Bai, Sandra Martinez-Morilla, Myrto Moutafi, John F. Thompson, Richard A. Scolyer, and David L. Rimm. Objective assessment of tumor infiltrating lymphocytes as a prognostic marker in melanoma using machine learning algorithms. *EBioMedicine*, 82:104143, August 2022.

[48] D. Baldi, M. Aiello, A. Duggento, M. Salvatore, and C. Cavaliere. MR Imaging-Histology Correlation by Tailored 3D-Printed Slicer in Oncological Assessment. *Contrast Media & Molecular Imaging*, 2019:1–9, May 2019.

[49] Maschenka CA. Balkenhol, Francesco Ciompi, Żaneta Świderska Chadaj, Rob van de Loo, Milad Intezar, Irene Otte-Höller, Daan Geijs, Johannes Lotz, Nick Weiss, Thomas de Bel, Geert Litjens, Peter Bult, and Jeroen AWM. van der Laak. Optimized tumour infiltrating lymphocyte assessment for triple negative breast cancer prognostics. *The Breast : Official Journal of the European Society of Mastology*, 56:78–87, February 2021.

[50] Roland Bammer. Basic principles of diffusion-weighted imaging. *European Journal of Radiology*, 45(3):169–184, March 2003.

[51] Peter Bankhead, Maurice B. Loughrey, José A. Fernández, Yvonne Dombrowski, Darragh G. McArt, Philip D. Dunne, Stephen McQuaid, Ronan T. Gray, Liam J. Murray, Helen G. Coleman, Jacqueline A. James, Manuel Salto-Tellez, and Peter W. Hamilton. QuPath: Open source software for digital pathology image analysis. *Scientific Reports*, 7(1):16878, December 2017.

[52] Laura Barisoni, Kyle J. Lafata, Stephen M. Hewitt, Anant Madabhushi, and Ulysses G. J. Balis. Digital pathology and computational image analysis in nephropathology. *Nature Reviews Nephrology*, 16(11):669–685, November 2020.

[53] Ajay Nagesh Basavanhally, Shridar Ganesan, Shannon Agner, James Peter Monaco, Michael D. Feldman, John E. Tomaszewski, Gyan Bhanot, and Anant Madabhushi. Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology. *IEEE transactions on bio-medical engineering*, 57(3):642–653, March 2010.

[54] Amrita Basu, Denise Warzel, Aras Eftekhari, Justin S. Kirby, John Freymann, Janice Knable, Ashish Sharma, and Paula Jacobs. Call for Data Standardization: Lessons Learned and Recommendations in an Imaging Study. *JCO Clinical Cancer Informatics*, (3):1–11, December 2019.

[55] Vipul Baxi, Robin Edwards, Michael Montalto, and Saurabh Saha. Digital pathology and artificial intelligence in translational medicine and clinical practice. *Modern Pathology: An Official Journal of the United States and Canadian Academy of Pathology, Inc*, 35(1):23–32, January 2022.

[56] Garrett F. Beeghly, Adrian A. Shimpi, Robert N. Riter, and Claudia Fischbach. Measuring and modelling tumour heterogeneity across scales. *Nature Reviews Bioengineering*, July 2023.

[57] Sara Belarouci and Mohammed Amine Chikh. Medical imbalanced data classification. *Advances in Science, Technology and Engineering Systems Journal*, 2(3):116–124, April 2017.

[58] Kaustav Bera, Nathaniel Braman, Amit Gupta, Vamsidhar Velcheti, and Anant Madabhushi. Predicting cancer outcomes with radiomics and artificial intelligence in radiology. *Nature Reviews Clinical Oncology*, 19(2):132–146, February 2022.

[59] Oskar Bergengren, Kelly R. Pekala, Konstantina Matsoukas, Jonathan Fainberg, Sean F. Mungovan, Ola Bratt, Freddie Bray, Otis Brawley, Amy N. Luckenbaugh, Lorelei Mucci, Todd M. Morgan, and Sigrid V. Carlsson. 2022 Update on Prostate Cancer Epidemiology and Risk Factors—A Systematic Review. *European Urology*, 84(2):191–206, August 2023.

[60] Subrata Bhattacharjee, Cho-Hee Kim, Hyeon-Gyun Park, Deekshitha Prakash, Nuwan Madusanka, Nam-Hoon Cho, and Heung-Kook Choi. Multi-Features Classification of Prostate Carcinoma Observed in Histological Sections: Analysis of Wavelet-Based Texture and Colour Features. *Cancers*, 11(12):1937, December 2019.

[61] Eric Bienefeld. euCanSHare: EU-Canada joint data platform to facilitate multi-study cardiovascular research. December 2018. Publisher: Zenodo.

[62] Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag, Berlin, Heidelberg, 2006.

[63] S. A. Bobholz, A. K. Lowman, M. Brehler, F. Kyereme, S. R. Duenweg, J. Sherman, S. D. McGarry, E. J. Cochran, J. Connelly, W. M. Mueller, M. Agarwal, A. Banerjee, and P. S. LaViolette. Radio-Pathomic Maps of Cell Density Identify Brain Tumor Invasion beyond Traditional MRI-Defined Margins. *AJNR. American journal of neuroradiology*, 43(5):682–688, May 2022.

[64] Samuel A. Bobholz, Allison K. Lowman, Alexander Barrington, Michael Brehler, Sean McGarry, Elizabeth J. Cochran, Jennifer Connelly, Wade M. Mueller, Mohit Agarwal, Darren O'Neill, Andrew S. Nencka, Anjishnu Banerjee, and Peter S. LaViolette. Radiomic Features of Multiparametric MRI Present Stable Associations With Analogous Histological Features in Patients With Brain Cancer. *Tomography (Ann Arbor, Mich.)*, 6(2):160–169, June 2020.

[65] Zuhir Bodalal, Stefano Trebeschi, and Regina Beets-Tan. Radiomics: a critical step towards integrated healthcare. *Insights into Imaging*, 9(6):911–914, December 2018.

[66] Kevin M. Boehm, Pegah Khosravi, Rami Vanguri, Jianjiong Gao, and Sohrab P. Shah. Harnessing multimodal data integration to advance precision oncology. *Nature Reviews Cancer*, October 2021.

[67] Giuseppina Bonizzi, Lorenzo Zattoni, Maria Capra, Cristina Cassi, Giulio Taliento, Mariia Ivanova, Elena Guerini-Rocco, Marzia Fumagalli, Massimo Monturano, Adriana Albini, Giuseppe Viale, Roberto Orecchia, and Nicola Fusco. Standard operating procedures for biobank in oncology. *Frontiers in Molecular Biosciences*, 9:967310, 2022.

[68] Luca Bonomi, Yingxiang Huang, and Lucila Ohno-Machado. Privacy challenges and research opportunities for genomic data sharing. *Nature Genetics*, 52(7):646–654, July 2020.

[69] Valentina Brancato, Marco Aiello, Luca Basso, Serena Monti, Luigi Palumbo, Giuseppe Di Costanzo, Marco Salvatore, Alfonso Ragozzino, and Carlo Cavaliere. Evaluation of a multiparametric MRI radiomic-based approach for stratification of equivocal PI-RADS 3 and upgraded PI-RADS 4 prostatic lesions. *Scientific Reports*, 11(1):643, January 2021.

[70] Valentina Brancato, Carlo Cavaliere, Nunzia Garbino, Francesco Isgrò, Marco Salvatore, and Marco Aiello. The relationship between radiomics and pathomics in Glioblastoma patients: Preliminary results from a cross-scale association study. *Frontiers in Oncology*, 12, 2022.

[71] Valentina Brancato, Marco Cerrone, Marialuisa Lavitrano, Marco Salvatore, and Carlo Cavaliere. A Systematic Review of the Current Status and Quality of Radiomics for Glioma Differential Diagnosis. *Cancers*, 14(11):2731, May 2022.

[72] Valentina Brancato, Giuseppina Esposito, Luigi Coppola, Carlo Cavaliere, Peppino Mirabelli, Camilla Scapicchio, Rita Borgheresi, Emanuele Neri, Marco Salvatore, and Marco Aiello. Standardizing digital biobanks: integrating imaging, genomic, and clinical data for precision medicine. *Journal of Translational Medicine*, 22(1):136, February 2024.

[73] Valentina Brancato, Nunzia Garbino, Lorenzo Mannelli, Marco Aiello, Marco Salvatore, Monica Franzese, and Carlo Cavaliere. Impact of radiogenomics in esophageal cancer on clinical outcomes: A pilot study. *World Journal of Gastroenterology*, 27(36):6110–6127, September 2021.

[74] Valentina Brancato, Silvia Nuzzo, Liberatore Tramontano, Gerolama Condorelli, Marco Salvatore, and Carlo Cavaliere. Predicting Survival in Glioblastoma Patients Using Diffusion MR Imaging Metrics-A Systematic Review. *Cancers*, 12(10):E2858, October 2020.

[75] Alvis Brazma. Minimum Information About a Microarray Experiment (MIAME)–successes, failures, challenges. *TheScientificWorldJournal*, 9:420–423, May 2009.

[76] Wouter Bulten, Kimmo Kartasalo, Po-Hsuan Cameron Chen, Peter Ström, Hans Pinckaers, Kunal Nagpal, Yuannan Cai, David F. Steiner, Hester van Boven, Robert Vink, Christina Hulsbergen-van de Kaa, Jeroen van der Laak, Mahul B. Amin, Andrew J. Evans, Theodorus van der Kwast, Robert Allan, Peter A. Humphrey, Henrik Grönberg, Hemamali Samaratunga, Brett Delahunt, Toyonori Tsuzuki, Tomi Häkkinen, Lars Egevad, Maggie Demkin, Sohier Dane, Fraser Tan, Masi Valkonen, Greg S. Corrado, Lily Peng, Craig H. Mermel, Pekka Ruusuvuori, Geert Litjens, and Martin Eklund. Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge. *Nature Medicine*, 28(1):154–163, January 2022. Number: 1 Publisher: Nature Publishing Group.

[77] Nicole Bussola, Alessia Marcolini, Valerio Maggio, Giuseppe Jurman, and Cesare Furlanello. AI Slipping on Tiles: Data Leakage in Digital Pathology. In Alberto Del Bimbo, Rita Cucchiara, Stan Sclaroff, Giovanni Maria Farinella, Tao Mei, Marco Bertini, Hugo Jair Escalante, and Roberto Vezzani, editors, *Pattern Recognition. ICPR International Workshops and Challenges*, Lecture Notes in Computer Science, pages 167–182, Cham, 2021. Springer International Publishing.

[78] Gabriele Campanella, Matthew G. Hanna, Luke Geneslaw, Allen Miraflor, Vitor Werneck Krauss Silva, Klaus J. Busam, Edi Brogi, Victor E. Reuter, David S. Klimstra, and Thomas J. Fuchs. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nature Medicine*, 25(8):1301–1309, August 2019. Number: 8 Publisher: Nature Publishing Group.

[79] Lior Carmi, Mishael Zohar, and Gianluigi M. Riva. The European General Data Protection Regulation (GDPR) in mHealth: Theoretical and practical aspects for practitioners' use. *Medicine, Science, and the Law*, 63(1):61–68, January 2023.

[80] Alexandre Carré, Guillaume Klausner, Myriam Edjlali, Marvin Lerousseau, Jade Briend-Diop, Roger Sun, Samy Ammari, Sylvain Reuzé, Emilie Alvarez Andres, Théo Estienne, Stéphane Niyoteka, Enzo Battistella, Maria Vakalopoulou, Frédéric Dhermain, Nikos Paragios, Eric Deutsch, Catherine Oppenheim, Johan Pallud, and Charlotte Robert. Standardization of brain MR images across machines and protocols: bridging the gap for MRI-based radiomics. *Scientific Reports*, 10(1):12340, December 2020.

[81] Ahmad Chaddad, Michael Jonathan Kucharczyk, Paul Daniel, Siham Sabri, Bertrand J. Jean-Claude, Tamim Niazi, and Bassam Abdulkarim. Radiomics in Glioblastoma: Current Status and Challenges Facing Clinical Implementation. *Frontiers in Oncology*, 9:374, 2019.

[82] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16:321–357, June 2002.

[83] Shifu Chen, Tanxiao Huang, Tiexiang Wen, Hong Li, Mingyan Xu, and Jia Gu. MutScan: fast detection and visualization of target mutations by scanning FASTQ data. *BMC bioinformatics*, 19(1):16, January 2018.

[84] Stephen A. Chervitz, Eric W. Deutsch, Dawn Field, Helen Parkinson, John Quackenbush, Phillipe Rocca-Serra, Susanna-Assunta Sansone, Christian J. Stoeckert, Chris F. Taylor, Ronald Taylor, and Catherine A. Ball. Data standards for Omics data: the basis of data sharing and reuse. *Methods in Molecular Biology (Clifton, N.J.)*, 719:31–69, 2011.

[85] Geetha Soujanya Chilla, Cher Heng Tan, Chenjie Xu, and Chueh Loo Poh. Diffusion weighted magnetic resonance imaging and its recent trend-a survey. *Quantitative Imaging in Medicine and Surgery*, 5(3):407–422, June 2015.

[86] Kenneth Clark, Bruce Vendt, Kirk Smith, John Freymann, Justin Kirby, Paul Koppel, Stephen Moore, Stanley Phillips, David Maffitt, Michael Pringle, Lawrence Tarbox, and Fred Prior. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *Journal of Digital Imaging*, 26(6):1045–1057, December 2013.

[87] David Clunie, Dan Hosseinzadeh, Mikael Wintell, David De Mena, Nieves Lajara, Marcial Garcia-Rojo, Gloria Bueno, Kiran Saligrama, Aaron Stearrett, David Toomey, Esther Abels, Frank Van Apeldoorn, Stephane Langevin, Sean Nichols, Joachim Schmid, Uwe Horchner, Bruce Beckwith, Anil Parwani, and Liron Pantanowitz. Digital Imaging and Communications in Medicine Whole Slide Imaging Connectathon at Digital Pathology Association Pathology Visions 2017. *Journal of Pathology Informatics*, 9:6, 2018.

[88] David A. Clunie. *DICOM structured reporting*. PixelMed Pub, Bangor, Pa, 2000.

[89] Peter J. A. Cock, Christopher J. Fields, Naohisa Goto, Michael L. Heuer, and Peter M. Rice. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research*, 38(6):1767–1771, April 2010.

[90] G. Collewet, M. Strzelecki, and F. Mariette. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. *Magnetic Resonance Imaging*, 22(1):81–91, January 2004.

[91] Lee AD Cooper, Jun Kong, David A Gutman, William D Dunn, Michael Nalisnik, and Daniel J Brat. Novel genotype-phenotype associations in human cancers enabled by advanced molecular platforms and computational analysis of whole slide images. *Laboratory Investigation*, 95(4):366–376, April 2015.

[92] Luigi Coppola, Alessandra Cianflone, Anna Maria Grimaldi, Mariarosaria Incoronato, Paolo Bevilacqua, Francesco Messina, Simona Baselice, Andrea Soricelli, Peppino Mirabelli, and Marco Salvatore. Biobanking in health care: evolution and future directions. *Journal of Translational Medicine*, 17(1):172, December 2019.

[93] Germán Corredor, Xiangxue Wang, Cheng Lu, Vamsidhar Velcheti, Eduardo Romero, and Anant Madabhushi. A watershed and feature-based approach for automated detection of lymphocytes on lung cancer images. In *Medical Imaging 2018: Digital Pathology*, volume 10581, pages 213–218. SPIE, March 2018.

[94] Germán Corredor, Xiangxue Wang, Yu Zhou, Cheng Lu, Pingfu Fu, Konstantinos Syrigos, David L. Rimm, Michael Yang, Eduardo Romero, Kurt A. Schalper, Vamsidhar Velcheti, and Anant Madabhushi. Spatial architecture and arrangement of tumor-infiltrating lymphocytes for predicting likelihood of recurrence in early-stage non-small cell lung cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 25(5):1526–1534, March 2019.

[95] National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC). Radiology Data from the Clinical Proteomic Tumor Analysis Consortium Glioblastoma Multiforme [CPTAC-GBM] collection, 2018.

[96] Oscar Cuadros Linares, Aurea Aurea Soriano-Vargas, Bruno S. Faiçal, Bernd Hamann, Alexandre T. Fabro, and Agma J. M. Traina. Efficient Segmentation of Cell Nuclei in Histopathological Images. In *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*, pages 47–52, July 2020.

[97] Miao Cui and David Y. Zhang. Artificial intelligence and computational pathology. *Laboratory Investigation*, 101(4):412–422, April 2021. Number: 4 Publisher: Nature Publishing Group.

[98] V. Curcin, S. Miles, R. Danger, Y. Chen, R. Bache, and A. Taweel. Implementing interoperable provenance in biomedical research. *Future Generation Computer Systems*, 34:1–16, May 2014.

[99] Marina D'Amato, Przemysław Szostak, and Benjamin Torben-Nielsen. A Comparison Between Single- and Multi-Scale Approaches for Classification of Histopathology Images. *Frontiers in Public Health*, 10:892658, July 2022.

[100] Christel Daniel, Marcial García Rojo, Jacques Klossa, Vincenzo Della Mea, David Booker, Bruce A. Beckwith, and Thomas Schrader. Standardizing the use of whole slide images in digital pathology. *Computerized Medical Imaging and Graphics*, 35(7):496–505, October 2011.

[101] Pasquale De Blasio and Ida Biunno. New Challenges for Biobanks: Accreditation to the New ISO 20387:2018 Standard Specific for Biobanks. *Biotech (Basel (Switzerland))*, 10(3):13, July 2021.

[102] Marleen de Bruijne. Machine learning approaches in medical image analysis: From detection to diagnosis. *Medical Image Analysis*, 33:94–97, October 2016.

[103] Scott Doyle, Michael Feldman, John Tomaszewski, and Anant Madabhushi. A boosted Bayesian multiresolution classifier for prostate cancer detection from digitized needle biopsies. *IEEE transactions on bio-medical engineering*, 59(5):1205–1218, May 2012.

[104] Marta Drake-Pérez, Jose Boto, Aikaterini Fitsiori, Karl Lovblad, and Maria Isabel Vargas. Clinical applications of diffusion weighted imaging in neuroradiology. *Insights into Imaging*, 9(4):535–547, August 2018.

[105] Isabel Dregely, Davide Prezzi, Christian Kelly-Morland, Elisa Roccia, Radhouene Neji, and Vicky Goh. Imaging biomarkers in oncology: Basics and application to MRI: MRI Biomarkers in Oncology. *Journal of Magnetic Resonance Imaging*, 48(1):13–26, July 2018.

[106] Petra Duhm-Harbeck and Jens K. Habermann. Data Protection in Healthcare-Integrated Biobanking. *Innovations in Digital Health, Diagnostics, and Biomarkers*, 3(2023):1–7, January 2023.

[107] Loïc Duron, Daniel Balvay, Saskia Vande Perre, Afef Bouchouicha, Julien Savatovsky, Jean-Claude Sadik, Isabelle Thomassin-Naggara, Laure Fournier, and Augustin Lecler. Gray-level discretization impacts reproducible MRI radiomics texture features. *PloS One*, 14(3):e0213459, 2019.

[108] Oliver Eidel, Jan-Oliver Neumann, Sina Burth, Pascal J. Kieslich, Christine Jungk, Felix Sahm, Philipp Kickingereder, Karl Kiening, Andreas Unterberg, Wolfgang Wick, Heinz-Peter Schlemmer, Martin Bendszus, and Alexander Radbruch. Automatic Analysis of Cellularity in Glioblastoma and Correlation with ADC Using Trajectory Analysis and Automatic Nuclei Counting. *PLOS ONE*, 11(7):e0160250, July 2016.

[109] Niina Eklund, Ny Haingo Andrianarisoa, Esther van Enckevort, Gabriele Anton, Annelies Debucquoy, Heimo Müller, Linda Zaharenko, Cäcilia Engels, Lars Ebert, Michael Neumann, Joachim Geeraert, Veronique T'Joen, Hans Demski, Élodie Caboux, Rumyana Proynova, Barbara Parodi, Sebastian Mate, Erik van Iperen, Roxana Merino-Martinez, Philip R. Quinlan, Petr Holub, and Kaisa Silander. Extending the Minimum Information About BIobank Data Sharing Terminology to Describe Samples, Sample Donors, and Events. *Biopreservation and Biobanking*, 18(3):155–164, June 2020.

[110] Khalid El Bairi, Harry R. Haynes, Elizabeth Blackley, Susan Fineberg, Jeffrey Shear, Sophia Turner, Juliana Ribeiro de Freitas, Daniel Sur, Luis Claudio Amendola, Masoumeh Gharib, Amine Kallala, Indu Arun, Farid Azmoudeh-Ardalan, Luciana Fujimoto, Luz F. Sua, Shi-Wei Liu, Huang-Chun Lien, Pawan Kirtani, Marcelo Balancin, Hicham El Attar, Prerna Guleria, Wenxian Yang, Emad Shash, I.-Chun Chen, Veronica Bautista, Jose Fernando Do Prado Moura, Bernardo L. Rapoport, Carlos Castaneda, Eunice Spengler, Gabriela Acosta-Haab, Isabel Frahm, Joselyn Sanchez, Miluska Castillo, Najat Bouchmaa, Reena R. Md Zin, Ruohong Shui, Timothy Onyuma, Wentao Yang, Zaheed Husain, Karen Willard-Gallo, An Coosemans, Edith A. Perez, Elena Provenzano, Paula Gonzalez Ericsson, Eduardo Richardet, Ravi Mehrotra, Sandra Sarancone, Anna Ehinger, David L. Rimm, John M. S. Bartlett, Giuseppe Viale, Carsten Denkert, Akira I. Hida, Christos Sotiriou, Sibylle Loibl, Stephen M. Hewitt, Sunil Badve, William Fraser Symmans, Rim S. Kim, Giancarlo Pruneri, Shom Goel, Prudence A. Francis, Gloria Inurrigarro, Rin Yamaguchi, Hernan Garcia-Rivello, Hugo Horlings, Said Afqir, Roberto Salgado, Sylvia Adams, Marleen Kok, Maria Vittoria Dieci, Stefan Michiels, Sandra Demaria, and Sherene Loi. The tale of TILs in breast cancer: A report from The International Immuno-Oncology Biomarker Working Group. *npj Breast Cancer*, 7(1):1–17, December 2021.

[111] Benjamin M Ellingson, Patrick Y Wen, and Timothy F Cloughesy. Evidence and context of use for contrast enhancement as a surrogate of disease burden and treatment response in malignant glioma. *Neuro-Oncology*, 20(4):457–471, March 2018.

[112] Jonathan I Epstein. Prostate cancer grading: a decade after the 2005 modified system. *Modern Pathology*, 31:47–63, January 2018.

[113] Jonathan I. Epstein, Michael J. Zelefsky, Daniel D. Sjoberg, Joel B. Nelson, Lars Egevad, Cristina Magi-Galluzzi, Andrew J. Vickers, Anil V. Parwani, Victor E. Reuter, Samson W. Fine, James A. Eastham, Peter Wiklund, Misop Han, Chandana A. Reddy, Jay P. Ciezki, Tommy Nyberg, and

Eric A. Klein. A Contemporary Prostate Cancer Grading System: A Validated Alternative to the Gleason Score. *European Urology*, 69(3):428–435, March 2016.

[114] Giuseppina Esposito, Giulio Pagliari, Marco Randon, Peppino Mirabelli, Marialuisa Lavitrano, Marco Aiello, and Marco Salvatore. BCU Imaging Biobank, an Innovative Digital Resource for Biomedical Research Collecting Imaging and Clinical Data From Human Healthy and Pathological Subjects. *Open Journal of Bioresources*, 8:4, September 2021.

[115] Brancato V. et al. The relationship between radiomics and pathomics in Glioblastoma patients: Preliminary results from a cross-scale association study-Supplementary Materials.

[116] Graham J. Etherington, Ricardo H. Ramirez-Gonzalez, and Dan MacLean. bio-samtools 2: a package for analysis and visualization of sequence and alignment data with SAMtools in Ruby. *Bioinformatics (Oxford, England)*, 31(15):2565–2567, August 2015.

[117] European Society of Radiology (ESR). ESR Position Paper on Imaging Biobanks. *Insights into Imaging*, 6(4):403–410, August 2015.

[118] Harriet Evans, Emily Hero, Fayyaz Minhas, Noorul Wahab, Katherine Dodd, Harvir Sahota, Ratnadeep Ganguly, Andrew Robinson, Manjuvani Neerudu, Elaine Blessing, Pallavi Borkar, and David Snead. Standardized Clinical Annotation of Digital Histopathology Slides at the Point of Diagnosis. *Modern Pathology*, 36(11):100297, November 2023.

[119] Andrey Fedorov, Matthew Hancock, David Clunie, Mathias Brochhausen, Jonathan Bona, Justin Kirby, John Freymann, Steve Pieper, Hugo J. W. L. Aerts, Ron Kikinis, and Fred Prior. DICOM re-encoding of volumetrically annotated Lung Imaging Database Consortium (LIDC) nodules. *Medical Physics*, 47(11):5953–5965, November 2020.

[120] Andrew H. Fischer, Kenneth A. Jacobson, Jack Rose, and Rolf Zeller. Hematoxylin and eosin staining of tissue and cell sections. *CSH protocols*, 2008:pdb.prot4986, May 2008.

[121] Isabella Fornacon-Wood, Hitesh Mistry, Christoph J. Ackermann, Fiona Blackhall, Andrew McPartlin, Corinne Faivre-Finn, Gareth J. Price, and James P. B. O'Connor. Reliability and prognostic value of radiomic features are highly dependent on choice of feature extraction platform. *European Radiology*, 30(11):6241–6250, November 2020.

[122] Francesca Fornasa. Diffusion-weighted Magnetic Resonance Imaging: What Makes Water Run Fast or Slow? *Journal of Clinical Imaging Science*, 1:27, May 2011.

[123] Isabel Fortier, Paul R. Burton, Paula J. Robson, Vincent Ferretti, Julian Little, Francois L'Heureux, Mylène Deschênes, Bartha M. Knoppers, Dany Doiron, Joost C. Keers, Pamela Linksted, Jennifer R. Harris, Geneviève Lachance, Catherine Boileau, Nancy L. Pedersen, Carol M. Hamilton, Kristian Hveem, Marilyn J. Borugian, Richard P. Gallagher, John McLaughlin, Louise Parker, John D. Potter, John Gallacher, Rudolf Kaaks, Bette Liu, Tim Sprosen, Anne Vilain, Susan A. Atkinson, Andrea Rengifo, Robin Morton, Andres Metspalu, H. Erich Wichmann, Mark Tremblay, Rex L. Chisholm, Andrés Garcia-Montero, Hans Hillege, Jan-Eric Litton, Lyle J. Palmer, Markus Perola, Bruce H. R. Wolffenbuttel, Leena Peltonen, and Thomas J. Hudson. Quality, quantity and harmony: the DataSHaPER approach to integrating data across bioclinical studies. *International Journal of Epidemiology*, 39(5):1383–1393, October 2010.

[124] Chiara Frascarelli, Giuseppina Bonizzi, Camilla Rosella Musico, Eltjona Mane, Cristina Cassi, Elena Guerini Rocco, Annarosa Farina, Aldo Scarpa, Rita Lawlor, Luca Reggiani Bonetti, Stefania Caramaschi, Albino Eccher, Stefano Marletta, and Nicola Fusco. Revolutionizing Cancer Research: The Impact of Artificial Intelligence in Digital Biobanking. *Journal of Personalized Medicine*, 13(9):1390, September 2023.

[125] Michela Gabelloni, Lorenzo Faggioni, Rita Borgheresi, Giuliana Restante, Jorge Shortrede, Lorenzo Tumminello, Camilla Scapicchio, Francesca Coppola, Dania Cioni, Ignacio Gómez-Rico, Luis Martí-Bonmatí, and Emanuele Neri. Bridging gaps between images and data: a systematic update on imaging biobanks. *European Radiology*, 32(5):3173–3186, May 2022.

[126] Galen F. Gao, Joel S. Parker, Sheila M. Reynolds, Tiago C. Silva, Liang-Bo Wang, Wanding Zhou, Rehan Akbani, Matthew Bailey, Saianand Balu, Benjamin P. Berman, Denise Brooks, Hu Chen, Andrew D. Cherniack, John A. Demchok, Li Ding, Ina Felau, Sharon Gaheen, Daniela S. Gerhard, David I. Heiman, Kyle M. Hernandez, Katherine A. Hoadley, Reyka Jayasinghe, Anab Kemal, Theo A. Knijnenburg, Peter W. Laird, Michael K. A. Mensah, Andrew J. Mungall, A. Gordon Robertson, Hui Shen, Roy Tarnuzzer, Zhining Wang, Matthew Wyczalkowski, Liming Yang, Jean C. Zenklusen, Zhenyu Zhang, Genomic Data Analysis Network, Han Liang, and Michael S. Noble. Before and After: Comparison of Legacy and Harmonized TCGA Genomic Data Commons' Data. *Cell Systems*, 9(1):24–34.e10, July 2019.

[127] Maria Rosaria Giovagnoli and Daniele Giansanti. Artificial Intelligence in Digital Pathology: What Is the Future? Part 1: From the Digital Slide Onwards. *Healthcare (Basel, Switzerland)*, 9(7):858, July 2021.

[128] Ross Girshick. Fast R-CNN. 2015. Publisher: arXiv Version Number: 2.

[129] Concepción González Hernando, Laura Esteban, Teresa Cañas, Enrique Van den Brule, and Miguel Pastrana. The role of magnetic resonance imaging in oncology. *Clinical & Translational Oncology: Official Publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico*, 12(9):606–613, September 2010.

[130] Julien Guiot, Akshayaa Vaidyanathan, Louis Deprez, Fadila Zerka, Denis Danthine, Anne-Noelle Frix, Philippe Lambin, Fabio Bottari, Nathan Tsoutzidis, Benjamin Miraglio, Sean Walsh, Wim Vos, Roland Hustinx, Marta Ferreira, Pierre Lovinfosse, and Ralph T. H. Leijenaar. A review in radiomics: Making personalized medicine a reality via routine imaging. *Medicinal Research Reviews*, 42(1):426–440, January 2022.

[131] Lu Guo, Gang Wang, Yuanming Feng, Tonggang Yu, Yu Guo, Xu Bai, and Zhaoxiang Ye. Diffusion and perfusion weighted magnetic resonance imaging for tumor volume definition in radiotherapy of brain tumors. *Radiation Oncology*, 11(1):123, December 2016.

[132] Laxmi Gupta, Barbara Mara Klinkhammer, Claudia Seikrit, Nina Fan, Nassim Bouteldja, Philipp Gräbel, Michael Gadermayr, Peter Boor, and Dorit Merhof. Large-scale extraction of interpretable features provides new insights into kidney histopathology – A proof-of-concept study. *Journal of Pathology Informatics*, 13:100097, January 2022.

[133] Pramod Kumar Gupta, Rishi Awasthi, Shalini Singh, Sanjay Behari, K. J. Maria Das, Rakesh Kumar Gupta, and Shaleen Kumar. Value of Minimum Apparent Diffusion Coefficient on Magnetic Resonance Imaging as a Biomarker for Predicting Progression of Disease Following Surgery and Radiotherapy in Glial Tumors from a Tertiary Care Center in Northern India. *Journal of Neurosciences in Rural Practice*, 8(2):185–193, June 2017.

[134] Rajarsi Gupta, Tahsin Kurc, Ashish Sharma, Jonas S. Almeida, and Joel Saltz. The Emergence of Pathomics. *Current Pathobiology Reports*, 7(3):73–84, September 2019.

[135] Louis Guttman. Some necessary conditions for common-factor analysis. *Psychometrika*, 19(2):149–161, June 1954.

[136] Akihiro Haga, Wataru Takahashi, Shuri Aoki, Kanabu Nawa, Hideomi Yamashita, Osamu Abe, and Keiichi Nakagawa. Standardization of imaging features for radiomics analysis. *The Journal of Medical Investigation*, 66(1.2):35–37, February 2019.

[137] Robert M. Haralick, K. Shanmugam, and Its'Hak Dinstein. Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6):610–621, November 1973.

[138] Tomohiro Hayakawa, V. B. Surya Prasath, Hiroharu Kawanaka, Bruce J. Aronow, and Shinji Tsuruoka. Computational Nuclei Segmentation Methods in Digital Pathology: A Survey. *Archives of Computational Methods in Engineering*, 28(1):1–13, January 2021.

[139] Markus D. Herrmann, David A. Clunie, Andriy Fedorov, Sean W. Doyle, Steven Pieper, Veronica Klepeis, Long P. Le, George L. Mutter, David S. Milstone, Thomas J. Schultz, Ron Kikinis, Gopal K. Kotecha, David H. Hwang, Katherine P. Andriole, A. John Iafrate, James A. Brink, Giles W. Boland, Keith J. Dreyer, Mark Michalski, Jeffrey A. Golden, David N. Louis, and Jochen K. Lennerz. Implementing the DICOM Standard for Digital Pathology. *Journal of Pathology Informatics*, 9:37, 2018.

[140] Christian Herz, Jean-Christophe Fillion-Robin, Michael Onken, Jörg Riesmeier, Andras Lasso, Csaba Pinter, Gabor Fichtinger, Steve Pieper, David Clunie, Ron Kikinis, and Andriy Fedorov. *dcmqi* : An Open Source Library for Standardized Communication of Quantitative Image Analysis Results Using DICOM. *Cancer Research*, 77(21):e87–e90, November 2017.

[141] Petr Holub, Florian Kohlmayer, Fabian Prasser, Michaela Th Mayrhofer, Irene Schlünder, Gillian M. Martin, Sara Casati, Lefteris Koumakis, Andrea Wutte, Łukasz Kozera, Dominik Strapagiel, Gabriele Anton, Gianluigi Zanetti, Osman Ugur Sezerman, Maimuna Mendy, Dalibor Valík, Marialuisa Lavitrano, Georges Dagher, Kurt Zatloukal, GertJan B. van Ommen, and Jan-Eric Litton. Enhancing Reuse of Data and Biological Material in Medical Research: From FAIR to FAIR-Health. *Biopreservation and Biobanking*, 16(2):97–105, April 2018.

[142] Xinrui Huang, Zhaotong Li, Minghui Zhang, and Song Gao. Fusing hand-crafted and deep-learning features in a convolutional neural network model to identify prostate cancer in pathology images. *Frontiers in Oncology*, 12, 2022.

[143] Xinmi Huo, KokHaur Ong, Kah Weng Lau, Laurent Gole, David Young, Char Loo Tan, Chongchong Zhang, Yonghui Zhang, Xiaohui Zhu, Longjie Li, Hao Han, Haoda Lu, Gabriel Marini, Jun Xu, Kun Gui, Bingxian Chen, Wenzhao Shi, Wanyuan Chen, Stephan Sanders, Hwee Kuan Lee, Susan Swee-Shan Hue, Weimiao Yu, and Soo Yong Tan. Comprehensive AI Model Development for Gleason Grading:From Scanning, Cloud-based Annotation to Pathologist-AI Interaction. preprint, In Review, September 2022.

[144] Shah Hussain, Iqra Mubeen, Niamat Ullah, Syed Shahab Ud Din Shah, Bakhtawar Abduljalil Khan, Muhammad Zahoor, Riaz Ullah, Farhat Ali Khan, and Mujeeb A. Sultan. Modern Diagnostic Imaging Technique Applications and Risk Factors in the Medical Field: A Review. *BioMed Research International*, 2022:5164970, June 2022.

[145] A. Ibrahim, S. Primakov, M. Beuque, H.C. Woodruff, I. Halilaj, G. Wu, T. Refaee, R. Granzier, Y. Widaatalla, R. Hustinx, F.M. Mottaghy, and P. Lambin. Radiomics for precision medicine: Current challenges, future prospects, and the proposal of a new framework. *Methods*, 188:20–29, April 2021.

[146] Mariarosaria Incoronato, Marco Aiello, Teresa Infante, Carlo Cavaliere, Anna Grimaldi, Peppino Mirabelli, Serena Monti, and Marco Salvatore. Radiogenomic Analysis of Oncological Data: A Technical Survey. *International Journal of Molecular Sciences*, 18(4):805, April 2017.

[147] Massimiliano Izzo. *Biomedical Research and Integrated Biobanking: An Innovative Paradigm for Heterogeneous Data Management*. Springer Theses. Springer International Publishing, Cham, 2016.

[148] Massimiliano Izzo, Francesco Mortola, Gabriele Arnulfo, Marco M. Fato, and Luigi Varesio. A digital repository with an extensible data model for biobanking and genomic analysis management. *BMC genomics*, 15 Suppl 3(Suppl 3):S3, 2014.

[149] Mariam Jamal-Hanjani, Sergio A. Quezada, James Larkin, and Charles Swanton. Translational Implications of Tumor Heterogeneity. *Clinical Cancer Research*, 21(6):1258–1266, March 2015.

[150] Andrew Janowczyk and Anant Madabhushi. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *Journal of Pathology Informatics*, 7:29, 2016.

[151] Harold Jeffreys. *Theory of probability*. Oxford classic texts in the physical sciences. Clarendon Press ; Oxford University Press, Oxford [Oxfordshire] : New York, 3rd ed edition, 1998.

[152] Xiaoyu Jiang, Hua Li, Jingping Xie, Eliot T. McKinley, Ping Zhao, John C. Gore, and Junzhong Xu. In vivo imaging of cancer cell size and cellularity using temporal diffusion spectroscopy: Cancer Cell Size and Cellularity Using IMPULSED. *Magnetic Resonance in Medicine*, 78(1):156–164, July 2017.

[153] Laura Gálvez Jiménez, Lucile Dierckx, Maxime Amodei, Hamed Razavi Khosroshahi, Natarajan Chidambaran, Anh-Thu Phan Ho, and Alberto

Franzin. Computational Evaluation of the Combination of Semi-Supervised and Active Learning for Histopathology Image Segmentation with Missing Annotations. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, page 2552, 2023.

[154] Thomas Kalinski, Ralf Zwönitzer, Mathias Roßner, Harald Hofmann, Albert Roessner, and Thomas Guenther. Digital Imaging and Communications in Medicine (DICOM) as standard in digital pathology. *Histopathology*, 61(1):132–134, July 2012.

[155] Aditya Kashyap, Maria Anna Rapsomaniki, Vesna Barros, Anna Fomitcheva-Khartchenko, Adriano Luca Martinelli, Antonio Foncubierta Rodriguez, Maria Gabrani, Michal Rosen-Zvi, and Govind Kaigala. Quantification of tumor heterogeneity: from data acquisition to metric generation. *Trends in Biotechnology*, 40(6):647–676, June 2022.

[156] Robert E. Kass and Adrian E. Raftery. Bayes Factors. *Journal of the American Statistical Association*, 90(430):773–795, June 1995.

[157] Jakob Nikolas Kather, Narmin Ghaffari Laleh, Sebastian Foersch, and Daniel Truhn. Medical domain knowledge in domain-agnostic generative AI. *npj Digital Medicine*, 5(1):90, July 2022.

[158] Jakob Nikolas Kather, Cleo-Aron Weis, Francesco Bianconi, Susanne M. Melchers, Lothar R. Schad, Timo Gaiser, Alexander Marx, and Frank Gerrit Zöllner. Multi-class texture analysis in colorectal cancer histology. *Scientific Reports*, 6(1):27988, June 2016. Number: 1 Publisher: Nature Publishing Group.

[159] Philipp Kickingereder, Ulf Neuberger, David Bonekamp, Paula L. Piechotta, Michael Götz, Antje Wick, Martin Sill, Annekathrin Kratz, Russell T. Shinohara, David T. W. Jones, Alexander Radbruch, John Muschelli, Andreas Unterberg, Jürgen Debus, Heinz-Peter Schlemmer, Christel Herold-Mende, Stefan Pfister, Andreas von Deimling, Wolfgang Wick, David Capper, Klaus H. Maier-Hein, and Martin Bendszus. Radiomic subtyping improves disease stratification beyond key molecular, clinical, and standard imaging characteristics in patients with glioblastoma. *Neuro-Oncology*, 20(6):848–857, May 2018.

[160] Tim-Rasmus Kiehl. Digital and Computational Pathology: A Specialty Reimagined. In Sepehr Ehsani, Patrick Glauner, Philipp Plugmann, and Florian M. Thieringer, editors, *The Future Circle of Healthcare*, pages 227–250. Springer International Publishing, Cham, 2022. Series Title: Future of Business and Finance.

[161] Tim-Rasmus Kiehl. Digital and Computational Pathology: A Specialty Reimagined. In Sepehr Ehsani, Patrick Glauner, Philipp Plugmann, and Florian M. Thieringer, editors, *The Future Circle of Healthcare: AI, 3D Printing, Longevity, Ethics, and Uncertainty Mitigation*, Future of Business and Finance, pages 227–250. Springer International Publishing, Cham, 2022.

[162] Aaro Kiuru, Rudy Mattheus, and Katrine Weisteen Bjerde. Standardization of medical imaging in Europe: an integrated initiative. *Computer Methods and Programs in Biomedicine*, 43(1-2):21–26, May 1994.

[163] F. Klauschen, K.-R. Müller, A. Binder, M. Bockmayr, M. Hägele, P. Seegerer, S. Wienert, G. Pruneri, S. de Maria, S. Badve, S. Michiels, T. O. Nielsen, S. Adams, P. Savas, F. Symmans, S. Willis, T. Gruosso, M. Park, B. Haibe-Kains, B. Gallas, A. M. Thompson, I. Cree, C. Sotiriou, C. Solinas, M. Preusser, S. M. Hewitt, D. Rimm, G. Viale, S. Loi, S. Loibl, R. Salgado, C. Denkert, and International Immuno-Oncology Biomarker Working Group. Scoring of tumor-infiltrating lymphocytes: From visual estimation to machine learning. *Seminars in Cancer Biology*, 52(Pt 2):151–157, October 2018.

[164] S. Klein, M. Staring, K. Murphy, M.A. Viergever, and J. Pluim. elastix: A Toolbox for Intensity-Based Medical Image Registration. *IEEE Transactions on Medical Imaging*, 29(1):196–205, January 2010.

[165] Stefan Klein, Marius Staring, Patrik Andersson, and Josien P. W. Pluim. Preconditioned Stochastic Gradient Descent Optimisation for Monomodal Image Registration. In Gabor Fichtinger, Anne Martel, and Terry Peters, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2011*, volume 6892, pages 549–556. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011. Series Title: Lecture Notes in Computer Science.

[166] Stefan Klein, Uulke A. van der Heide, Irene M. Lips, Marco van Vulpen, Marius Staring, and Josien P. W. Pluim. Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information. *Medical Physics*, 35(4):1407–1417, April 2008.

[167] Bartha Maria Knoppers, Alexander Bernier, Sarion Bowers, and Emily Kirby. Open Data in the Era of the GDPR: Lessons from the Human Cell Atlas. *Annual Review of Genomics and Human Genetics*, 24(1):369–391, August 2023.

[168] Dow-Mu Koh, Nickolas Papanikolaou, Ulrich Bick, Rowland Illing, Charles E. Kahn, Jayshree Kalpathi-Cramer, Celso Matos, Luis Martí-Bonmatí, Anne Miles, Seong Ki Mun, Sandy Napel, Andrea Rockall, Evis

Sala, Nicola Strickland, and Fred Prior. Artificial intelligence and machine learning in cancer imaging. *Communications Medicine*, 2(1):133, October 2022.

[169] Haridimos Kondylakis, Varvara Kalokyri, Stelios Sfakianakis, Kostas Marias, Manolis Tsiknakis, Ana Jimenez-Pastor, Eduardo Camacho-Ramos, Ignacio Blanquer, J. Damian Segrelles, Sergio López-Huguet, Caroline Barelle, Magdalena Kogut-Czarkowska, Gianna Tsakou, Nikolaos Siopis, Zisis Sakellariou, Paschalis Bizopoulos, Vicky Drossou, Antonios Lalas, Konstantinos Votis, Pedro Mallol, Luis Marti-Bonmati, Leonor Cerdá Alberich, Karine Seymour, Samuel Boucher, Esther Ciarrocchi, Lauren Fromont, Jordi Rambla, Alexander Harms, Andrea Gutierrez, Martijn P. A. Starmans, Fred Prior, Josep Ll. Gelpi, and Karim Lekadir. Data infrastructures for AI in medical imaging: a report on the experiences of five EU projects. *European Radiology Experimental*, 7(1):20, May 2023.

[170] Jun Kong, Lee A. D. Cooper, Fusheng Wang, Jingjing Gao, George Teodoro, Lisa Scarpace, Tom Mikkelsen, Matthew J. Schniederjan, Carlos S. Moreno, Joel H. Saltz, and Daniel J. Brat. Machine-Based Morphologic Analysis of Glioblastoma Using Whole-Slide Pathology Images Uncovers Clinically Relevant Molecular Correlates. *PLoS ONE*, 8(11):e81049, November 2013.

[171] Yan Kong, Georgi Z. Genchev, Xiaolei Wang, Hongyu Zhao, and Hui Lu. Nuclear Segmentation in Histopathological Images Using Two-Stage Stacked U-Nets With Attention Mechanism. *Frontiers in Bioengineering and Biotechnology*, 8:573866, October 2020.

[172] Zuzana Kos, Elvire Roblin, Rim S. Kim, Stefan Michiels, Brandon D. Gallas, Weijie Chen, Koen K. van de Vijver, Shom Goel, Sylvia Adams, Sandra Demaria, Giuseppe Viale, Torsten O. Nielsen, Sunil S. Badve, W. Fraser Symmans, Christos Sotiriou, David L. Rimm, Stephen Hewitt, Carsten Denkert, Sibylle Loibl, Stephen J. Luen, John M. S. Bartlett, Peter Savas, Giancarlo Pruneri, Deborah A. Dillon, Maggie Chon U. Cheang, Andrew Tutt, Jacqueline A. Hall, Marleen Kok, Hugo M. Horlings, Anant Madabhushi, Jeroen van der Laak, Francesco Ciompi, Anne-Vibeke Laenkholm, Enrique Bellolio, Tina Gruosso, Stephen B. Fox, Juan Carlos Araya, Giuseppe Floris, Jan Hudeček, Leonie Voorwerk, Andrew H. Beck, Jen Kerner, Denis Larsimont, Sabine Declercq, Gert Van den Eynden, Lajos Pusztai, Anna Ehinger, Wentao Yang, Khalid AbdulJabbar, Yinyin Yuan, Rajendra Singh, Crispin Hiley, Maise al Bakir, Alexander J. Lazar, Stephen Naber, Stephan Wienert, Miluska Castillo, Giuseppe Curigliano, Maria-Vittoria Dieci, Fabrice André, Charles Swanton, Jorge Reis-Filho,

Joseph Sparano, Eva Balslev, I-Chun Chen, Elisabeth Ida Specht Stovgaard, Katherine Pogue-Geile, Kim R. M. Blenman, Frédérique Penault-Llorca, Stuart Schnitt, Sunil R. Lakhani, Anne Vincent-Salomon, Federico Rojo, Jeremy P. Braybrooke, Matthew G. Hanna, M. Teresa Soler-Monsó, Daniel Bethmann, Carlos A. Castaneda, Karen Willard-Gallo, Ashish Sharma, Huang-Chun Lien, Susan Fineberg, Jeppe Thagaard, Laura Comerma, Paula Gonzalez-Ericsson, Edi Brogi, Sherene Loi, Joel Saltz, Frederick Klaushen, Lee Cooper, Mohamed Amgad, David A. Moore, and Roberto Salgado. Pitfalls in assessing stromal tumor infiltrating lymphocytes (sTILs) in breast cancer. *NPJ Breast Cancer*, 6:17, May 2020.

[173] Katsumi Kose. Physical and technical aspects of human magnetic resonance imaging: present status and 50 years historical review. *Advances in Physics: X*, 6(1):1885310, January 2021. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/23746149.2021.1885310.

[174] Kory Kreimeyer, Matthew Foster, Abhishek Pandey, Nina Arya, Gwendolyn Halford, Sandra F Jones, Richard Forshee, Mark Walderhaug, and Taxiarchis Botsis. Natural language processing systems for capturing and standardizing unstructured clinical information: A systematic review. *Journal of Biomedical Informatics*, 73:14–29, September 2017.

[175] Christiane K. Kuhl and Daniel Truhn. The Long Route to Standardized Radiomics: Unraveling the Knot from the End. *Radiology*, 295(2):339–341, May 2020.

[176] Neeraj Kumar, Ruchika Verma, Sanuj Sharma, Surabhi Bhargava, Abhishek Vahadane, and Amit Sethi. A Dataset and a Technique for Generalized Nuclear Segmentation for Computational Pathology. *IEEE Transactions on Medical Imaging*, 36(7):1550–1560, July 2017.

[177] Yung-Lung Kuo, Chien-Chuan Ko, and Ming-Ji Lee. LYMPHATIC INFILTRATION DETECTION IN BREAST CANCER H&E IMAGE PRIOR TO LYMPHADENECTOMY. *Biomedical Engineering: Applications, Basis and Communications*, 26(04):1440007, August 2014.

[178] Sasagu Kurozumi, Kenichi Inoue, Hiroshi Matsumoto, Takaaki Fujii, Jun Horiguchi, Tetsunari Oyama, Masafumi Kurosumi, and Ken Shirabe. Prognostic utility of tumor-infiltrating lymphocytes in residual tumor after neoadjuvant chemotherapy with trastuzumab for HER2-positive breast cancer. *Scientific Reports*, 9(1):1583, February 2019.

[179] Kyle J. Lafata, Yuqi Wang, Brandon Konkel, Fang-Fang Yin, and Mustafa R. Bashir. Radiomics: a primer on high-throughput image phenotyping. *Abdominal Radiology*, 47(9):2986–3002, September 2022.

[180] Andrew Lagree, Majidreza Mohebpour, Nicholas Meti, Khadijeh Saednia, Fang-I. Lu, Elzbieta Slodkowska, Sonal Gandhi, Eileen Rakovitch, Alex Shenfield, Ali Sadeghi-Naini, and William T. Tran. A review and comparison of breast tumor cell nuclei segmentation performances using deep convolutional neural networks. *Scientific Reports*, 11(1):8025, April 2021.

[181] Philippe Lambin, Ralph T. H. Leijenaar, Timo M. Deist, Jurgen Peerlings, Evelyn E. C. de Jong, Janita van Timmeren, Sebastian Sanduleanu, Ruben T. H. M. Larue, Aniek J. G. Even, Arthur Jochems, Yvonka van Wijk, Henry Woodruff, Johan van Soest, Tim Lustberg, Erik Roelofs, Wouter van Elmpt, Andre Dekker, Felix M. Mottaghy, Joachim E. Wildberger, and Sean Walsh. Radiomics: the bridge between medical imaging and personalized medicine. *Nature Reviews. Clinical Oncology*, 14(12):749–762, December 2017.

[182] Philippe Lambin, Emmanuel Rios-Velazquez, Ralph Leijenaar, Sara Carvalho, Ruud G. P. M. van Stiphout, Patrick Granton, Catharina M. L. Zegers, Robert Gillies, Ronald Boellard, André Dekker, and Hugo J. W. L. Aerts. Radiomics: extracting more information from medical images using advanced feature analysis. *European Journal of Cancer (Oxford, England: 1990)*, 48(4):441–446, March 2012.

[183] Drew Landis, William Courtney, Christopher Dieringer, Ross Kelly, Margaret King, Brittny Miller, Runtang Wang, Dylan Wood, Jessica A. Turner, and Vince D. Calhoun. COINS Data Exchange: An open platform for compiling, curating, and disseminating neuroimaging data. *NeuroImage*, 124:1084–1088, January 2016.

[184] Han Le, Rajarsi Gupta, Le Hou, Shahira Abousamra, Danielle Fassler, Luke Torre-Healy, Richard A. Moffitt, Tahsin Kurc, Dimitris Samaras, Rebecca Batiste, Tianhao Zhao, Arvind Rao, Alison L. Van Dyke, Ashish Sharma, Erich Bremer, Jonas S. Almeida, and Joel Saltz. Utilizing Automated Breast Cancer Detection to Identify Spatial Distributions of Tumor-Infiltrating Lymphocytes in Invasive Breast Cancer. *The American Journal of Pathology*, 190(7):1491–1504, July 2020.

[185] Sara Leibfarth, David Mönnich, Stefan Welz, Christine Siegel, Nina Schwenzer, Holger Schmidt, Daniel Zips, and Daniela Thorwarth. A strategy for multimodal deformable image registration to integrate PET/MR into radiotherapy treatment planning. *Acta Oncologica*, 52(7):1353–1359, October 2013.

[186] Karim Lekadir, Richard Osuala, Catherine Gallin, Noussair Lazrak, Kaisar Kushibar, Gianna Tsakou, Susanna Aussó, Leonor Cerdá Alberich, Kostas

Marias, Manolis Tsiknakis, Sara Colantonio, Nickolas Papanikolaou, Zohaib Salahuddin, Henry C. Woodruff, Philippe Lambin, and Luis Martí-Bonmatí. FUTURE-AI: Guiding Principles and Consensus Recommendations for Trustworthy Artificial Intelligence in Medical Imaging, September 2021.

[187] Joshua J. Li, Julia Y. Tsang, and Gary M. Tse. Tumor Microenvironment in Breast Cancer—Updates on Therapeutic Implications and Pathologic Assessment. *Cancers*, 13(16):4233, January 2021.

[188] Sheng Li, Scott W. Tighe, Charles M. Nicolet, Deborah Grove, Shawn Levy, William Farmerie, Agnes Viale, Chris Wright, Peter A. Schweitzer, Yuan Gao, Dewey Kim, Joe Boland, Belynda Hicks, Ryan Kim, Sagar Chhangawala, Nadereh Jafari, Nalini Raghavachari, Jorge Gandara, Natàlia Garcia-Reyero, Cynthia Hendrickson, David Roberson, Jeffrey Rosenfeld, Todd Smith, Jason G. Underwood, May Wang, Paul Zumbo, Don A. Baldwin, George S. Grills, and Christopher E. Mason. Multi-platform assessment of transcriptome profiling using RNA-seq in the ABRF next-generation sequencing study. *Nature Biotechnology*, 32(9):915–925, September 2014.

[189] Xiangrui Li, Paul S. Morgan, John Ashburner, Jolinda Smith, and Christopher Rorden. The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *Journal of Neuroscience Methods*, 264:47–56, May 2016.

[190] Xiaohong Li, Guy N. Brock, Eric C. Rouchka, Nigel G. F. Cooper, Dongfeng Wu, Timothy E. O'Toole, Ryan S. Gill, Abdallah M. Eteleeb, Liz O'Brien, and Shesh N. Rai. A comparison of per sample global scaling and per gene normalization methods for differential expression analysis of RNA-seq data. *PloS One*, 12(5):e0176185, 2017.

[191] Zhi-Cheng Li, Hongmin Bai, Qiuchang Sun, Qihua Li, Lei Liu, Yan Zou, Yinsheng Chen, Chaofeng Liang, and Hairong Zheng. Multiregional radiomics features from multiparametric MRI for prediction of MGMT methylation status in glioblastoma multiforme: A multicentre study. *European Radiology*, 28(9):3640–3650, September 2018.

[192] Wenyi Lin, Kyle Hasenstab, Guilherme Moura Cunha, and Armin Schwartzman. Comparison of handcrafted features and convolutional neural networks for liver MR image adequacy assessment. *Scientific Reports*, 10(1):20336, November 2020. Number: 1 Publisher: Nature Publishing Group.

[193] Ali Hasan Md. Linkon, Md. Mahir Labib, Tarik Hasan, Mozammal Hossain, and Marium-E Jannat. Deep learning in prostate cancer diagnosis and

Gleason grading in histopathology images: An extensive study. *Informatics in Medicine Unlocked*, 24:100582, January 2021.

[194] A.A. Litvin, D.A. Burkin, A.A. Kropinov, and F.N. Paramzin. Radiomics and Digital Image Texture Analysis in Oncology (Review). *Modern Technologies in Medicine*, 13(2):97–104, 2021.

[195] Lei Liu, Dong Zhao, Fanhua Yu, Ali Asghar Heidari, Jintao Ru, Huiling Chen, Majdi Mafarja, Hamza Turabieh, and Zhifang Pan. Performance optimization of differential evolution with slime mould algorithm for multilevel breast cancer image segmentation. *Computers in Biology and Medicine*, 138:104910, November 2021.

[196] Roberto Lo Gullo, Isaac Daimiel, Elizabeth A. Morris, and Katja Pinker. Combining molecular and imaging metrics in cancer: radiogenomics. *Insights into Imaging*, 11(1):1, January 2020.

[197] S. Loft and H. E. Poulsen. Cancer risk and oxidative DNA damage in man. *Journal of Molecular Medicine*, 74(6):297–312, June 1996.

[198] Cheng Lu, Rakesh Shiradkar, Zaiyi Liu, Biomedical Engineering Department, Case Western Reserve University, Cleveland 44106, OH, USA, and Department of Radiology, Guangzhou First People's Hospital, School of Medicine, South China University of Technology, Guangzhou 510080, China. Integrating pathomics with radiomics and genomics for cancer prognosis: A brief review. *Chinese Journal of Cancer Research*, 33(5):563–573, 2021.

[199] Zixiao Lu, Siwen Xu, Wei Shao, Yi Wu, Jie Zhang, Zhi Han, Qianjin Feng, and Kun Huang. Deep-Learning-Based Characterization of Tumor-Infiltrating Lymphocytes in Breast Cancers From Histopathology Images and Multiomics Data. *JCO clinical cancer informatics*, 4:480–490, May 2020.

[200] Giovanni Lucignani and Emanuele Neri. Integration of imaging biomarkers into systems biomedicine: a renaissance for medical imaging. *Clinical and Translational Imaging*, 7(2):149–153, April 2019.

[201] Chao Luo, Xiaojie Li, Lutao Wang, Jia He, Denggao Li, and Jiliu Zhou. How Does the Data set Affect CNN-based Image Classification Performance? In *2018 5th International Conference on Systems and Informatics (ICSAI)*, pages 361–366, November 2018.

[202] Alessio Luschi, Camilla Petraccone, Giuseppe Fico, Leandro Pecchia, and Ernesto Iadanza. Semantic Ontologies for Complex Healthcare Structures: A Scoping Review. *IEEE Access*, 11:19228–19246, 2023.

[203] Marc Macenko, Marc Niethammer, J. S. Marron, David Borland, John T. Woosley, Xiaojun Guan, Charles Schmitt, and Nancy E. Thomas. A method for normalizing histology slides for quantitative analysis. In *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 1107–1110, June 2009.

[204] Anant Madabhushi, Shannon Agner, Ajay Basavanhally, Scott Doyle, and George Lee. Computer-aided prognosis: Predicting patient and disease outcome via quantitative fusion of multi-scale, multi-modal data. *Computerized Medical Imaging and Graphics*, 35(7-8):506–514, October 2011.

[205] Anant Madabhushi and George Lee. Image analysis and machine learning in digital pathology: Challenges and opportunities. *Medical Image Analysis*, 33:170–175, October 2016.

[206] Shino Magaki, Seyed A. Hojat, Bowen Wei, Alexandra So, and William H. Yong. An Introduction to the Performance of Immunohistochemistry. *Methods in molecular biology (Clifton, N.J.)*, 1897:289–298, 2019.

[207] Norberto Malpica, Carlos Ortiz de Solórzano, Juan José Vaquero, Andrés Santos, Isabel Vallcorba, José Miguel García-Sagredo, and Francisco del Pozo. Applying watershed algorithms to the segmentation of clustered nuclei. *Cytometry*, 28(4):289–297, 1997.

[208] Kristina Malsagova, Artur Kopylov, Alexander Stepanov, Tatyana Butkova, Alexandra Sinitsyna, Alexander Izotov, and Anna Kaysheva. Biobanks—A Platform for Scientific and Biomedical Research. *Diagnostics*, 10(7):485, July 2020.

[209] MAQC Consortium, Leming Shi, Laura H. Reid, Wendell D. Jones, Richard Shippy, Janet A. Warrington, Shawn C. Baker, Patrick J. Collins, Francoise de Longueville, Ernest S. Kawasaki, Kathleen Y. Lee, Yuling Luo, Yongming Andrew Sun, James C. Willey, Robert A. Setterquist, Gavin M. Fischer, Weida Tong, Yvonne P. Dragan, David J. Dix, Felix W. Frueh, Frederico M. Goodsaid, Damir Herman, Roderick V. Jensen, Charles D. Johnson, Edward K. Lobenhofer, Raj K. Puri, Uwe Schrf, Jean Thierry-Mieg, Charles Wang, Mike Wilson, Paul K. Wolber, Lu Zhang, Shashi Amur, Wenjun Bao, Catalin C. Barbacioru, Anne Bergstrom Lucas, Vincent Bertholet, Cecilie Boysen, Bud Bromley, Donna Brown, Alan Brunner, Roger Canales, Xiaoxi Megan Cao, Thomas A. Cebula, James J. Chen, Jing Cheng, Tzu-Ming Chu, Eugene Chudin, John Corson, J. Christopher Corton, Lisa J. Croner, Christopher Davies, Timothy S. Davison, Glenda Delenstarr, Xutao Deng, David Dorris, Aron C. Eklund, Xiao-hui Fan, Hong Fang, Stephanie Fulmer-Smentek, James C. Fuscoe, Kathryn Gallagher,

Weigong Ge, Lei Guo, Xu Guo, Janet Hager, Paul K. Haje, Jing Han, Tao Han, Heather C. Harbottle, Stephen C. Harris, Eli Hatchwell, Craig A. Hauser, Susan Hester, Huixiao Hong, Patrick Hurban, Scott A. Jackson, Hanlee Ji, Charles R. Knight, Winston P. Kuo, J. Eugene LeClerc, Shawn Levy, Quan-Zhen Li, Chunmei Liu, Ying Liu, Michael J. Lombardi, Yunqing Ma, Scott R. Magnuson, Botoul Maqsodi, Tim McDaniel, Nan Mei, Ola Myklebost, Baitang Ning, Natalia Novoradovskaya, Michael S. Orr, Terry W. Osborn, Adam Papallo, Tucker A. Patterson, Roger G. Perkins, Elizabeth H. Peters, Ron Peterson, Kenneth L. Philips, P. Scott Pine, Lajos Pusztai, Feng Qian, Hongzu Ren, Mitch Rosen, Barry A. Rosenzweig, Raymond R. Samaha, Mark Schena, Gary P. Schroth, Svetlana Shchegrova, Dave D. Smith, Frank Staedtler, Zhenqiang Su, Hongmei Sun, Zoltan Szallasi, Zivana Tezak, Danielle Thierry-Mieg, Karol L. Thompson, Irina Tikhonova, Yaron Turpaz, Beena Vallanat, Christophe Van, Stephen J. Walker, Sue Jane Wang, Yonghong Wang, Russ Wolfinger, Alex Wong, Jie Wu, Chunlin Xiao, Qian Xie, Jun Xu, Wen Yang, Liang Zhang, Sheng Zhong, Yaping Zong, and William Slikker. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nature Biotechnology*, 24(9):1151–1161, September 2006.

[210] Daniel S. Marcus, Timothy R. Olsen, Mohana Ramaratnam, and Randy L. Buckner. The Extensible Neuroimaging Archive Toolkit: an informatics platform for managing, exploring, and sharing neuroimaging data. *Neuroinformatics*, 5(1):11–34, 2007.

[211] Paula Martin-Gonzalez, Mireia Crispin-Ortuzar, Leonardo Rundo, Maria Delgado-Ortet, Marika Reinius, Lucian Beer, Ramona Woitek, Stephan Ursprung, Helen Addley, James D. Brenton, Florian Markowetz, and Evis Sala. Integrative radiogenomics for virtual biopsy and treatment monitoring in ovarian cancer. *Insights into Imaging*, 11(1):94, August 2020.

[212] Hansmann Martin-Leo, Klauschen Frederick, Samek Wojciech, Müller Klaus-Robert, Donnadieu Emmanuel, Scharf Sonja, Hartmann Sylvia, Koch Ina, Ackermann Jörg, Pantanowitz Liron, Schäfer Hendrik, and Wurzel Patrick. Imaging bridges pathology and radiology. *Journal of Pathology Informatics*, 14:100298, January 2023.

[213] Luis Martí-Bonmatí, Ángel Alberich-Bayarri, Ruth Ladenstein, Ignacio Blanquer, J. Damian Segrelles, Leonor Cerdá-Alberich, Polyxeni Gkontra, Barbara Hero, J. M. García-Aznar, Daniel Keim, Wolfgang Jentner, Karine Seymour, Ana Jiménez-Pastor, Ismael González-Valverde, Blanca Martínez de Las Heras, Samira Essiaf, Dawn Walker, Michel Rochette,

Marian Bubak, Jordi Mestres, Marco Viceconti, Gracia Martí-Besa, Adela Cañete, Paul Richmond, Kenneth Y. Wertheim, Tomasz Gubala, Marek Kasztelnik, Jan Meizner, Piotr Nowakowski, Salvador Gilpérez, Amelia Suárez, Mario Aznar, Giuliana Restante, and Emanuele Neri. PRIMAGE project: predictive in silico multiscale analytics to support childhood cancer personalised evaluation empowered by imaging biomarkers. *European Radiology Experimental*, 4(1):22, April 2020.

[214] David S. McClintock, Jacob T. Abel, and Toby C. Cornish. Whole Slide Imaging Hardware, Software, and Infrastructure. In Anil V. Parwani, editor, *Whole Slide Imaging*, pages 23–56. Springer International Publishing, Cham, 2022.

[215] Anna Sophia McKenney, Emily Weg, Tejus A. Bale, Aaron T. Wild, Hyemin Um, Michael J. Fox, Andrew Lin, Jonathan T. Yang, Peter Yao, Maxwell L. Birger, Florent Tixier, Matthew Sellitti, Nelson S. Moss, Robert J. Young, and Harini Veeraraghavan. Radiomic Analysis to Predict Histopathologically Confirmed Pseudoprogression in Glioblastoma Patients. *Advances in Radiation Oncology*, page 100916, February 2022.

[216] Claire McQuin, Allen Goodman, Vasiliy Chernyshev, Lee Kamentsky, Beth A. Cimini, Kyle W. Karhohs, Minh Doan, Liya Ding, Susanne M. Rafelski, Derek Thirstrup, Winfried Wiegraebe, Shantanu Singh, Tim Becker, Juan C. Caicedo, and Anne E. Carpenter. CellProfiler 3.0: Next-generation image processing for biology. *PLoS biology*, 16(7):e2005970, July 2018.

[217] Juan S. Medina-Martínez, Juan E. Arango-Ossa, Max F. Levine, Yangyu Zhou, Gunes Gundem, Andrew L. Kung, and Elli Papaemmanuil. Isabl Platform, a digital biobank for processing multimodal patient data. *BMC Bioinformatics*, 21(1):549, December 2020.

[218] Carmelo Messina, Rodolfo Bignone, Alberto Bruno, Antonio Bruno, Federico Bruno, Marco Calandri, Damiano Caruso, Pietro Coppolino, Riccardo De Robertis, Francesco Gentili, Irene Grazzini, Raffaele Natella, Paola Scalise, Antonio Barile, Roberto Grassi, and Domenico Albano. Diffusion-Weighted Imaging in Oncology: An Update. *Cancers*, 12(6):1493, June 2020.

[219] Alexandros Mittos, Bradley Malin, and Emiliano De Cristofaro. Systematizing Genome Privacy Research: A Privacy-Enhancing Technologies Perspective. 2017. Publisher: arXiv Version Number: 2.

[220] Diana Montezuma, Sara P. Oliveira, Pedro C. Neto, Domingos Oliveira, Ana Monteiro, Jaime S. Cardoso, and Isabel Macedo-Pinto. Annotating

for Artificial Intelligence Applications in Digital Pathology: A Practical Guide for Pathologists and Researchers. *Modern Pathology*, 36(4), April 2023. Publisher: Elsevier.

[221] Clara Mosquera-Lopez, Sos Agaian, Alejandro Velez-Hoyos, and Ian Thompson. Computer-Aided Prostate Cancer Diagnosis From Digitized Histopathology: A Review on Texture-Based Systems. *IEEE Reviews in Biomedical Engineering*, 8:98–113, 2015.

[222] Ramakrishnan Mukundan. Image Features Based on Characteristic Curves and Local Binary Patterns for Automated HER2 Scoring. *Journal of Imaging*, 4(2):35, February 2018.

[223] Heimo Müller, Georges Dagher, Martina Loibner, Cornelia Stumptner, Penelope Kungl, and Kurt Zatloukal. Biobanks for life sciences and personalized medicine: importance of standardization, biosafety, biosecurity, and data management. *Current Opinion in Biotechnology*, 65:45–51, October 2020.

[224] Kristen Naegle, Nancy R. Gough, and Michael B. Yaffe. Criteria for biological reproducibility: what does "n" mean? *Science Signaling*, 8(371):fs7, April 2015.

[225] Kunal Nagpal, Davis Foote, Fraser Tan, Yun Liu, Po-Hsuan Cameron Chen, David F. Steiner, Naren Manoj, Niels Olson, Jenny L. Smith, Arash Mohtashamian, Brandon Peterson, Mahul B. Amin, Andrew J. Evans, Joan W. Sweet, Carol Cheung, Theodorus van der Kwast, Ankur R. Sangoi, Ming Zhou, Robert Allan, Peter A. Humphrey, Jason D. Hipp, Krishna Gadepalli, Greg S. Corrado, Lily H. Peng, Martin C. Stumpe, and Craig H. Mermel. Development and Validation of a Deep Learning Algorithm for Gleason Grading of Prostate Cancer From Biopsy Specimens. *JAMA oncology*, 6(9):1372–1380, September 2020.

[226] Yang Nan, Javier Del Ser, Simon Walsh, Carola Schönlieb, Michael Roberts, Ian Selby, Kit Howard, John Owen, Jon Neville, Julien Guiot, Benoit Ernst, Ana Pastor, Angel Alberich-Bayarri, Marion I. Menzel, Sean Walsh, Wim Vos, Nina Flerin, Jean-Paul Charbonnier, Eva van Rikxoort, Avishek Chatterjee, Henry Woodruff, Philippe Lambin, Leonor Cerdá-Alberich, Luis Martí-Bonmatí, Francisco Herrera, and Guang Yang. Data harmonisation for information fusion in digital healthcare: A state-of-the-art systematic review, meta-analysis and future research directions. *Information Fusion*, 82:99–122, June 2022.

[227] Emanuele Neri and Daniele Regge. Imaging biobanks in oncology: European perspective. *Future Oncology (London, England)*, 13(5):433–441, February 2017.

[228] Kien Nguyen, Anindya Sarkar, and Anil K. Jain. Prostate cancer grading: use of graph cut and spatial arrangement of nuclei. *IEEE transactions on medical imaging*, 33(12):2254–2270, December 2014.

[229] Stephanie Nougaret, Hichem Tibermacine, Marion Tardieu, and Evis Sala. Radiomics: an Introductory Guide to What It May Foretell. *Current Oncology Reports*, 21(8):70, June 2019.

[230] Rita Noumeir. Benefits of the DICOM Structured Report. *Journal of Digital Imaging*, 19(4):295–306, December 2006.

[231] James P. B. O'Connor, Eric O. Aboagye, Judith E. Adams, Hugo J. W. L. Aerts, Sally F. Barrington, Ambros J. Beer, Ronald Boellaard, Sarah E. Bohndiek, Michael Brady, Gina Brown, David L. Buckley, Thomas L. Chenevert, Laurence P. Clarke, Sandra Collette, Gary J. Cook, Nandita M. deSouza, John C. Dickson, Caroline Dive, Jeffrey L. Evelhoch, Corinne Faivre-Finn, Ferdia A. Gallagher, Fiona J. Gilbert, Robert J. Gillies, Vicky Goh, John R. Griffiths, Ashley M. Groves, Steve Halligan, Adrian L. Harris, David J. Hawkes, Otto S. Hoekstra, Erich P. Huang, Brian F. Hutton, Edward F. Jackson, Gordon C. Jayson, Andrew Jones, Dow-Mu Koh, Denis Lacombe, Philippe Lambin, Nathalie Lassau, Martin O. Leach, Ting-Yim Lee, Edward L. Leen, Jason S. Lewis, Yan Liu, Mark F. Lythgoe, Prakash Manoharan, Ross J. Maxwell, Kenneth A. Miles, Bruno Morgan, Steve Morris, Tony Ng, Anwar R. Padhani, Geoff J. M. Parker, Mike Partridge, Arvind P. Pathak, Andrew C. Peet, Shonit Punwani, Andrew R. Reynolds, Simon P. Robinson, Lalitha K. Shankar, Ricky A. Sharma, Dmitry Soloviev, Sigrid Stroobants, Daniel C. Sullivan, Stuart A. Taylor, Paul S. Tofts, Gillian M. Tozer, Marcel van Herk, Simon Walker-Samuel, James Wason, Kaye J. Williams, Paul Workman, Thomas E. Yankeelov, Kevin M. Brindle, Lisa M. McShane, Alan Jackson, and John C. Waterton. Imaging biomarker roadmap for cancer studies. *Nature Reviews Clinical Oncology*, 14(3):169–186, March 2017.

[232] Nobuyuki Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, January 1979. Conference Name: IEEE Transactions on Systems, Man, and Cybernetics.

[233] Anwar R. Padhani. Dynamic contrast-enhanced MRI in clinical oncology: current status and future directions. *Journal of magnetic resonance imaging: JMRI*, 16(4):407–422, October 2002.

[234] Liron Pantanowitz, Navid Farahani, and Anil Parwani. Whole slide imaging in pathology: advantages, limitations, and emerging perspectives. *Pathology and Laboratory Medicine International*, page 23, June 2015.

[235] Liron Pantanowitz, Ashish Sharma, Alexis B. Carter, Tahsin Kurc, Alan Sussman, and Joel Saltz. Twenty Years of Digital Pathology: An Overview of the Road Travelled, What is on the Horizon, and the Emergence of Vendor-Neutral Archives. *Journal of Pathology Informatics*, 9(1):40, January 2018.

[236] Liron Pantanowitz, John H. Sinard, Walter H. Henricks, Lisa A. Fatheree, Alexis B. Carter, Lydia Contis, Bruce A. Beckwith, Andrew J. Evans, Avtar Lal, Anil V. Parwani, and College of American Pathologists Pathology and Laboratory Quality Center. Validating whole slide imaging for diagnostic purposes in pathology: guideline from the College of American Pathologists Pathology and Laboratory Quality Center. *Archives of Pathology & Laboratory Medicine*, 137(12):1710–1722, December 2013.

[237] Panagiotis Papadimitroulas, Lennart Brocki, Neo Christopher Chung, Wistan Marchadour, Franck Vermet, Laurent Gaubert, Vasilis Eleftheriadis, Dimitris Plachouris, Dimitris Visvikis, George C. Kagadis, and Mathieu Hatt. Artificial intelligence: Deep learning in oncological radiomics and challenges of interpretability and data harmonization. *Physica Medica*, 83:108–121, March 2021.

[238] Barbara Parodi. Biobanks: A Definition. In Deborah Mascalzoni, editor, *Ethics, Law and Governance of Biobanking*, volume 14, pages 15–19. Springer Netherlands, Dordrecht, 2015. Series Title: The International Library of Ethics, Law and Technology.

[239] Wiktor Paskal, Adriana M. Paskal, Tomasz Dębski, Maciej Gryziak, and Janusz Jaworowski. Aspects of Modern Biobank Activity - Comprehensive Review. *Pathology oncology research: POR*, 24(4):771–785, October 2018.

[240] L. Paul Chew. Constrained delaunay triangulations. *Algorithmica*, 4(1-4):97–108, June 1989.

[241] Jurgen Peerlings, Henry C. Woodruff, Jessica M. Winfield, Abdalla Ibrahim, Bernard E. Van Beers, Arend Heerschap, Alan Jackson, Joachim E. Wildberger, Felix M. Mottaghy, Nandita M. DeSouza, and Philippe Lambin. Stability of radiomics features in apparent diffusion coefficient maps from a multi-centre test-retest trial. *Scientific Reports*, 9(1):4800, March 2019.

[242] Carel F. W. Peeters, Caroline Übelhör, Steven W. Mes, Roland Martens, Thomas Koopman, Pim de Graaf, Floris H. P. van Velden, Ronald Boellaard, Jonas A. Castelijns, Dennis E. te Beest, Martijn W. Heymans, and Mark A. van de Wiel. Stable prediction with radiomics data. 2019. Publisher: arXiv Version Number: 1.

[243] Rémy Peyret, Ahmed Bouridane, Fouad Khelifi, Muhammad Atif Tahir, and Somaya Al-Maadeed. Automatic classification of colorectal and prostatic histologic tumor images using multiscale multispectral local binary pattern texture features and stacked generalization. *Neurocomputing*, 275:83–93, January 2018.

[244] John H. Phan, Chang F. Quo, Chihwen Cheng, and May Dongmei Wang. Multiscale Integration of -Omic, Imaging, and Clinical Data in Biomedical Informatics. *IEEE reviews in biomedical engineering*, 5:74–87, 2012.

[245] Matti Pietikäinen, Abdenour Hadid, Guoying Zhao, and Timo Ahonen. *Computer Vision Using Local Binary Patterns*, volume 40 of *Computational Imaging and Vision*. Springer London, London, 2011.

[246] Daniel Pinkel and Donna G. Albertson. Array comparative genomic hybridization and its applications in cancer. *Nature Genetics*, 37 Suppl:S11–17, June 2005.

[247] Whitney B. Pope and Garth Brandal. Conventional and advanced magnetic resonance imaging in patients with high-grade glioma. *The Quarterly Journal of Nuclear Medicine and Molecular Imaging*, 62(3), September 2018.

[248] A.J. Prager, N. Martinez, K. Beal, A. Omuro, Z. Zhang, and R.J. Young. Diffusion and Perfusion MRI to Differentiate Treatment-Related Changes Including Pseudoprogression from Recurrent Tumors in High-Grade Gliomas with Histopathologic Evidence. *American Journal of Neuroradiology*, 36(5):877–885, May 2015.

[249] G. Pruneri, A. Vingiani, V. Bagnardi, N. Rotmensz, A. De Rose, A. Palazzo, A. M. Colleoni, A. Goldhirsch, and G. Viale. Clinical validity of tumor-infiltrating lymphocytes analysis in patients with triple-negative breast cancer. *Annals of Oncology: Official Journal of the European Society for Medical Oncology*, 27(2):249–256, February 2016.

[250] Noémie Rabilloud, Pierre Allaume, Oscar Acosta, Renaud De Crevoisier, Raphael Bourgade, Delphine Loussouarn, Nathalie Rioux-Leclercq, Zineeddine Khene, Romain Mathieu, Karim Bensalah, Thierry Pecot, and Solene-Florence Kammerer-Jacquet. Deep Learning Methodologies Applied to Digital Pathology in Prostate Cancer: A Systematic Review. *Diagnostics*, 13(16):2676, January 2023. Number: 16 Publisher: Multidisciplinary Digital Publishing Institute.

[251] J. Rajeswari and M. Jagannath. Advances in biomedical signal and image processing – A systematic review. *Informatics in Medicine Unlocked*, 8:13–19, 2017.

[252] Saima Rathore, Muhammad A. Iftikhar, Metin N. Gurcan, and Zissimos Mourelatos. Radiopathomics: Integration of radiographic and histologic characteristics for prognostication in glioblastoma, September 2019. arXiv:1909.07581 [cs, eess].

[253] Saima Rathore, Tamim Niazi, Muhammad Aksam Iftikhar, and Ahmad Chaddad. Glioma Grading via Analysis of Digital Pathology Images Using Machine Learning. *Cancers*, 12(3):578, March 2020.

[254] Muhammad Imran Razzak, Saeeda Naz, and Ahmad Zaib. Deep Learning for Medical Image Processing: Overview, Challenges and the Future. In Nilanjan Dey, Amira S. Ashour, and Surekha Borra, editors, *Classification in BioApps: Automation of Decision Making*, Lecture Notes in Computational Vision and Biomechanics, pages 323–350. Springer International Publishing, Cham, 2018.

[255] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You Only Look Once: Unified, Real-Time Object Detection. 2015. Publisher: arXiv Version Number: 5.

[256] Beatriz Remeseiro and Veronica Bolon-Canedo. A review of feature selection methods in medical applications. *Computers in Biology and Medicine*, 112:103375, September 2019.

[257] Stefania Rizzo, Francesca Botta, Sara Raimondi, Daniela Origgi, Cristiana Fanciullo, Alessio Giuseppe Morganti, and Massimo Bellomi. Radiomics: the facts and the challenges of image analysis. *European Radiology Experimental*, 2(1):36, November 2018.

[258] Fernando Roberti de Siqueira, William Robson Schwartz, and Helio Pedrini. Multi-scale gray level co-occurrence matrices for texture description. *Neurocomputing*, 120:336–345, November 2013.

[259] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, volume 9351, pages 234–241. Springer International Publishing, Cham, 2015. Series Title: Lecture Notes in Computer Science.

[260] Jeffrey N. Rouder, Paul L. Speckman, Dongchu Sun, Richard D. Morey, and Geoffrey Iverson. Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16(2):225–237, April 2009.

[261] Somak Roy, Christopher Coldren, Arivarasan Karunamurthy, Nefize S. Kip, Eric W. Klee, Stephen E. Lincoln, Annette Leon, Mrudula Pullambhatla, Robyn L. Temple-Smolkin, Karl V. Voelkerding, Chen Wang, and

Alexis B. Carter. Standards and Guidelines for Validating Next-Generation Sequencing Bioinformatics Pipelines. *The Journal of Molecular Diagnostics*, 20(1):4–27, January 2018.

[262] Cynthia Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5):206–215, May 2019.

[263] Aaron Michael Rulseh, Jiří Keller, Jan Klener, Jan Šroubek, Vladimír Dbalý, Martin Syrůček, František Tovaryš, and Josef Vymazal. Long-term survival of patients suffering from glioblastoma multiforme treated with tumor-treating fields. *World Journal of Surgical Oncology*, 10(1):220, December 2012.

[264] Vincenzo Salemme, Giorgia Centonze, Federica Cavallo, Paola Defilippi, and Laura Conti. The Crosstalk Between Tumor Cells and the Immune Microenvironment in Breast Cancer: Implications for Immunotherapy. *Frontiers in Oncology*, 11, 2021.

[265] R. Salgado, C. Denkert, S. Demaria, N. Sirtaine, F. Klauschen, G. Pruneri, S. Wienert, G. Van den Eynden, F. L. Baehner, F. Penault-Llorca, E. A. Perez, E. A. Thompson, W. F. Symmans, A. L. Richardson, J. Brock, C. Criscitiello, H. Bailey, M. Ignatiadis, G. Floris, J. Sparano, Z. Kos, T. Nielsen, D. L. Rimm, K. H. Allison, J. S. Reis-Filho, S. Loibl, C. Sotiriou, G. Viale, S. Badve, S. Adams, K. Willard-Gallo, and S. Loi. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. *Annals of Oncology*, 26(2):259–271, February 2015.

[266] Joel Saltz, Jonas Almeida, Yi Gao, Ashish Sharma, Erich Bremer, Tammy DiPrima, Mary Saltz, Jayashree Kalpathy-Cramer, and Tahsin Kurc. Towards Generation, Management, and Exploration of Combined Radiomics and Pathomics Datasets for Cancer Research. *AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science*, 2017:85–94, 2017.

[267] Joel Saltz, Rajarsi Gupta, Le Hou, Tahsin Kurc, Pankaj Singh, Vu Nguyen, Dimitris Samaras, Kenneth R. Shroyer, Tianhao Zhao, Rebecca Batiste, John Van Arnam, Samantha J. Caesar-Johnson, John A. Demchok, Ina Felau, Melpomeni Kasapi, Martin L. Ferguson, Carolyn M. Hutter, Heidi J. Sofia, Roy Tarnuzzer, Zhining Wang, Liming Yang, Jean C. Zenklusen, Jiashan (Julia) Zhang, Sudha Chudamani, Jia Liu, Laxmi Lolla, Rashi Naresh, Todd Pihl, Qiang Sun, Yunhu Wan, Ye Wu, Juok Cho, Timothy DeFreitas, Scott Frazer, Nils Gehlenborg, Gad Getz, David I. Heiman,

Jaegil Kim, Michael S. Lawrence, Pei Lin, Sam Meier, Michael S. Noble, Gordon Saksena, Doug Voet, Hailei Zhang, Brady Bernard, Nyasha Chambwe, Varsha Dhankani, Theo Knijnenburg, Roger Kramer, Kalle Leinonen, Yuexin Liu, Michael Miller, Sheila Reynolds, Ilya Shmulevich, Vesteinn Thorsson, Wei Zhang, Rehan Akbani, Bradley M. Broom, Apurva M. Hegde, Zhenlin Ju, Rupa S. Kanchi, Anil Korkut, Jun Li, Han Liang, Shiyun Ling, Wenbin Liu, Yiling Lu, Gordon B. Mills, Kwok-Shing Ng, Arvind Rao, Michael Ryan, Jing Wang, John N. Weinstein, Jiexin Zhang, Adam Abeshouse, Joshua Armenia, Debyani Chakravarty, Walid K. Chatila, Ino de Bruijn, Jianjiong Gao, Benjamin E. Gross, Zachary J. Heins, Ritika Kundra, Konnor La, Marc Ladanyi, Augustin Luna, Moriah G. Nissan, Angelica Ochoa, Sarah M. Phillips, Ed Reznik, Francisco Sanchez-Vega, Chris Sander, Nikolaus Schultz, Robert Sheridan, S. Onur Sumer, Yichao Sun, Barry S. Taylor, Jioajiao Wang, Hongxin Zhang, Pavana Anur, Myron Peto, Paul Spellman, Christopher Benz, Joshua M. Stuart, Christopher K. Wong, Christina Yau, D. Neil Hayes, Joel S. Parker, Matthew D. Wilkerson, Adrian Ally, Miruna Balasundaram, Reanne Bowlby, Denise Brooks, Rebecca Carlsen, Eric Chuah, Noreen Dhalla, Robert Holt, Steven J. M. Jones, Katayoon Kasaian, Darlene Lee, Yussanne Ma, Marco A. Marra, Michael Mayo, Richard A. Moore, Andrew J. Mungall, Karen Mungall, A. Gordon Robertson, Sara Sadeghi, Jacqueline E. Schein, Payal Sipahimalani, Angela Tam, Nina Thiessen, Kane Tse, Tina Wong, Ashton C. Berger, Rameen Beroukhim, Andrew D. Cherniack, Carrie Cibulskis, Stacey B. Gabriel, Galen F. Gao, Gavin Ha, Matthew Meyerson, Steven E. Schumacher, Juliann Shih, Melanie H. Kucherlapati, Raju S. Kucherlapati, Stephen Baylin, Leslie Cope, Ludmila Danilova, Moiz S. Bootwalla, Phillip H. Lai, Dennis T. Maglinte, David J. Van Den Berg, Daniel J. Weisenberger, J. Todd Auman, Saianand Balu, Tom Bodenheimer, Cheng Fan, Katherine A. Hoadley, Alan P. Hoyle, Stuart R. Jefferys, Corbin D. Jones, Shaowu Meng, Piotr A. Mieczkowski, Lisle E. Mose, Amy H. Perou, Charles M. Perou, Jeffrey Roach, Yan Shi, Janae V. Simons, Tara Skelly, Matthew G. Soloway, Donghui Tan, Umadevi Veluvolu, Huihui Fan, Toshinori Hinoue, Peter W. Laird, Hui Shen, Wanding Zhou, Michelle Bellair, Kyle Chang, Kyle Covington, Chad J. Creighton, Huyen Dinh, HarshaVardhan Doddapaneni, Lawrence A. Donehower, Jennifer Drummond, Richard A. Gibbs, Robert Glenn, Walker Hale, Yi Han, Jianhong Hu, Viktoriya Korchina, Sandra Lee, Lora Lewis, Wei Li, Xiuping Liu, Margaret Morgan, Donna Morton, Donna Muzny, Jireh Santibanez, Margi Sheth, Eve Shinbrot, Linghua Wang, Min Wang, David A. Wheeler, Liu Xi, Fengmei Zhao, Julian Hess, Elizabeth L. Appelbaum, Matthew Bailey, Matthew G. Cordes, Li Ding, Catrina C. Fronick, Lucinda A. Fulton, Robert S. Fulton,

Cyriac Kandoth, Elaine R. Mardis, Michael D. McLellan, Christopher A. Miller, Heather K. Schmidt, Richard K. Wilson, Daniel Crain, Erin Curley, Johanna Gardner, Kevin Lau, David Mallery, Scott Morris, Joseph Paulauskis, Robert Penny, Candace Shelton, Troy Shelton, Mark Sherman, Eric Thompson, Peggy Yena, Jay Bowen, Julie M. Gastier-Foster, Mark Gerken, Kristen M. Leraas, Tara M. Lichtenberg, Nilsa C. Ramirez, Lisa Wise, Erik Zmuda, Niall Corcoran, Tony Costello, Christopher Hovens, Andre L. Carvalho, Ana C. de Carvalho, José H. Fregnani, Adhemar Longatto-Filho, Rui M. Reis, Cristovam Scapulatempo-Neto, Henrique C. S. Silveira, Daniel O. Vidal, Andrew Burnette, Jennifer Eschbacher, Beth Hermes, Ardene Noss, Rosy Singh, Matthew L. Anderson, Patricia D. Castro, Michael Ittmann, David Huntsman, Bernard Kohl, Xuan Le, Richard Thorp, Chris Andry, Elizabeth R. Duffy, Vladimir Lyadov, Oxana Paklina, Galiya Setdikova, Alexey Shabunin, Mikhail Tavobilov, Christopher McPherson, Ronald Warnick, Ross Berkowitz, Daniel Cramer, Colleen Feltmate, Neil Horowitz, Adam Kibel, Michael Muto, Chandrajit P. Raut, Andrei Malykh, Jill S. Barnholtz-Sloan, Wendi Barrett, Karen Devine, Jordonna Fulop, Quinn T. Ostrom, Kristen Shimmel, Yingli Wolinsky, Andrew E. Sloan, Agostino De Rose, Felice Giuliante, Marc Goodman, Beth Y. Karlan, Curt H. Hagedorn, John Eckman, Jodi Harr, Jerome Myers, Kelinda Tucker, Leigh Anne Zach, Brenda Deyarmin, Hai Hu, Leonid Kvecher, Caroline Larson, Richard J. Mural, Stella Somiari, Ales Vicha, Tomas Zelinka, Joseph Bennett, Mary Iacocca, Brenda Rabeno, Patricia Swanson, Mathieu Latour, Louis Lacombe, Bernard Têtu, Alain Bergeron, Mary McGraw, Susan M. Staugaitis, John Chabot, Hanina Hibshoosh, Antonia Sepulveda, Tao Su, Timothy Wang, Olga Potapova, Olga Voronina, Laurence Desjardins, Odette Mariani, Sergio Roman-Roman, Xavier Sastre, Marc-Henri Stern, Feixiong Cheng, Sabina Signoretti, Andrew Berchuck, Darell Bigner, Eric Lipp, Jeffrey Marks, Shannon McCall, Roger McLendon, Angeles Secord, Alexis Sharp, Madhusmita Behera, Daniel J. Brat, Amy Chen, Keith Delman, Seth Force, Fadlo Khuri, Kelly Magliocca, Shishir Maithel, Jeffrey J. Olson, Taofeek Owonikoko, Alan Pickens, Suresh Ramalingam, Dong M. Shin, Gabriel Sica, Erwin G. Van Meir, Hongzheng Zhang, Wil Eijckenboom, Ad Gillis, Esther Korpershoek, Leendert Looijenga, Wolter Oosterhuis, Hans Stoop, Kim E. van Kessel, Ellen C. Zwarthoff, Chiara Calatozzolo, Lucia Cuppini, Stefania Cuzzubbo, Francesco DiMeco, Gaetano Finocchiaro, Luca Mattei, Alessandro Perin, Bianca Pollo, Chu Chen, John Houck, Pawadee Lohavanichbutr, Arndt Hartmann, Christine Stoehr, Robert Stoehr, Helge Taubert, Sven Wach, Bernd Wullich, Witold Kycler, Dawid Murawa, Maciej Wiznerowicz, Ki Chung, W. Jeffrey Edenfield, Julie Martin, Eric Baudin, Glenn Bubley,

Raphael Bueno, Assunta De Rienzo, William G. Richards, Steven Kalkanis, Tom Mikkelsen, Houtan Noushmehr, Lisa Scarpace, Nicolas Girard, Marta Aymerich, Elias Campo, Eva Giné, Armando López Guillermo, Nguyen Van Bang, Phan Thi Hanh, Bui Duc Phu, Yufang Tang, Howard Colman, Kimberley Evason, Peter R. Dottino, John A. Martignetti, Hani Gabra, Hartmut Juhl, Teniola Akeredolu, Serghei Stepa, Dave Hoon, Keunsoo Ahn, Koo Jeong Kang, Felix Beuschlein, Anne Breggia, Michael Birrer, Debra Bell, Mitesh Borad, Alan H. Bryce, Erik Castle, Vishal Chandan, John Cheville, John A. Copland, Michael Farnell, Thomas Flotte, Nasra Giama, Thai Ho, Michael Kendrick, Jean-Pierre Kocher, Karla Kopp, Catherine Moser, David Nagorney, Daniel O'Brien, Brian Patrick O'Neill, Tushar Patel, Gloria Petersen, Florencia Que, Michael Rivera, Lewis Roberts, Robert Smallridge, Thomas Smyrk, Melissa Stanton, R. Houston Thompson, Michael Torbenson, Ju Dong Yang, Lizhi Zhang, Fadi Brimo, Jaffer A. Ajani, Ana Maria Angulo Gonzalez, Carmen Behrens, Jolanta Bondaruk, Russell Broaddus, Bogdan Czerniak, Bita Esmaeli, Junya Fujimoto, Jeffrey Gershenwald, Charles Guo, Alexander J. Lazar, Christopher Logothetis, Funda Meric-Bernstam, Cesar Moran, Lois Ramondetta, David Rice, Anil Sood, Pheroze Tamboli, Timothy Thompson, Patricia Troncoso, Anne Tsao, Ignacio Wistuba, Candace Carter, Lauren Haydu, Peter Hersey, Valerie Jakrot, Hojabr Kakavand, Richard Kefford, Kenneth Lee, Georgina Long, Graham Mann, Michael Quinn, Robyn Saw, Richard Scolyer, Kerwin Shannon, Andrew Spillane, onathan Stretch, Maria Synott, John Thompson, James Wilmott, Hikmat Al-Ahmadie, Timothy A. Chan, Ronald Ghossein, Anuradha Gopalan, Douglas A. Levine, Victor Reuter, Samuel Singer, Bhuvanesh Singh, Nguyen Viet Tien, Thomas Broudy, Cyrus Mirsaidi, Praveen Nair, Paul Drwiega, Judy Miller, Jennifer Smith, Howard Zaren, Joong-Won Park, Nguyen Phi Hung, Electron Kebebew, W. Marston Linehan, Adam R. Metwalli, Karel Pacak, Peter A. Pinto, Mark Schiffman, Laura S. Schmidt, Cathy D. Vocke, Nicolas Wentzensen, Robert Worrell, Hannah Yang, Marc Moncrieff, Chandra Goparaju, Jonathan Melamed, Harvey Pass, Natalia Botnariuc, Irina Caraman, Mircea Cernat, Inga Chemencedji, Adrian Clipca, Serghei Doruc, Ghenadie Gorincioi, Sergiu Mura, Maria Pirtac, Irina Stancul, Diana Tcaciuc, Monique Albert, Iakovina Alexopoulou, Angel Arnaout, John Bartlett, Jay Engel, Sebastien Gilbert, Jeremy Parfitt, Harman Sekhon, George Thomas, Doris M. Rassl, Robert C. Rintoul, Carlo Bifulco, Raina Tamakawa, Walter Urba, Nicholas Hayward, Henri Timmers, Anna Antenucci, Francesco Facciolo, Gianluca Grazi, Mirella Marino, Roberta Merola, Ronald de Krijger, Anne-Paule Gimenez-Roqueplo, Alain Piché, Simone Chevalier, Ginette McKercher, Kivanc Birsoy, Gene Barnett, Cathy

Brewer, Carol Farver, Theresa Naska, Nathan A. Pennell, Daniel Raymond, Cathy Schilero, Kathy Smolenski, Felicia Williams, Carl Morrison, Jeffrey A. Borgia, Michael J. Liptay, Mark Pool, Christopher W. Seder, Kerstin Junker, Larsson Omberg, Mikhail Dinkin, George Manikhas, Domenico Alvaro, Maria Consiglia Bragazzi, Vincenzo Cardinale, Guido Carpino, Eugenio Gaudio, David Chesla, Sandra Cottingham, Michael Dubina, Fedor Moiseenko, Renumathy Dhanasekaran, Karl-Friedrich Becker, Klaus-Peter Janssen, Julia Slotta-Huspenina, Mohamed H. Abdel-Rahman, Dina Aziz, Sue Bell, Colleen M. Cebulla, Amy Davis, Rebecca Duell, J. Bradley Elder, Joe Hilty, Bahavna Kumar, James Lang, Norman L. Lehman, Randy Mandt, Phuong Nguyen, Robert Pilarski, Karan Rai, Lynn Schoenfield, Kelly Senecal, Paul Wakely, Paul Hansen, Ronald Lechan, James Powers, Arthur Tischler, William E. Grizzle, Katherine C. Sexton, Alison Kastl, Joel Henderson, Sima Porten, Jens Waldmann, Martin Fassnacht, Sylvia L. Asa, Dirk Schadendorf, Marta Couce, Markus Graefen, Hartwig Huland, Guido Sauter, Thorsten Schlomm, Ronald Simon, Pierre Tennstedt, Oluwole Olabode, Mark Nelson, Oliver Bathe, Peter R. Carroll, June M. Chan, Philip Disaia, Pat Glenn, Robin K. Kelley, Charles N. Landen, Joanna Phillips, Michael Prados, Jeffry Simko, Karen Smith-McCune, Scott VandenBerg, Kevin Roggin, Ashley Fehrenbach, Ady Kendler, Suzanne Sifri, Ruth Steele, Antonio Jimeno, Francis Carey, Ian Forgie, Massimo Mannelli, Michael Carney, Brenda Hernandez, Benito Campos, Christel Herold-Mende, Christin Jungk, Andreas Unterberg, Andreas von Deimling, Aaron Bossler, Joseph Galbraith, Laura Jacobus, Michael Knudson, Tina Knutson, Deqin Ma, Mohammed Milhem, Rita Sigmund, Andrew K. Godwin, Rashna Madan, Howard G. Rosenthal, Clement Adebamowo, Sally N. Adebamowo, Alex Boussioutas, David Beer, Thomas Giordano, Anne-Marie Mes-Masson, Fred Saad, Therese Bocklage, Lisa Landrum, Robert Mannel, Kathleen Moore, Katherine Moxley, Russel Postier, Joan Walker, Rosemary Zuna, Michael Feldman, Federico Valdivieso, Rajiv Dhir, James Luketich, Edna M. Mora Pinero, Mario Quintero-Aguilo, Carlos Gilberto Carlotti, Jose Sebastião Dos Santos, Rafael Kemp, Ajith Sankarankuty, Daniela Tirapelli, James Catto, Kathy Agnew, Elizabeth Swisher, Jenette Creaney, Bruce Robinson, Carl Simon Shelley, Eryn M. Godwin, Sara Kendall, Cassaundra Shipman, Carol Bradford, Thomas Carey, Andrea Haddad, Jeffey Moyer, Lisa Peterson, Mark Prince, Laura Rozek, Gregory Wolf, Rayleen Bowman, Kwun M. Fong, Ian Yang, Robert Korst, W. Kimryn Rathmell, J. Leigh Fantacone-Campbell, Jeffrey A. Hooke, Albert J. Kovatich, Craig D. Shriver, John DiPersio, Bettina Drake, Ramaswamy Govindan, Sharon Heath, Timothy Ley, Brian Van Tine, Peter Westervelt, Mark A. Rubin, Jung Il Lee, Natália D. Aredes, Armaz Mariamidze,

Ilya Shmulevich, Arvind U. K. Rao, Alexander J. Lazar, Ashish Sharma, and Vésteinn Thorsson. Spatial Organization and Molecular Correlation of Tumor-Infiltrating Lymphocytes Using Deep Learning on Pathology Images. *Cell Reports*, 23(1):181–193.e7, April 2018.

[268] Massimo Salvi, U. Rajendra Acharya, Filippo Molinari, and Kristen M. Meiburger. The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis. *Computers in Biology and Medicine*, 128:104129, January 2021.

[269] Massimo Salvi and Filippo Molinari. Multi-tissue and multi-scale approach for nuclei segmentation in H&E stained images. *BioMedical Engineering OnLine*, 17(1):89, December 2018.

[270] Camilla Scapicchio, Michela Gabelloni, Sara Maria Forte, Leonor Cerdá Alberich, Lorenzo Faggioni, Rita Borgheresi, Paola Erba, Fabiola Paiar, Luis Marti-Bonmati, and Emanuele Neri. DICOM-MIABIS integration model for biobanks: a use case of the EU PRIMAGE project. *European Radiology Experimental*, 5(1):20, December 2021.

[271] Caroline A. Schneider, Wayne S. Rasband, and Kevin W. Eliceiri. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7):671–675, July 2012.

[272] Muhammad Shaban, Shan E. Ahmed Raza, Mariam Hassan, Arif Jamshed, Sajid Mushtaq, Asif Loya, Nikolaos Batis, Jill Brooks, Paul Nankivell, Neil Sharma, Max Robinson, Hisham Mehanna, Syed Ali Khurram, and Nasir Rajpoot. A digital score of tumour-associated stroma infiltrating lymphocytes predicts survival in head and neck squamous cell carcinoma. *The Journal of Pathology*, 256(2):174–185, February 2022.

[273] Lizhi Shao, Zhenyu Liu, Lili Feng, Xiaoying Lou, Zhenhui Li, Xiao-Yan Zhang, Xiangbo Wan, Xuezhi Zhou, Kai Sun, Da-Fu Zhang, Lin Wu, Guanyu Yang, Ying-Shi Sun, Ruihua Xu, Xinjuan Fan, and Jie Tian. Multiparametric MRI and Whole Slide Image-Based Pretreatment Prediction of Pathological Response to Neoadjuvant Chemoradiotherapy in Rectal Cancer: A Multicenter Radiopathomic Study. *Annals of Surgical Oncology*, 27(11):4296–4306, October 2020.

[274] D.M. Shaw, B.S. Elger, and F. Colledge. What is a biobank? Differing definitions among biobank stakeholders. *Clinical Genetics*, 85(3):223–227, March 2014.

[275] Zhenwei Shi, Alberto Traverso, Johan Van Soest, Andre Dekker, and Leonard Wee. Technical Note: Ontology-guided radiomics analysis workflow (O-RAW). *Medical Physics*, 46(12):5677–5684, December 2019.

[276] Ben Shofty, Moran Artzi, Shai Shtrozberg, Claudia Fanizzi, Francesco DiMeco, Oz Haim, Shira Peleg Hason, Zvi Ram, Dafna Ben Bashat, and Rachel Grossman. Virtual biopsy using MRI radiomics for prediction of BRAF status in melanoma brain metastasis. *Scientific Reports*, 10(1):6623, April 2020.

[277] Emma Short, Anne Y. Warren, and Murali Varma. Gleason grading of prostate cancer: a pragmatic approach. *Diagnostic Histopathology*, 25(10):371–378, October 2019.

[278] Xin Shu, Zhigang Song, Jinlong Shi, Shucheng Huang, and Xiao-Jun Wu. Multiple channels local binary pattern for color texture representation and classification. *Signal Processing: Image Communication*, 98:116392, October 2021.

[279] Bennett O.V. Shum, Ilya Henner, Daniele Belluoccio, and Marcus J. Hinchcliffe. Utility of NIST Whole-Genome Reference Materials for the Technical Validation of a Multigene Next-Generation Sequencing Test. *The Journal of Molecular Diagnostics*, 19(4):602–612, July 2017.

[280] Nikita Shvetsov, Morten Grønnesby, Edvard Pedersen, Kajsa Møllersen, Lill-Tove Rasmussen Busund, Ruth Schwienbacher, Lars Ailo Bongo, and Thomas Karsten Kilvaer. A Pragmatic Machine Learning Approach to Quantify Tumor-Infiltrating Lymphocytes in Whole Slide Images. *Cancers*, 14(12):2974, January 2022.

[281] Santa Slokenberga, Olga Tzortzatou, and Jane Reichel, editors. *GDPR and Biobanking: Individual Rights, Public Interest and Research Regulation across Europe*, volume 43 of *Law, Governance and Technology Series*. Springer International Publishing, Cham, 2021.

[282] Sirpa Soini. Finland on a Road towards a Modern Legal Biobanking Infrastructure. *European Journal of Health Law*, 20(3):289–294, 2013.

[283] James Sorace, Denise R Aberle, Dena Elimam, Silvana Lawvere, Ossama Tawfik, and W Dean Wallace. Integrating pathology and radiology disciplines: an emerging opportunity? *BMC Medicine*, 10(1):100, December 2012.

[284] Andrea Sottoriva, Inmaculada Spiteri, Sara G. M. Piccirillo, Anestis Touloumis, V. Peter Collins, John C. Marioni, Christina Curtis, Colin

Watts, and Simon Tavaré. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, 110(10):4009–4014, March 2013.

[285] Angelika M. Stefan, Quentin F. Gronau, Felix D. Schönbrodt, and Eric-Jan Wagenmakers. A tutorial on Bayes Factor Design Analysis using an informed prior. *Behavior Research Methods*, 51(3):1042–1058, June 2019.

[286] Taofik Ahmed Suleiman, Adetola Mary Tolulope, Funmilola Olanike Wuraola, Rotimi Olorunfemi, Wakeel Adekunle Kasali, Bright Odinaka Okorocha, Christian Dirisu, and Prisca Chinonso Njoku. Overview of Cancer Management—The Role of Medical Imaging and Machine Learning Techniques in Early Detection of Cancer: Prospects, Challenges, and Future Directions. *Open Access Library Journal*, 10(4):1–21, April 2023. Number: 4 Publisher: Scientific Research Publishing.

[287] Peng Sun, Jiehua He, Xue Chao, Keming Chen, Yuanyuan Xu, Qitao Huang, Jingping Yun, Mei Li, Rongzhen Luo, Jinbo Kuang, Huajia Wang, Haosen Li, Hui Hui, and Shuoyu Xu. A Computational Tumor-Infiltrating Lymphocyte Assessment Method Comparable with Visual Reporting Guidelines for Triple-Negative Breast Cancer. *EBioMedicine*, 70:103492, August 2021.

[288] Alexey Surov, Hans Jonas Meyer, and Andreas Wienke. Correlation between apparent diffusion coefficient (ADC) and cellularity is different in several tumors: a meta-analysis. *Oncotarget*, 8(35):59492–59499, August 2017.

[289] Reed T. Sutton, David Pincock, Daniel C. Baumgart, Daniel C. Sadowski, Richard N. Fedorak, and Karen I. Kroeker. An overview of clinical decision support systems: benefits, risks, and strategies for success. *npj Digital Medicine*, 3(1):1–10, February 2020. Number: 1 Publisher: Nature Publishing Group.

[290] Zaneta Swiderska-Chadaj, Hans Pinckaers, Mart van Rijthoven, Maschenka Balkenhol, Margarita Melnikova, Oscar Geessink, Quirine Manson, Mark Sherman, Antonio Polonia, Jeremy Parry, Mustapha Abubakar, Geert Litjens, Jeroen van der Laak, and Francesco Ciompi. Learning to detect lymphocytes in immunohistochemistry with deep learning. *Medical Image Analysis*, 58:101547, December 2019.

[291] Thomas D. V. Swinscow and Michael J. Campbell. *Statistics at square one*. BMJ Publ. Group, London, 10. ed., 3rd impr., repr edition, 2004.

[292] Filip Szczepankiewicz, Danielle van Westen, Elisabet Englund, Carl-Fredrik Westin, Freddy Ståhlberg, Jimmy Lätt, Pia C. Sundgren, and Markus Nilsson. The link between diffusion MRI and tumor heterogeneity: Mapping cell eccentricity and density by diffusional variance decomposition (DIVIDE). *NeuroImage*, 142:522–532, November 2016.

[293] Ali Tabesh and Mikhail Teverovskiy. Tumor Classification in Histological Images of Prostate Using Color Texture. In *2006 Fortieth Asilomar Conference on Signals, Systems and Computers*, pages 841–845, Pacific Grove, CA, USA, 2006. IEEE.

[294] Clare Tempany, Jagadeesan Jayender, Tina Kapur, Raphael Bueno, Alexandra Golby, Nathalie Agar, and Ferenc Jolesz. Multimodal Imaging for Improved Diagnosis and Treatment of Cancers. *Cancer*, 121(6):817–827, March 2015.

[295] Jeppe Thagaard, Elisabeth Specht Stovgaard, Line Grove Vognsen, Søren Hauberg, Anders Dahl, Thomas Ebstrup, Johan Doré, Rikke Egede Vincentz, Rikke Karlin Jepsen, Anne Roslind, Iben Kümler, Dorte Nielsen, and Eva Balslev. Automated Quantification of sTIL Density with H&E-Based Digital Image Analysis Has Prognostic Potential in Triple-Negative Breast Cancers. *Cancers*, 13(12):3050, January 2021.

[296] The Genome Standards Consortium, Robert M Bowers, Nikos C Kyrpides, Ramunas Stepanauskas, Miranda Harmon-Smith, Devin Doud, T B K Reddy, Frederik Schulz, Jessica Jarett, Adam R Rivers, Emiley A Eloe-Fadrosh, Susannah G Tringe, Natalia N Ivanova, Alex Copeland, Alicia Clum, Eric D Becraft, Rex R Malmstrom, Bruce Birren, Mircea Podar, Peer Bork, George M Weinstock, George M Garrity, Jeremy A Dodsworth, Shibu Yooseph, Granger Sutton, Frank O Glöckner, Jack A Gilbert, William C Nelson, Steven J Hallam, Sean P Jungbluth, Thijs J G Ettema, Scott Tighe, Konstantinos T Konstantinidis, Wen-Tso Liu, Brett J Baker, Thomas Rattei, Jonathan A Eisen, Brian Hedlund, Katherine D McMahon, Noah Fierer, Rob Knight, Rob Finn, Guy Cochrane, Ilene Karsch-Mizrachi, Gene W Tyson, Christian Rinke, Alla Lapidus, Folker Meyer, Pelin Yilmaz, Donovan H Parks, A Murat Eren, Lynn Schriml, Jillian F Banfield, Philip Hugenholtz, and Tanja Woyke. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nature Biotechnology*, 35(8):725–731, August 2017.

[297] Robert Tibshirani. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.

[298] Hamid Reza Tizhoosh and Liron Pantanowitz. Artificial Intelligence and Digital Pathology: Challenges and Opportunities. *Journal of Pathology Informatics*, 9:38, 2018.

[299] Yuri Tolkach, Tilmann Dohmgörgen, Marieta Toma, and Glen Kristiansen. High-accuracy prostate cancer pathology using deep learning. *Nature Machine Intelligence*, 2(7):411–418, July 2020.

[300] Michal R. Tomaszewski and Robert J. Gillies. The Biological Meaning of Radiomic Features. *Radiology*, 298(3):505–516, March 2021.

[301] Matteo Tortora, Ermanno Cordelli, Rosa Sicilia, Lorenzo Nibid, Edy Ippolito, Giuseppe Perrone, Sara Ramella, and Paolo Soda. RadioPathomics: Multimodal Learning in Non-Small Cell Lung Cancer for Adaptive Radiotherapy. *IEEE Access*, 11:47563–47578, 2023. Conference Name: IEEE Access.

[302] Pamela Tozzo, Arianna Delicati, Beatrice Marcante, and Luciana Caenazzo. Digital Biobanking and Big Data as a New Research Tool: A Position Paper. *Healthcare*, 11(13):1825, January 2023. Number: 13 Publisher: Multidisciplinary Digital Publishing Institute.

[303] William T. Tran, Katarzyna Jerzak, Fang-I. Lu, Jonathan Klein, Sami Tabbarah, Andrew Lagree, Tina Wu, Ivan Rosado-Mendez, Ethan Law, Khadijeh Saednia, and Ali Sadeghi-Naini. Personalized Breast Cancer Treatments Using Artificial Intelligence in Radiomics and Pathomics. *Journal of Medical Imaging and Radiation Sciences*, 50(4 Suppl 2):S32–S41, December 2019.

[304] Alberto Traverso, Leonard Wee, Andre Dekker, and Robert Gillies. Repeatability and Reproducibility of Radiomic Features: A Systematic Review. *International Journal of Radiation Oncology*Biology*Physics*, 102(4):1143–1158, November 2018.

[305] Eleftherios Trivizakis, Georgios S. Ioannidis, Ioannis Souglakos, Apostolos H. Karantanas, Maria Tzardi, and Kostas Marias. A neural pathomics framework for classifying colorectal cancer histopathology images based on wavelet multi-scale texture analysis. *Scientific Reports*, 11(1):15546, July 2021.

[306] Lambros Tselikas, Roger Sun, Samy Ammari, Laurent Dercle, Steve Yevich, Antoine Hollebecque, Maud Ngo-Camus, Claudio Nicotra, Eric Deutsch, Frederic Deschamps, and Thierry de Baere. Role of image-guided biopsy and radiomics in the age of precision medicine. *Chinese Clinical Oncology*, 8(6):57–57, December 2019.

[307] Olli Tynninen, Hannu J. Aronen, Marko Ruhala, Anders Paetau, Kristina Von Boguslawski, Oili Salonen, Juha Jääskeläinen, and Timo Paavonen. MRI Enhancement and Microvascular Density in Gliomas: Correlation with Tumor Cell Proliferation. *Investigative Radiology*, 34(6):427, June 1999.

[308] Jarle Urdal, Kjersti Engan, Vebjorn Kvikstad, and Emilius A.M. Janssen. Prognostic prediction of histopathological images by local binary patterns and RUSBoost. In *2017 25th European Signal Processing Conference (EUSIPCO)*, pages 2349–2353, Kos, August 2017. IEEE.

[309] Abhishek Vahadane, Tingying Peng, Amit Sethi, Shadi Albarqouni, Lichao Wang, Maximilian Baust, Katja Steiger, Anna Melissa Schlitter, Irene Esposito, and Nassir Navab. Structure-Preserving Color Normalization and Sparse Stain Separation for Histological Images. *IEEE Transactions on Medical Imaging*, 35(8):1962–1971, August 2016.

[310] Martin Vallières, Alex Zwanenburg, Bodgan Badic, Catherine Cheze Le Rest, Dimitris Visvikis, and Mathieu Hatt. Responsible Radiomics Research for Faster Clinical Translation. *Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine*, 59(2):189–193, February 2018.

[311] Joost J. M. van Griethuysen, Andriy Fedorov, Chintan Parmar, Ahmed Hosny, Nicole Aucoin, Vivek Narayan, Regina G. H. Beets-Tan, Jean-Christophe Fillion-Robin, Steve Pieper, and Hugo J. W. L. Aerts. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Research*, 77(21):e104–e107, November 2017.

[312] Janita E. Van Timmeren, Davide Cester, Stephanie Tanadini-Lang, Hatem Alkadhi, and Bettina Baessler. Radiomics in medical imaging—"how-to" guide and critical reflection. *Insights into Imaging*, 11(1):91, December 2020.

[313] Rami S. Vanguri, Jia Luo, Andrew T. Aukerman, Jacklynn V. Egger, Christopher J. Fong, Natally Horvat, Andrew Pagano, Jose De Arimateia Batista Araujo-Filho, Luke Geneslaw, Hira Rizvi, Ramon Sosa, Kevin M. Boehm, Soo-Ryum Yang, Francis M. Bodd, Katia Ventura, Travis J. Hollmann, Michelle S. Ginsberg, Jianjiong Gao, MSK MIND Consortium, Rami Vanguri, Matthew D. Hellmann, Jennifer L. Sauter, and Sohrab P. Shah. Multimodal integration of radiology, pathology and genomics for prediction of response to PD-(L)1 blockade in patients with non-small cell lung cancer. *Nature Cancer*, 3(10):1151–1164, August 2022.

[314] Mario Verdicchio, Valentina Brancato, Carlo Cavaliere, Francesco Isgrò, Marco Salvatore, and Marco Aiello. A pathomic approach for tumor-infiltrating lymphocytes classification on breast cancer digital pathology images. *Heliyon*, 9(3):e14371, March 2023.

[315] Gregory Verghese, Jochen K Lennerz, Danny Ruta, Wen Ng, Selvam Thavaraj, Kalliopi P Siziopikou, Threnesan Naidoo, Swapnil Rane, Roberto Salgado, Sarah E Pinder, and Anita Grigoriadis. Computational pathology in cancer diagnosis, prognosis, and prediction – present day and prospects. *The Journal of Pathology*, 260(5):551–563, 2023. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/path.6163.

[316] Charles Vesteghem, Rasmus Froberg Brøndum, Mads Sønderkær, Mia Sommer, Alexander Schmitz, Julie Støve Bødker, Karen Dybkær, Tarec Christoffer El-Galaly, and Martin Bøgsted. Implementing the FAIR Data Principles in precision oncology: review of supporting initiatives. *Briefings in Bioinformatics*, 21(3):936–945, May 2020.

[317] Mitko Veta, Paul J. van Diest, Robert Kornegoor, André Huisman, Max A. Viergever, and Josien P. W. Pluim. Automatic nuclei segmentation in H&E stained breast cancer histopathology images. *PloS One*, 8(7):e70221, 2013.

[318] Benjamin Voigt, Oliver Fischer, Bruno Schilling, Christian Krumnow, and Christian Herta. Investigation of semi- and self-supervised learning methods in the histopathological domain. *Journal of Pathology Informatics*, 14:100305, January 2023.

[319] Eric-Jan Wagenmakers, Richard D. Morey, and Michael D. Lee. Bayesian Benefits for the Pragmatic Researcher. *Current Directions in Psychological Science*, 25(3):169–176, June 2016.

[320] Liang-Bo Wang, Alla Karpova, Marina A. Gritsenko, Jennifer E. Kyle, Song Cao, Yize Li, Dmitry Rykunov, Antonio Colaprico, Joseph H. Rothstein, Runyu Hong, Vasileios Stathias, MacIntosh Cornwell, Francesca Petralia, Yige Wu, Boris Reva, Karsten Krug, Pietro Pugliese, Emily Kawaler, Lindsey K. Olsen, Wen-Wei Liang, Xiaoyu Song, Yongchao Dou, Michael C. Wendl, Wagma Caravan, Wenke Liu, Daniel Cui Zhou, Jiayi Ji, Chia-Feng Tsai, Vladislav A. Petyuk, Jamie Moon, Weiping Ma, Rosalie K. Chu, Karl K. Weitz, Ronald J. Moore, Matthew E. Monroe, Rui Zhao, Xiaolu Yang, Seungyeul Yoo, Azra Krek, Alexis Demopoulos, Houxiang Zhu, Matthew A. Wyczalkowski, Joshua F. McMichael, Brittany L. Henderson, Caleb M. Lindgren, Hannah Boekweg, Shuangjia Lu, Jessika Baral, Lijun Yao, Kelly G. Stratton, Lisa M. Bramer, Erika Zink, Sneha P. Couvillion, Kent J. Bloodsworth, Shankha Satpathy, Weiva Sieh, Simina M. Boca, Stephan Schürer, Feng Chen, Maciej Wiznerowicz, Karen A. Ketchum, Emily S. Boja, Christopher R. Kinsinger, Ana I. Robles, Tara Hiltke, Mathangi Thiagarajan, Alexey I. Nesvizhskii, Bing Zhang, D. R. Mani, Michele Ceccarelli, Xi S. Chen, Sandra L. Cottingham, Qing Kay Li, Albert H. Kim, David Fenyö, Kelly V. Ruggles, Henry Rodriguez, Mehdi Mesri, Samuel H.

Payne, Adam C. Resnick, Pei Wang, Richard D. Smith, Antonio Iavarone, Milan G. Chheda, Jill S. Barnholtz-Sloan, Karin D. Rodland, Tao Liu, Li Ding, and Clinical Proteomic Tumor Analysis Consortium. Proteogenomic and metabolomic characterization of human glioblastoma. *Cancer Cell*, 39(4):509–528.e20, April 2021.

[321] Lu Wang, Xinyi Chen, Lu Zhang, Long Li, YongBiao Huang, Yinan Sun, and Xianglin Yuan. Artificial intelligence in clinical decision support systems for oncology. *International Journal of Medical Sciences*, 20(1):79–86, January 2023.

[322] Xiangxue Wang, Vamsidhar Velcheti, Pranjal Vaidya, Kaustav Bera, Anant Madabhushi, Arjun Khunger, Pradnya Patil, and Humberto Choi. RaPtomics: integrating radiomic and pathomic features for predicting recurrence in early stage lung cancer. In Metin N. Gurcan and John E. Tomaszewski, editors, *Medical Imaging 2018: Digital Pathology*, page 21, Houston, United States, March 2018. SPIE.

[323] H-Erich Wichmann, Klaus A Kuhn, Melanie Waldenberger, Dominik Schmelcher, Simone Schuffenhauer, Thomas Meitinger, Sebastian H R Wurst, Gregor Lamla, Isabel Fortier, Paul R Burton, Leena Peltonen, Markus Perola, Andres Metspalu, Peter Riegman, Ulf Landegren, Michael J Taussig, Jan-Eric Litton, Martin N Fransson, Johann Eder, Anne Cambon-Thomsen, Jasper Bovenberg, Georges Dagher, Gert-Jan Van Ommen, Michael Griffith, Martin Yuille, and Kurt Zatloukal. Comprehensive catalog of European biobanks. *Nature Biotechnology*, 29(9):795–797, September 2011.

[324] Mark D. Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E. Bourne, Jildau Bouwman, Anthony J. Brookes, Tim Clark, Mercè Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris T. Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, Alasdair J. G. Gray, Paul Groth, Carole Goble, Jeffrey S. Grethe, Jaap Heringa, Peter A. C. 't Hoen, Rob Hooft, Tobias Kuhn, Ruben Kok, Joost Kok, Scott J. Lusher, Maryann E. Martone, Albert Mons, Abel L. Packer, Bengt Persson, Philippe Rocca-Serra, Marco Roos, Rene van Schaik, Susanna-Assunta Sansone, Erik Schultes, Thierry Sengstag, Ted Slater, George Strawn, Morris A. Swertz, Mark Thompson, Johan van der Lei, Erik van Mulligen, Jan Velterop, Andra Waagmeester, Peter Wittenburg, Katherine Wolstencroft, Jun Zhao, and Barend Mons. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1):160018, March 2016. Number: 1 Publisher: Nature Publishing Group.

[325] Marvin N. Wright and Andreas Ziegler. ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software*, 77:1–17, March 2017.

[326] Vicki Xafis and Markus K. Labude. Openness in Big Data and Data Repositories: The Application of an Ethics Framework for Big Data in Health and Research. *Asian Bioethics Review*, 11(3):255–273, September 2019.

[327] Yi-Bin Xi, Fan Guo, Zi-Liang Xu, Chen Li, Wei Wei, Ping Tian, Ting-Ting Liu, Lin Liu, Gang Chen, Jing Ye, Guang Cheng, Long-Biao Cui, Hong-Juan Zhang, Wei Qin, and Hong Yin. Radiomics signature: A potential biomarker for the prediction of MGMT promoter methylation in glioblastoma. *Journal of magnetic resonance imaging: JMRI*, 47(5):1380–1387, May 2018.

[328] Hongming Xu, Yoon Jin Cha, Jean R. Clemenceau, Jinhwan Choi, Sung Hak Lee, Jeonghyun Kang, and Tae Hyun Hwang. Spatial analysis of tumor-infiltrating lymphocytes in histological sections using deep learning techniques predicts survival in colorectal carcinoma. *The Journal of Pathology. Clinical Research*, 8(4):327–339, July 2022.

[329] Hongming Xu, Sunho Park, and Tae Hyun Hwang. Computerized Classification of Prostate Cancer Gleason Scores from Whole Slide Images. *IEEE/ACM transactions on computational biology and bioinformatics*, 17(6):1871–1882, 2020.

[330] Yongjun Xu, Xin Liu, Xin Cao, Changping Huang, Enke Liu, Sen Qian, Xingchen Liu, Yanjun Wu, Fengliang Dong, Cheng-Wei Qiu, Junjun Qiu, Keqin Hua, Wentao Su, Jian Wu, Huiyu Xu, Yong Han, Chenguang Fu, Zhigang Yin, Miao Liu, Ronald Roepman, Sabine Dietmann, Marko Virta, Fredrick Kengara, Ze Zhang, Lifu Zhang, Taolan Zhao, Ji Dai, Jialiang Yang, Liang Lan, Ming Luo, Zhaofeng Liu, Tao An, Bin Zhang, Xiao He, Shan Cong, Xiaohong Liu, Wei Zhang, James P. Lewis, James M. Tiedje, Qi Wang, Zhulin An, Fei Wang, Libo Zhang, Tao Huang, Chuan Lu, Zhipeng Cai, Fang Wang, and Jiabao Zhang. Artificial intelligence: A powerful paradigm for scientific research. *The Innovation*, 2(4):100179, November 2021.

[331] Faliu Yi, Lin Yang, Shidan Wang, Lei Guo, Chenglong Huang, Yang Xie, and Guanghua Xiao. Microvessel prediction in H&E Stained Pathology Images using fully convolutional neural networks. *BMC Bioinformatics*, 19(1):64, December 2018.

[332] Stephen S F Yip and Hugo J W L Aerts. Applications and limitations of radiomics. *Physics in Medicine and Biology*, 61(13):R150–R166, July 2016.

[333] Guangyuan Yu, Xuefei Li, Ting-Fang He, Tina Gruosso, Dongmei Zuo, Margarita Souleimanova, Valentina Muñoz Ramos, Atilla Omeroglu, Sarkis Meterissian, Marie-Christine Guiot, Li Yang, Yuan Yuan, Morag Park, Peter P. Lee, and Herbert Levine. Predicting Relapse in Patients With Triple Negative Breast Cancer (TNBC) Using a Deep-Learning Approach. *Frontiers in Physiology*, 11, 2020.

[334] Kun-Hsing Yu, Ce Zhang, Gerald J. Berry, Russ B. Altman, Christopher Ré, Daniel L. Rubin, and Michael Snyder. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nature Communications*, 7(1):12474, August 2016.

[335] Muhammad Mohsin Zafar, Zunaira Rauf, Anabia Sohail, Abdul Rehman Khan, Muhammad Obaidullah, Saddam Hussain Khan, Yeon Soo Lee, and Asifullah Khan. Detection of tumour infiltrating lymphocytes in CD3 and CD8 stained histopathological images using a two-phase deep CNN. *Photodiagnosis and Photodynamic Therapy*, 37:102676, March 2022.

[336] Zanfardino, Pane, Mirabelli, Salvatore, and Franzese. TCGA-TCIA Impact on Radiogenomics Cancer Research: A Systematic Review. *International Journal of Molecular Sciences*, 20(23):6033, November 2019.

[337] Mario Zanfardino, Monica Franzese, Katia Pane, Carlo Cavaliere, Serena Monti, Giuseppina Esposito, Marco Salvatore, and Marco Aiello. Bringing radiomics into a multi-omics framework for a comprehensive genotype–phenotype characterization of oncological diseases. *Journal of Translational Medicine*, 17(1):337, December 2019.

[338] Mark D. Zarella, Douglas Bowman, Famke Aeffner, Navid Farahani, Albert Xthona, Syeda Fatima Absar, Anil Parwani, Marilyn Bui, and Douglas J. Hartman. A Practical Guide to Whole Slide Imaging: A White Paper From the Digital Pathology Association. *Archives of Pathology & Laboratory Medicine*, 143(2):222–234, February 2019.

[339] Junfeng Zhang, Heng Liu, Haipeng Tong, Sumei Wang, Yizeng Yang, Gang Liu, and Weiguo Zhang. Clinical Applications of Contrast-Enhanced Perfusion MRI Techniques in Gliomas: Recent Advances and Current Challenges. *Contrast Media & Molecular Imaging*, 2017:1–27, 2017.

[340] Sen Zhao, Oleg Agafonov, Abdulrahman Azab, Tomasz Stokowy, and Eivind Hovig. Accuracy and efficiency of germline variant calling pipelines for human genome data. *Scientific Reports*, 10(1):20222, November 2020.

[341] Kelly H. Zou, Simon K. Warfield, Aditya Bharatha, Clare M. C. Tempany, Michael R. Kaus, Steven J. Haker, William M. Wells, Ferenc A. Jolesz, and

Ron Kikinis. Statistical validation of image segmentation quality based on a spatial overlap index. *Academic Radiology*, 11(2):178–189, February 2004.

[342] Alex Zwanenburg, Stefan Leger, Martin Vallières, and Steffen Löck. Image biomarker standardisation initiative. *Radiology*, 295(2):328–338, May 2020. arXiv:1612.07003 [cs, eess].

Author's publications

1. Brancato V, Cavaliere C, Garbino N, Isgrò F, Salvatore M, Aiello M. The relationship between radiomics and pathomics in Glioblastoma patients: Preliminary results from a cross-scale association study. Front Oncol. 2022 Oct 6;12:1005805. doi: 10.3389/fonc.2022.1005805.

2. Valentino M, Bianco V, Miccio L, Memmolo P, Brancato V, Libretti P, Gambacorta M, Salvatore M, Ferraro P. Beyond conventional microscopy: Observing kidney tissues by means of fourier ptychography. Front Physiol. 2023 Feb 13;14:1120099. doi: 10.3389/fphys.2023.1120099.

3. Verdicchio M, Brancato V, Cavaliere C, Isgrò F, Salvatore M, Aiello M. A pathomic approach for tumor-infiltrating lymphocytes classification on breast cancer digital pathology images. Heliyon. 2023 Mar 9;9(3):e14371. doi: 10.1016/j.heliyon.2023.e14371. PMID: 36950640.

4. Brancato V, Esposito G, Coppola L, Cavaliere C, Mirabelli P, Scapicchio C, Borgheresi R, Neri E, Salvatore M, Aiello M. Standardizing digital biobanks: integrating imaging, genomic, and clinical data for precision medicine. J Transl Med. 2024 Feb 5;22(1):136. doi: 10.1186/s12967-024-04891-8.