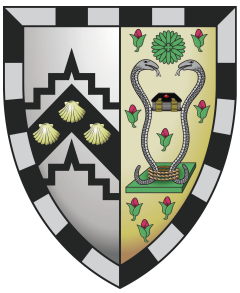
Novel multi-parametric, multi-modality imaging for the assessment of tumour biology in renal cell carcinoma



Stephan Ursprung

Department of Radiology
University of Cambridge

This thesis is submitted for the degree of Doctor of Philosophy

Gonville and Caius College August 2021

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit of 60,000 words excluding appendices, bibliography, tables and figures.

Stephan Ursprung
August 2021

Abstract

Novel multi-parametric, multi-modality imaging for the assessment of tumour biology in renal cell carcinoma

Stephan Ursprung

Renal cancer (RCC) is the 7th commonest cancer in the UK and clinically, morphologically, genetically and metabolically heterogeneous. Inaccurate patient stratification limits the treatment of RCC. This thesis investigates imaging techniques and image-analysis methods ranging from easily transferrable to highly specialised and dependent on dedicated infrastructure.

Artificial intelligence (AI) for diagnosis, prediction and prognostication of RCC experiences increasing scientific interest. However, methodological challenges prevent the validation and qualification of algorithms as imaging biomarkers. Texture features, quantitative descriptors of image composition, are sensitive to minor variations in the tumour segmentation, which future modelling approaches should consider. Twenty-one and 47% of features were poorly reproducible after manual and algorithmic re-segmentation.

Measuring response to systemic treatment becomes increasingly important with more therapeutic options available. The poor sensitivity of size-based criteria in targeted and immuno-oncology treatment is well-documented. Three clinical trials investigated morphological and physiological ^{1}H-MRI to detect early response and predict outcome. Multiparametric MRI demonstrated physiological changes after 12 days of anti-vascular therapy, which were compatible with its mechanism of action. The reduction in tumour diffusivity correlated with the long-term volumetric response and progression-free survival. Two trials demonstrated the ability of neoadjuvant antiangiogenic treatment to reduce the extent of venous tumour thrombi in 10/35 patients and enable a less morbid surgical approach in 41%.

The most common genetic alteration in RCC impacts metabolism including glycolysis. Hyperpolarised [1-^{13}C]pyruvate MRI (hpMRI) probed the spatially heterogeneous lactic acid fermentation in nine patients with renal tumours. Paired multi-regional tissue samples and imaging afforded a biological understanding of the hpMRI signal. The apparent reaction

rate constant (k_{PL}) correlated with the tumour grade, expression of MCT1, the membrane transporter taking up pyruvate, and total LDH, catalysing pyruvate-to-lactate conversion.

Improved image analysis tools, physiological and novel metabolic MRI can support precision oncology through better patient stratification and early treatment response detection. The WIRE window-of-opportunity trial of novel treatments in RCC has adopted computational and imaging techniques investigated here as primary, secondary and exploratory endpoints.

Acknowledgements

It takes a whole village to raise a child.

A PhD is hardly the achievement of a single person but the joint product of the dedication of countless people to my personal and academic development over the past 28 years. I am forever grateful to my parents for encouraging my scientific curiosity, to my teachers for sharpening my critical thinking, and to my mentors and colleagues for their advice and help. While I cannot possibly thank all the people whose support has enabled this PhD, I would like to highlight some who were instrumental in this chapter of my life.

Professor Ferdia Gallagher has been an outstanding supervisor and mentor for the past four years. I am immensely grateful for his guidance and insight which have shaped my PhD and will define my future academic and clinical career. I thank Professor Grant Stewart for co-supervising my research and helping me build the bridge between two areas of fascinating medical research. I count myself incredibly fortunate to have Professor Evis Sala as my mento, whose vision and advice will impact my work for years to come. Dr Ramona Woitek, thank you for your invaluable support and friendship over the past four years. I could not have wished for a kinder and more supportive advisor and colleague on this journey. The staff and scientists at the Department of Radiology and on the wider Cambridge Biomedical Campus have been a rich source of advice and knowledge which I was fortunate to benefit from and hopefully able to contribute to as well. I am particularly indebted to Dr Mireia Crispin Ortuzar, Dr Lorena Escudero Sanchez, Dr Marcel Gehrung, Dr Andrew Gill, Dr Mary McLean, Dr Andrew Priest, Dr Leonardo Rundo and Dr Anne Warren for supporting my work with their expertise. I thank Professor Tim Eisen and Dr Tristan Barrett for being my first-year examiners and helping me shape my PhD project. Finally, I am grateful to all the patients who have considered taking part in the research which led to this PhD. Finally, I thank my parents, Heidi and Ralph Ursprung, for their unwithering support throughout the past 24 years of my education.

Table of contents

List of figures

List of tables

Nomenclature

Gene/Protein Names

EGFR Epidermal Growth Factor Receptor

FH Fumarate Hydratase

FLCN Folliculin

GLUT1 Glucose Transporter 1

HIF Hypoxia-induced Factor

HK HexoKinase

MCT MonoCarboxylate Transporte

PDGF Platelet-derived Growth Factor

VEGF(R) Vascular Endothelial Growth Factor (Receptor)

VHL von Hippel-Lindau

Acronyms / Abbreviations

AI Artificial Intelligence

AIF Arterial Input Function

AMLwvf Angiomyolipoma without visible fat

BHD Birt-Hogg-Dubé Syndrome

BMI Body Mass Index

BOLD Blood Oxygen Level-Dependent

CAD Computer Assisted Detection/Diagnosis

ccRCC clear cell Renal Cell Carcinoma

CDC Center for Disease Control

IMDC International Metastatic RCC Database Consortium

ceCT Contrast-Enhanced Computed Tomography

CEUS Contrast-Enhanced Ultrasound

cfDNA cell-free DNA

chRCC chromophobe Renal Cell Carcinoma

CI Confidence Interval

CN Cytoreductive Nephrectomy

CR Complete Response

ctDNA circulating tumour DNA

D_0 Diffusivity

DCE Dynamic Contrast-Enhanced

DDR(i) DNA Damage Response inhibitor

DICOM Digital Imaging and Communications in Medicine

DL Deep Learning

DMSA DiMercapto Succinic Acid

DSC Sørensen-Dice similarity coefficient

DWI Diffusion-Weighted Imaging

EPI Echo Planar Imaging

FDG Fluorodeoxyglucose

FFPE Formalin-Fixed and Paraffin-Embedded

FOV Field-of-View

f_p Perfusion Fraction

GDI Gross Domestic Income

GLRLM Grey Level Run Length Matrix

GLSZM Grey Level Size Zone Matrix

hp hyperpolarised

HU Houndsfield Unit

$iAUC_{90}$ Area under the contrast-enhancement curve in the initial 90 seconds

ICC Interclass Correlation Coefficient

IMDC International Metastatic Renal-Cell Carcinoma Database Consortium

IO Immuno-oncology

LDH Lactate Dehydrogenase

IQR Interquartile Range

irRC immune-related Response Criteria

ISUP International Society of Urological Pathology

IVC Inferior Vena Cava

IVIM Intravoxel Incoherent Motion

Lac Lactate

Lac/Pyr Lactate-to-Pyruvate Ratio

LC-MS Liquid Chromatography-Mass Spectrometry

MCT1/4 Monocarboxylate Transporter 1/4

mRCC metastatic Renal Cell Carcinoma

MRI Magnetic Resonance Imaging

MRS Magnetic Resonance Spectroscopy

NADPH Nicotinamide Adenine Dinucleotide Phosphate

NMR Nuclear Magnetic Resonance

PARPi Poly (ADP-ribose) polymerase (inhibitor)

iCPD immune Confirmed Progressive Disease

iUPD immune Unconfirmed Progressive Disease

PD Progressive Disease

(i)PERCIST (immune) PET Response Criteria in Solid Tumours

PET Positron Emission Tomography

PFS Progression-free survival

p.i. post-injection

PN Partial Nephrectomy

PPP Pentose Phosphate Pathway

pRCC papillary Renal Cell Carcinoma

PR Partial Response

Pyr Pyruvate

QUADAS Quality Assessment of Diagnostic Accuracy Studies

RCC Renal Cell Carcinoma

(i)RECIST (immune)Response Evaluation Criteria in Solid Tumours

RF Pulse Radio Frequency Pulse

RN Radical Nephrectomy

ROC Receiver Operating Curve

ROI Regioin of Interest

ROS Reactive Oxygen Species

RQS Radiomics Quality Score

SACT Systemic Anticancer Therapy

SD Stable Disease

SNR Signal-to-Noise Ratio

SpSp Spectral Spatial

SUV Standardised Uptake Value

TCGA The Cancer Genome Atlas

TIL Tumour-Infiltrating Lymphocytes

TKI Tyrosine Kinase Inhibitor

VTT Venous Tumour Thrombus

WHO World Health Organisation

Chapter 1

Introduction

Renal cell cancer (RCC) is the 8th and 12th most common solid cancer type globally in men and women, respectively [1]. The heterogeneous appearance and variable clinical behaviour continue to pose challenges in the management of patients with renal cancer [2, 3]. Increasing knowledge on tumour genomics and metabolomics has contributed to our understanding of renal tumour biology. Concurrently, new operating techniques, minimally invasive treatment modalities and new systemic anti-tumour therapeutics have fundamentally changed how renal cancer is being treated [4, 5]. Imaging plays a crucial role throughout the renal cancer pathway. Particularly critical is its involvement in the diagnosis, treatment assessment and follow-up of these patients [6]. The co-evolution of new imaging techniques, molecular biomarkers and new therapeutic strategies towards personalised healthcare can improve patient outcomes.

1.1 Epidemiology

Globally, 403,202 people, of whom 245,507 (63%) were men, were newly diagnosed with renal cancer in 2018. Concurrently, 175,098 patients (65% male) died due to kidney cancer. The age-standardised incidence rate varies widely between 16.9/100,000 in Belarus and 0.17/100,000 in Equatorial Guinea, reaching a global average of 4.2/100,000. No genetic, lifestyle or environmental factors explaining the elevated rate of RCC in Eastern Europe could be identified [7]. In the United Kingdom, 13,683 new diagnoses, corresponding to an age-standardised incidence of 10.2/100,000, were observed [1]. Furthermore, countries with higher gross national income per capita (GDI) according to the world bank observe higher age-standardised incidence rates, ranging between 9.4/100,000 in countries with a GDI of $\geqslant$ USD 12,376 and 1.5/100,000 in countries with a GDI of $\leqslant$ USD 1.025. While the

incidence rate in higher-income countries is increased by $\geqslant$500%, the mortality rate is only double the one registered in low-income countries. Finally, the incidence rate of renal cancer is age-dependent, peaking between the age of 85 and 89 in the UK [8, 9].

Over recent years, the incidence of RCC has increased in many countries. This was particularly pronounced in Brazil where a year-on-year increase in renal cancer incidence of > 10% was observed between 1998 and 2008 [10]. In the UK, the RCC incidence increased by 87% between 1993 and 2017, equivalent to an average annual increase of 2.9% with those aged 80 years and above seeing the most dramatic increase [9]. This increase in incidence can be partially attributed to the increasing use of cross-sectional imaging and more frequent diagnosis of RCC as an incidental finding [11]. This observation is supported by the decreasing stage at diagnosis over time until recently and the lower stage in patients diagnosed incidentally compared to symptomatically [11, 12]. Increased use of medical imaging and earlier diagnosis, together with improved access to treatment, may also explain the disproportionately increased incidence of RCC in high-income countries compared to the increase in mortality [13]. It is expected that the incidence in RCC will continue to rise by another 28% until 2035 which is a greater increase than for other common cancer types like prostate, breast, lung, colorectal or pancreatic cancer [9].

Cancer-specific survival in RCC is heavily dependent on age and stage at diagnosis. While 87% of patients diagnosed with stage I RCC in the UK (48% of all diagnoses) survive for at least five years, 5-year overall survival falls to 12.4% in the 24% diagnosed with stage IV RCC. (RCC staging see: Table 1.1, page 3). Meanwhile, cancer-specific survival decreases from 84% in patients diagnosed before the age of 45 to 48% in patients diagnosed after the age of 75. Overall, 64% of patients diagnosed with RCC in the UK between 2013 and 2017 survived for 5 years or longer [14]. Compared to the ten previous years, this represents an increase in 5-year survival by 14%. All age groups experienced an improvement in cancer-specific survival, however, this was more pronounced in younger patients [15].

The increase in survival is attributed to better treatment with the availability of new targeted systemic therapeutics, immuno-oncology agents but also the increasing diagnosis of earlier stage RCC [12]. However, the earlier diagnosis also increased the risk of treating tumours that would not have become symptomatic during a patients lifetime. This is referred to as overdiagnosis and overtreatment [16].

Several independent risk factors for the development of renal cancer have been established in addition to patient age. Combined, they are thought to contribute to 34% of all RCC in the UK. Obesity has been associated with the development of clear cell renal cell carcinoma and chromophobe renal cell carcinoma [17] and each unit increase in body mass index

Table 1.1 TNM staging of renal cell carcinoma. IVC: Inferior Vena Cava. (Adapted from: Brierley *et al.*, TNM Classification of Malignant Tumours, 8th edition, 2018, Wiley-Blackwell, New York.)

Primary Tumour (T)			
TX/T0	Primary tumour cannot be assessed / No primary tumour		
T1	Greatest diameter $<$ 7 cm and no extrarenal extension		
T1a	Greatest diameter $<$ 4 cm		
T1b	Greatest diameter $>$ 4 and $\leq$ 7cm		
T2	Greatest diameter $>$ 7cm and no extrarenal extension		
T2a	Greatest diameter $>$ 7 and $\leq$ 10cm		
T2b	Greatest diameter $>$ 10 cm		
T3	Invasion of major veins and perinephric tissue excl. adrenal glands and not beyond Gerota's fascia		
T3a	Gross invasion of renal vein, muscle containing branches, perinephric or renal sinus fat		
T3b	Extension to infra-diaphragmatic IVC		
T3c	Extension to supra-diaphragmatic IVC or invasion of IVC wall		
T4	Invasion beyond Gerota's fascia		
Lymph nodes (N)			
NX/N0	Lymph node metastasis cannot be assessed / not present		
N1	Metastasis in one or more regional lymph nodes		
Distant metastasis (M)			
M0	No distant metastasis		
M1	Distant metastasis		
UICC cancer stage			
Stage I	T1	N0	M0
Stage II	T2	N0	M0
Stage III	T1-T2	N1	M0
	T3	any N	M0
Stage IV	T4	any N	any M
	any T	any N	M1

(BMI) confers a risk increase of 34% in men and 24% women in a dose-response fashion [18]. Consequently, 31.1% and 42.5% of all renal cancers in the European Union and the United States of America respectively are thought to be attributable to excess weight. The population attributable fraction in renal cancer is higher than in other cancer types commonly associated with obesity such as colorectal or breast cancer but lower than in oesophageal and endometrial cancer [19].

Similarly, hypertension predisposes to the development of RCC in a dose-dependent manner with a relative risk increase of $\leqslant 2.3$ in hypertensive patients with a systolic blood pressure $\leqslant$ 160mmHg or a diastolic blood pressure $\leqslant$ 100mmHg [20]. The relative risk is thought to increase by 12% and 22% for every 10mmHg of systolic and diastolic hypertension [21].

Smoking increases the risk of kidney cancer by 37% and 61% in women and men, respectively. There is a positive dose-response relationship between the number of cigarettes smoked and the excess risk of developing kidney cancer. Long term smoking cessation ($\geq$ 10 years) is associated with a risk reduction compared to continued smoking [22]. However, smoking is not only a risk factor for the development of renal cancer, smoking is also associated with shorter overall survival in patients with metastatic RCC (mRCC) undergoing cytoreductive nephrectomy [23] or targeted therapy [24].

Finally, patients with end-stage renal failure requiring haemodialysis and renal transplant recipients experience an increased incidence of renal cancer [25]. Genetic risk factors for the development of renal cancer will be discussed in more detail in section 1.3 Genomics and metabolomics in RCC.

1.2 Diagnostic and therapeutic approaches to RCC

1.2.1 Diagnostic workup

Most patients access renal cancer care following an incidental diagnosis on cross-sectional abdominal imaging which is partly explained by the increasing number of imaging procedures [11, 26]. Patients presenting with gross haematuria, flank pain or a palpable mass may be referred specifically for the investigation of these symptoms. Nonetheless, this classical clinical triad is observed in fewer than 10% of all patients with renal cancer and is associated with an inferior prognosis resulting from advanced disease and aggressive histology [27]. More frequently, in up to one-third of patients, RCC presents with paraneoplastic symptoms including fever, anaemia, fatigue, and weight loss, which are, however, non-specific [28].

Table 1.2 Bosniak classification of cystic renal lesions (Adapted from: Silverman *et al.*, Management of the Incidental Renal Mass, 2008, Radiology.

Bosniak Category	Description	Management
I	Thin wall without septa, calcification or solid components. Water enhancement and no enhancement.	No follow-up
II	Few thin septa without measurable contrast enhancement, fine calcification or a short segment of slightly thickened calcification. Homogeneously high-attenuating, well-defined masses ($\geqslant$3 cm) with no enhancement.	No follow-up
IIF	Multiple thin septa and minimal smooth thickening of wall or septa without measurable contrast enhancement. Calcifications may be thick or nodular. Intrarenal non-enhancing high-attenuation masses $\leqslant$3 cm.	5-year follow-up
III	Thickened (irregular) wall or septa with measurable contrast-enhancement.	Intervention
IV	Containing enhancing soft-tissue component.	Intervention

Renal lesions are a very common finding on cross-sectional abdominal imaging and can be found in over 60% of patients above the age of 80 [29]. The vast majority of these are simple renal cysts with a very low risk of harbouring a malignancy [30]. The Bosniak classification defines morphological criteria which are associated with malignant histology and serve as a means of stratifying patients according to the risk of malignancy (Table 1.2) [31]. Malignant histology is present in 0-5% of patients with Bosniak category I and II cysts, whereas category III and IV cysts show malignant histology in ~50% and ~90%, respectively [30, 32]. However, estimating the true rate of malignancy in the lower Bosniak categories is difficult as research is primarily based on retrospective surgical cohorts and CT/MRI datasets. This excludes the large group of patients with simple cysts found on ultrasound who were discharged and did not undergo any further imaging [32]. Recently, an updated classification algorithm has been proposed to include magnetic resonance imaging (MRI), improving inter-reader agreement and reducing the rate of benign histology in Bosniak III lesions. This updated version of the Bosniak classification is awaiting validation (Table 1.3) [33].

Standard of care imaging in RCC

The central role of imaging in the diagnosis, staging, surgical planning, response monitoring and follow-up is emphasised by the absence of validated diagnostic and prognostic molecular biomarkers specific to RCC [5, 34]. Therefore, non-invasive imaging tests with the ability to differentiate benign from malignant lesions could avoid unnecessary biopsies, overdiagnosis and overtreatment.

The differentiation of benign simple cysts and potentially malignant complex cystic or solid renal tumours is the initial step in the workup of patients with renal lesions and can be achieved with contrast-enhanced ultrasound (CEUS) or CT (ceCT) [35]. The most recent appropriateness criteria of the American College of Radiology suggest that CEUS, ceMRI or ceCT are equally suitable for the investigation of renal masses which were indeterminate on initial diagnostic imaging [36]. Meanwhile, current guidelines of the European Society of Oncology [4] and the European Society of Urology [5] recommend the use of abdominal ceCT or MRI. Additional CEUS or MRI may be considered for indeterminate lesions on CT [37] and MRI can be used to further characterise inferior vena cava (IVC) thrombus [38]. Furthermore, ceCT of the chest, abdomen and pelvis is appropriate for local, regional and distant staging in patients with renal masses [39].

Multiphase ceCT, as a primary tool for the diagnosis, classification and follow-up of renal lesions, consists of an unenhanced phase followed by an intravenous injection of iodinated contrast media and the acquisition of one or more contrast-enhanced phases. Guidelines recommend the acquisition of an unenhanced phase followed by a corticomedullary, nephrographic and delayed phase 25-30 s, 60-80 s, and 5-10 min after the injection of 100 -120 ml of 350 mg/ml iodinated contrast agent [40]. However, in the clinical setting four phases are rarely acquired. A slice thickness of 3 mm improves the diagnostic confidence in the detection of small renal lesions compared to 5 mm slices [41]. The unenhanced scan serves as a baseline to determine relative enhancement after contrast injection and for the differentiation of solid and cystic lesions as well as haemorrhage. During the corticomedullary phase, the renal arteries, veins, and cortex show maximum enhancement which aids surgical planning through the detection of multiple renal arteries and of the presence and extension of a potential tumour thrombus in the inferior vena cava. Most renal masses show increased early enhancement which is particularly apparent in ccRCC. Additionally, hypervascular metastases in the liver, spleen, and pancreas are best visualised in the corticomedullary phase. Most renal lesions show the best conspicuity during the nephrographic phase when small tumours are best-visualised [42]. The delayed phase allows the assessment of the collecting system which is essential before surgery and in the diagnosis of urothelial malignancies [40].

MRI may be used in the diagnostic workup of renal tumours for example in the presence of a tumour thrombus in the inferior vena cava, indeterminate lesions or contraindications to iodinated contrast agents. The Society of Abdominal Radiology proposes a renal mass MRI protocol consisting of an axial and/or coronal acquisition of a T_1w sequence, T_2w sequence, T_1w contrast-enhanced sequence with fat saturation and diffusion-weighted imaging (DWI) [43].

RCC metastasises most frequently to the lung, bone, lymph nodes, liver, adrenal glands, and brain (in order of descending frequency) [44]. Therefore, chest imaging with CT (European Association of Urology, National Comprehensive Cancer Network) or x-ray (American College of Radiology) is indicated as part of the staging to detect lung and skeletal metastases [5, 45, 36]. Chest CT is generally preferred over chest x-ray due to its increased sensitivity. However, universal chest CT is associated with a significant radiation exposure and algorithms for risk stratification based on tumour stage, systemic symptoms, and the platelet to haemoglobin ratio have been proposed. The number of chest CTs could be reduced by 37% and a sensitivity of 97% retained when only patients with a risk for pulmonary metastasis of $>1\%$ were scanned [46].

In patients who are presenting with symptomatic skeletal disease, a Technetium-99m (^{99m}Tc) methylene diphosphonate bone scan may be appropriate for the detection of bone metastases [5]. However, osteolytic bone metastases from a renal primary may show no increased tracer uptake, resulting in ^{99m}Tc bone scans only detecting 50-78% of the bone metastases detectable by ^{18}F-fluorodeoxyglucose positron emission tomography (FDG-PET) [47, 48] and 62% detectable by CT follow-up for at least 12 months [49]. In the same cohort, whole-body MRI with a T_1w and STIR sequence achieved a sensitivity of 94%. Furthermore, ^{18}F-sodium PET/CT has shown superior sensitivity in the detection of bone metastases compared to ^{99m}Tc and CT [50]. Consensus opinion suggests that bone metastases are generally symptomatic and, therefore, bone imaging is not recommended in asymptomatic patients. However, this is based on studies before chest CT was the standard of care and no recent, high-quality data is available to support this consensus [51].

Finally, brain imaging with CT or preferentially MRI is only recommended in symptomatic patients [4, 5]. Similar to bone metastasis, this recommendation is based on historical information, where 79% of patients with brain metastases observed on CT showed clinical symptoms and asymptomatic brain metastasis was only detected in 3% of the patients [52]. More recent retrospective analyses found occult brain metastasis in 4.67% of patients undergoing screening for various clinical trials [53].

In patients where surgical management is considered and the residual renal function is a concern owing to a to pre-existing reduction in the glomerular filtration rate or comorbidities likely leading to future renal impairment, ^{99m}Tc dimercaptosuccinic acid (^{99m}Tc-DMSA) renal scintigraphy can inform on split renal function [54, 5]. A reduced predicted renal function may direct therapy towards nephron-sparing partial nephrectomy or ablative treatment options.

Differentiation of RCC subtypes

Renal tumours are currently classified into 51 subtypes according to the 2016 WHO classification of Tumours of the Urinary System and Male Genital Organs [55]. The most frequent malignant primary tumour entities, the clear cell renal cell carcinoma (ccRCC), papillary renal cell carcinoma (pRCC), chromophobe renal cell carcinoma (chRCC) and collecting duct renal cell carcinoma (CDC), account for 70%, 10-15%, 4-6% and <1% respectively and arise from different cells within the nephron [56–58]. In addition to these epithelial tumours, mesenchymal, neuro-endocrine and nephrogenic blastemal cells can give rise to less common renal tumour subtypes [55]. As well as on tumour subtype, pathological diagnosis informs on tumour WHO/ISUP grade [59], tumour stage (Table 1.1) and prognostic factors like vascular invasion, necrosis, sarcomatoid features, as well as invasion of the collecting system and perinephric fat. The histological subtype is an independent predictor of survival and chRCC carries a better prognosis than pRCC which is prognostically superior to ccRCC [60–62]. While imaging characteristics of renal lesions may be suggestive of a specific histological tumour type, a definitive diagnosis can only be reached after histological confirmation. Therefore, biopsies should be performed in patients undergoing systemic anticancer therapy (SACT) or local ablative therapy and should be considered in patients with small lesions scheduled for surgery to avoid overtreatment [5, 34]. Renal mass biopsies achieve a diagnostic rate of 92% and exceed the sensitivity and specificity of fine-needle aspiration. The biopsy accurately captures the final histological subtype in 90.3%, however, under-grading due to sampling bias is common and the concordance rate for tumour grade reaches only 62%. Complications, most commonly bleeding, occur in about 8% of all renal mass biopsies and 99% of these are self-limiting (Clavien Dindo Grade 1) [63]. Routine biopsy of renal masses <4 cm can reduce the rate of benign histology following (partial) nephrectomy [64]. In contrast, biopsies are not recommended in patients who are too frail to receive any form of treatment.

Table 1.3 2019 adaptation of the Bosniak classification including criteria for MRI (Adapted from: Silverman *et al.*, Bosniak Classification of Cystic Renal Masses, Version 2019: An Update Proposal and Needs Assessment, 2019, Radiology.)

Bosniak Category	Description	
I	CT:	Smooth wall < 2mm without septa or calcification, homogeneous (-9 - 20 HU) attenuation. The wall may enhance.
	MRI:	Smooth wall < 2mm without septa or calcification, homogeneous simple fluid. The wall may enhance.
II	CT:	Smooth walls $\leqslant$ 2 mm and: (1) 1-3 thin ($\leqslant$ 2 mm) septa which may enhance and calcifications of any type. (2) homogeneus hyperattenuation $\leqslant$ 70 HU at non-contrast CT. (3) Homogeneous non-enhancing mass $>$ 20 HU with calcification of any type. (4) Homogeneous mass (-9 - 20 HU) at non-contrast CT. (5) Homogeneous mass (21 - 30 HU) at portal venous CT. (6) Homogeneous low-attenuation mass and too small to characterise.
	MRI:	Smooth walls $\leqslant$ 2 mm and: (1) 1-3 < 2mm enhancing septa, any non-enhancing septa, calcifications of any type. (2) homogeneous hyperintense T_2w signal. (3) homogeneous hyperintense T_1w signal (2.5 x kidney).
IIF	CT:	Smooth enhancing wall or septa ~3 mm or $\geqslant$ 4 smooth thin ($\leqslant$ 2 mm) enhancing septa .
	MRI:	(1) Smooth enhancing wall or septa ~3 mm or $\geqslant$ 4 smooth thin ($\leqslant$ 2 mm) enhancing septa . (2) Heterogeneous hyperintensity on fat-saturated T_1w sequence.
III	CT/ MRI:	Enhancing wall or septa $\geqslant$ 4 mm with $\leqslant$ 3 mm protrusions.
IV	CT/ MRI:	Enhancing nodule protruding $\geqslant$ 4 mm or any convex protrusion with acute margins.

1.2.2 Treatment of localised RCC

In RCC amenable to curative treatment, surgery is the treatment of choice in fit patients [65]. Radical (RN) or partial nephrectomy (PN) may be offered to patients as curative treatment options. Partial nephrectomy is the treatment option of choice in patients with small (T1) tumours or pre-existing renal function impairment as the preservation of renal function may prevent cardiovascular complications, delay the development of end-stage renal disease and the need for haemodialysis [66, 67]. The improved preservation of renal function comes at the expense of higher peri-operative complication rates [68]. Importantly, the outcome in terms of overall and cancer-specific survival following RN and PN is thought to be equivalent in the absence of any conclusive evidence suggesting otherwise [69]. However, PN is associated with a higher rate of local, ipsilateral recurrence [70]. Positive surgical margins at PN are associated with a 16% risk for local recurrence whereas 3% of the patients with a negative margin will experience local recurrence [71]. Additionally, PN may be a viable treatment option in selected patients with T2 RCC, offering comparable oncological outcomes to RN [72]. Strong evidence suggests ipsilateral adrenalectomy in patients where involvement is clinically evident whereas evidence for lymphadenectomy in patients with adverse features is weak [5].

Pre-surgical imaging plays an important role in personalised surgical planning and several scoring systems have been developed to select the most appropriate operative approach. The PADUA (preoperative aspects and dimensions used for anatomical classification in renal tumours) and R.E.N.A.L. nephrometry scores take the maximal tumour radius (R), its exophytic growth (E), how near it comes to the collecting system (N), its position relative to the kidney (A) and the interpolar lines (L) into consideration for a standardised description of renal tumour anatomy. Both scores are predictive of perioperative complications during partial nephrectomy [73–75]. Furthermore, the nephrometry score can inform treatment decisions [76]. Other scoring systems assess the intra-renal vascular involvement [77], the contact area between the tumour and surrounding [78], normal-appearing tissue or a size-weighted distance from the collecting system [79].

Percutaneous thermal ablation (cryoablation or radiofrequency ablation) is a minimally invasive treatment option particularly suited to patients unfit for surgery. While the oncological outcome in small (T1a) tumours appears to be equivalent to surgery, patients with T1b tumours experience shorter cancer-specific survival compared to patients undergoing partial nephrectomy [80].

Following nephrectomy with curative intent, 20-40% of patients experience either local or distant recurrence [81]. The invasion of draining vessels and surrounding fat, increased

tumour size, grade, necrosis and sarcomatoid differentiation are known risk factors for disease recurrence, some of which are accounted for in the Leibovich score which is a validated tool for the prediction of cancer-specific survival in ccRCC patients [82–84].

In contrast to other cancer types, neoadjuvant or adjuvant SACT is only recommended as part of clinical trials in RCC. Initial evidence from a phase II clinical trial (AXIPAN) indicates that neoadjuvant axitinib may allow nephron-sparing surgery in patients with T2-stage RCC which would otherwise have been candidates for RN [85]. Others have found neoadjuvant sunitinib [86] or pazopanib [87] may facilitate PN in patients where preservation of renal function is imperative.

1.2.3 Treatment of metastatic RCC

Treatment strategies for metastatic RCC have evolved rapidly. The introduction of vascular endothelial growth factor (VEGF) targeted therapy including the receptor tyrosine kinase inhibitors (TKI) sunitinib [88] and pazopanib [89] as well as immune-checkpoint inhibitors like ipilimumab, nivolumab [90] and pembrolizumab [91] have improved overall survival in patients with mRCC (Table 1.4). As a result, the increase in survival in recent years is primarily attributable to improvements in the treatment of patients with advanced disease [12].

Table 1.4 Selection of recent trials leading to a change in the guidelines for the treatment of mRCC.

Trial Name	Intervention	Control	mOS [months]	mPFS [months]	Ref
none	Sunitinib	IFN-α	26.4 / 21.8	11 / 5	[88, 92]
VEG105192	Pazopanib	Placebo	22.9 / 20.5	9.2 / 4.2	[89, 93]
CheckMate 214	Ipi/Nivo	Sunitinib	nr / 26	11.6 / 8.4	[90]
KEYNOTE-426	Pembro/Ax	Sunitinib	nr / nr	15.1 / 11.1	[91]

Ax: Axitinib, IFN-α: Interferon alpha, Ipi: ipilimumab, mOS: median overall survival, mPFS: median progression-free survival, Nivo: nivolumab, nr: not reached, Pembro: pembrolizumab

Despite extensive research, particularly in ccRCC, the ideal time point to start treatment remains unclear and an initial period of observation appears safe in patients with low metastatic burden and few symptoms [94]. As a result of the KEYNOTE-426 trial, the combination of pembrolizumab and axitinib is the recommended front-line treatment in patients with ccRCC, irrespective of the International Metastatic Renal Cell Carcinoma Database Consortium (IMDC) risk group [95, 96]. Pembrolizumab is a monoclonal antibody inhibiting the PD-1 receptor expressed on T-cells, leading to an enhanced immune response to tumours

[97] while axitinib is a multi-targeted TKI with an improved safety profile compared to sunitinib [98]. The cellular receptor of the vascular endothelial growth factor (VEGFR) is, among other receptors, a target of TKI. Binding of VEGFR and inhibition of the downstream signalling pathway through phosphatidylinositol 3-kinase / protein kinase B (PI3K/Akt) and mitogen-activated protein kinase (MAPK) leads to reduced proliferation, migration and survival [99].

Meanwhile, the improved survival of patients with IMDC intermediate- and poor-risk metastatic ccRCC receiving the combination of Ipilimumab and Nivolumab compared to Sunitinib in the CheckMate 214 trial has led to the recommendation of this combination as a first-line treatment in these patients [4] (Table 1.4). Ipilimumab is a monoclonal antibody directed against CTLA-4, which is expressed on activated T-cells and regulatory T-cells and leads to a down-regulation of the T-cell response [100]. Nivolumab, like Pembrolizumab, is a monoclonal antibody that is directed against PD-1 [101].

The IMDC risk groups were defined in an attempt to predict survival in patients with mRCC receiving VEGF-targeted therapy (Table 1.5) [96]. A retrospective validation in patients receiving immune checkpoint inhibitors revealed that higher IMDC risk groups continue to experience shorter survival [102]. Few data exist on the optimal second-line treatment after progression on ipilimumab/nivolumab or pembrolizumab/axitinib, however, VEGF-targeted therapy is recommended (Level/Grade of evidence III B, evidence levels see section A.1) [95]. Equally, recommendations on the treatment of non-ccRCC are based on weaker evidence. As a result, the inclusion of patients into clinical trials is encouraged. VEGF-targeting remains a first-line treatment option for all sub-types, while cisplatin-based chemotherapy is an option in collecting duct carcinoma and nivolumab/ipilimumab in predominantly sarcomatoid tumours [4].

Another shift in the treatment of patients with mRCC has occurred in the recommendations for cytoreductive nephrectomy (CN). The CARMENA [105] trial showed that sunitinib alone is non-inferior to CN plus sunitinib in patients with high-volume, intermediate- to poor-risk mRCC. Furthermore, initial sunitinib treatment followed by deferred CN may confer increased OS compared to immediate CN followed by sunitinib as suggested in the SURTIME trial [106]. However, patients with low-volume, favourable or intermediate-risk metastatic disease and a good performance status may benefit from CN and, where possible, complete metastasectomy may prolong survival [107, 108]. Therefore, CN remains an option in these patients who would otherwise be candidates for surveillance [4]. However, the role of metastasectomy and radiotherapy in patients with oligometastatic RCC remains controversial. The literature suggests an improvement in OS after complete metastasectomy [109]. In the

Table 1.5 International Metastatic Renal Cell Carcinoma Database Consortium risk score for survival of patients with mRCC.

Parameter	Score
Clinical Parameters	
Karnofsky Performance Status $< 80\%$ [103]	1
Time from diagnosis to treatment < 1 year	1
Laboratory Parameters	
Haemoglobin $<$ 120 g/l	1
Calcium $>$ upper limit of the norm*	1
Neutrophil count $>$ upper limit of the norm*	1
Platelet count $>$ upper limit of the norm*	1

IMDC Risk Score	Risk Group	Median Survival
0	Favourable	43.2 months
1-2	Intermediate	22.5 months
$\leqslant 3$	Poor	7.8 months

Adapted from Heng *et al.* [96, 104]. *The upper limit of the norm is laboratory dependent. Generally, the normal range is Calcium: 2.12-2.54 mmol/L, Neutrophil count: 2.0-7.0 $\times$ 10^9/L, Platelet: 150-400 $\times$ 10^9/L .

case of brain metastases, surgical resection with or without conventional radiotherapy and stereotactic radiosurgery (SRS) achieve comparable survival outcomes [110]. Radiotherapy of bone metastases may be used in the palliation of tumour pain [111] and interventional embolisation of tumour-feeding vessels can help alleviate symptoms such as gross haematuria or pain from the primary tumour [112]. However, all these studies were retrospective and at a high risk of selection bias [108].

1.2.4 Imaging follow-up

The most appropriate imaging follow-up after (partial) nephrectomy for RCC remains contentious. It has to balance improved oncological outcomes with cost, radiation exposure from follow-up CTs and competing risks. Current guidelines suggest a risk-adapted follow-up. Table 1.6 summarises recommendations for post-surgical follow-up in patients with localised renal cancer from the European Association of Urology, the Canadian Urological Association, the National Comprehensive Cancer Network [5, 113, 45] .

Tumour response to systemic treatment in clinical trials is typically assessed with the response evaluation criteria in solid tumours (RECIST) version 1.1 [114] which are based on

unidimensional size measurements of tumour manifestations. RECIST 1.1 defines measurable and non-measurable lesions as well as target and non-target lesions which aims to facilitate a representative selection of tumour deposits and increase the reproducibility of measurements. Measurable lesions have a long-axis diameter of $\geq$10 mm on CT/MR with a slice thickness $\leq$5 mm (short-axis $\geq$15 mm for lymph nodes) and in bone lesions, only non-calcified parts are measurable. Solid lesions should be prefered over cystic lesions. Up to five target lesions, no more than two of which may be in the same organ, are defined and the evolution of the sum of diameters is assessed. All other tumour manifestations are non-target lesions and recorded as present or absent. If all target and non-target lesions have disappeared and lymph nodes reached a diameter of $<$10 mm at a follow-up time point, the patient has achieved a complete response (CR). However, if the sum of the diameters has decreased by $\geq$30% compared to the initial sum of diameters, the patient has achieved a partial response (PR). An increase in the sum of diameters by $\geq$20% and at least 5mm over the lowest sum measured during the current line of treatment (the nadir), an unequivocal progression of non-target lesions or the appearance of a new lesion is defined as progressive disease (PD)s. Finally, all other constellations designate stable disease (SD) Table 1.7. In clinical practice, lesion size is taken into account when assessing treatment response.

Table 1.6 Guideline recommendations for imaging follow-up of localised renal cancer. Adapted from the EAU, CUA and NCCN guidelines.

Risk/Guideline	Months post surgery					
	3	6	12	18	24	Follow-up
Low-Risk						
EAU	-	US	CT	-	US	CT every 2 years starting at 36 months
CUA	-	-	CXR	-	CXR	continue annual CXR
					CMU	CT-A at 60 months
NCCN			CTC		CTC	CTC annually for 5 years
		CMU		CMU		CMU annually for 3 yr
Intermediate-Risk						
EAU	-	CT	CT	-	CT	CT every 2 years starting at 36 months
CUA	-	XR/CTC	XR/CTC	XR/CTC	XR/CTC	XR/CTC at 30, 36, 48 and 60 months
			CMU		CMU	CMU at 36 and 60 months
NCCN	CTC	CTC	CTC	CTC	CTC	CTC 3-6 mthly to 3 yr, yearly to 5 yr
	CMU	CMU	CMU	CMU	CMU	CMU 3-6 mthly to 3 yr, yearly to 5 yr
High-Risk						
EAU				See above		
CUA	-	XR/CTC	XR/CTC	XR/CTC	XR/CTC	XR/CTC at 30, 36, 48 and 60 months
		CMU	CMU	CMU	CMU	CMU at 36 and 60 months
NCCN				See above		
Very high-Risk						
EAU				Undefined		
CUA	XR/CTC	XR/CTC	XR/CTC	XR/CTC	XR/CTC	XR/CTC at 30, 36, 48 and 60 months
		CMU		CMU		CMU at 36 and 60 months
NCCN				Undefined		

CM: CT or MRI of the abdomen, CMU: CT, MRI or ultrasound of the abdomen, CT: CT of the chest and abdomen, CTC: CT of the chest, CXR: Chest x-ray, US: Ultrasound of the abdomen. For information on the follow-up after ablative treatment and further details, please refer to the full guidelines referenced in section 1.2.4.

Table 1.7 Summary of RECIST 1.1 treatment response

Response Group	Criteria
Complete Response (CR)	Disappearance of all lesions (including non-target lesions) and no new lesions.
Partial Response (PR)	Reduction in the sum of diameters >30% compared to baseline, no unequivocal progressionof non-target lesions and no new lesions.
Stable Disease (SD)	Neither CR nor PR nor PD
Progressive Disease (PD)	Increase in the sum of diameters >20% above the nadir, unequivocal Progression of non-target lesions, appearance of new lesions.

Adapted from Eisenhauer *et al.* [114].

RECIST 1.1 has been defined in the era of cytotoxic and cytostatic chemotherapy. As a consequence of the mechanism of action of these systemic anti-cancer agents, a size reduction was an adequate means for early treatment response detection [115]. However, RECIST has been criticised for the arbitrary selection of cut-off values between response categories, disregard for changes in tumour attenuation and inconsistencies in determining the relevant lesion diameters [115]. Furthermore, the subjectivity in the selection of target lesions leads to different response assessments. Inter-reader agreement on RECIST response reached 96% on a study and 92% on a patient level in a multi-reader trial if readers selected identical target lesions. However, if discordant lesions were selected, the agreement was poor at 53% and 37%, respectively, demonstrating the subjectivity of the RECIST criteria [116]. The superior inter-reader agreement of tumour volume measurements compared to the selection of the largest cross-section has been reported repeatedly [117, 118]. However, without substantial computational support, lesion segmentation is very time consuming, limiting the applicability of volumetric assessments.

RECIST is suboptimal for early response assessment of anti-angiogenic therapy [119]. Therefore, alternative measures taking attenuation, marked central fill-in, and new central necrosis into account have been defined. While RECIST remains the most used and best-assessed response metric, the MASS, Choi, mChoi and SACT criteria are potential alternatives [120]. Even these response criteria have been defined before immune checkpoint inhibitors and targeted treatment agents were applied broadly in the standard of care practice. Consequently, they capture response patterns that are specific to these new treatment strategies insufficiently. A significant number of patients starting immune checkpoint inhibition experience an initial increase in tumour size or appearance of new lesions on early follow-up imaging, followed by stabilisation or decrease in size. Psuedoprogression occurs in 5 – 15% of patients with RCC treated with Nivolumab [121] and in 6% (95%-CI: 5 – 7 %) in a

meta-analysis of multiple disease sites in 3402 patients [122]. Hyperprogression is a more recently identified and still poorly understood response pattern where patients are progressive at the first follow-up and tumour growth rates double with the start of immunotherapy [123]. Data on the prevalence of hyperprogression in RCC is insufficient, however, it is thought to occur in 1 – 30% of patients treated with immunotherapy [124]. The question of whether the phenomenon could represent the natural course of the disease in few non-responders rather than a specific type of response to immunotherapy remains to be answered [125]. These new patterns of treatment response have led to the development of adapted immune response criteria. In the absence of consensus criteria, irRC [126] and irRECIST [127] based on the WHO and RECIST classification, respectively, have been proposed. In 2017 these were summarised in the iRECIST criteria in the hope to establish a unified system for response assessment in immuno-oncology [128]. With the addition of iUPD (immune unconfirmed progressive disease) requiring confirmation through further progression after 4-8 weeks to reach iCPD (immune confirmed progressive disease), the iRECIST criteria allow patients to reach iSD; iPR or iCR (immune stable disease, partial response or complete response) even after pseudoprogression.

In addition to these morphological criteria, PERCIST (Positron Emission Tomography Response Criteria in Solid Tumours) and their extension iPERCIST define metabolic tumour response on FDG-PET [129, 130]. However, due to the limitations of FDG-PET in RCC outlined below (section 1.4), these criteria have not found widespread application in RCC.

1.3 Genomics and metabolomics in RCC

The diverse histological subtypes of RCC are accompanied by an equally varied clinical course of the disease, response to treatment and genetic background. The identification of germline mutations causing hereditary RCC in von Hippel-Lindau (VHL) disease, hereditary papillary RCC or hereditary leiomyomatosis and renal cell carcinoma (HLRCC), has led to the discovery of genes that are also frequently mutated in sporadic kidney cancer [131–133]. Sporadic ccRCC is strongly associated with the inactivation of *VHL*, occurring in 80 – 90% of ccRCC patients due to inactivating mutations or promoter hypermethylation [134, 135]. *VHL* inactivation leads to decreased degradation of the transcription factor hypoxia-inducible factor (HIF-1α and HIF-2α) and subsequently increased transcription of growth factors (VEGF, PDGF), growth factor receptors (EGFR) and glucose transporters (GLUT1) which promote survival and proliferation of cancer cells [136, 137]. Mutations in the *MET* oncogene commonly found in hereditary pRCC are shared among a subset of

sporadic type I pRCC. Copy number alterations of *MET* are found in 81% and 46% of sporadic type I and II pRCC, respectively [138]. Patients with "hereditary leiomyomatosis and renal cell carcinoma" or Reed's syndrome carry a germline mutation in one of the enzymes of the citric acid cycle (fumarate hydratase, *FH*) and develop aggressive type II pRCC [139, 140]. Birt-Hogg-Dubé (BHD) syndrome follows an autosomal dominant pattern of inheritance and is characterised by mutations of the folliculin gene (FLCN) which acts as a tumour suppressor gene [134]. Patients with BHD are at increased risk of developing renal tumours, predominantly oncocytomas and chRCC alongside other benign and malignant tumours [141]. However, mutations in *FLCN* are not commonly observed in sporadic oncocytomas or chRCC and its role remains to be determined [142]. Mutations in *p53*, a tumour suppressor gene involved in cell cycle regulation, are shared among more than half of human cancers. Its level of expression is of prognostic value in different cancer types including breast, colorectal and cervical cancer [143–145]. A recent meta-analysis found that accumulation of *p53* in immunohistochemistry, which is commonly observed in mutated *p53*, was associated with significantly inferior OS in RCC patients (HR = 2.17, 95% CI: 1.51 – 3.13) [146].

1.3.1 Genetic intratumoral heterogeneity in RCC

Recently, the high degree of genetic heterogeneity in lesions within one patient has received much attention. In the seminal paper, Gerlinger *et al.* describe how one in five mutations in an mRCC patient is private to a single biopsy in a multi-regional sequencing approach. Furthermore, 63–69% of all somatic mutations were heterogeneously distributed and hence not detectable in some of the tissue samples [2]. The clinically most common approach of analysing a single tumour biopsy is at risk of underestimating the true genetic heterogeneity since it captures only 55% of the mutational burden on average. Obtaining a complete picture of the mutational landscape in a patient will be crucial once therapy decisions are based on the mutational status. Otherwise, minority clones resistant to the therapy of choice and not adequately represented in a single biopsy may cause treatment failure [147]. The TRACERx renal consortium, in their multi-region sequencing approach, describe seven genetic trajectories along which RCC evolve and which may be prognostically relevant. They show that high genetic heterogeneity and genome instability may occur independently of each other and both are associated with rapid disease progression [148]. Moreover, they address the question of the number of biopsies that would be needed to cover the intratumoral genetic heterogeneity reliably: on average, seven biopsies would be needed to detect 75% of

driver events (mutations or somatic copy number alterations). However, fewer biopsies are required in evolutionary subtypes with low intratumoral heterogeneity such as the “multiple clonal driver” or “VHL monodriver” types, and most biopsies are required in the highly heterogeneous “PBRM1-SETD2” tumours. *PBRM1* encodes a component of the chromatin-remodelling complex and is a tumour suppressor gene frequently mutated in ccRCC [149]. In their comparison of the genetic landscape of primary lesions and metastases, they show evidence of genetic bottlenecking, i.e., only a subset of clones in a primary tumour appears to metastasise. Additionally, only few driver mutations occur newly in metastatic tumour tissue. The evolutionary trajectories described above are predictive of the rate of progression in the metastatic setting; patients with "multiple clonal drivers", "BAP1 driven" and "VHL wild-type" tumours are at increased risk of rapid progression [150]. While the large genetic heterogeneity of RCC has long been recognised, neither diversity on a patient-level nor the intra-patient heterogeneity are considered in current treatment recommendations [4].

Although multiple biopsies offer an improved understanding of the genetic landscape of RCC in patients, this has to be balanced with the risk of complications from renal tumour biopsies. Minor bleeding detected in follow-up imaging is the most frequent complication, meanwhile, major blood loss requiring therapeutic intervention occurs in <1% of patients [151]. Seeding of the tumour along the biopsy path was observed very rarely in the past and has not been reported since the introduction of coaxial biopsy techniques [152]. Pneumothorax, intestinal perforations and arteriovenous fistulas are rare but important complications of renal biopsies [153]. Another procedural risk is the rate of non-diagnostic biopsies in about 20% of patients [154, 155]. The overall risk of non-diagnostic biopsies increases with the number and decreasing size of tumour regions targeted.

1.3.2 Cell-free DNA in RCC

Tissue genomics requires access to tissue samples either through surgery or invasive biopsy. Therefore, cell-free DNA (cfDNA) which includes circulating tumour DNA (ctDNA) as well as non-tumour DNA in patient serum and DNA in urine is an attractive tool for screening, diagnosis and treatment monitoring in cancer. In contrast to a conventional biopsy taken from a single lesion, ctDNA may provide genomic information on all the lesions in a patient and thereby guide targeted therapy. However, in melanoma, only 64% of shared mutations and 14% of private mutations were detected in ctDNA [10]. Patients with RCC harbour larger amounts of cfDNA compared to benign tumours (mean ± SD: 3319 ± 3181 genome equivalents/ml vs. 1288 ± 913 genome equivalents/ml, $p < 0.001$) and VHL methylation

was more frequent (50.3% vs. 9.3%) [156]. The methylation state is subtype dependent and most frequently found in ccRCC [156]. Higher levels of cfDNA at baseline are predictive of metastatic disease and postoperative recurrence [157]. While these analyses targeted any cfDNA and measured copy numbers of genes not generally mutated in RCC (housekeeping genes), detecting tumour-specific mutations has proven much more challenging. Smith *et al.* only detected ctDNA in 27.5% using an untargeted approach. Targeting to mutations previously detected in tissue biopsies increased this rate to ~50% [158]. This mirrors findings by others [159, 160]. Meanwhile detection of ctDNA and response to mostly VEGF-directed therapy were uncorrelated in a small cohort [161]. Recent advances in the detection and classification of differential ctDNA methylation patterns between patients with RCC and healthy controls suggest that machine learning models may correctly identify 97% of patients with RCC [162]. In summary, cfDNA analysis is a promising tool for minimally invasive diagnosis and treatment response monitoring in RCC. However, further validation in larger and more diverse cohorts is required before cfDNA may be used as a clinical biomarker in RCC.

1.3.3 RCC - a metabolic disease

RCC is driven by metabolic changes in addition to genetic alternations. Multiple mutations commonly found in RCC directly or indirectly alter the cellular metabolism in RCC. The loss of *VHL*, the most common mutation in ccRCC, profoundly alters energy metabolism [163]. In normoxic conditions, von Hippel Lindau tumour suppressor protein (pVHL) mediates the proteasomal degradation of hypoxia-inducible factors (HIFs) [164] after their hydroxylation by prolyl hydroxylase (PHD) [165]. In hypoxic conditions or after a loss of *VHL*, HIF-1α and, more importantly in renal tumorigenesis HIF-2α [166] are stabilised and up-regulate the transcription of proteins involved in cellular energy metabolism including glucose transporters, hexokinase and pyruvate dehydrogenase kinase [167]. Together, the transcriptional effects of HIF lead to a shift towards aerobic glycolysis which makes RCC a prime example for dysregulated cellular energetics, a hallmark of cancer [168]. The shift towards the energetically inefficient glycolytic metabolism even in the presence of oxygen, also called the Warburg effect, promotes immune evasion [169] and tumour invasion [170] through the acidification of the tumour microenvironment and supports sustained replication through the maintenance of the redox homeostasis [171]. In addition to suppressing immune functions of tumour infiltrating lymphocytes (TILs) [169], lactic acid promotes the polarisation of tumour-associated macrophages to an M2-like state [172] which leads to a

suppression of the tumour immune response and facilitates invasion, neo-angiogenesis and metastasis [173, 174].

Apart from alterations in the glycolytic pathway, RCC commonly presents with an up-regulation of the pentose phosphate pathway (PPP) and fatty acid synthesis which are associated with an inferior prognosis [175]. The PPP generates nicotinamide adenine dinucleotide phosphate (NADPH) which provides reducing equivalents for the protection against reactive oxygen species (ROS) and anabolic reactions including nucleotide synthesis [176].

Germline mutations in hereditary forms of RCC involving metabolic genes result in specific metabolic phenotypes. ccRCC associated with succinate dehydrogenase (*SDH*) mutation is characterised by an accumulation of succinate while fumarate hydratase (*FH*) mutations lead to an accumulation of fumarate in patients with hereditary leiomyoma renal cell carcinoma (HLRCC) [177]. The metabolic alterations in both conditions ultimately lead to an overexpression of HIF [178].

The idea of renal cancer as a metabolic disease has led to increased interest in tumour metabolomics [163]. Metabolic profiling of ccRCC and the healthy human kidney reveals an increase in intermediates of the carbohydrate metabolism and a decrease in amino acids [179]. Artificial intelligence-based analysis of metabolomic data identifies distinct metabolic profiles in histological subtypes and may identify RCC based on urine and plasma metabolomics [180–182]. RNA sequencing reveals an almost universal downregulation of metabolic pathways in ccRCC which is the opposite of the transcriptomic landscape observed in most other cancers [183]. Strikingly, gene expression and metabolomics show no correlation, highlighting the limited bearing of transcriptomics as a surrogate marker for metabolism. Metabolic signatures differentiating normal kidney and RCC, histological RCC subtypes as well as RCC stages have been identified [184]. Pathway analyses demonstrated that, while amino acid concentrations are generally diminished in RCC compared to normal kidney, the pattern of amino acid concentrations is associated with the histological subtype [184].

Traditionally, invasive tissue sampling was a pre-condition for investigating tumour metabolism. Mass spectrometry (MS) and nuclear magnetic resonance (NMR) can quantify the concentrations of many metabolites in little tissue. However, the sampling process itself introduces bias and prevents kinetic studies in patients. Metabolic imaging has the potential to fill this unmet need.

1.4 Metabolic Imaging in RCC

Radiological imaging sets itself apart from histological, genomic and metabolic analyses in its ability to provide spatially resolved information from entire organs and large regions of the body. Additionally, its non-invasive and non-destructive nature make it ideally suited for repeated examinations. However, none of the standard of care imaging techniques introduced above exploits the specific metabolic alterations found in ccRCC.

The clinically most broadly applied metabolic imaging technique is [^{18}F]-fluorodeoxyglucose positron emission tomography (^{18}F-FDG-PET), using a radioactively labelled glucose analogue. Following uptake into the cell through GLUT1 transporters, FDG undergoes hexokinase (HK) mediated phosphorylation [185]. Due to the lack of the C2-oxygen, FDG-6PO_4 is trapped intracellularly in the pentose phosphate pathway [186]. Among other areas including infectious, inflammatory, cardiac and neurologic diseases, FDG-PET finds broad application in oncology. Many cancer types show an increased avidity for FDG, which is primarily mediated by the overexpression of GLUT1 and HK. Furthermore, the FDG-signal measured as standardised uptake value (SUV) is related to tumour aggressiveness in many cancers. Therefore, FDG-PET is commonly used in the diagnostic workup or to follow systemic treatment in lung cancer, lymphoma, breast cancer, malignant melanoma, head and neck cancer and others [187]. Its use in RCC is limited by the renal excretion of FDG, resulting in a poor tumour-to-background contrast. There is probably no routine clinical role for FDG-PET in the detection of renal masses, due to its inferior sensitivity compared to ceCT [188]. Meanwhile, one report suggests FDG-PET may differentiate low- and high-grade histology [189]. Furthermore, evidence suggests that FDG-PET can detect RCC metastases more sensitively than ceCT and bone scintigraphy [190]. Current clinical guidelines for the treatment of metastatic RCC see no role for routine FDG-PET [4].

Clinical MRI is based on the detection of electromagnetic waves emitted by the nuclear spin of protons, the nuclei of hydrogen atoms, which were polarised in an external magnetic field and excited with radio-frequency pulses [191]. Protons in water and fat generate the bulk of the MRI signal. However, less abundant molecules carrying hydrogen atoms also emit nuclear resonance signals at frequencies that are specific to the chemical surroundings of these hydrogen atoms. Magnetic resonance spectroscopy (MRS) records these signals from a volume of tissue at a range of frequencies and is a non-invasive way of investigating metabolite concentrations *in vivo*. It is the clinical equivalent to nuclear magmatic resonance for the identification of molecules *in vitro* which was first described in 1938 [192]. The low concentration of most molecules other than water (~10^{-4}-times lower, water: 33 M,

glucose: 0.1 – 5 mM, lactate: up to 10 – 30 mM in cancer) [193] results in lower signal intensity, longer acquisition times and decreased spatial resolution [194]. NMR has found limited clinical applications, mostly in neuroradiology, for the differentiation of cerebral lymphoma, gliomas and metastasis, diagnosis of WHO tumour grade, neurodegenerative and mitochondrial disorders amongst other indications [195]. While ^{1}H-spectroscopy has found some application in the diagnosis of prostate cancer, it is no longer mentioned in the most recent European guidelines [196]. In renal imaging, however, MRS is restricted to the research setting where its potential to differentiate lesion types has been investigated in small patient cohorts [197, 198]. Sevcenco *et al.* describe an increased Choline signal-to-noise ratio (SNR) in malignant renal masses compared to benign lesions including cysts. However, the choline signal was positively associated with tumour size in RCC and benign lesions were smaller than malignant lesions. Additionally, the SNR, being dependent on numerous factors like the voxel size, coil architecture, field strength and number of signal averages, is not generalisable and, therefore, unsuitable as an imaging biomarker [197]. Therefore, this MRS approach appears to contribute little to the differentiation of small, indeterminate renal lesions. Furthermore, respiratory motion in conjunction with long acquisition times and the low spatial resolution of the technique have so far prevented a broader application.

MR-activity is not exclusive to ^{1}H nuclei. All atomic nuclei with an odd number of protons or neutrons have a nuclear spin and can emit an MR signal [195]. The MR signal is dependent on the concentration of an atom, the abundance of the MR-active isotope, its relative sensitivity, and quadrupolar relaxation for nuclei with spins $I > ½$. The low concentration of non-hydrogen nuclei coupled with the low sensitivity of MRI has long limited their use in MRS. Hyperpolarisation techniques, achieving a transiently increased alignment of nuclear spins with an external magnetic field over the thermal equilibrium, have made it possible to image ^{13}C, ^{3}He and ^{129}Xe. Among these, ^{13}C is of particular interest to metabolic cancer imaging since carbon occurs in many molecules involved in pathways commonly disturbed in cancer. [1-^{13}]C-pyruvate (^{13}C-pyruvate) was the first such molecule becoming available for hyperpolarised imaging in patients. It is located at the crossroads of glycolysis, citric acid cycle and amino acid metabolism (Figure 1.1).

The feasibility of hyperpolarised ^{13}C-pyruvate MRI in renal cancer has been demonstrated in a case report and case series [199, 200] and *in vitro* studies of RCC suggest increased pyruvate to lactate conversion and more rapid lactate efflux with increasing tumour aggressiveness [201]. Modelling of ^{13}C-label exchange between pyruvate and lactate on hyperpolarised MRI in breast cancer is associated with tumour aggressiveness and response to systemic anti-cancer treatment [202, 203]. Similarly, increasing Gleason tumour grade in

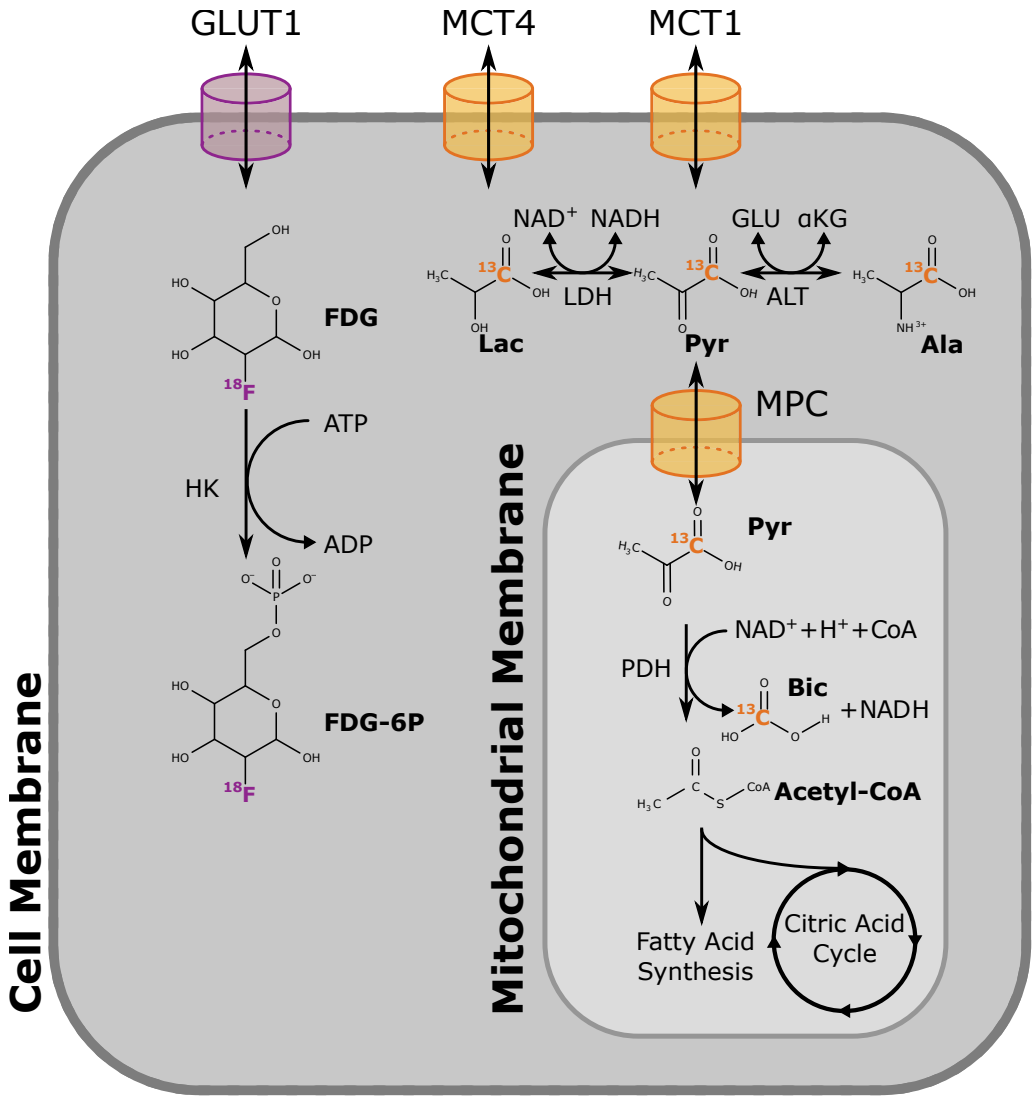


Fig. 1.1 Hyperpolarised ^{13}C-pyruvate is taken up into the cell through the monocarboxylate transporter 1 (MCT1) and rapidly metabolised in the cytosol to lactate and alanine. Following mitochondrial uptake through the mitochondrial pyruvate carrier (MPC), pyruvate dehydrogenase (PDH) catalysis the formation of Acetyl-CoA and transfer of the ^{13}C-label to bicarbonate. In contrast, FDG does not undergo any further metabolism after phosphorylation. Ala: Alanine, ALT: Alanine Aminotransferase, Bic: Bicarbonate, CoA: Coenzyme A, FDG:Fluorodeoxyglucose, FDG-6P: Fluorodeoxyglucose 6-phosphate, GLUT: Glucose Transporter, HK: Hexokinase, Lac: Lactate, LDH: Lactate Dehydrogenase, MCT: Monocarboxylate Transporter, MPC: Mitochondrial Pyruvate Carrier, Pyr: Pyruvate.

prostate cancer is associated with an increased metabolic activity measured on hyperpolarised ^{13}C-pyruvate MRI in patients and *in vitro* [204–206]. Finally, pyruvate metabolism has been demonstrated in primary and secondary brain tumours, the normal human brain and the diabetic and healthy heart [207–209]. Further details on the technique will be provided in chapter 5.

1.5 Research question and structure of the thesis

This PhD aimed to improve the understanding of the imaging phenotype of renal cell carcinoma and investigate correlations between heterogeneity observed on multimodality, multiparametric imaging, metabolic, transcriptional and genetic heterogeneity in renal tumours.

Chapter two describes the use of computational tumour heterogeneity metrics derived from standard of care imaging for the prediction of histological subtypes, aggressiveness and outcome. It critically examines the existing literature on computer-assisted diagnostic systems for renal masses, describes a computational approach to the development of robust and well-generalisable radiomics models, and sets out a strategy for the development and validation of these on a multi-institutional dataset. Based on readily available imaging data, these results can be externally validated in any other healthcare setting.

Chapter three assesses the ability of multiparametric, physiological ^{1}H-MRI to detect early response to systemic anti-cancer treatment in prospective clinical trials and compares this to conventional response criteria. Early changes in tumour physiology will be compared to the long term benefit from treatment.

Chapter four describes the development of an image-guided approach to acquiring tissue samples from nephrectomy specimens registered to the pre-operative MRI. Accurate registration of imaging data and tissue samples enabled multi-omics data integration and was a pre-condition for multi-regional assessment and correlation of intratumoral metabolic heterogeneity on imaging with tissue genomics, transcriptomics and metabolomics.

Chapter five investigates the relationship of imaging biomarkers derived from hyperpolarised ^{13}C-pyruvate metabolic imaging with clinical markers of tumour aggressiveness and analyses the complementary value of hyperpolarised pyruvate MRI to information on tumour perfusion, diffusion impedance and oxygenation on physiological ^{1}H-MRI. It describes the association of the tumour metabolic imaging phenotype with tissue metabolomics, protein expression.

Chapter six describes the application of insight gained from ^{1}H-MRI (chapter 3), image-guided tissue sampling (chapter 4) and ^{13}C-MRI (chapter 5) in a prospective, multi-arm, multi-site window of opportunity trial in RCC and summarises initial results.

Finally, chapter seven provides an overview and clinical perspective on the findings and highlights areas for future research.

Chapter 2

Applications of artificial intelligence in renal cell carcinoma

2.1 Introduction

The radiological practice relies largely on the subjective interpretation of imaging data by an expert radiologist. Reports will therefore be dependent on reader experience. Quantitative, reader independent imaging markers may supplement expert opinion and increase diagnostic, predictive and prognostic accuracy and generalisability [210]. Radiomics is an ensemble of variables that can be calculated from medical images, providing quantitative, mineable, high-dimensional data. These include histogram, texture and shape analysis, that extract information from imaging data that may not be visible to the human eye [211, 212]. The differentiation of RCC from benign renal lesions, especially oncocytoma and angiomyolipoma without visible fat (AMLwvf), can be challenging by subjective radiological image interpretation [213]. Quantitative image analysis may reveal radiomic signatures that can classify renal tumour subtype and aggressiveness or predict response to targeted treatment, therefore, aiding treatment stratification and decision making.

Texture analysis describes the mathematical quantification of the distribution of grey levels in an ROI which may de on their spatial relationship [211]. While statistical texture analysis methods have received the most attention, they are complemented with structural, spectral and model-based methods [214]. First-order texture features are defined by their spatial independence and are derived from the histogram of grey levels present in an image. Based on their distribution, the mean, median, standard deviation, skewness, and kurtosis are calculated amongst others. Second-order texture features are sensitive to the relative distribution of grey levels [215]. Examples of this are statistical measures based on the

grey level co-occurrence matrix (GLCM) which quantifies the frequency of a pair of grey levels being found at a predefined distance averaged over all directions. Different parameters emphasise a homogeneous or heterogeneous distribution of grey levels [216]. Finally, third-order texture features take the directionality of multiple (i.e. >2) voxels into account [217]. The grey level run length matrix (GLRLM) or size zone matrix (GLSZM) quantify the length of lines or the size of zones with identical grey levels. Parameters calculated based on these matrices may emphasise small or large features as well as bright or dark features [218].

Texture analysis has been applied primarily to CT and less frequently to MR images of RCC for diagnostic, prognostic and predictive purposes. The question of malignancy in small renal lesions has been repeatedly investigated with CT and MR texture analysis. Texture information based on unenhanced and contrast-enhanced CT may aid the differentiation between oncocytoma, angiomyolipoma without visible fat (AML.wovf) and RCC [219–221]. MR-based texture features show a moderate to high accuracy in the differentiation of small papillary RCC, clear cell RCC and oncocytoma [222]. Multiple additional papers address the ability to differentiate histologic subtypes of RCC based on image texture in non-size-restricted cohorts and propose texture-based models achieving areas under the receiver operating characteristic curve (ROC) for the recognition of the histologic subtype over 0.90 [223, 224]. Furthermore, texture analysis of the primary tumour and change in texture values after treatment with sunitinib combined with the IMDC model are better predictors of OS and PFS than IMDC alone [225, 226]. The IMDC is a prognostic model for patients with mRCC undergoing first-line targeted therapy predicting survival. It is based on performance status, time since diagnosis, haemoglobin, calcium, neutrophil count and platelet count and has been validated recently for prognosis in the post-first-line setting [96, 227]. Finally, the image texture of metastatic sites appears to change as a result of successful treatment with VEGF inhibitors [228].

The universal applicability of texture analysis to digital images makes it an attractive tool for quantitative image analysis in research and as a potential clinical tool. As it is based on the reconstructed image, it can be easily implemented in the diagnostic pathway. Notwithstanding these advantages, its sensitivity to acquisition parameters must not be ignored. Phantom studies show that texture may not be reproducible across scanners and many features are dependent on acquisition parameters in addition to intra- and inter-reader variation. In CT texture analysis, the reconstruction kernel and tube voltage have been shown to impact most texture parameters strongly. The slice thickness, field of view and tube current affect some metrics while most are robust to changes in pitch [229]. MR texture analysis appears even more demanding regarding technical standardization of texture analysis workflows because

of the non-standard nature of the recorded signal intensity. In addition to image acquisition, feature extraction specifications, such as grey level quantisation or the application of filters, introduce heterogeneity to texture analysis [230, 231]. The current body of literature on CT and MR texture analysis in renal cancer mainly relies on retrospective datasets, regularly derived from routine clinical examinations, ignoring substantial technical variability. The lack of validated models and the poor reproducibility of texture-based algorithms have so far prevented their application as a clinical decision support tool. Critically, Cancer Research UK (CRUK), the European Organisation for Research and Treatment of Cancer (EORTC) and others have called for a more efficient and rigorous validation of new imaging biomarkers to aid their translation from pre-clinical over research to clinical tools [232–234, 210, 235]. In the case of texture-based predictive models, this includes the precise reporting of acquisition, reconstruction, and feature extraction parameters as well as the assessment of repeatability, reproducibility and cost-effectiveness. The use of inter-reader comparisons, internal and external validation datasets should form part of such an effort [236]. To date, most studies are retrospective and perform poorly regarding repeatability, validation, and adjustment for multiple testing. The Image Biomarker Standardization Initiative undertook efforts to provide guidance to unify the definition and extraction of radiomics features [237].

This PhD will investigate the value of CT and MR texture analysis for the prediction of tumour metastasis, response to therapy and survival in patients with RCC.

2.2 Review and meta-analysis of existing literature on radiomics in renal cancer

Parts of this work have been published:
Ursprung S, Beer L, Bruining A, Woitek R, Stewart GD, Gallagher FA, Sala E. Radiomics of computed tomography and magnetic resonance imaging in renal cell carcinoma—a systematic review and meta-analysis. European Radiology. 2020; 30, 3558-3566.
Prof Evis Sala and Prof Grant Stewart have conceived the idea for this review. Stephan Ursprung has defined the research question, developed the methodology, conducted the literature search, reviewed the studies, analysed the data and interpreted the results.

2.2.1 Introduction

The increasing affordability of computing power has supported the widespread application of radiomic analyses in oncological imaging research to generate diagnostic, predictive and prognostic models [238]. While reports linking CT and MR radiomics to histology, survival or treatment response are numerous and available for many tumour types, adoption in clinical practice has been slow [239].

As this challenge of overcoming the translational gap between a research tool and clinical practice is not unique to radiomics, CRUK and EORTC have published a roadmap to highlight critical steps in the development of imaging biomarkers. Key steps include assessing the technical and biological validity, qualification, and cost-effectiveness (fig. 2.1) [232].

Additionally, a comparative radiomics quality score (RQS) [236] was proposed to assess the methodological quality of radiomic research. The score assesses data selection, preparation, feature extraction and reduction, model generation, validation and data sharing. The score is summarised in table 2.1. The authors see the main application of this score in the hands of critical editors, reviewers and readers who wish to ascertain how well radiomic research follows best-practice guidelines and if deviations from these are sufficiently justified. As such, the RQS has found application in multiple reviews on radiomic analyses [241–243]. While two of these reviews focused on oncological radiomic research more broadly, one focused specifically on neuro-oncology.

Standard of care imaging, primarily contrast-enhanced CT complemented by MRI as a problem-solving tool as described in section 1.2.1, provides essential differential diagnostic

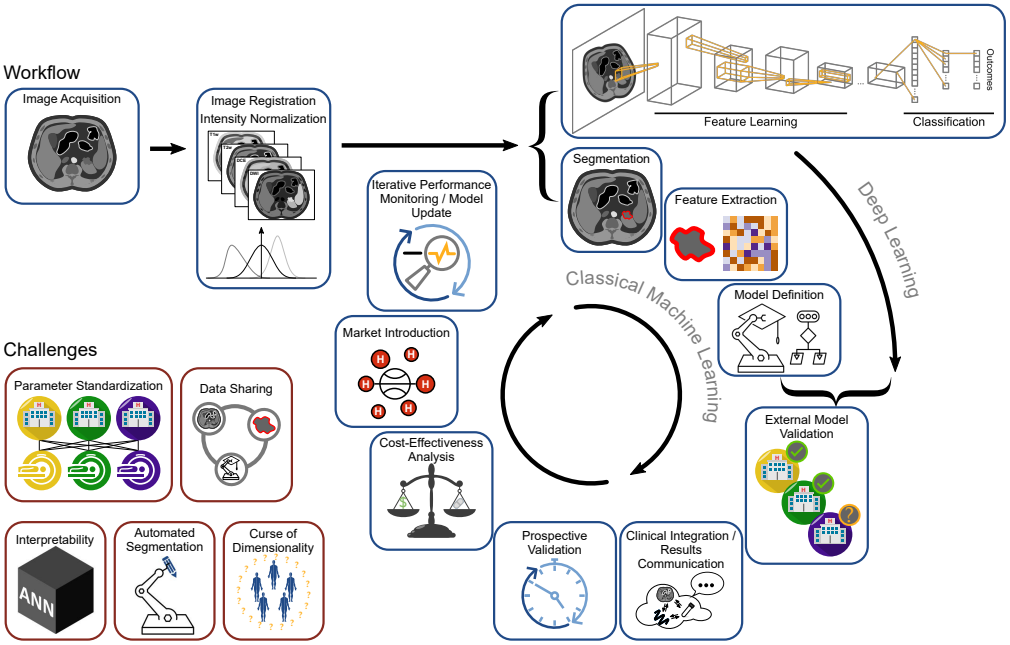


Fig. 2.1 The development of artificial intelligence models as clinical decision support systems in radiology involves multiple steps. Image retrieval and pre-processing are common to any machine learning pipeline. Pre-processing can consist of intensity normalisation or image registration. Classical machine learning techniques such as support vector machines or random forests will be preceded by image segmentation and the automated extraction of handcrafted features. Meanwhile, deep learning may use entire images, image patches or segmented images to learn features. Any model will undergo external and ideally prospective validation. After the introduction to the market, continued model monitoring and re-training will ensure that AI solutions evolve to maintain their performance. ANN: Artificial Neural Network (Reproduced from [240] under the CC-BY licence)

Table 2.1 Radiomics Quality Score. Adapted from [236] and [240]

.

RQS Item	Definition	Avg.
Image Protocol	$+1$ well documented, $+1$ for publicly available protocols	24%
Multiple Segmentations	$+1$ multiple segmentations (different physicians, algorithms or perturbation of regions of interest)	38%
Phantom Study	$+1$ if texture phantomsfor feature robustness assessment	0%
Multiple Time Points	$+1$ multiple timepoints for feature robustness assessment	1%
Feature Reduction	-3 if nothing, $+3$ if either feature reduction or correction for multiple comparisons	8%
Non Radiomics	$+1$ if multivariable analysis with non-radiomic features	15%
Biological Correlates	$+1$ if present	98%
Cut-off	$+1$ if cutoff pre-defined/median/continuous risk variable	11%
Discrimination/ Resampling	$+1$ discrimination statistic and statistical significance $+1$ if resampling applied	46%
Calibration	$+1$ calibration statistic and statistical significance	2%
Prospective	$+7$ prospective validation within a registered study	14%
Validation	-5 no validation, $+2/+3$ internal/external validation, $+4$ two external datasets or previously published signature $+5$ validation on $\geqslant 3$ datasets from >1 institute	0%
Gold Standard	$+2$ for comparison to gold standard	87%
Clinical Utility	$+2$ for reporting potential clinical utility	96%
Cost-effectiveness	$+1$ for cost-effectiveness analysis	0%
Open Science	$+1$ for open source scans, $+1$ for open source segmentations, $+1$ for open source code, $+1$ open source representative segmentations and features	1%

RQS: Radiomics Quality Score, Avg: Average percentage of points achieved for each RQS item.

information for treatment planning in patients with renal lesions. However, the identification of benign renal lesions, especially oncocytoma and angiomyolipoma without visible fat, can be difficult [213]. Artificial intelligence (AI) techniques such as artificial neural networks or machine learning methods from radiomic features have been proposed to address such diagnostic problems and even expand the realm of the radiological assessment to new areas like identifying tumour histology and predicting treatment response or survival [244, 245, 226].

This systematic review investigates whether methodological characteristics of prospective and retrospective studies of radiomics in renal tumours prevent effective clinical translation. As per the pre-specified protocol, a meta-analysis of radiomic models proposed for the differentiation of AMLwvf and RCC was conducted.

2.2.2 Methods

Literature search. PRISMA-DTA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis for Diagnostic Test Accuracy) guidelines were followed during the preparation of this systematic review [246] and the research protocol was published on PROSPERO (CRD 42018115263) before the review process commenced. PubMed, Web of Science and EMBASE were used to identify English, primary publications after 01/01/2000. The search was last performed on 30/10/2018. The databases were searched for literature on radiomics, texture analysis and histogram analysis in kidney and renal diseases performed on MRI or CT. Histogram analysis was included during an update of the research protocol as the review of the initially included manuscripts revealed a shift in terminology over time where more recent publications called "histogram analysis" "texture analysis" without any major change in methodology.

Titles and Abstracts were screened and all potentially eligible articles were downloaded as full-texts for secondary screening. Authors were contacted where full-texts were not available online. Articles were included if they employed radiomics analysis on CT and/or MR for diagnosis, prediction or prognosis in patients with renal tumours. Case reports and conference abstracts were excluded since their brevity precludes an adequate description of the methodology. A backward literature search was performed on the reference list of the included manuscripts to identify additional literature. Uncertainties in the inclusion of articles were resolved in consensus with the co-reviewers.

Data extraction. The methodological quality and the risk of bias were assessed with the RQS and the framework defined in Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) [247], respectively. The purpose of the RQS has been defined above and the distribution of the 36 points over the 16 dimensions is detailed in table 2.1. The score ranges from -8 to 36, while an RQS of $\leqslant 0$ points is defined as 0% and an RQS 36 points as 100%. In contrast to the RQS, QUADAS-2 is qualitative rather than quantitative. It employs signalling questions that can be adapted to the individual research questions to allow identification of bias in four domains: patient selection, index test, reference standard and flow and timing.

At least two reviewers independently extracted the information of the RQS and QUADAS-2 assessment for each study into a structured data collection instrument. Two of the articles were used as training cases to develop a shared understanding of the data points among the reviewers. The reviewers analysed the two manuscripts independently, discussed discrepant ratings and agreed on the interpretation of ambiguous scoring items.

Statistical analysis. All analyses were performed in the R language for statistical computing [248]. Ratings were averaged across the reviewers for reporting. The RQS was treated as a continuous variable and agreement between raters was calculated using the interclass correlation coefficient (ICC). Ten versions of the ICC with different assumptions and preconditions have been defined. A single source, two-way random-effects model for absolute agreement was employed. The two-way random effects model is appropriate for studies where a random selection from a group of raters analyses each observation. Absolute agreement was measured because no systematic difference would be expected between raters. Ratings from all raters were treated equally and, therefore, the mean was employed as the assessment basis [249]. The calculation of the ICC was implemented in the "irr" package (Version 0.84.1)

Individual items of the RQS were treated as ordinal variables. Inter-rater agreement was assessed using a modified version of Fleiss' kappa statistic. Fleiss' kappa statistic assesses the agreement of a constant number of raters in defining a categorical variable for each observation. The standard Fleiss' kappa statistic is only applicable to categorical data. Because the RQS contains elements with more than two possible ratings and Fleiss' kappa would not penalise a disagreement with ratings of 0 and 2 more than 0 and 1, the modified version of the statistic described by Marasini *et. al.* [250], which takes ordinal data into account, was applied (S*). Besides, the modified Fleiss' kappa statistic does not depend on the prevalence of the outcome in the study population or the number of values a rating can take [251, 252]. Confidence intervals for inter-reader agreements were estimated using a bootstrap approach. Random sampling was repeated 1000-times. The significance of inter-reader agreement was assessed using a Monte Carlo simulation. One thousand simulations were performed to determine the distribution of the inter-reader agreement under the null hypothesis of agreement by chance alone. The p-value of the observed S* under the null hypothesis was calculated [253]. The implementation of the modified Fleiss' kappa in the "raters" package (Version 2.0.1) for the R language for statistical computing was used.

A meta-analysis was performed on studies assessing the ability of radiomics to differentiate AMLwvf from RCC as pre-specified in the review protocol. For the purpose of this meta-analysis, fat poor AML, AMLwvf and AML without macroscopic fat were aggregated as they did not differ in their radiological definition. Two-by-two contingency tables were either recorded from the publications or reconstructed if sufficient information was provided in other forms. Only the model with the highest AUC or the highest Youden's J statistic was included for each publication. The meta-analysis was conducted using the "metafor" package (Version 4.2-0) and the summary effect size was calculated from the per-study odds ratio with

a random-effects model. The odds ratios were calculated according to eq. (2.1) where true positive (TP) denotes RCC recognised as RCC by the model, false negative (FN) are RCC misclassified as AMLwvf, TN are AMLwvf classified as AMLwvf and FP are AMLwvf misclassified as RCC.

$$OddsRatio = \frac{TP \cdot TN}{FN \cdot FP} \tag{2.1}$$

The random-effects model was chosen to account for the methodological heterogeneity among the studies. Inter-study dispersion was investigated with Cochran's Q and I^2 metrics. Cochran's Q assesses the hypothesis that the distribution of results is homogenous and p-values < .05 would generally lead to the rejection of this null-hypothesis [254]. As with a small number of studies Cochran's Q can be distorted, I^2, an estimate of the percentage of the variability in effect estimators which can be attributed to true heterogeneity between studies rather than chance through sampling error, was also reported. I^2 values of 25% and less are usually considered to be low or unimportant, 25% - 50% moderate and values above 75% are considered high [255]. Measuring inter-study dispersion assumes that, if all studies were methodologically identical and variation in results were only due to the random selection of study participants, the effect sizes would follow a chi-squared distribution. The results of a trim and fill analysis were represented as a funnel plot in the assessment of the risk of publication bias: The logarithmic diagnostic odds ratio of an individual study is plotted against its standard error which is inversely related to sample size. Asymmetric distribution of studies on the funnel plot can be due to preferential reporting of significant studies and may therefore indicate publication bias. The dotted diagonal lines border the sector where 95% of all studies were assumed to lie [68]. Trim-and-fill analysis aims at identifying and correcting funnel plot asymmetry resulting from publication bias by removing small studies which contribute to the asymmetry and replacing them with their missing counterpart to achieve symmetry. This method performs better with decreasing inter-study dispersion [69].

2.2.3 Results

Seven hundred and seventy-six publications including 263 duplicates were identified. Four-hundred-and-fifty-four of the remaining 513 articles were rejected after screening for the title and abstract. Finally, 57 of 59 manuscripts were included after a review of the full-texts (fig. 2.2).

Radiomics Quality Score and QUADAS-2 assessment

The RQS of the included studies averaged for all reviewers ranged from -4 to 16.6 (mean $\pm$ standard deviation: 3.41 $\pm$ 4.43, median/interquartile range: 4.5/6.17). While the best study achieved an RQS of 46%, the average reached only 9.4%. Average ratings resolved for individual dimensions within RQS are summarised in table 2.1. Study-level results are summarised in table B.1, page 238. Discrimination statistics, biological correlates and suggestions of clinical usefulness featured in most publications. However, the use of texture phantoms, cost-effectiveness analyses, sharing of segmentations and code were absent from all. The repeatability of radiomic measurements and generalisability of models to other patient cohorts was only assessed infrequently. Wang *et al.* were the only to image patients twice and define the repeatability of histogram analysis for DCE-MRI [256]. When patients were re-imaged 48-72 hours after the first scan, the mean and mode of K^{trans} showed moderate agreement between the two scans while skewness and kurtosis showed poor agreement. In contrast, histogram parameters of v_e were well reproducible. The majority of studies (61% 35/57) included only a single or few slices of the tumour in their analysis. Five studies (9%) employed some form of semi-automated segmentation such as interpolation, region growing or a random walker algorithm. One-third of the studies reported inter-reader agreement for the measurement of texture features and extraction of radiomic signatures, ranging from moderate to excellent.

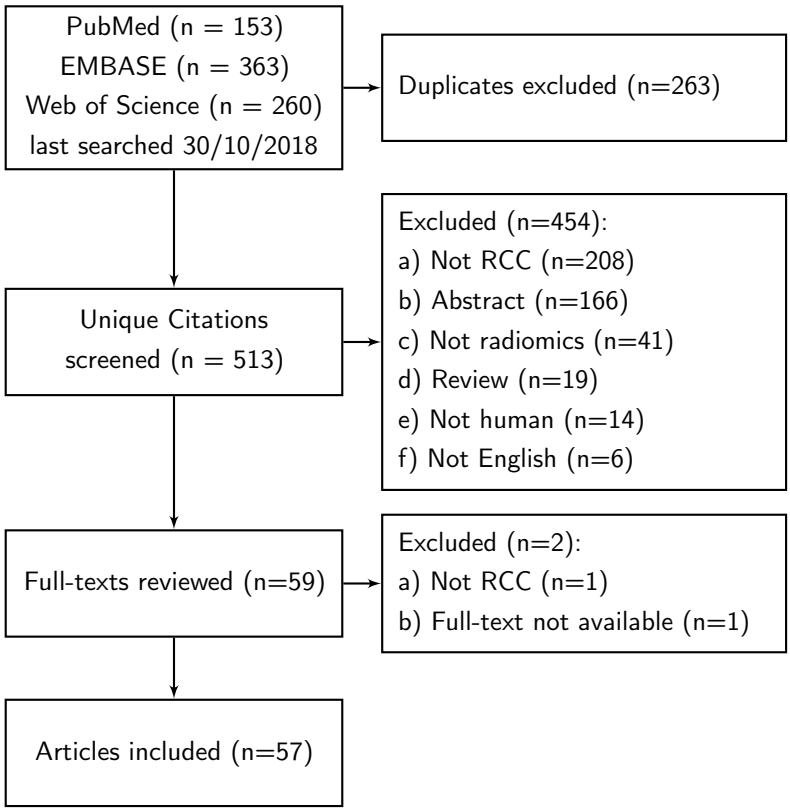


Fig. 2.2 CONSORT diagram of studies included in the systematic review.

Each study included a median of 61 patients (range $2-249$) and extracted a median of 24 radiomic features (range $4-18,720$). The ratio of patients to features ranged from 0.004 to 25 (median 2.2). Just over half of all articles (29/57) described either compensation for multiple comparisons or feature reduction to some degree. While none of the studies was planned and registered prospectively, images in eight (14%) were collected within prospective studies. Three (5%) algorithms underwent independent validation, one of which was validated on data from an external institution that was not part of the training data.

QUADAS-2 identified an elevated risk of bias in the studies included in this review. Individual risk assessments are summarised in table B.2, page 242. Insufficient description of the index test, in particular incomplete information on image acquisition parameters, was the single most prevalent factor increasing the risk of bias (92% missed CT kernel). A retrospective study design (86%) and requirement for participants to have undergone surgery (74%) were also common risk factors of bias in participant selection. Overall, five studies were considered to be at low risk of bias in all four domains.

Reviewers underwent training in the use of the RQS and QUADAS-2 questionnaires as described above. Before the training, reviewers varied in their interpretation on how

biological correlates should be discussed to be rated as present. The reviewers decided to follow existing literature on the RQS and accept clinical endpoints such as survival, tumour type or tumour grade as biological correlates [241]. After the training phase, the modified Fleiss' kappa statistic showed substantial to almost perfect agreement ($0.76 \leqslant S^* \leqslant 1.0$). The agreement was only moderate for the quality of the description of the imaging protocol (S*: 0.45), the discussion of clinical utility (S*: 0.60) and the inclusion of non-radiomic factors in the modelling (S*: 0.67, details in table B.3, page 246). Nonetheless, the summed RQS correlated strongly among the three reviewers ($ICC = 0.92$ [95% confidence interval: $0.80 - 0.98$]). The inter-reader correlation coefficient increased to 0.96 (95% confidence interval: $0.93 - 0.98$) when studies that were rated by only two of the three readers were also considered. Inter-reader agreement in the risk of bias analysis was analysed individually for each of the seven indicator questions in QUADAS-2. The reviewers disagreed strongest on whether patient selection criteria were free from bias (absolute agreement 58%). The agreement was $\geqslant 75\%$ for all other signalling questions.

Meta-Analysis

Positive study outcomes are significantly more likely to be published than inconclusive or negative outcomes [257]. In radiomic research, this is a concern as parameters can be adapted, leading to multiple comparisons, until a positive result is found. None of the studies included in this review followed a study protocol that would prevent this optimisation of results. Seven per cent of the studies under review (4/57) reported non-significant outcomes. All of them reported on the differentiation of RCC and angiomyolipoma.

As pre-specified in the review protocol, we conducted a meta-analysis of the largest group of studies answering a homologous question. The differentiation of malignant tumours of the kidney from AMLwvf was addressed by 13 studies (23%). Ten of these included sufficient information to derive 2×2 contingency tables. Odds ratios, used as the effect size for the meta-analysis, were calculated from the contingency tables (eq. (2.1)). The random-effects model showed a summary effect size of all radiomics approaches to differentiating RCC and AMLwvf of 5.89 (95% confidence-interval: 4.02-8.23 $P < .001$; fig. 2.3a). Inter-study dispersion was moderate with $I^2 = 33.5\%$ and Cochran's $Q = 13.42$ with 9 degrees of freedom failed to reach statistical significance ($P = .15$). The asymmetric distribution of effect sizes and standard errors suggested possible publication bias (fig. 2.3b). However, the effect size would remain significant ($OR = 5.55$ (95% confidence-interval: $3.77 - 8.16$, $P < .001$) even after a symmetric distribution was achieved by trim and fill analysis. Definitions of radiomic

features and features selected for the final signature differed between the studies and no feature was employed in more than two studies.

2.2.4 Discussion

In this systematic review, we described that common methodological characteristics delay the translation of many proposed radiomic algorithms into clinical practice. The scarcity of external validation, insufficient description of the methods and reluctance to share data and code delay the generalisation of algorithms to other patient cohorts. Machine learning studies, even though being particularly difficult to reproduce without access to the code, achieved a higher average rating than studies not involving any machine learning methods (5.16 ± 3.66 vs. 0.83 ± 4.27, $P < .001$). This may have been related to them being published more recently. These challenges will need to be addressed for radiomics to fulfil its promise of improving diagnostic performance with existing hardware and protocols.

Radiomic research is a classical example of a large p, small N problem. This describes questions where statistical inference is attempted with a large number of variables and relatively few observations. The multiplicity of comparisons increases the risk of detecting spurious correlations occurring by chance alone. Traditional methods to correct for multiple testing like the Bonferroni correction are well suited for a modest number of comparisons but can be overly conservative for large numbers of parameters and when these are not fully independent. The Holm procedure or false discovery rate (Benjamini-Hochberg) can limit the risk of discovering spurious correlations and are commonly used in genomics research with thousands of parameters. A rule of thumb suggests that ten events per fitted variable are required for logistic regression and proportional hazard models [258]. However, in radiomic research as well as genetic correlation studies this rule is frequently violated. The intrinsic dependency of radiomic features is another reason why feature reduction is crucial in radiomic studies. Commonly used techniques to remove strongly correlated radiomic features include principal component analysis (PCA) and the least absolute shrinkage and selection operator (LASSO). Additionally, poorly reproducible features could be identified in phantom studies and removed from the analysis. Ultimately, independent validation will have to demonstrate the value of ML algorithms in clinical practice.

Strikingly, none of the multifactorial radiomic models reviewed here has been independently validated or achieved clinical translation despite showing potential to answer new questions not commonly answerable by radiology. Furthermore, the exclusively retrospective design risks introducing reporting bias as researchers may adapt the methodology and

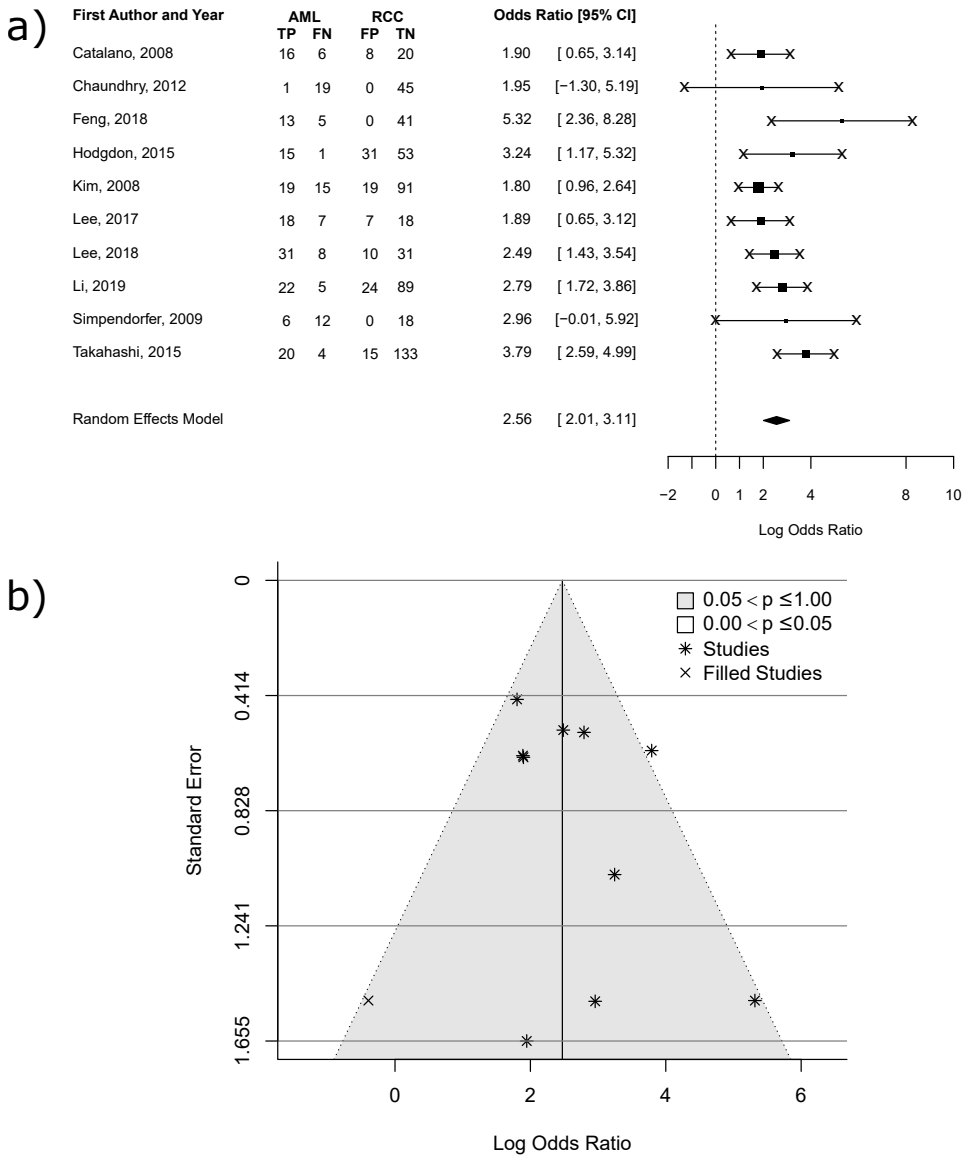


Fig. 2.3 (a) Forest plot depicting the effect size as log-transformed odds ratios for the differentiation of AML without visible fat (AMLwvf) and renal cell carcinoma (RCC) using radiomics. True positives: AMLwvf correctly diagnosed, False negatives: AMLwvf diagnosed as RCC. False positives: RCC diagnosed as AML, True negatives: RCC correctly diagnosed. For the calculation of the summary effect, a random-effects model was used. (b) Funnel plot comparing the log-transformed odds ratios to the standard error (stars). Trim and fill analysis was employed to identify missing studies to obtain a symmetrical distribution (cross). A slight asymmetry with one missing publication reporting a lower than average odds ratio was identified. This can indicate a publication bias.

research questions until statistical significance is achieved while non-significant results are not reported. Additionally, most trials focused on surgical patient cohorts. These will be enriched in larger and malignant renal tumours, causing selection bias.

The critical appraisal of radiomic studies is one key factor in the selection of potentially clinically useful, high-quality models which should be prioritised for independent validation. The RQS is one tool that may facilitate this. However, due to its relative novelty, the score has not been broadly validated yet. The first application of the score showed varying interrater agreement [241]. Here we used two manuscripts to harmonise the application of the scoring system in this review. As a result, a very high agreement in the overall score and high agreement in all its items was achieved. Compared to the first application of the RQS, the average RQS rating was lower (9.4% vs 21.9%) as was the rating for the best-performing study (48% vs 55.5%). Another recently published review employing the RQS did not report interrater agreement [259]. Only few systematic reviews in radiomics literature have been published and even fewer assessed methodological quality systematically and quantitatively. As a result, the RQS has not yet found widespread application.

The appearance of radiological images varies widely depending on manufacturers, scanner models, acquisition and reconstruction settings. Only every second article in this review described the most important image acquisition parameters, despite evidence that they significantly affect many texture parameters [260, 229, 261]. One of the challenges of developing well-generalisable radiomic models is the standardisation of imaging to a point where the signal from tumour biology is not drowned by noise introduced by technical heterogeneity. Where standardisation is not feasible, the feature selection should take variations in acquisition parameters into account and correct the input data or texture features accordingly.

Texture analysis on MRI faces additional challenges as T1- and T2-weighted sequences are inherently non-quantitative. MR parametric maps may provide quantitative information on e.g. diffusivity, oxygenation or perfusion. In theory, these are standardisable. However, data on the repeatability and reproducibility of the ADC demonstrated coefficients of variation of 8.7% – 10.1% in liver metastases which were similar in magnitude to changes induced by 12 – 14 days of chemotherapy [262, 263]. Even greater coefficients of variation were demonstrated for DCE-MRI and K^{trans} measurements, reaching 21.8% in pancreatic adenocarcinoma, 15.6% in RCC, 13.9% in glioblastoma and 12% in prostate cancer [264–267]. Consequently, even these parametric maps pose difficulties for texture analysis. The majority of investigations on MR radiomics in this review (11/17) included parametric maps, most commonly the ADC or K^{trans}.

More often than not, texture models fail to consider intra-tumoral heterogeneity, despite its prognostic value [2, 150]. First, regions of interest frequently encompass only a single slice or few slices of the tumour. Second, averages of radiomic features for entire tumours are commonly used in the generation of radiomic models. Both factors may miss small highly aggressive tumour regions and are at risk of decreased prognostic ability. Few studies compared the diagnostic performance of whole lesion ROIs and single 2D ROIs and no conclusion on their diagnostic equivalence was reached. Spatially resolved analysis of radiomic data would also open further possibilities for multi-regional integration of imaging data with tissue biomarkers of RCC from the genomic, transcriptomic, proteomic and metabolomic analyses. This could improve the understanding of the representation of tumour biology in imaging phenotypes and the complementary value of imaging and molecular tissue analysis in the diagnosis and treatment planning of RCC.

This systematic review and meta-analysis have some limitations: first, the differentiation of oncocytomas from RCC would be clinically more impactful than the differentiation of RCC and AMLwvf. Among small renal masses, oncocytomas account for 60% of benign, resected tumours whereas angiomyolipomas account for 27% [64]. However, oncocytomas were assessed in only eight articles compared to 13 assessing the clinically less relevant differentiation of fat-poor angiomyolipomas and RCC. Second, the publications included in the meta-analysis exhibited some methodological variability which limits the direct comparability of their results. The inclusion criteria were particularly variable, where the control group consisted of RCC of multiple histological subtypes in six and ccRCC only in four studies. Different imaging equipment and the vast choice of filtration and radiomic features introduce further variation among studies. The selection of a random-effects model for the meta-analysis can compensate for some of this variability and offer relevant information on the consistency of results and the effect size.

The availability of the RQS to improve manuscripts that were submitted for publication after 2017 may have introduced some bias. However, there was no trend towards an improved rating over time. Finally, the tools used to assess the methodological quality of manuscripts, the RQS and QUADAS-2, have some limitations as well. The RQS is relatively recent and therefore not well-established yet. Furthermore, it is unclear, how the weighing of individual factors in the score was determined. QUADAS-2, on the other hand, is well-established but qualitative in nature, which complicates the direct comparison of manuscripts.

In summary, this systematic review has identified characteristics shared among much of the radiomic research in renal cancer which will complicate translation into clinical applications. Prospective studies and well-designed models will be required to demonstrate

where radiomic algorithms can supplement clinical practice. Finally, strategies to produce generalisable and robust radiomics models need to be found.

2.3 Strategies for the development of robust and generalisable radiomic models

Acknowledgements:
The work presented in this section received substantial support from Dr Lorena Escudero Sanchez and Dr Leonardo Rundo who developed the code for the automated perturbation of tumour segmentations, implemented the pipeline for texture feature extraction, and conducted the statistical analysis. Prof Grant Stewart and Mr Tobias Klatte kindly provided access to the surgical database of patients undergoing nephrectomy.
Stephan Ursprung has conceived the study in collaboration with Dr Rundo, Dr Escudero, Prof Evis Sala and Prof Stewart, defined the research questions, retrieved and anonymised the imaging data, extracted the clinical information from the participants' electronic medical records, segmented the tumours, performed the statistical analysis jointly with Dr Escudero and Dr Rundo and interpreted the results.

2.3.1 Introduction

Radiomics are potentially very useful in the diagnosis and treatment stratification of renal cancer [240]. However, widespread clinical application is lacking despite active research initiatives. The hesitance to undertake the labour intensive and costly validation of these novel imaging biomarkers are reasons for this. Emerging imaging biomarkers first need to undergo technical validation and demonstrate their repeatability and reproducibility. A better understanding of the vulnerabilities of the technique is crucial to increase the likelihood of data-driven imaging biomarkers to generalise well regarding time and acquisition centre [232, 268]. Unless imaging biomarkers are technically valid, they are unlikely to perform favourably in the clinical validation step.

The development of well-generalisable texture models requires a careful selection of parameters and validation in a heterogeneous dataset. However, depencencies between radiomic features on the one hand and acquisition parameters, feature extraction and segmentation on the other hand are poorly understood. We hypothesised that an improved understanding of

these confounding factors would increase the generalisability of models answering clinically relevant questions in RCC imaging. Therefore, a dataset consisting of patients undergoing nephrectomy at Addenbrooke's hospital encompassing all imgaing studies from referring hospitals was collected. The aims of this retrospective study were severalfold: First, to explore the sensitivity of CT radiomic features in renal cancer to variations in the tumour segmentation. Second, to determine the robustness of radiomic features to varying acquisition parameters. Third, to identify features that are sensitive to the underlying tissue biology. Finally, to define radiomic models for the biological characterisation of renal cancer on a subset of robust features. This section discusses the results of the first aim and the plans for the remaining aims.

2.3.2 Methods

Participants and ethics. This retrospective analysis included patients who underwent tumour nephrectomy at Addenbrooke's hospital between 01/01/2015 and 31/07/2017. Consecutive participants were identified from the prospectively maintained surgical database. The last pre-surgical CT of the chest, abdomen and pelvis was retrieved from the institutional PACS system and anonymised for the analysis. Addenbrooke's is the regional referral centre for nephrectomies, therefore, this study included CT scans from Addenbrooke's hospital and its referral centres in the East Anglia region. Patients had to be >18 years of age at the time of surgery, be diagnosed with a benign or malignant renal tumour on post-surgical histology, be treatment naïve and have no underlying renal disease like autosomal dominant polycystic kidney disease (ADPKD).

Patients were excluded from the analysis if no cortico-medullary phase CT of the primary tumour was available, if a split bolus contrast administration scheme was utilised or if tumours were purely cystic and contained no solid tumour component with a maximum axial diameter $\geqslant$1 cm in any slice. Similarly, tumours $\leqslant$1 cm in diameter were excluded. Furthermore, patients were excluded if artefacts (e.g. beam hardening artefacts from metal implants) covered the renal tumour or if tumours were incompletely imaged. Patients were excluded if they underwent tumour biopsy immediately before CT. Split bolus contrast administration serves to achieve opacification of the urinary tract and a portal venous phase-like enhancement in a single acquisition. Split bolus CT scans show equivalent diagnostic value to two separate acquisitions and reduce the radiation dose to patients. Simultaneously, the procedure time and need for repeating the excretory phase image owing to insufficient opacification are higher with a split bolus [269, 270]. This retrospective study was approved

by the institutional research ethics committee of the University of Cambridge and Cambridge University Hospital's Research and Development (R&D) department that waived the requirement for written informed consent.

Image acquisition and segmentation. Contrast-enhanced cortico-medullary phase CT scans were acquired on the scanner hardware and with the standard protocol of the referring centre. CT scans were acquired 70-100s after the intravenous injection of iodine-containing contrast agents (e.g. Omnipaque 300 mg I/L, GE Healthcare) dosed by weight. Typical acquisition and reconstruction parameters were (mean [range; $\pm$ S.D.]): slice thickness of 4.4 mm [1.0 – 5.0 mm; $\pm$ 0.96 mm], in-plane voxel size: 0.77 mm [0.54 – 0.98 mm; $\pm$ 0.1 mm], tube voltage of 116 kVp [100 – 140 kVp; $\pm$ 11 kVp], automatic tube current modulation 269 mAs [51 – 663 mAs; $\pm$ 178 mAs], soft tissue kernel, pitch factor 1.06 [0.6 – 1.4; $\pm$ 0.2]. The two most common combinations of acquisition parameters are summarised in table 2.2.

Images were retrieved in the DICOM format and segmented semi-automatically using Microsoft Radiomics software (Version 1.0.30558.1, Microsoft, Cambridge UK.) This software was developed as part of the InnerEye project (https://www.microsoft.com/en-us/research/blog/project-innereye-open-source-deep-learning-toolkit-democratizing-medical-imaging-ai/, accessed: 22/06/2021). The renal tumour, contralateral normal kidney, major vessels and right-sided psoas muscle were outlined (fig. 2.4). A proprietary software algorithm first determined the structures' outer margins based on a rough manual segmentation. This segmentation was refined by the user and after a few slices of the CT had been segmented, the algorithm interpolated the remaining slices. This interpolation minimised a combined term of the euclidean distance and a distance metric based on grey level gradients between a new and existing contour. Interpolated segmentations were adjusted manually which resulted in an iterative correction of the automated segmentation. The finalised segmentation was exported in the DICOM RTSTRUCT (radiotherapy structure) format. With the support of the Microsoft Radiomics tool, an abdominal CT series could be segmented in 20 – 30 minutes. If more than one renal tumour was present in a dataset, all were segmented but only the largest was considered for further analysis. For the assessment of intra-reader variability in texture parameters, the segmentation of 141 randomly selected cases was repeated after a 4 month washout period.

External validation dataset. The Kidney Tumour Segmentation Challenge run in 2019 (KiTS19) made a dataset of 210 CT scans and corresponding segmentations of patients who underwent nephrectomy at the University of Minnesota Medical Center between 2010

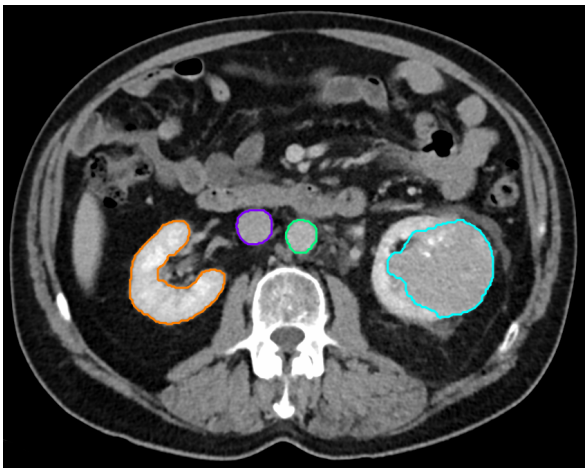


Fig. 2.4 The renal tumour (cyan), normal kidney (orange), abdominal aorta (green) and inferior vena vaca (violet) were segmented semiautomatically.

Table 2.2 Acquisition parameters of the CT scans used in the radiomic analyses.

Parameter	Combination 1	Combination 2
Vendor	Siemens	GE
Model	Somatom Definition AS(+)	Optima CT660
Reconstruction kernel	I30f/3	Standard
Slice thickness [mm]	5	3.75
Tube voltage [kVp]	100	120
Tube current [mAs]	349 [63 – 642; ±178] (CARE Dose)	152 [79 – 351; ±72]
Contrast delay [s]	100	70
In-plane resolution [mm]	0.7 [0.54 – 0.87; ±0.08]	0.80 [0.70 – 0.98; ±0.07]
Pitch factor	1.15 [0.6 – 1.4; ±0.26]	0.98

Values indicate the mean. Range and standard deviation in square brackets.

and 2018 publicly available. These CT scans were acquired in the late arterial phase, and tumours and the adjacent normal-appearing kidney were segmented semi-automatically [271]. Patients were excluded from further analysis if they presented multiple tumours and the histological data could not be assigned to one lesion unambiguously, lesions which were smaller than 5 cm^3 or if the segmentation was inaccurate and involved structures besides the tumour. The cutoff of 5 cm^3 was selected to reach a cohort comparable to the internal dataset. Sixty-six scans were excluded and 144 patients were included.

Morphological operations. Morphological operations were performed on the 3D segmentation of renal tumours to approximate the effects of inter-reader variability. It has been

shown in previous studies, for example in carotic plaques [272], that the variation of the ROIs obtained after performing automatic morphological operations is comparable to the differences observed with multiple segmentations by different radiologists. These operations are applied here to the internal and external validation datasets.

Transformations were applied with a transformation element of one voxel radius in all three dimensions. Four operations were investigated:

- Dilation of the ROIs defined a voxel at the centre of the transformation element as belonging to the ROI if one voxel in the transformation element belonged to the ROI.
- Erosion of the ROI defined a voxel at the centre of the transformation element as belonging to the background if one voxel in the transformation element belonged to the background.
- Opening performed an erosion followed by a dilation operation. This removed small protrusions in the ROI and provoked the confluence of small but adjacent gaps in the ROI. The opening operation was performed once and twice to achieve a gradual increase in its effect.
- Closing performed a dilation followed by an erosion, which removed small areas of background and led to the confluence of small protrusions in the ROI. The closing operation was not considered in the final analysis because segmentations remained effectively unaltered unless many iterations were performed.

Radiomic feature extraction. First, scans and segmentations were converted from the DICOM to the Neuroimaging Informatics Technology Initiative (NIfTI) format [273] using in-house software implemented in MATLAB R2019b. Subsequently, 107 3D radiomic features were computed with the PyRadiomics package (version 3.0.1) for the Python environment. No image filters were applied and 14 shape, 18 first-order, 24 GLCM [216, 274], 14 Grey Level Dependence Matrix (GLDM) [275], 16 GLRLM [276], 16 GLSZM [277] and 5 Neighbouring Grey Tone Difference Matrix (NGTDM) [278] features were calculated. A list of the computed features can be found in table B.4 (page 247) and a detailed description of the features on https://pyradiomics.readthedocs.io/en/latest/features.html (accessed: 27/06/2021).

Scans were quantized into 64 grey levels for the calculation of the 2^{nd} order texture features. As alterations of features were compared within cases, no spatial resampling was applied.

The ROOT open-source data analysis framework developed at CERN (Geneva, Switzerland) was used *via* its Python interface to integrate with other Python libraries, such as NumPy, SciPy and scikit-learn for the data analysis [279].

Clinical data. Clinical data were extracted from the surgical database and included: age, sex, type of surgery, date of diagnosis, date of surgery, date of recurrence, date of death, histological subtype, TNM stage at diagnosis, WHO/ISUP tumour grade, presence of tumour necrosis, sarcomatoid or rhabdoid features, Leibovich risk score, presurgical and postsurgical renal function.

Data analysis. The intra-reader agreement and the agreement between the original, manual segmentation and its automated deformations was assessed with the Sørensen-Dice similarity coefficient (DSC, eq. (2.2)). Subsequently, the relative difference in texture features extracted with the respective segmentations was computed according to eq. (2.3) where f is the texture feature. The subscript denotes the first and second measurements of the feature. Features where the median Δ_r across patients exceeds 20% or where the median Δ_r across patients exceeds 10% and the 3$^{\text{rd}}$ quartile exceeds 25% will be deemed non-robust to variations in the tumour segmentation.

$$\text{DSC} = \frac{2\,|X \cap Y|}{|X| + |Y|} \tag{2.2}$$

$$\Delta_r(\%) = 100\frac{|f_1 - f_2|}{|f_1|}. \tag{2.3}$$

2.3.3 Results

Dataset. The prospectively collected surgical database identified 405 patients who underwent tumour nephrectomy between January 2015 and July 2017. Ninety-eight of the identified datasets were excluded from further analysis. The main reasons were the non-availability of a nephrographic phase (n = 67) and insufficient solid tissue for segmentation and radiomic feature extraction (n = 13) (fig. 2.6). Finally, scans of 307 patients were segmented. Figure 2.5 visualises the distribution of the most important acquisition parameters and table 2.3 summarises the patient characteristics.

For the study of the intra-reader agreement and morphological operations, 141 cases which had been segmented twice were available. However, 18 tumours were so small that masks were empty after the opening operation. Therefore, these were excluded from this analysis and the remaining 123 cases were investigated alongside the 144 cases from the KiTS19 dataset.

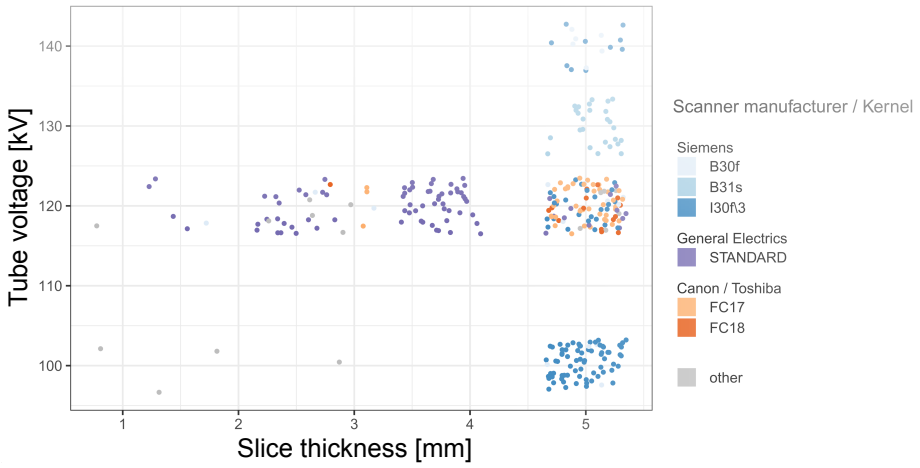


Fig. 2.5 Distribution of the slice thickness, tube voltage, scanner manufacturer and reconstruction kernel among the scans included in the dataset.

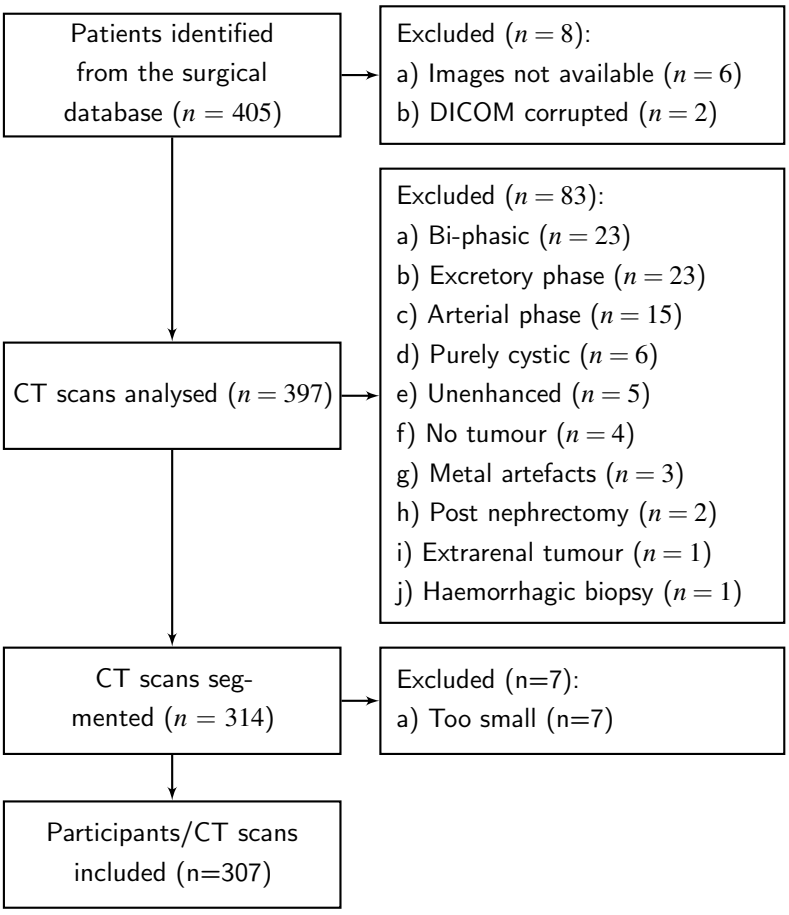


Fig. 2.6 CONSORT diagram of patients included in the RCC radiomics study

Intra-reader agreement. Intra-reader agreement in the segmentation of renal tumours was high with a median DSC of 0.952, (IQR: 0.936 – 0.966, fig. 2.7). A subset of texture features showed a high relative variation between the first and the second segmentation: 23/107 features were poorly reproducible according to the criteria outlined above (fig. 2.8). While all shape features were well reproducible, the first-order features of the minimum intensity and skewness of the grey-level histogram varied considerably between the two segmentations. Eight of the poorly reproducible features were features emphasising low grey levels while only two features emphasised high grey levels (table B.5, page 257).

Sensitivity to morphological operations. Figure 2.9a shows an example of the morphological operations that were applied to the tumour ROI to simulate their robustness to inter-reader variation. The opening operation had only a minor effect on texture features. All features remained reproducible after one iteration of the opening and only a single feature, cluster

Table 2.3 Characteristics of the patients involved in the radiomic analyses.

Patient characteristic	n Total: 307	%
Age (Median ± IQR)	65 ± 15.5	
Sex (female/male)	109/198	35.5/64.5
Histology		
clear cell RCC	228	74.3
papillary RCC	34	11.1
Oncocytoma	18	5.9
chromophobe RCC	16	5.2
other malignant	4	1.3
other benign	4	1.3
Angiomyolipoma	3	1.0
WHO/ISUP Grade (ccRCC and pRCC)		
1	0	0
2	112	
3	97	
4	52	
NA	1	
Tumour Stage (only malignant)		
T1a	67	23.9
T1b	38	13.6
T2a	9	3.2
T2b	4	1.4
T3a	149	53.2
T3b	7	2.5
T3c	0	0
T4	6	2.1
Treatment		
radical nephrectomy	233	75.9
partial nephrectomy	74	24.1
Leibovich score (ccRCC)		
0	31	13.6
1	16	7.0
2	23	10.1
3	10	4.4
4	32	14.0
5	29	12.7
6	32	14.0
7	14	6.1
8	22	10.5
9	18	7.9
10	1	0.4
11	3	1.3

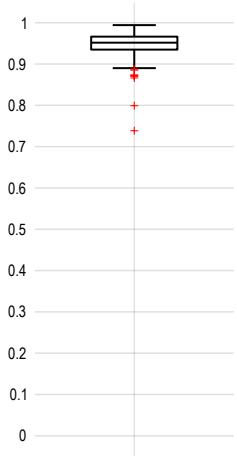


Fig. 2.7 Intra-reader agreement in the segmentation of renal tumours quantified with the Dice similarity coefficient.

shade, which also exceeded the variability threshold in the study on the intra-reader agreement, was poorly reproducible after two iterations (figs. B.1 and B.2, page 251). The erosion and particularly dilation operations were more aggressive as observed in the DSC (fig. 2.9) and introduced higher variations. After the erosion, 50/107 features were irreproducible, which included all of the irreproducible features from the intra-reader agreement. Following the dilation operation, 63 features were irreproducible, again including all the irreproducible features from the intra-reader agreement. Most features which were poorly reproducible following an erosion were also poorly reproducible after the dilation. The only two exceptions were the surface area and GLCM difference average (figs. B.3 and B.4, page 253). All features with high grey level emphasis were sensitive to these more aggressive morphological operations (table B.6, page 260).

External validation. The morphological operations were applied to the publicly available segmentations of the KiTS19 dataset, which yielded DSCs comparable to the internal dataset (fig. 2.10a). However, the variation in the volumes of ROIs after the application of the morphological operations was much greater for a given baseline volume in the KiTS19 validation dataset compared to the internal dataset (figs. 2.10b and 2.10c). The reproducibility of texture features was comparable between the internal and the KiTS19 datasets. Erosion and dilation tended to introduce less variation in some texture parameters in the KiTS19 dataset (figs. B.5 and B.6). Table B.7 on page 264 details the median values for each feature.

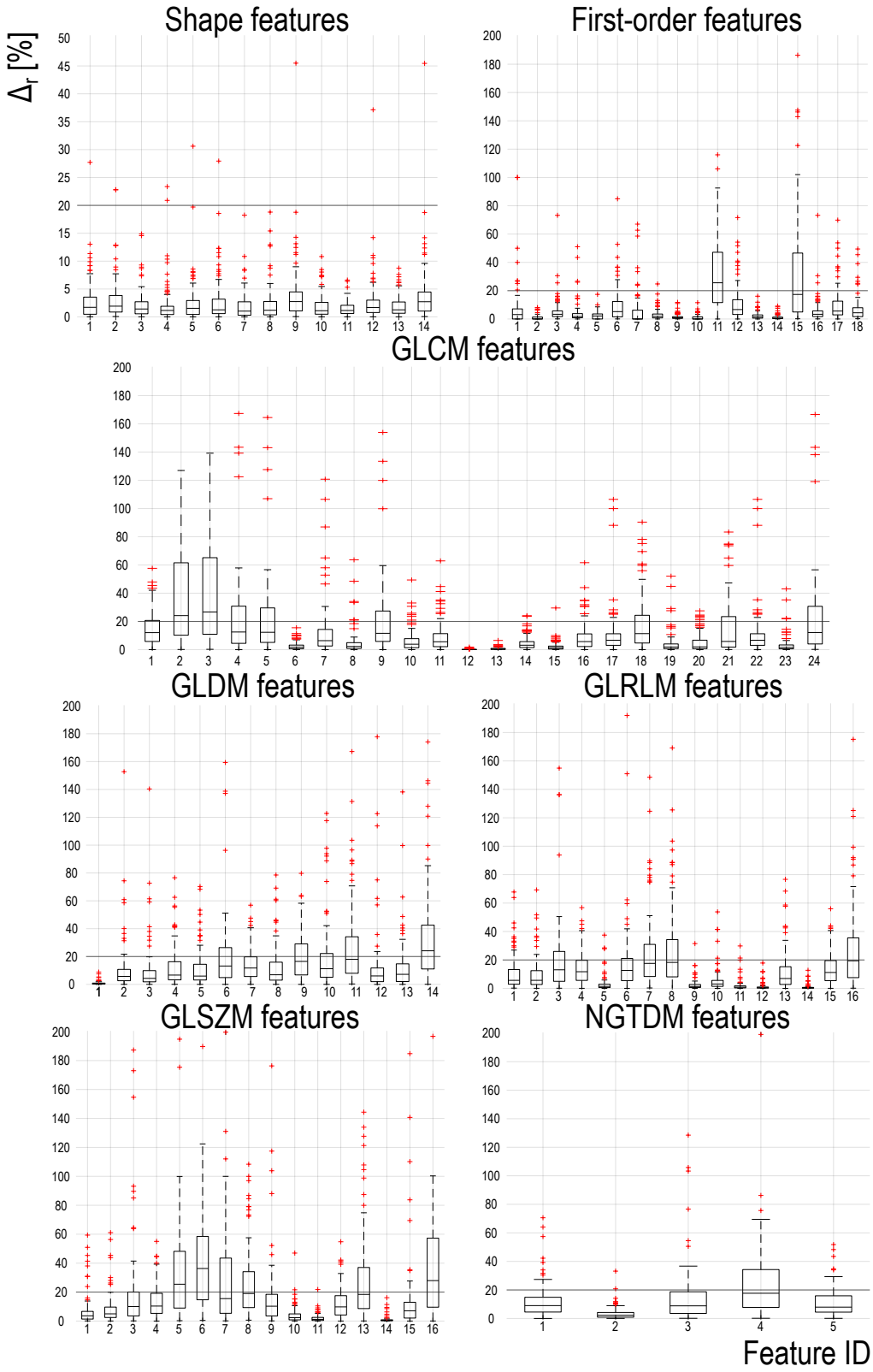
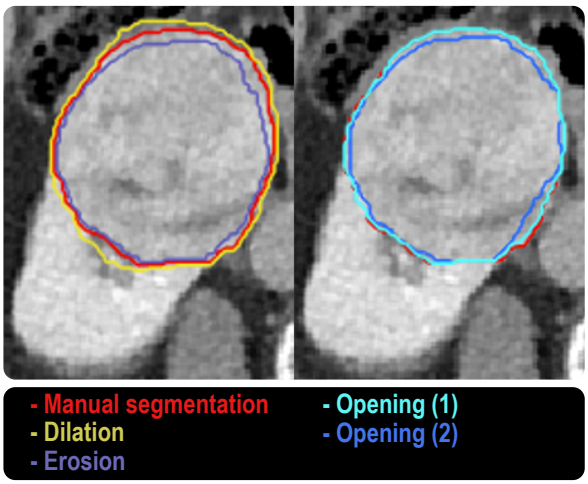
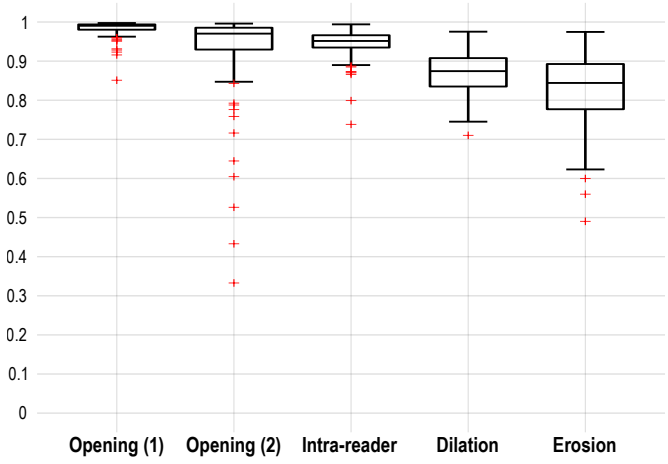


Fig. 2.8 Intra-reader variation in texture feature measurements. Boxplots reach from the 25thto the 75th percentile. Whiskers extend to the 5th and 95th percentile. Please note the different scaling of the y-axis in the top left panel. Please refer to table B.4, page 247 for the IDs of the texture features. GLCM: Grey level coocurrence matrix,GLDM: Grey-level density matrix, GLRLM: Grey-level runlength matrix, GLSZM:Grey-level size zone matrix, NGTDM: Neighbourhood grey-tone difference matrix.

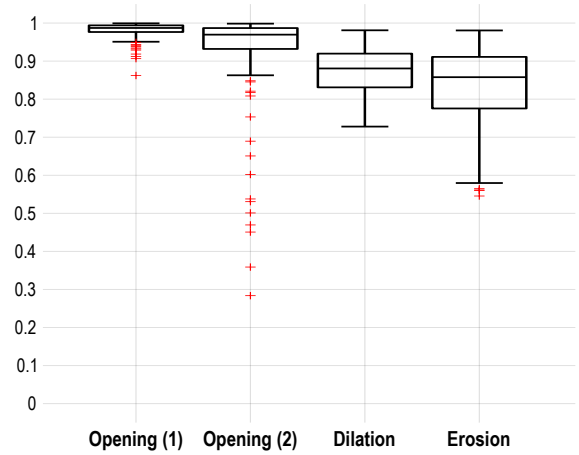


(a)

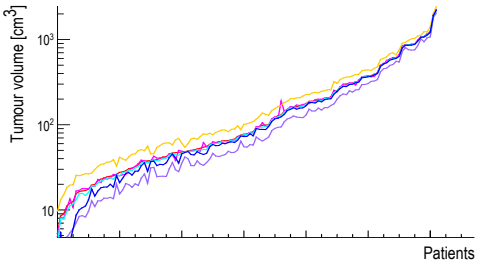


(b)

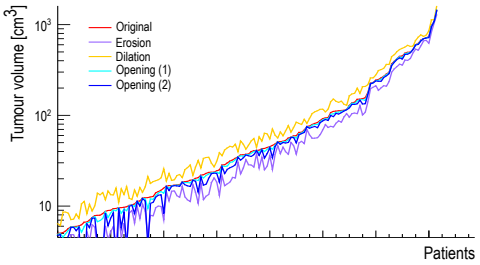
Fig. 2.9 Example of the morphological operations (erosion, dilation and opening) performed on the original, manually segmented region of interest (a). Dice Similarity Coefficient (DSC) for the manual re-segmentation and the automated morphological operations compared to the ground truth segmentation, ordered by decreasing DSC (b).



(a) DSC for the KiTS validation dataset



(b) Internal dataset



(c) KiTS validation dataset

Fig. 2.10 Dice Similarity Coefficient (DSC) for the automated morphological operations on the KiTS19 dataset (a). Comparison of the tumour volume before and after the morphological operations in the internal and the KiTS19 datasets (b and c).

2.3.4 Discussion

The initial results of this study highlight important considerations for the rational selection of texture parameters to be included in CT radiomics-based diagnostic and prognostic models. Some texture features are highly sensitive to minor variations in the segmentation of renal tumours. Fewer features remain reproducible as perturbations of the segmentation become more aggressive. Furthermore, similar features were irreproducible in the internal and external (KiTS19) cohorts, suggesting that features are well or poorly reproducible regardless of the segmenting clinician, the acquisition hardware, software and other particularities in the selected cohorts of patients.

The variability of features in the context of perturbed segmentations will be central to the assessment of the robustness of radiomic algorithms. While features with a high relative difference in this study are not automatically unsuitable for inclusion in models, special care needs to be taken in using them for a classification or prediction purpose. For example, if a model appears to differentiate between two tumour types or prognosticate outcomes based on a small (e.g. 20%) difference in a feature with a large inter-reader variation (e.g. 30%), that model is unlikely to generalise well, potentially producing different results for different expert segmentations. It will be the next step of this study to evaluate the impact of less reproducible features on a model's performance.

This retrospective study had limitations that are common to all retrospective radiomic research. Retrospective inclusion of patients risked introducing selection bias as patients who were ineligible for surgery owing to comorbidities were not included in the surgical database used to identify the patients for this study. On the other hand, tumours of patients who did not undergo surgery were not phenotyped equally well and would reduce the overall quality of the data. However, a second cohort of 164 participants who underwent primary oncological treatment including follow-up imaging is available to study in a next step. Furthermore, image acquisition parameters differed between sites. However, this will always be the case for real-world datasets where different vendors use different filtration, reconstruction kernels and detector architectures, requiring robustness analyses of features in heterogeneous cohorts. Radiomic algorithms which are robust to such differences are most likely to be well-generalisable.

Ongoing research divided the dataset into subsets that are identical in all but one imaging parameter to study which features are correlated with the reconstruction kernel, tube voltage and slice thickness. Furthermore, the segmentation of the normal kidney and psoas muscle is being utilised to identify and exclude features from further analysis which are insensitive to phenotypically vastly different tissue types. A subset of features that is sufficiently robust to

variations in acquisition parameters and segmentation as well as sensitive to tissue phenotype will be used to derive machine learning models for the differentiation of benign and malignant tumours, tumour subtypes and aggressiveness. Furthermore, models to predict the risk of tumour recurrence and renal failure will be explored.

In conclusion, some texture features are extremely sensitive to variations in the tumour segmentation of different RCC CT datasets. This work has identified features that require particular care and evaluation when included in radiomic models. It complements future sensitivity analyses investigating other vulnerabilities of imaging biomarkers.

Chapter 3

Response to tyrosine kinase inhibitors in RCC assessed using conventional imaging

3.1 Background

The standard of care follow-up for RCC undergoing systemic treatment is described above (section 1.2.4, page 13). However, the unidimensional measurement of tumour size, the most commonly used imaging response parameter, is a late sign of treatment response and failure. Additionally, the specific mechanisms of action of immunotherapy can mask treatment response based on size. Peudoprogression describes an initial increase in tumour volume or appearance of new lesions followed by stabilisation or response, which can be mistaken for progressive disease. Increased T-cell infiltration following the initiation of immunotherapy is thought to explain this phenomenon [280]. The use of volumetric response criteria is justified in settings where the reduction in tumour size predicts the outcome. Neoadjuvant systemic treatment in patients with venous tumour thrombi may facilitate the surgical resection of the tumour and decrease morbidity. In this context, it is the reduction in tumour size that will determine the success of the treatment.

New response parameters in clinical research and practice are, therefore, needed to identify treatment success, minimise the time patients spend on ineffective treatment and accelerate the progression onto a new line of therapy.

3.1.1 Physiological MR imaging in RCC as a marker of treatment response

MRI can provide information on physiological parameters in addition to anatomical structures.

Diffusion-weighted imaging (DWI)

DWI measures the Brownian motion of water, which is commonly restricted in neoplasms, aiding diagnosis, prognosis, and therapy response prediction [281, 282]. The concept of imaging diffusion emerged in 1985, and the same pulse sequence has remained the backbone of DWI [283, 284]. Pulsed gradient spin-echo sequences are most commonly employed to acquire diffusion sensitive MR images. Two inverse gradient pulses supplement a spin-echo sequence that could otherwise serve for the acquisition of T_2-weighted images. The first gradient pulse de-phases spins based on their position along the gradient. The second inverse gradient pulse re-phases the spins that have remained at the same position along the gradient axis. Nuclei that have diffused away from their original location will not experience the same inverse pulse and retain a phase shift, leading to signal attenuation. The gradients are applied over three orthogonal directions, and the signal is averaged to reduce the effect of tissue anisotropy [285]. Regions of high diffusion restriction containing mainly immobile protons will appear bright on DWI, and unrestricted diffusion will appear dark. The b-value (Equation 3.1 where: γ denotes the gyromagnetic ratio, G the amplitude of the diffusion gradient pulse, δ the duration of the pulse and Δ the interval between the pulses) quantifies the strength of the diffusion sensitivity. The signal intensity S in DWI is dependent on the signal intensity without the application of diffusion gradients S_0, the b-value and the diffusion coefficient D indicating the velocity of diffusion in the tissue (Equations 3.2 and 3.3) [286].

$$b = (\gamma G \delta)^2 (\Delta - \frac{\delta}{3}) \tag{3.1}$$

$$\ln\left(\frac{S}{S_0}\right) = -bD \tag{3.2}$$

$$General: \quad \ln\left(\frac{S}{S'}\right) = -(b - b')D \tag{3.3}$$

DWI sequences are intrinsically T_2-weighted, especially at low b-values. The estimation of the apparent diffusion coefficient (ADC) produces diffusion maps independent of T_2 [287]. For the calculation of the ADC, the experimentally determined value of the diffusivity (D),

at least two b-values are needed as per equation 3.4 [288]. Some found the acquisition of more than two b-values to reduce the variation of the results and therefore leads to a more accurate estimation of the ADC. However, results are conflicting; a consensus remains to be established [289, 290]. The ADC may aid the discrimination of ccRCC, pRCC and chRCC [291, 292]. However, its ability to distinguish benign from malignant lesions and lesion subtypes is inferior to contrast-enhanced MR [293–295]. A recent systematic review and meta-analysis of 2588 renal lesions reported a sensitivity and specificity for the differentiation of ccRCC from other tumour types of 80% and 78% respectively [296]. Additionally, the ADC correlates inversely with the nuclear grade in ccRCC [297, 298] and tumour growth rate in VHL-associated ccRCC [299].

$$ADC = \frac{\ln(\frac{S'}{S})}{b - b'} \tag{3.4}$$

More advanced modelling of the diffusion signal may provide additional information on tumour physiology. In contrast to regular DWI, diffusion kurtosis imaging (DKI) bases on a more general model which does not assume a Gaussian distribution of the diffusion of water molecules and captures the difference from a normal distribution with the kurtosis term [300]. Therefore, an excess kurtosis term supplements equation 3.2 to quantify the difference of the diffusion from a normal distribution (Equation 3.5). The renal architecture, consisting of directional structures like the loop of Henle and the collecting ducts in the medulla, results in a non-Gaussian distribution of the diffusion and, consequently, apparent diffusion (D_{app}) and kurtosis (K_{app}) may represent tissue microstructure better than ADC alone. Although evidence suggests that DKI parameters may distinguish high- from low-grade ccRCC [301, 302], the utility of DKI in renal imaging remains to be established [303].

$$\ln\left(\frac{S}{S_0}\right) = -bD_{app} + \frac{1}{6}b^2 D_{app}^2 K_{app} \tag{3.5}$$

Depending on the b-value and tissue composition, the diffusion of intracellular water, extravascular extracellular water, and the microcirculation in the capillary bed make up the diffusion signal. The concept of intravoxel incoherent motion (IVIM) accounts for these processes, and imaging techniques are available to disentangle the contributions of blood flow and molecular diffusion to the DWI signal. The speed of the directional motion of water

molecules in the microperfusion is one order of magnitude larger than the non-directional motion of water in the stroma. As a result, the MR signal from intracapillary water decreases more rapidly with increasing b-values (diffusion-weighting) than tissue water. Bi-exponential modelling of the diffusion signal may disentangle the contribution of microperfusion with the associated pseudo-diffusion coefficient D^* and true tissue diffusion with the diffusion coefficient D_0 to the total tissue diffusion signal. The relative contribution of each can be quantified with the perfusion fraction f_p denoting the relative tissue volume occupied by microvessels [304, 305]. The equation for the IVIM type diffusion signal can be obtained by adding a pseudo diffusion term to equation (3.2) resulting in equation (3.6) where f_p designates the perfusion fraction, D_0 the apparent diffusion coefficient and D^* the pseudo diffusion coefficient). In renal diffusion imaging, tubular flow contributes to the diffusion signal in addition to the aforementioned phenomena [306, 307]. However, it is currently impossible to separate blood flow from tubular excretion in IVIM DWI of the kidney. IVIM parameters improve the ability of DWI to discriminate between ccRCC, pRCC and chRCC compared to ADC alone [308]. While ADC and D derived from IVIM show a similar ability to distinguish between ccRCC and AML without visible fat, the discrimination of non-ccRCC and fat poor AML improves significantly with the application of IVIM [309]. IVIM-type DWI shows promise for the differentiation of multiple benign and malignant renal masses [310].

$$\frac{S}{S_0} = f_p e^{-b(D_0+D^*)} + (1-f_p)e^{-bD_0} \tag{3.6}$$

Blood oxygenation dependent MRI

The uncontrolled proliferation of tumours causes them to outgrow their oxygen supply, leading to tissue hypoxia frequently observed in cancer [311]. Polarographic measurement of tissue oxygen pressure in the kidney reveals a pO_2 of 30 mmHg in the normal organ and 13.3 mmHg in RCC [312]. Blood oxygen level-dependent (BOLD) MRI exploits the paramagnetic properties of deoxyhaemoglobin to probe the tissue oxygen pressure non-invasively. The presence of paramagnetic metalloproteins accelerates the dephasing of the MR signal of surrounding water molecules. BOLD images are acquired using a gradient echo sequence, which does not contain the 180°-pulse found in spin-echo sequences. Therefore, the MR signal decays with the susceptibility dependent $T_2{}^*$ rather than T_2. Multiple repetitions with varying echo times allow the absolute quantification of the $T_2^* = \frac{1}{R_2^*}$ relaxation time [313]. A function of the echo time (ET) and the relaxation constant (R_2^*) describes the signal

(eq. (3.7)). BOLD imaging was first applied in the human kidney in 1996 [314] and has since been shown to differentiate benign from malignant and low-grade from high-grade lesions [315, 316, 84]. Furthermore, BOLD imaging may aid subtype differentiation, as R_2^* is significantly higher in chromophobe compared to clear cell and papillary RCC [317].

$$S(\mathrm{TE}) = S_0 e^{-R2^*\mathrm{TE}} \tag{3.7}$$

Dynamic contrast-enhanced MRI

Dynamic contrast-enhanced MRI (DCE-MRI) describes the rapid acquisition of multiple series of T_1-weighted MR images following the administration of an intra-vascular low molecular weight contrast agent. The technique assesses tissue perfusion while the contrast agent is mainly within the blood vessels and vascular permeability when it starts to pass into the extravascular space. The model-free analysis of the signal intensity time curve is mathematically simpler and more reproducible than the analytical quantification of parametric values [318]. The onset time, area under the contrast curve, time-to-peak, absolute and relative enhancement describe properties of the signal curve [319]. On the other hand, parametric models may infer information on physiological processes such as blood flow, mean transit time, extraction fraction and the volume of the extracellular space from the DCE measurements [320].

A large number of free parameters requires a high SNR and a high temporal resolution of the DCE data of the order of 1.5 seconds [321]. This limits the spatial resolution and field-of-view (FOV), leading to a preference for a less demanding compartmental model in many studies. The most frequently used model by Tofts *et al.* describes the contrast concentration-time curve as a function of three free parameters (Equation 3.8) [322, 323]. K^{trans} designates the diffusion constant from the plasma into the interstitial space, v_{e} the interstitial space, and v_{p} the fractional plasma volume. However, through the incorporation of plasma flow and tissue permeability into the transfer constant K^{trans}, these cannot be resolved separately. Therefore, the correct physiological interpretation of K^{trans} depends on the relationship of capillary permeability and blood flow. In the presence of highly permeable blood vessels, the blood flow limits the flux of gadolinium into the extravascular space, whereas in a low permeability setting, K^{trans} is determined by vessel permeability [324]. Many more models with differing assumptions, numbers of free parameters and requirements regarding input data quality exist [319]. The Tofts model requires an estimation of the arrival of the contrast agent in the tissue in the form of an arterial input function (AIF) based on measurements

in a feeding vessel. Alternatives are an AIF from a reference tissue or a population-based estimate. DCE-MRI using the (extended) Tofts model discriminates ccRCC from non-ccRCC with moderate sensitivity and specificity [325]. Furthermore, K^{trans} is significantly associated with PFS in sorafenib treated RCC [326]. Alternatively, non-contrast-enhanced perfusion imaging is possible with arterial spin labelling, using blood labelled with RF pulses as an endogenous contrast medium [327].

$$\text{Tofts Model:} \quad C_t(t) = K^{\text{trans}} \int_0^t C_p e^{\frac{-K^{\text{trans}}}{v_e}(t-t')} dt' \tag{3.8}$$

$$\text{Extended Tofts Model:} \quad C_t(t) = v_p C_p(t) + K^{\text{trans}} \int_0^t C_p e^{\frac{-K^{\text{trans}}}{v_e}(t-t')} dt' \tag{3.9}$$

$$\text{DCE rate constant:} \quad k_{ep} = \frac{K^{\text{trans}}}{v_e} \tag{3.10}$$

where C_p designates the plasma concentration of the contrast agent and C_t the tissue concentration.

3.2 Assessment of early response to neoadjuvant tyrosine kinase inhibitor treatment in RCC using multiparametric physiological MRI

Parts of this work have been published:
Ursprung S*, Priest AN* *et al.* Multiparametric MRI for assessment of early response to neoadjuvant sunitinib in renal cell carcinoma. PLoS One. 2021;16(10):e0258988.
*Shared first authorship

Acknowledgements:
Professor Tim Eisen and Dr Sarah Welsh have conceived the study, developed the trial protocol and recruited the participants. Dr Andrew Priest, Prof Ferdia Gallagher and Dr Tristan Barrett have defined and developed the MRI protocol. Dr Priest has processed the imaging data. Dr Barrett and Dr Fulvio Zaccagna have outlined the tumours. I thank Professor Eisen for allowing me to access the imaging data. Stephan Ursprung has performed the image analysis, conducted the statistical analysis, visualised the results and written the manuscript.

3.2.1 Introduction

Despite a rapid expansion of the therapeutic spectrum for advanced and metastatic RCC, TKIs play a significant role in many therapeutic regimens. Recent research established the combination treatments consisting of a TKI (axitinib or cabozantinib) and an IO agent (nivolumab or pembrolizumab) as the standard of care in patients with advanced RCC of any risk profile [4, 95]. Sunitinib is used in all lines of therapy: as a first-line treatment in patients with contraindications for IO treatment and as a second-line treatment after TKI/IO combinations. As more treatment options become available, patient stratification and treatment response monitoring become increasingly important.

Routinely, response assessments are carried out on contrast-enhanced CT every eight to sixteen weeks and their evaluation is restricted to subjective or semi-quantitative measurements of volumetric tumour response similar to the RECIST 1.1 criteria [114]. However, the poor sensitivity of these criteria to both early response and progression is widely documented, and the clinical evidence and biological validity of the threshold for disease progression remain elusive [4]. While a reduction in tumour volume and the resulting survival benefit are

the ultimate goals in the treatment of mRCC, long-term stabilisation of the tumour volume is simultaneously beneficial and challenging to capture [328, 329].

Image-guided treatment planning in the era of personalised cancer care requires early markers of treatment response to direct patients towards an effective treatment before volumetric progression occurs. TKI treatment targeting VEGF inhibits tumour angiogenesis and reduces vascular maintenance, which may cause physiological changes before the tumour volume changes [330]. MRI is potentially sensitive to these early changes in tumour physiology. In particular, the perfusion fraction from IVIM-DWI and K^{trans} from DCE-MRI may measure early changes in vascular permeability and architecture [304]. Furthermore, BOLD, which is sensitive to changes in tissue oxygenation, may detect an increase in tissue hypoxia similar to the one induced in melanoma xenografts treated with sunitinib [331, 314].

Here we investigate the early effects of sunitinib treatment on RCC primaries using morphological and physiological MRI. Additionally, we correlate findings from multiparametric MRI with the long term outcome. The NeoSun trial of neoadjuvant sunitinib in patients with advanced renal cancer requiring surgery collected the data analysed here. Welsh *et al.* have reported the primary oncological results and the safety profile of sunitinib in the pre-surgical window elsewhere [332].

3.2.2 Methods

Study design and ethical approval. NeoSun was a prospective, single-arm, single-centre, open-label phase II clinical trial of neoadjuvant sunitinib in patients with advanced ccRCC scheduled to undergo nephrectomy. The national research ethics committee in the United Kingdom approved this prospective study (REC ref: 09/HO304/69), and all participants provided written informed consent.

Inclusion and exclusion criteria. Participants had to have histologically confirmed ccRCC with evidence of metastasis, age >18 years and ECOG performance status 0 to 1. Previous treatment for mRCC was not permitted, and participants had to be suitable for MRI. Participants entered the imaging sub-study of NeoSun if they completed both MRI assessments. Section C.1 on page 269 lists detailed inclusion and exclusion criteria.

Treatment. Participants received 50 mg sunitinib o.d. for 12 days before undergoing radical cytoreductive nephrectomy. One patient also underwent adrenalectomy due to an adrenal metastasis, which resulted in the resection of all measurable lesions according to RECIST 1.1. Following surgery, patients recommenced sunitinib treatment on a four weeks

on and two weeks off six-weekly cycle until disease progression. Patients were followed up with 12-weekly ceCT of the chest, abdomen and pelvis, which was assessed according to RECIST 1.1 to identify disease progression. In the case of unacceptable toxicity determined with the Common Terminology Criteria for Adverse Events, Version 4.0, treatment was discontinued [333].

Survival. Progression-free and overall survival were measured between the date of registration and progression according to RECIST 1.1 or death from any cause, respectively. Participants were censored on the date when they were last seen in the clinic if they were lost to follow-up. Surviving patients were censored on 01/03/2020. Follow-up ranged from seven years and three months to nine years and nine months.

Image acquisition. Participants were scanned on a 1.5 T Discovery MR450 system (GE Healthcare, Waukesha WI, USA) before and after 12 days of neoadjuvant sunitinib. Additionally, participants underwent a base-line contrast-enhanced CT scan of the chest, thorax and abdomen at baseline, which was used for treatment response assessment in the metastatic deposits in the adjuvant treatment phase.

The multiparametric MRI used an 8-channel cardiac receive array coil. Detailed acquisition parameters can be found in table C.1 on page 272. The IVIM-type DWI was acquired in a sagittal orientation and b-values of 0, 150, 500, 700 and 900 s/mm^2 were used. The dual-spin-echo echo-planar imaging sequence used the following parameters (table C.1): echo time (TE) 73.4 ms; repetition time (TR) 4000 ms; field of view (FoV) 35×35 cm^2; slice thickness/gap 4/1 mm; acquisition matrix 104×104; number of excitations (Nex) 6; receiver bandwidth $\pm$50 kHz; ASSET factor 2; scan time 5 minutes 16 seconds. Trace-weighted images were acquired by averaging 3 orthogonal diffusion directions for the non-zero b-values. Saturation bands placaed around the FoV reduced wrapping artefacts.

T_2-weighted morphological images in three orthogonal planes were acquired using a respiratory triggered fast-recovery Fast Spin-Echo pulse sequence, with these parameters: TE 48-69 ms; TR 1 breath; echo train length 10-13; FoV 35×35 cm^2; slice thickness/gap=4/1 mm; acquisition matrix 320×224; 2 Nex with no phase-wrap to remove aliasing.

T_1-weighted morphological images were acquired using a 2D fast spoiled gradient-echo (FSPGR) sequence with these parameters: TE 4.8 ms; TR 139 ms; flip angle 70°; FoV 35×35 cm^2; slice thickness/gap 4/1 mm; acquisition matrix 256×256; 0.75 Nex (partial Fourier); parallel imaging (ASSET) factor 2.

R_2^* mapping of the renal tumour was performed in sagittal orientation using a multi-echo gradient-echo sequence and ten echo times between 4.76 and 47.6 ms with 4.76 ms echo spacing. During each breath-hold, all echo times for a single slice were acquired. Other parameters were: TR 100 ms; flip angle 25°; FoV 40×40 cm^2; slice thickness/gap 4/1 mm; acquisition matrix 128×128; receiver bandwidth ±31.25 kHz; ASSET factor 2. Two slices were acquired in each 7-second breath-hold, with multiple breath-holds used to cover the entire tumour volume.

DCE-MRI data were acquired using a fast spoiled gradient-echo (FSPGR) sequence in coronal oblique orientation with the following parameters: TE 1.6 ms; TR 3.9 ms; flip angle 18°; FoV 35×35 cm^2; slice thickness 5 mm; acquisition matrix 160×160×20–30; receiver bandwidth ±41.67 kHz; 0.5 Nex (elliptical k-space coverage); ASSET factor 2; temporal resolution 4.3–6.4 s; 94–140 dynamic phases. In each case, the total scan time was approximately 10 minutes. The number of slices (and consequently the temporal resolution) was varied to image the entire tumour volume. Gd-DOTA 0.1 mmol/kg (Dotarem, Guerbet) was administered during the dynamic series.

The DCE-MRI acquisition was preceded by the acquisition of T_1 mapping data using a multiple flip-angle FSPGR technique (flip angles 1°, 3°, 5°, 10°, 15°, 20°) with 1 Nex; no parallel imaging; scan time 12.5 s (breath-hold) for each flip angle, other parameters the same as for the DCE-MRI series.

Image Analysis. IVIM DWI data was processed using in-house software developed in MATLAB (The Mathworks) and run on the scanner hardware. The diffusivity D was derived from the mono-exponential fitting of the signal from the non-zero b-values. Subsequently, the difference between the measurements at b = 0 s/mm^2 and the extrapolated signal from the higher b-values served to calculate the perfusion fraction f_p. A change in the diffusivity or perfusion fraction of 20% or more exceeded the test-re-test variation and indicated a treatment effect in the individual case. This corresponds to literature values for the repeatability of ADC measurements in hepatic metastases [262, 263].

Custom programmed MATLAB code generated R_2^* and T_2^* based on mono-exponential fitting of the multi-echo GRE sequence using the nonlinear Levenberg–Marquardt algorithm. A log-linear approximation computed the initial values of the fits.

MIStar (Apollo Medical Imaging Technology) processed the T_1 mapping and DCE-MRI data. Each dataset was pre-processed using a 3×3 median filter. The co-registration within and between the datasets removed spatial misregistrations between the T_1 mapping and DCE datasets caused by motion as far as possible. Together they were used to calculate

dynamic maps of Gadolinium concentration which were fitted using the Tofts model with a population-averaged Arterial Input Function [334, 335] to calculate maps of the transfer constant K^{trans} and quantify the contrast-concentration-versus-time curve for 90 seconds after the arrival of the contrast bolus (iAUC_{90}). Based on existing literature on the repeatability of K^{trans} measurements, a threshold for detecting treatment response above the variation which can occur by chance was set at 30% [334, 336].

One of two consultant radiologists with eight and five years of experience segmented the tumour, its necrotic parts and the normal-appearing kidney on the K^{trans}, R_2^* and diffusivity maps as well as on the coronal T_2w FRFSE sequence in the ImageSetViewer Software, version 1.7 (University Health Network Toronto). Radiologists were blinded to the clinical outcome. In-house MATLAB code derived the 3D volume of the normal kidney, whole tumour, viable tumour and tumour necrosis from the individual 2D segmentations. The segmentation of the D and K^{trans} maps transferred directly to the co-registered parametric maps and median values for the viable tumour ROI were recorded for analysis.

Statistical analysis. A biomedical statistician developed the statistical analysis plan and analysed the data in the SAS suite. The figures for this thesis were produced in R (the R foundation for statistical computing, Vienna, Austria) using the ggplot2 (3.3.3) and hrbthemes (0.8.0) packages for visualisation. The Mann-Whitney U and Wilcoxon signed-rank tests quantified differences in unpaired and paired samples. Patient survival was analysed visually with Kaplan-Meier curves (survival [3.2-7], survminer [0.4.8]) and quantitatively with log-rank tests. Analyses were exploratory, and no correction for multiple testing was applied. P values > 0.05 were accepted as statistically significant.

3.2.3 Results

Of the 22 patients who underwent screening between April 2011 and January 2014, six had to be excluded because they either did not meet the inclusion criteria or declined to participate. After enrollment, four patients had to be excluded for reasons detailed in fig. 3.1. Overall, imaging data from thirteen participants were available, one of whom was unable to complete both imaging time points. The reduced dataset in this patient was of poor quality and excluded from the analysis. Eventually, twelve complete datasets were available for the study of imaging outcomes. Welsh *et al.* have reported the primary endpoints of drug safety, oncological efficacy, and surgical complications separately [337].

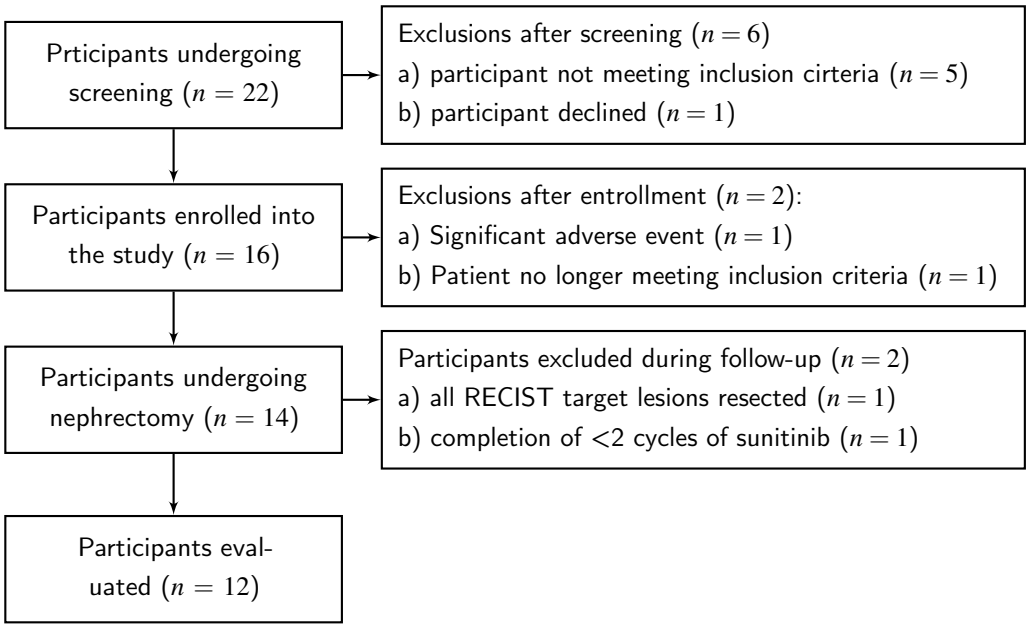


Fig. 3.1 NeoSun CONSORT Diagram

Volumetric findings. The mean total tumour volume reduced from 633 ml (S.D. ± 386 ml, range 175 – 1369 ml) to 507 ml (± 351 ml, 90 – 1213 ml, -25%, $P < 0.001$) after 12 days of sunitinib. A greater, non-significant reduction was measured for the solid tumour volume (-28%) from 388 ml (± 249 ml, 104 – 903 ml) to 295 ml (± 217 ml, 81 – 749 ml). The baseline solid tumour component amounted to 68% (15 – 95%) of the total tumour volume. The necrotic and cystic tumour component, which accounted for 32% of the total tumour volume at baseline, reduced by only 17% from 245 ml (± 263 ml, 10 – 885 ml) to 212 ml (± 251 ml, 8 – 830 ml, $P = 0.005$). The reduction in the cystic/necrotic component was significantly inferior to the reduction in the solid component ($P = 0.042$). The reduction in the overall tumour volume was negatively correlated with the relative volume of tumour necrosis at baseline: ρ = -0.69, 95%-CI: -0.10 – -0.97, $P = 0.02$. Neither the reduction in whole nor solid tumour volume was associated with long term survival (PFS or OS) or the RECIST response during the post-surgical follow-up (figs. 3.2a, 3.2b, 3.3a and 3.3b). Participants with partial or complete response according to RECIST (n = 9/12) tended to experience a longer PFS than patients with stable or progressive disease (Median PFS: 77 vs. 41 weeks, $P = 0.064$, fig. 3.2e).

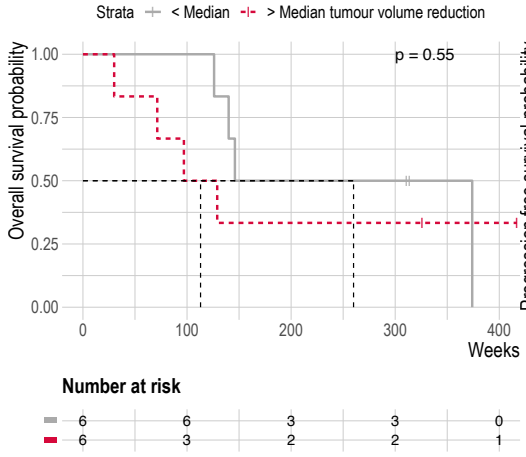
Physiological imaging. A significant reduction in the median solid tumour diffusivity ($P = 0.02$) was observed after neoadjuvant sunitinib treatment (1298×10^{-6} mm^2/s before

Table 3.1 NeoSun: Patient Characteristics

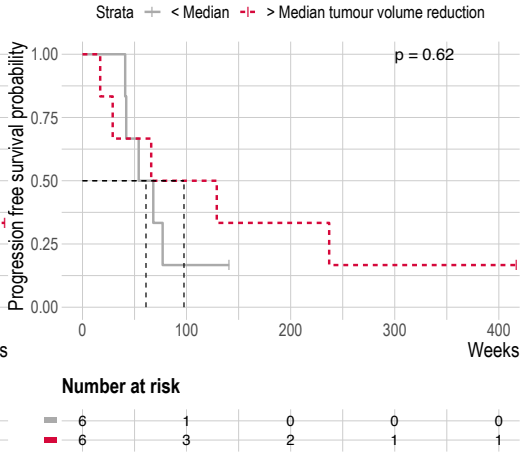
Participant feature	
Number of patients	12
Age at consent	mean: 61 yr
	SD: 7.6 yr
Sex (m / f)	11 / 1
Stage at diagnosis:	
pT1a	0
pT1b	1
pT2a	0
pT2b	0
pT3a	10
pT3b	1
pT4	0
Metastatic deposits:	
Lung	10
Lymph Nodes	5
Adrenal	2
Pleura	2
Bone	1
Pancreas	1
Overall survival	median: 138 wk
	IQR: 119 - 317 wk
Progression-free survival	median: 67 wk
	IQR: 42 - 132 wk

IQR: interquartile range, SD: standard deviation, wk: week(s), yr: year(s)

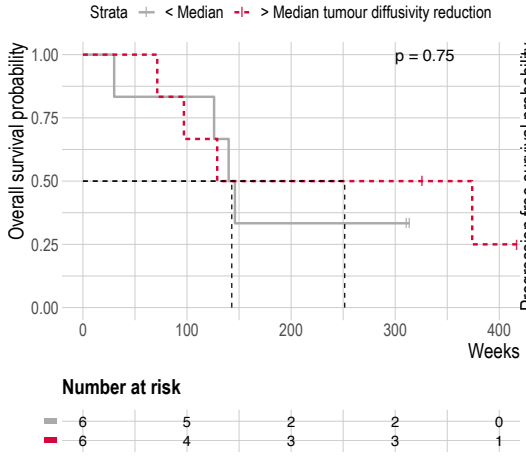
(a) Overall survival vs. Volume



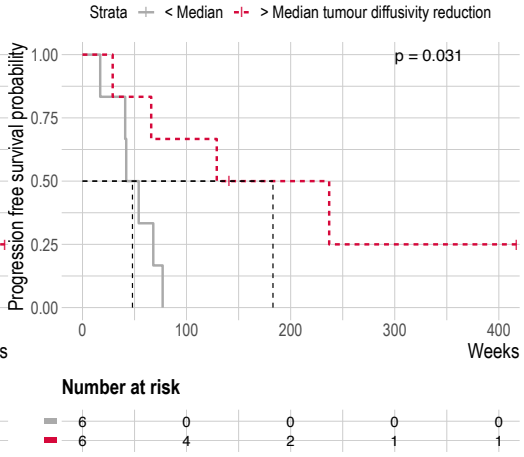
(b) Progression-free survival vs. Volume



(c) Overall survival vs. D_0



(d) Progression-free survival vs. D_0



(e) Progression-free survival vs. Objective response

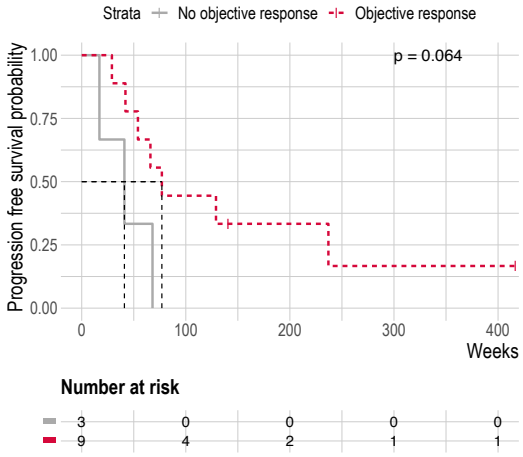


Fig. 3.2 While the reduction in the primary tumour volume following 12 days of neoadjuvant sunitinib treatment was not associated with progression-free or overall survival (a and b), an early reduction in D_0 above the median was predictive of longer progression-free survival (c and d). An objective response according to RECIST was borderline significantly associated with progression-free survival.

vs. 1200×10^{-6} mm^2/s after treatment, -7.7%). However, a reduction greater than the pre-specified cut-off was observed in only 2/12 participants. Participants with a reduction in median D_0 greater than the median of all participants experienced a longer PFS ($P = 0.031$, 183 weeks vs. 48 weeks, figs. 3.2c and 3.2d) but similar OS. The level of diffusion restriction at baseline was not predictive of OS or PFS. Furthermore, participants with an objective response according to RECIST showed a greater reduction in D compared to participants with stable or progressive disease (-11% vs. -3%, fig. 3.3d).

Simultaneously, the median $R_2{}^*$ of the solid tumour rose from 19 s^{-1} to 28 s^{-1} ($P = 0.001$) which was, however, unrelated to participants' survival or RECIST response.

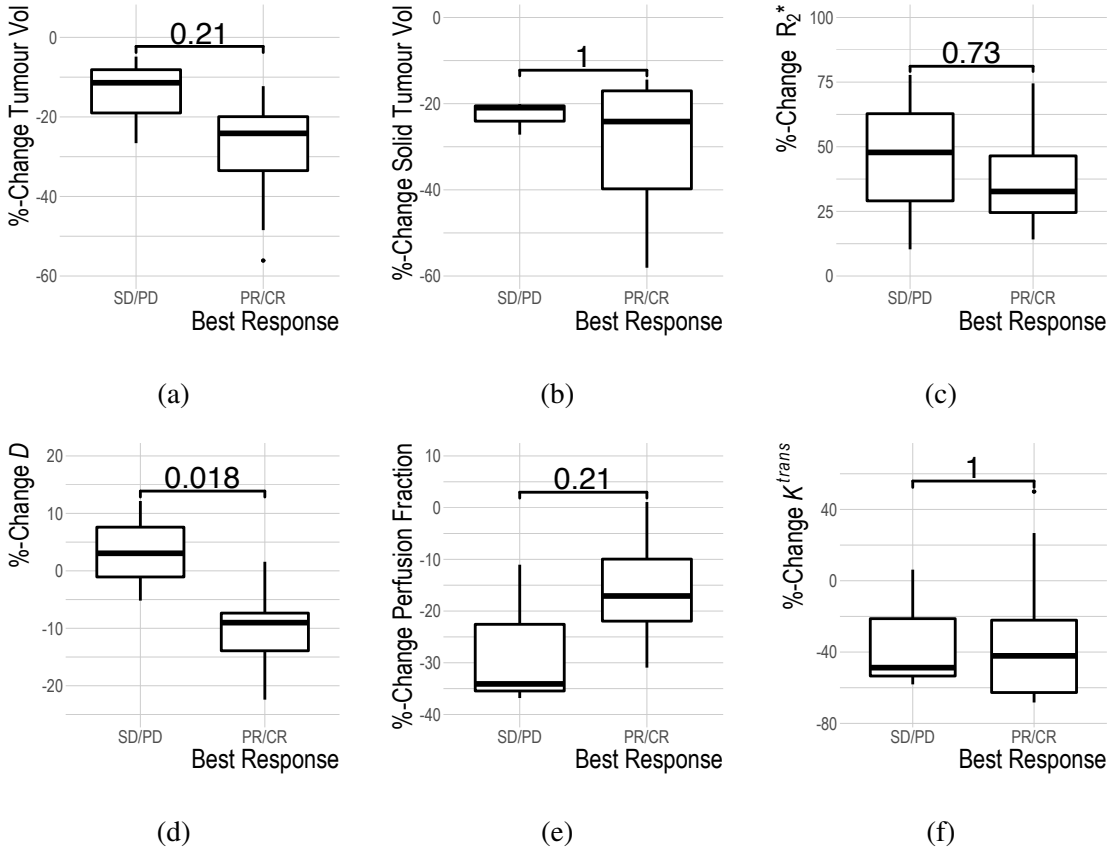


Fig. 3.3 Changes in the primary tumour of participants in NeoSun relative to the best response according to RECIST in the post-surgical follow-up of the metastatic deposits.

A reduction in median K^{trans} and iAUC$_{90}$ was apparent in the solid tumor. The median K^{trans} fell by 28% from 0.415 min^{-1} to 0.305 min^{-1} ($P = 0.01$)and the iAUC$_{90}$ decreased by 38% from 0.55 to 0.36 mM min ($P = 0.002$). While 9/12 participants achieved a reduction in K^{trans} exceeding the pre-determined cut-off of 30% signifying treatment response. However,

they could not derive a survival benefit from this. Figure 3.4 represents an example of imaging changes resulting from 12 days of neoadjuvant sunitinib treatment.

3.2.4 Discussion

Here we demonstrate that MRI is sensitive to changes in tumour physiology 12 days after starting sunitinib treatment in ccRCC. Furthermore, early changes in tumour diffusivity correlated with the best achieved volumetric response according to the RECIST criteria and progression-free survival.

Participants with the largest proportion of solid tumour benefited most from the neoadjuvant treatment, providing a potential rationale for the selection of patients for neoadjuvant chemoreduction: participants with primarily viable tumours where the tumour size increases the surgical complexity and risk may benefit most from a short course of pre-surgical treatment. In the adjuvant setting in patients with advanced ccRCC not eligible for nephrectomy, a reduction in the solid tumour volume may provide a more sensitive measure of early treatment response than the whole tumour volume or a RECIST diameter. The reduction in diffusivity was uniquely associated with PFS in this cohort, even though substantial changes exceeding 20% were rare. Further research is warranted to establish the generalisability of alternative tumour volume measurements and physiological parameters as markers of early treatment response.

Despite the small patient cohort, the changes in the physiological imaging parameters were plausible, given the known mechanism of action of the treatment. Sunitinib reduces the activation of the vascular endothelial growth factor receptors 1-3 (VEGFR1-3) among other receptor tyrosine kinases. This inhibition results in reduced angiogenesis and vascular maintenance, which may decrease the tumour capillary volume measured by the perfusion fraction on IVIM-DWI [338]. Furthermore, it may increase R_2^*, representing decreased tissue oxygenation and vascularisation. The simultaneous decrease in K^{trans} and iAUC_{90} are consistent with these drug effects and decreased vascular permeability. The reduction in tumour microvascularity was confirmed histologically [332]. The MRI techniques to measure these physiological parameters are widely available and could be valuable non-invasive markers of response to TKI. Less established imaging parameters like the time-to-peak from CEUS have been suggested as early surrogates of RECIST response after 12 weeks, PFS and OS in a study on 37 patients [339]. However, these initial results did not generalise in a multi-centre study of 157 patients with mRCC [340, 341]. In contrast, NeoSun was the first drug trial with a controlled treatment regimen in patients receiving sunitinib as a first-line

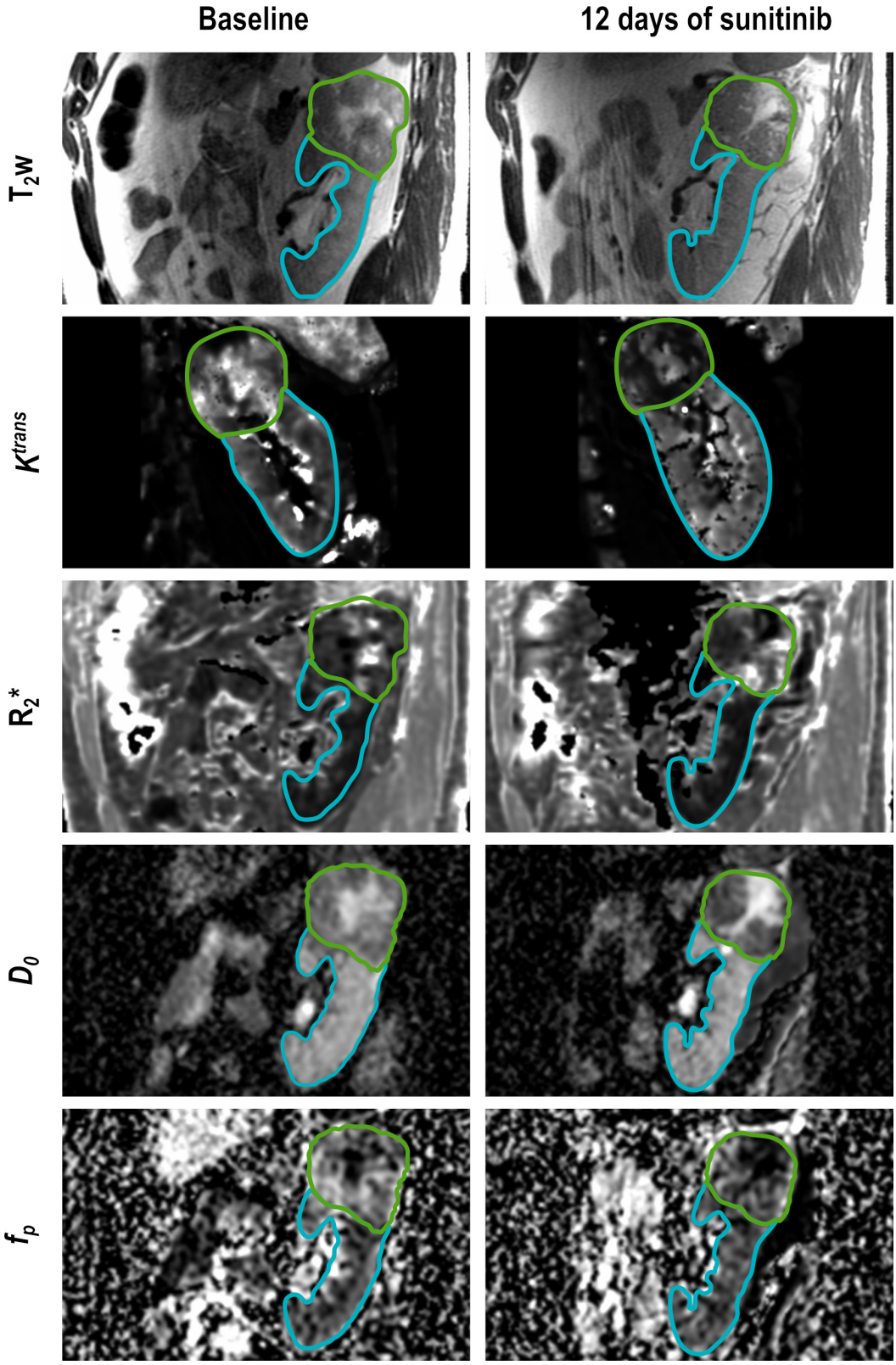


Fig. 3.4 Example of physiological MRI maps of a patient in NeoSun. A reduction in K^{trans}, D_0 and f_p is apparent following treatment while R_2^* was increased in this patient. See text for an explanation of the imaging parameters.

treatment. Unlike CEUS, MRI is operator independent, examines a larger volume including the primary tumour and abdominal metastatic deposits, and can accommodate varying patient habitus. Furthermore, techniques to compensate for respiratory motion are well established in MRI, which remains challenging in CEUS. The superior anatomical coverage of MRI is a decisive advantage in a disease where treatment response is spatially variable [342].

In the NeoSun trial, the diffusion restriction increased following 12 days of treatment which correlated with a longer PFS and better volumetric response following nephrectomy. Histologically, increasing necrosis accompanied the increase in diffusion restriction [332]. ccRCC diffusivity shows a complex response pattern at the beginning of sunitinib treatment: the ADC increases at three days, normalises after ten days and remains unchanged at 3 – 4.5 months [343, 245]. However, NeoSun differentiates itself from these earlier studies by measuring the tumour diffusivity in viable tumour components alone, limiting the risk of diluting the signal from the most aggressive part of the tumour. Furthermore, IVIM-DWI, as used in the NeoSun trial, affords information on tumour diffusivity (D) and the vascular contribution to the motion of water molecules in the tissue (f_p). In contrast, the ADC is a composite term that includes both tissue diffusion and vascular perfusion. Future studies could include multiple MRIs between the time points used in this study to determine potentially sequential changes in the physiology of ccRCC treated with TKI.

The NeoSun trial has several limitations. The number of participants was relatively small and fell short of the recruitment target. Changes in clinical guidelines resulted in considerably narrower criteria for cytoreductive nephrectomies in patients with metastatic ccRCC, which limited the potential for recruitment. Recent evidence suggested that only patients with a low metastatic burden may benefit from surgical resection of the primary [4]. The prospective data collection and standardised treatment according to a research protocol are a strength of this trial. Consistent timing and better-controlled treatment and imaging conditions were specified as part of this trial. Furthermore, existing studies of multiparametric MRI in RCC have investigated similar numbers of participants [315, 344]. While analyses in this study remain exploratory, confirmatory evidence from tissue analysis support biologically plausible findings on physiological imaging. Further research is required to determine the generalisability of these results and to assess the clinical impact of early treatment response detection using MRI.

In summary, NeoSun provided evidence for the possibility to detect early physiological changes in ccRCC undergoing TKI treatment on multiparametric MRI. The antiangiogenic effects of sunitinib caused significant changes in DCE, BOLD and IVIM-type DWI parameters after 12 days of treatment, which makes them candidates for imaging biomarker develop-

ment. The imaging response to TKI treatment was biologically plausible and associated with the volumetric tumour response and progression-free survival. Further research into physiological MRI biomarkers of early treatment response is warranted.

3.3 Response of venous tumour thrombi to tyrosine kinase inhibition

3.3.1 Introduction

Extension of RCC into the venous vasculature is frequent, 30–60% may present with renal venous (RV) involvement, and in 3–10% the thrombus extends into the IVC [345]. Multiple classifications of the extent of thrombotic involvement have been proposed. Most orient themselves at anatomical landmarks like the liver and the diaphragm (table 3.2). Neves and Zincke proposed the classification system which will be used in this thesis [346]. The level of the tumour thrombus in the IVC is independent of lymphatic involvement, metastatic dissemination and survival in patients undergoing nephrectomy [347]. 10-year cancer-specific overall survival is poorer in patients with IVC involvement (29%) compared to renal vein involvement only (66%) [347, 348]. Furthermore, a venous tumour thrombus (VTT) is a significant adverse prognostic factor in patients who are not eligible for surgery (17%) [349, 350]. Vessel wall invasion of the RV complicates surgical resection and is a risk factor for poor survival outcomes [351].

Table 3.2 NAXIVA and A-PREDICT: Classification systems for IVC thrombi

Classification	RV	<2cm RVO	Inf.hep.	Int.hep below HV	Int.hep. above HV	Sup.hep. inf.dia	Sup.dia	Atrial
Novick [352]	I	II					III	IV
Ciancio [353]	I	II		IIIa	IIIb	IIIc	IIId	IV
Moinzadeh [347]		I					II	III
Mandhani [354]		I			II			III
Neves [346]		I	II			III	IV	
Leibovich [355]	0		I*	II		III	IV	
WHO/UICC	T3a	T3b					T3c	

* below the caudate lobe, Dia: Diaphragmatic, Hep: Hepatic, HV: Hepatic vein, Inf: infra, Int: intra, RV(O): Renal vein (ostium), Sup: supra, UICC: Union for International Cancer Control, WHO: World Health Organisation.

The presence of a tumour thrombus increases the risk of surgical complications; 18.8% experience perioperative complications; arrhythmia, bleeding and pneumonia being the most common ones [356]. The risk of these complications increases with the level of the VTT thrombus [357]. Additionally, the procedure is associated with increased perioperative mortality, which is highly dependent on the surgeon's experience and increases with the VTT level (1.5-10% in high volume centres) [357–360]. In comparison, 11.9% of patients with T1b and T2 RCC develop perioperative complications after radical nephrectomy [72]. In a prospective study, 13.9% of patients undergoing radical nephrectomy experienced complications [361]. Estimates of the perioperative mortality among all patients undergoing nephrectomy reach 1.59% in a recent meta-analysis and 0.5% in a registry study in the United States [362, 363]. Furthermore, the development of chronic renal insufficiency is a major long term risk [357]. Distention of the IVC $\geqslant$24 mm at the level of the renal vein ostium is a moderately accurate marker of wall invasion [364]. However, pre-operative MRI may support the prediction of vessel wall invasion with good sensitivity and specificity of 92.3% and 86.4%, respectively [365]. This could facilitate surgical planning as vascular wall involvement might necessitate resection and reconstruction with a patch or tube graft to achieve negative surgical margins [366].

As increasing VTT levels lead to more complex surgical approaches, neo-adjuvant treatment could potentially decrease surgical morbidity. Initial, retrospective analyses of patients treated with a range of targeted therapies show a reduction in thrombus levels in only 4–7% of the patients [367, 368]. In contrast, 42.1% of patients treated with a median of three cycles of sunitinib experienced a reduction in their VTT level [369]. However, only level IV and V evidence of the effectiveness of neoadjuvant treatment with the intent of reducing surgical complexity in patients with VTT is available (Evidence levels: table A.1, page 235).

Axitinib is a highly potent inhibitor of VEGFR 1-3 [370] which, in combination with the IO agent Pembrolizumab, is indicated as the first-line treatment in advanced ccRCC of any MSKCC risk group [371]. Its half-maximal inhibitory concentration (IC_{50}) is five- to 500-fold lower compared to sunitinib, sorafenib and pazopanib [372]. In isolation, axitinib elicits an objective treatment response according to RECIST 1.0 in 13/23 responding patients within two months [373]. The tumour size reduction with axitinib exceeds the reduction with either sunitinib or sorafenib [374]. Additionally, a greater proportion of patients achieve an objective response with axitinib (32%) than sorafenib (15%) [375]. In a phase II trial of patients with localised ccRCC, the median size of the primary tumour reduced from 10.0 to 6.9 cm (median reduction: 28.3%) within 12 weeks. The majority of the reduction occurred by week seven [376].

The value of axitinib as a neoadjuvant treatment to reduce VTT extent, facilitate nephrectomy and improve surgical outcomes in patients with RCC is unclear. This chapter presents the imaging results from the NAXIVA trial of neoadjuvant axitinib in patients with VTT undergoing surgery and the A-PREDICT trial of axitinib in patients with VTT who were not initially eligible for nephrectomy.

3.3.2 NAXIVA - Phase II Neoadjuvant study of axitinib for reducing the extent of venous tumour thrombus in clear cell renal cell cancer with venous invasion

Parts of this work have been presented at ASC-GU 2021:
Stewart GD, Welsh SJ, Ursprung S, Gallagher FA, *et. al.* NAXIVA: A phase II neoadjuvant study of axitinib for reducing extent of venous tumor thrombus in clear cell renal cell cancer (RCC) with venous invasion. Journal of Clinical Oncology 39, 6 (2021) 275-275.

Acknowledgements:
Professor Grant Stewart and Professor Robert Jones conceived the study and developed the trial protocol. Dr Sarah Welsh led the translational aspects of the trial and Prof Ferdia Gallagher, Dr Iosif Mendichovszky and Dr Andrew Priest defined and developed the MRI protocol. The NAXIVA trial group recruited the patients and acquired the trial data. I thank Professor Stewart for allowing me to access the imaging data for the joint central review with Professor Gallagher.
Stephan Ursprung has performed the image analysis, conducted the statistical analysis, visualised and interpreted the results.

Methods

Study design and ethical approval. The response of tumour thrombi to neoadjuvant axitinib treatment was assessed in the NAXIVA cohort. NAXIVA was a multi-centre, single-arm, single-agent, open-label, phase II clinical trial of axitinib in patients with clear cell RCC and venous tumour involvement at diagnosis. The study obtained ethical approval from the Cambridge South Research Ethics Committee, REC-number: 17/EE/0240 and participants provided written informed consent. (ClinicalTrials.gov Identifier: NCT03494816)

Participants. NAXIVA included patients suitable for surgical resection of the primary who presented with biopsy-proven ccRCC extending into the main renal vein or IVC. Patients were eligible for the trial regardless of lymph node or distant metastasis if surgical treatment was appropriate (T3, N0/1, M0/1). Participants had to be ECOG performance status <2 and have adequate renal function measured by urinalysis ($<$2+ protein or urinary protein is $<$2g per 24 hours or protein-creatinine-ratio $<$200mg/mmol) and serum creatinine ($\leqslant$1.5xULN or estimated creatinine clearance $\geqslant$30mL/min as calculated using the Cockcroft-Gault equation). All female patients with reproductive potential had to have a negative pregnancy test no longer than 14 days before starting trial treatment. Patients with distant metastasis were excluded if they had MSKCC poor-risk disease. Furthermore, patients with previous invasive malignancies in the past two years, who were breastfeeding, unable to take or absorb the medication or suffered cardiovascular complications in the past year were excluded from the trial. For detailed inclusion and exclusion criteria, please refer to section C.3 on page 273.

Trial treatment. Patients received eight weeks of trial treatment starting with 5 mg of axitinib b.d. for 14 days. In the absence of CTCAE grade 3 or higher adverse events, the dose was escalated to 7mg b.d. and 10mg b.d. after two weeks each. If patients developed treatment-refractory hypertension or CTCAE grade 3 adverse events, the treatment was interrupted, and the dosing could be reduced to 3 mg b.d. and 2 mg b.d. if necessary. Patients completed the treatment between 36 hours and seven days before nephrectomy in week 9.

Imaging and image evaluation. Patients underwent ceCT and a baseline multiparametric MRI up to four weeks before enrollment in the study. The tumour burden on the CT was assessed according to RECIST 1.1 and the tumour thrombus level on the MRI according to the Mayo/Neves' classification (table 3.3). A second MRI acquired after 14-21 days of treatment served to re-assess the tumour thrombus length. Additionally, patients with M0 disease at recruitment received a CT of the thorax. A final radiological assessment involving an MRI scan and a CT of the thorax in patients with M0 disease and a CT of the thorax, abdomen and pelvis in patients with M1 disease occurred in the 9^{th} week after starting treatment. Whenever possible, the extent of the VTT was assessed on the MRI. Subsequently, patients underwent nephrectomy, and a standard of care restaging CT was carried out 12 weeks after surgery.

MRI scans were acquired on 1.5 T scanners from different manufacturers depending on on-site availability. The protocol included an axial T_1w dual-echo in-phase/opposed-phase GRE sequence, axial and coronal balanced SSFP sequences, axial and coronal or sagittal T_2w

sequences, axial T_2w TSE with fat suppression and spin-echo EPI DWI at b-values of 0 and 600 s/mm^2. ADC maps were reconstructed using mono-exponential fitting as implemented by the scanner manufacturer. Following the injection of 0.1 ml of 1M Gadobutrol (Gadovist, Bayer, Reading, UK) per kilogram of patient body weight, eight phases of a coronal multi-phase Dixon sequence were acquired with a temporal resolution of approximately 15 s. Section C.4 on page 276 lists the imaging parameters of one of the participating sites.

RECIST measurements and VTT level assessments were carried out at the recruiting site and confirmed centrally. An imaging researcher with four years of experience and a consultant radiologist completed the central VTT review in consensus. The tumour thrombus length was defined as the sum of the length of the thrombus in the renal vein, the thrombus in the IVC below the renal vein ostium and the thrombus above the renal vein ostium. The percentage change in VTT length was derived according to eq. (3.11).

$$\%-Change = 1 - \frac{sum_t}{sum_{BL}} \cdot 100 \tag{3.11}$$

where sum_t is the summed length of the tumour thrombus at time point t, and sum_{BL} is the summed length of the tumour thrombus at baseline.

The distance between the renal vein ostium and the end of the VTT was a complementary measure in patients with an RV thrombus only. An increase in this distance was a sign of response. This measurement is potentially more reproducible than measuring the length of the thrombus itself. All image analysis was performed in OsiriX 10.0 (Pixemo SARL, Switzerland).

The Mayo/Neves' classification described the extent of the VTT (table 3.3 and fig. 3.5) [346]. An extension to the classification further differentiated RV thrombi into surgically meaningful categories. Left-sided RV thrombi were subdivided based on the location of the tip of the VTT relative to the insertion of the gonadal vein. RV thrombi on both sides were further classified depending on the involvement of the main renal vein or branches only.

Imaging endpoints. The primary endpoint was the change in Mayo tumour thrombus level between baseline and the follow-up MRI at nine weeks. Secondary imaging endpoints consisted of the change in tumour thrombus length.

Statistics. The trial required at least 20 patients, employed a Simon two-stage minimax design and was powered to distinguish an improvement in the Mayo classification in <5%

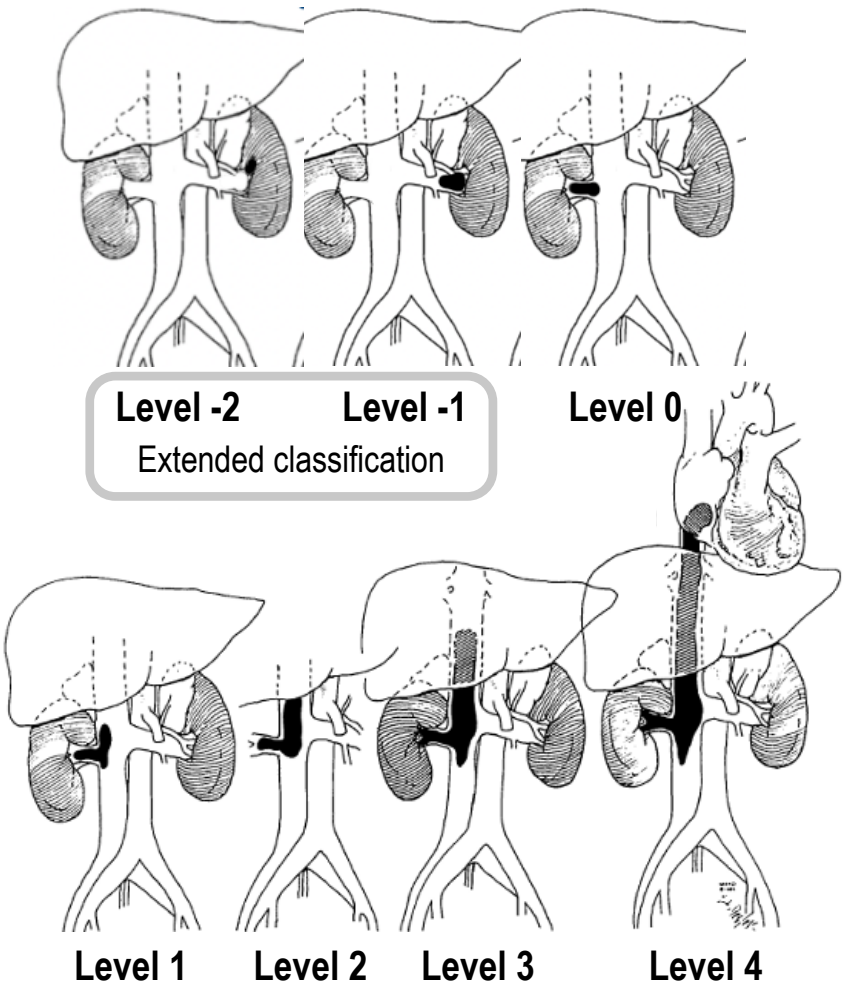


Fig. 3.5 Definition of tumour thrombus level according to the extended Mayo/Neves' classification. Levels 0 to 4 were defined in the original classification. The levels are described in table 3.3. Figure adapted from: Neves RJ and Zincke H. BJU. 1987; 59: 390-395.

Table 3.3 Extended Mayo/Neves' Classification

Level	Description
-2	Left: Thrombus not extending beyond branches of the renal vein
-1	Left: Thrombus not extending beyond insertion of gonadal vein into renal vein Right: Thrombus not extending beyond branches of the renal vein
0	Thrombus extending beyond gonadal vein but not into the IVC
1	Thrombus extending up to 2 cm into the IVC
2	Thrombus extending more than 2 cm into the IVC but not beyond the hepatic veins
3	Thrombus extending beyond hepatic veins but not beyond diaphragm
4	Thrombus extending beyond diaphragm

patients from a >25% improvement with 90% power at a 10% one-sided significance level. The initial stage included 13 patients, and the trial would have been stopped for futility if no patient had achieved a reduction in the Mayo classification at this point. Furthermore, if at any point during the study three or more patients experienced disease progression while under treatment, the trial would have been stopped because the likelihood of seeing at least three progressors if fewer than 5% of patients progress is <7.5%. Simultaneously, at least three patients will have to show a reduction in the VTT Mayo level for the trial to reach its primary endpoint.

All analyses were conducted in the treated patient cohort, and patients who had not received a single dose of trial treatment were excluded. The statistical analyses in this thesis have been performed in R (version 4.0.0, R foundation for statistical computing, Vienna, Austria) using the ggplot2 (3.3.3), hrbthemes (0.8.0) and ggchicklet (0.5.0) packages.

Results

Twenty-one participants (15 male, 6 female) were recruited into the NAXIVA trial between December 2017 and January 2020 across five sites in the UK. The median age at recruitment was 61 years, and 10/21 patients presented with metastatic disease. All participants were evaluable, of whom 17 progressed to surgery. Ten participants (48%) showed Mayo level 2 VTT, reaching more than two centimetres into the IVC but not above the hepatic veins. One patient had an extended Mayo level -2, one -1, and three 0 VTT; all would classify as Mayo 0 according to the original classification. Two patients presented with levels 1, 3 and 4 VTT each.

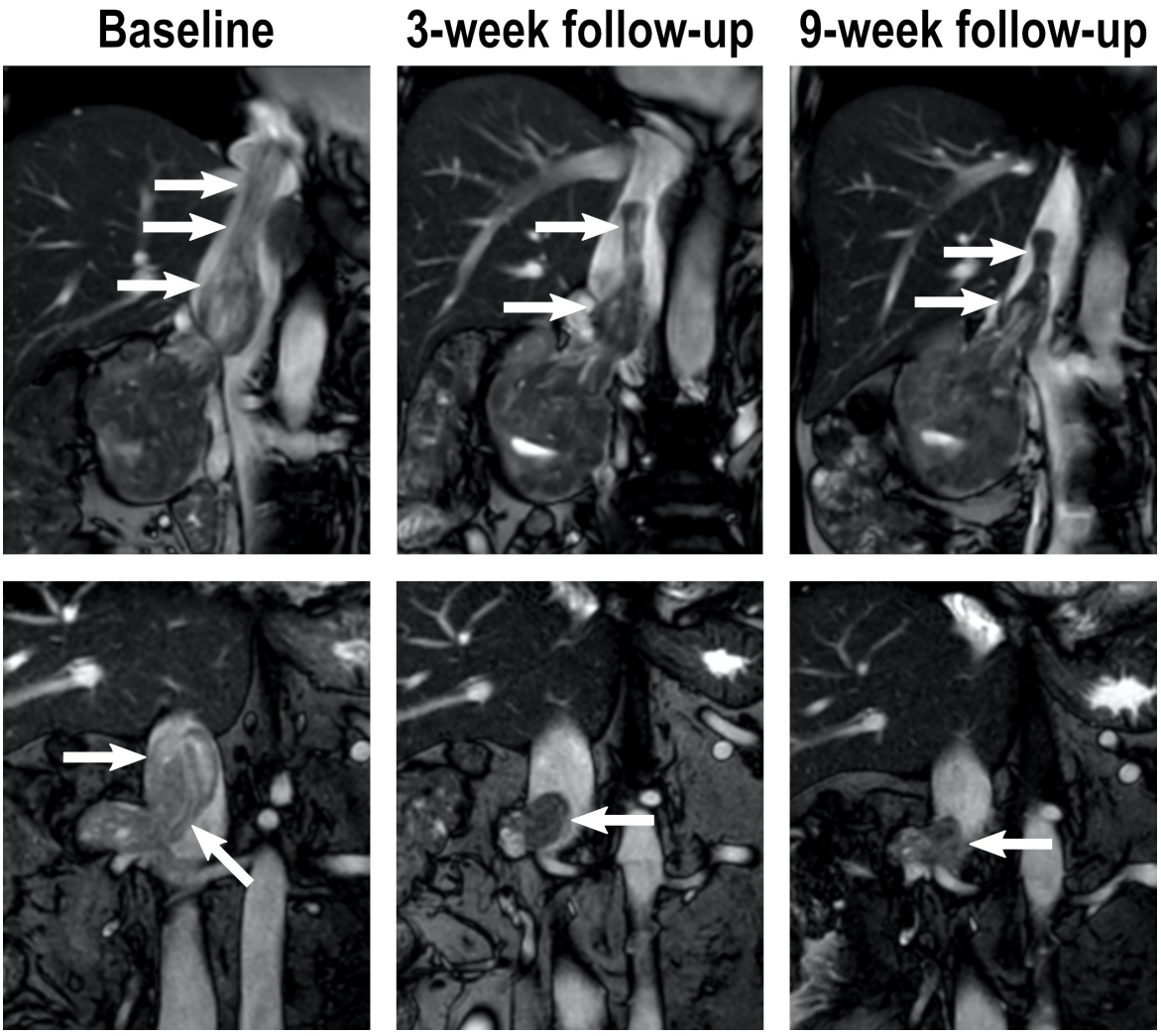


Fig. 3.6 Example of two participants in the NAXIVA trial responding to neoadjuvant axitinib. The tumour thrombus in the patient at the top reduced from a Mayo level 4 to level 2 (9.3 cm in IVC to 5.2 cm). The second thrombus reduced from a Mayo level 2 to level 1 (4.8 cm in IVC to 1.7 cm). On these T_2w images, the loss in thrombus girth and T_2w signal intensity are apparent.

Response of tumour thrombus length to axitinib treatment. Six patients (27%) experienced a reduction in tumour thrombus length by at least one Mayo level, while seven patients (33%) achieved a response according to the extended Mayo classification. A patient with a VTT extending to the right atrium showed the most striking response with a reduction in length to below the hepatic veins (level 4 to level 2). A second patient with a level 4 VTT achieved a reduction to the level of the hepatic veins (level 2/3). The VTT level reduced from 2 cm and more above the RV ostium but below the hepatic veins to less than 2 cm above the RV ostium in three patients (level 2 to level 1), while one tumour thrombus of similar height shrunk to an RV thrombus (level 2 to level 0). The thrombus with a response only according to the extended Mayo classification retracted within the renal vein past the insertion of the gonadal vein (level 0 to level -1). In 6/7 participants, the reduction in the Mayo level was measurable at the scan in weeks 3. One patient experienced a progression in tumour thrombus level from the extended Mayo level -1 to 0 between weeks 3 and 9. All responding participants at week three maintained their response until week nine. Figure 3.6 shows an example of two participants with a reduction in tumour thrombus level after axitinib treatment.

The median reduction in the length of the tumour thrombi was 8 mm after 3 weeks (IQR: -23.3 to +0.5, $P = 0.003$) and 12 mm after 9 weeks (IQR: -29.0 to -2.0, $P = 0.008$, no significant difference between weeks 3 and 9). Among the 16/21 (76%) participants with an extension of the tumour thrombus into the IVC, the median reduction in the length of the portion of the thrombus in the IVC was 17.5 mm (IQR: -1.5 mm to -25.0 mm) at three weeks and 17.5 mm (IQR: -4.5 mm to -30.5 mm) at nine weeks. Meanwhile, the VTT length remained unchanged in participants with thrombi of the renal vein only (median: 0.0 mm, range: -9 to +10 mm at 3 weeks and median: 0.5 mm, range: -12 to +16 mm at 9 weeks (fig. 3.8b).

The median length of the tumour thrombus, including the RV portion, the portion above and below the renal vein ostium, reduced from 102 mm (IQR: 73 to 131 mm) to 80.5 mm (IQR: 63 to 116 mm, $P = 0.008$) after 3 and 73 mm (IQR: 63 to 86 mm, $P = 0.002$) after 9 weeks, respectively. This corresponded to a median decrease in VTT length by 15.7% (IQR: 4.7 to 26.5%) and 20.4% (IQR: 5.3 to 39.0%) after 3 and 9 weeks respectively. By week 9, a non-significant median decrease in VTT length relative to week 3 of 5.1% (IQR: increase by 3.2% to decrease by 16.1%, $P = 0.13$) had occured (fig. 3.8a). The absolute decrease in the tumour thrombus length at 9 weeks was positively correlated with the initial VTT length ($\rho = 0.54$, $P = 0.02$, fig. 3.9. No correlation was observed between the absolute

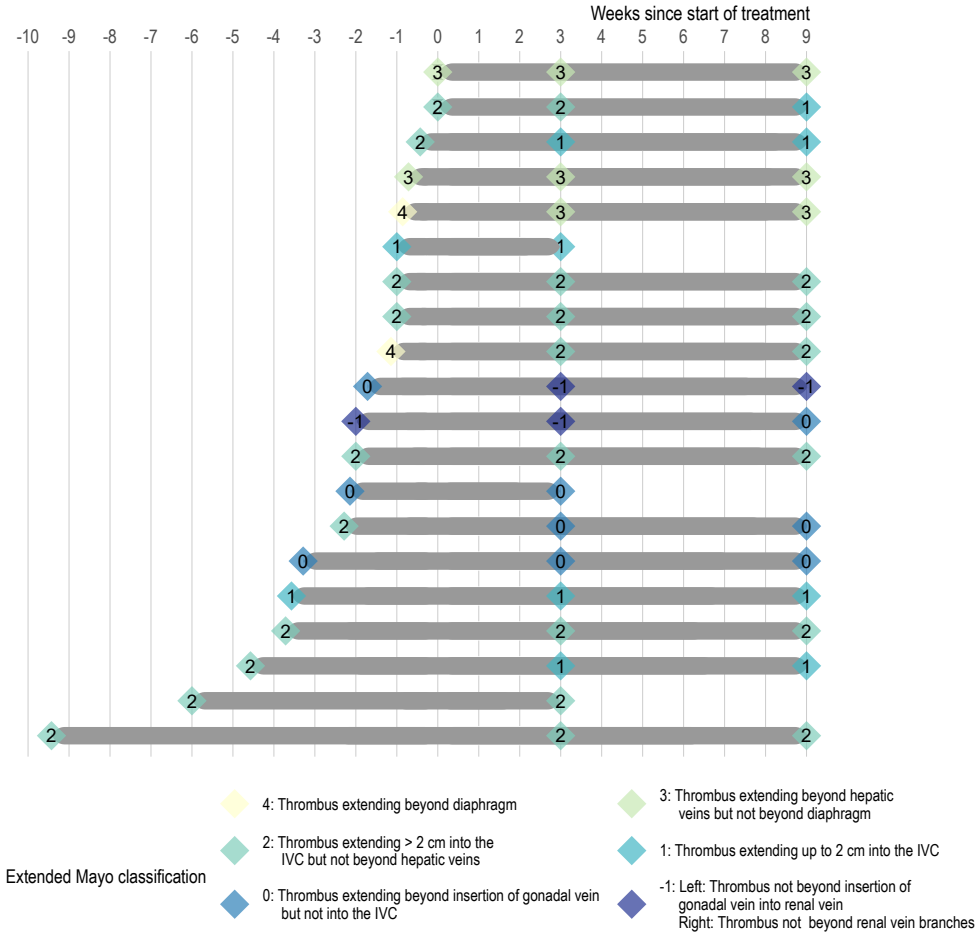
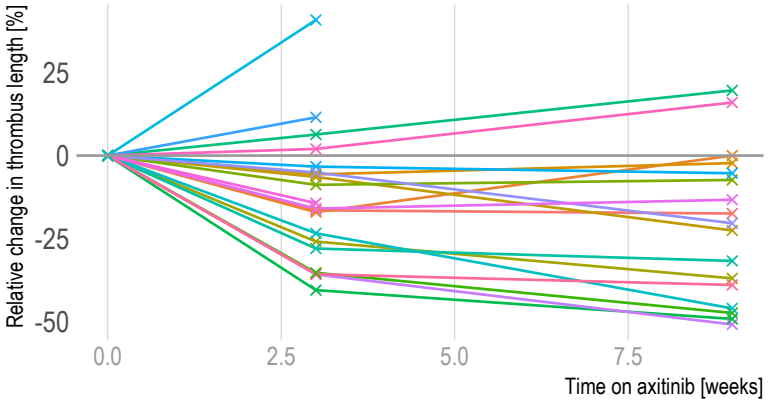


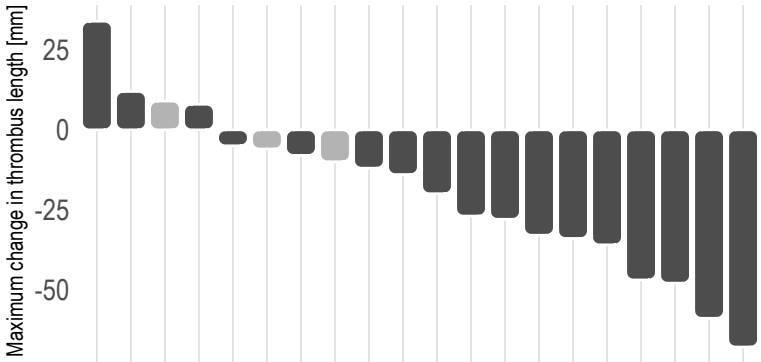
Fig. 3.7 Venous tumour thrombus time lines. Patients commence axitinib treatment at week 0.

($\rho = 0.37$, $P = 0.08$) and relative ($\rho = 0.08$, $P = 0.74$) decrease in VTT length at 3 as well as the relative decrease at 9 weeks ($\rho = 0.17$, $P = 0.51$) and the VTT length.

Unfortunately, the variable acquisition methods among the participating sites and lack of consistency within patients prevented a meaningful comparison of signal intensities and ADC within and between patients.



(a) Relative change in tumour thrombus length.



(b) Maximum absolute change in tumour thrombus length.

Fig. 3.8 Changes in tumour thrombus length during the axitinib treatment. In participants who achieved a reduction in the thrombus length, this was mostly achieved by week 3 (a). Thrombi extending into the inferior vena cava (black) experienced a larger absolute reduction in length compared to renal vein thrombi (grey).

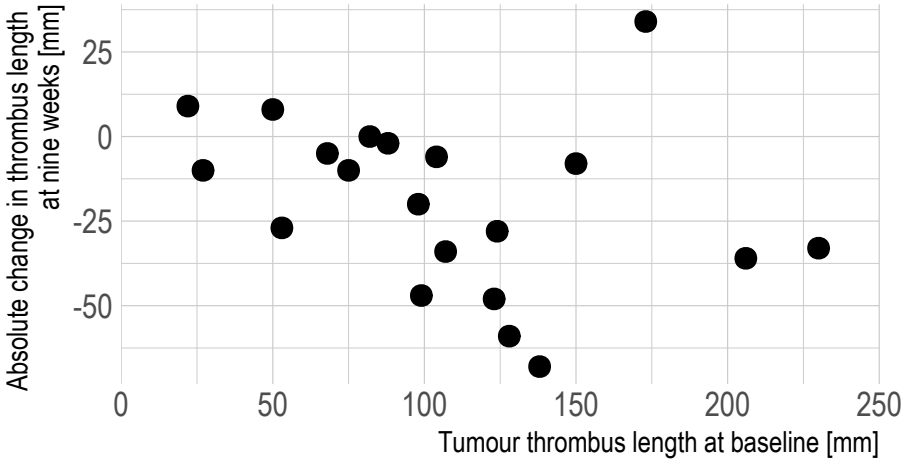


Fig. 3.9 Correlation between tumour thrombus length at baseline and absolute change in thrombus length.

3.3.3 Inferior vena cava tumour thrombus response in a phase II study of axitinib in patients with metastatic renal cell carcinoma unsuitable for nephrectomy (A-PREDICT)

Acknowledgements:
The work presented in this chapter received substantial support from Professor James Larkin who is the principal investigator of the A-PREDICT trial and allowed access to the imaging data to conduct this subgroup analysis. Dr Ben Jenkins developed the statistical analysis plan and performed some of the analysis. Professor Grant Stewart oversaw the sub-study and Dr Ramona Woitek and Dr Lucian Beer reviewed the measurements taken for this study.
Stephan Ursprung has defined the research question of the sub-study in collaboration with Professor Stewart, defined the sub-study methodology, developed parts of the statistical analysis plan, performed the image analysis, conducted some of the statistical analysis, visualised and interpreted the results. In addition, Stephan Ursprung performed the second central review for the A-PREDICT trial.

Methods

Ethics. The A-PREDICT Trial (A Phase II Study of Axitinib in Patients With Metastatic Renal Cell Cancer Unsuitable for Nephrectomy) received approval from the regional ethics committee (08/05/2012 Ref: 12/LO/0639, ClinicalTrials.gov Identifier: NCT01693822) and all patients provided written informed consent. This study reports on a subgroup analysis of participants in A-PREDICT with a VTT; the primary and secondary endpoints of the trial will be reported elsewhere.

Participant selection. A-PREDICT recruited participants with previously untreated, predominantly ccRCC unsuitable for either nephrectomy or a watch and wait approach. A-PREDICT was a single-arm, single-agent, open-label, multicentre phase II trial of axitinib. Inclusion criteria were the presence of measurable metastatic disease according to RECIST 1.1, adequate organ function, ECOG performance status of 0 or 1 and no uncontrolled hypertension. Active intracranial disease, active secondary malignancies, pregnancy, conditions or concurrent medication affecting the absorption or metabolism of the trial treatment or recent cardiovascular complications precluded participation in the trial.

Participants were included in this substudy on treatment response of the VTT if they presented with RV or IVC tumour thrombus or if they developed VTT while undergoing trial treatment. The presence of VTT was assessed centrally on the collected CT scans by an imaging researcher with three years of experience and two consultant radiologists. All available reconstructions were consulted to identify venous involvement. The results of the central assessment were correlated with the initial radiology reports and discrepancies resolved in consensus.

Trial treatment. Participants in this trial received 5 mg axitinib b.d. followed by a dose escalation to a maximum of 10 mg b.d. in the absence of side effects. The trial treatment continued until the patients ceased to derive clinical benefit. The treatment was interrupted for the week leading up to week nine at which point tumours were biopsied. Patients could undergo nephrectomy if the treating physician assessed them as suitable.

Imaging. Contrast-enhanced computed tomography was carried out at baseline and eight weekly during treatment for the first six months. After this, participants continued to receive 3-monthly CT follow-up. Disease response was assessed on CT scans according to RECIST 1.1 by the local radiology team, with central review by a consultant uro-radiologist at the end

of the trial. In case of discrepancies, these were resolved in consensus with a second central reviewer.

CT acquisitions followed the standard operating procedures of the recruiting hospital with a maximum slice thickness of 5 mm. The supplementary materials report detailed acquisition parameters of the participants included in the VTT sub-study (section C.5 on page 277).

VTT assessment. VTT were assessed as described for the NAXIVA trial (section 3.3.2, page 81). Additionally, the length of the VTT was determined. For VTT reaching into the IVC, the length of the thrombus was determined on coronal reconstructions of the thinnest available slices. The thrombus length was measured from the cranial aspect of the ostium of the renal vein to the top of the thrombus. For RV thrombi, the distance from the end of the thrombus to the RV ostium was measured on the thinnest available axial slices and recorded as negative length. If an IVC tumour thrombus reduced in length to become an RV thrombus only or an RV thrombus increased in size and extended into the IVC, the thrombus length was measured as described above depending on the length in the individual CT scan. If a VTT reached the renal vein ostium but not beyond, its length was recorded as 0 cm. An imaging scientist with three years of experience in renal imaging took the measurements, and two consultant radiologists reviewed them.

Statistical analysis. The following study endpoints were defined before access to any data was obtained:

- The primary endpoint of this study was the proportion of patients who achieve a reduction in VTT length according to the Mayo/Neves' classification after nine weeks of treatment. A sensitivity analysis will be conducted and proportions including 95% confidence intervals computed.
- The best difference in VTT length at any follow-up will be recorded as a secondary endpoint for this study and assessed graphically in a waterfall plot as well as with the median and interquartile range.
- The exploratory analysis assessed the response of patients' VTT according to the Mayo/Neves' classification at 16 weeks, 6 months and 1 year. Additionally, it investigated the ability of response to predict progression-free survival using a Kaplan Meier curve.

Statistical analyses were conducted using Stata 15.1 (Stata Corp., Texas, USA).

Results

Out of 65 participants recruited to the A-PREDICT trial, 15 fulfilled the inclusion criteria for this sub-study (fig. 3.10). In 12 patients, locally advanced disease precluded surgical treatment, while the high burden of metastatic disease was a factor in 13 patients. Eight of 15 patients (53%) presented with a VTT restricted to the RV. One patient had no follow-up scans after starting treatment owing to symptomatic progression. The demographics of the participants are summarised in table 3.4. Table 3.5 summarises the characteristics of the tumour thrombus. None of the patients developed VTT *de novo* during the trial treatment. The median tumour thrombus length in patients with IVC thrombus reached 3.1 cm (IQR: 2.9 to 3.7) and -3.0 cm (-4.6 to -2.2) in patients with renal vein only thrombus.

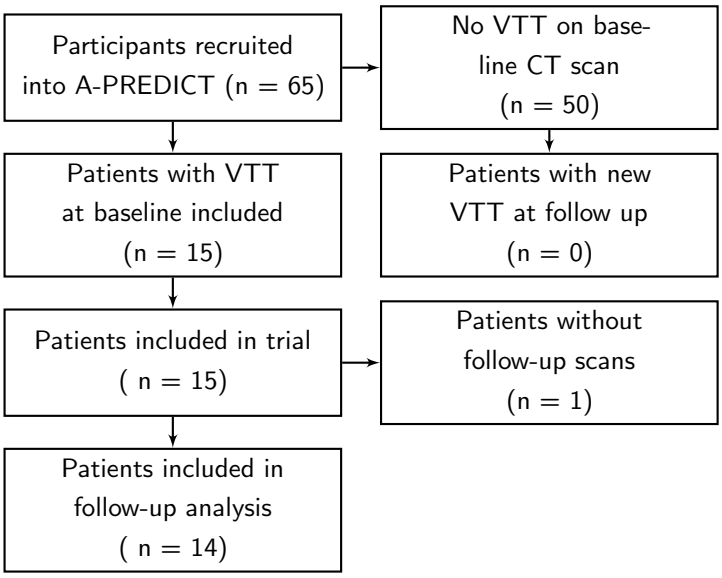


Fig. 3.10 CONSORT Diagram. This only represents the selection of the patients for the VTT sub-study. The consort diagram for the main study will be added above.

Primary endpoint. The median follow-up of the 14 patients who had more than a single scan was 42 weeks (range: 7 - 95 weeks, median 6 imaging follow-up time points (range: 1 - 9)). Four of 15 patients in the intention to treat cohort (26.7%; 95% confidence interval [CI] 7.8% to 55.1%) reached the primary endpoint of a decrease in tumour thrombus size according to the extended Mayo/Neves' Classification after nine weeks of treatment (Week 9 Day 1 scan). In the evaluable population, 4/14 (28.6%; 95% CI 8.4% to 58.1%) achieved a response at nine weeks. Figure 3.11 shows an example of a responding and non-responding patient.

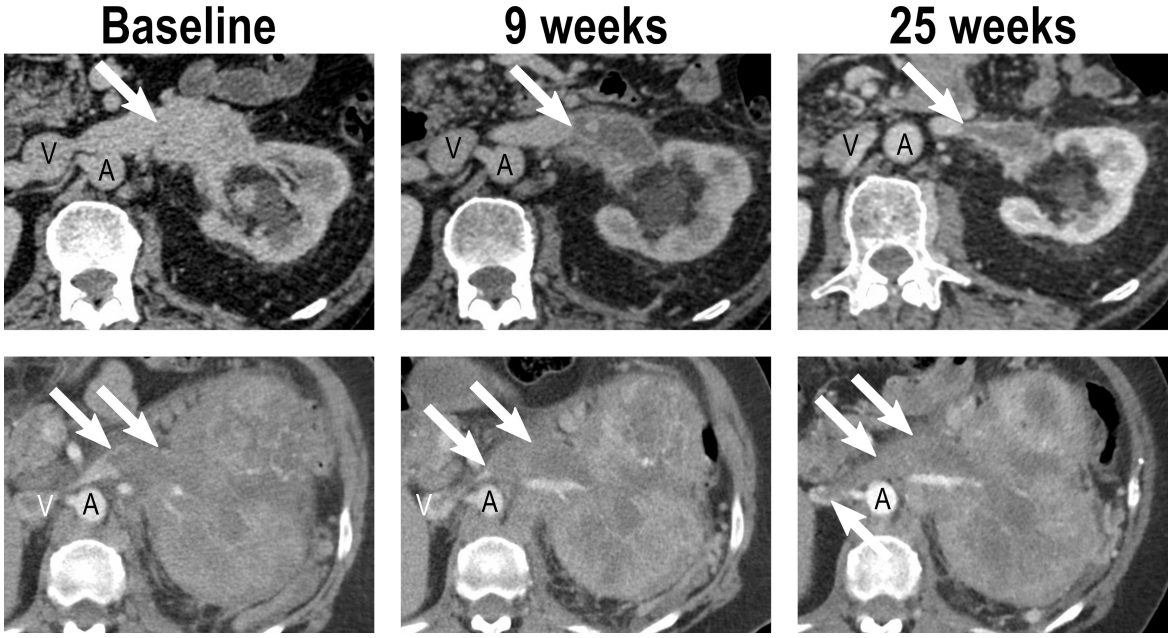


Fig. 3.11 The participant at the top showed a reduction in the extended Mayo classification from 0 to -1 which they maintained for one year. The tumour thrombus reduced in length by 2.9 and 2.4 cm at 9 and 25 weeks on axitinib, respectively. A reduction in girth and contrast enhancement was also apparent. The patient at the bottom progressed under axitinib by 0.7 and 5.1 cm at 9 and 25 weeks, respectively. Please note the contrast inflow artefacts in the inferior vena cava (V) which do not represent tumour thrombus. This was verified in the portal venous phase. A: Aorta.

Table 3.4 Patient characteristics of the VTT A-PREDICT cohort

Patient characteristic	VTT present (N = 15)		VTT absent (N = 50)		Total (N = 65)	
	n	%	n	%	n	%
Age (Median±IQR)	67.2 (53.9 – 73.9)		62.6 (56.7 – 68.6)		63.5 (56.7 – 69.0)	
Sex (female/male)	9/6*	60/40	13/37	24/74	22/43	33.8/66.2
Histology						
clear cell RCC	15	100	47	94.0	62	95.4
clear cell + chromophobe	0	0.0	2	4.0	2	3.1
clear cell + sarcomatoid	0	0.0	1	2.0	1	1.5
Tumour stage						
T1a	1	6.7	1	2.0	2	3.2
T1b	0	0.0	6	12.0	6	9.2
T2	0	0.0	11	22.0	11	16.9
T2a	0	0.0	4	8.0	4	6.2
T2b	0	0.0	4	8.0	4	6.2
T3	1	6.7	5	10.0	6	9.2
T3a	2	13.3	7	14.0	9	13.8
T3b	5	33.3	2	4.0	7	10.8
T4	2	13.3	3	6.0	5	7.7
Tx	4	26.7	7	14.0	11	16.9
Nodal stage						
N0	7	46.7	24	48.0	31	47.7
N1	4	26.7	19	38.0	23	35.4
NX	4	26.7	7	14.0	11	16.9
Metastatic stage						
M1	15	100	50	100	65	100
Fuhrman grade						
1	0	0.0	2	4.0	2	3.1
2	8	53.3	18	36.0	26	40.0
3	4	26.7	17	34.0	21	32.3
4	0	0.0	2	4.0	2	3.1
Unknown	3	20.0	11	22.0	14	21.5
Cocomitant radiotherapy						
no	14	93.3	41	82.0	55	84.6
yes	1**	6.7	9	18.0	10	15.4

*VTT were significanlty more common in female compared to male patients (Fisher's exact texts: $P = 0.027$) and had a significantly higher T-stage ($P = 0.008$) ** Patient received radiotherapy to spine.

Across all time points, a reduction in tumour thrombus length according to the Mayo/Neves' classification during axitinib treatment was observed in seven patients. Overall, the first benefit was measurable at a median of 17 weeks (range 7 to 58 weeks) after the first scan. Four

Table 3.5 Thrombus characteristics of the VTT A-PREDICT cohort

VTT characteristic	Total (N = 15) n	%
Location of VTT		
IVC	7	46.7
Left	1	14.3
Right	6	85.7
RV	8	53.3
Left	7	87.5
Right	1	12.5
Scan available at:		
Baseline	15	100
Week 9	14	93.3
Week 17	12	80
Week 25*	12	80
Week 37	9	60
Week 49	6	40
Week 61	3	20
Week 73	3	20
Week 85	2	13.3
Week 97	1	6.7
Baseline Mayo/Neves' Score		
-1	4	26.7
0	4	26.7
2	6	40
3	1	6.7

* One paitent underwent a nephrectomy, therefore, the tumour thrombus could no longer be followed. ** One scan was not evaluable due to poor contrast enhancement. IVC: inferior vena cava, RV: renal vein and VTT: venous tumour thrombus.

patients showed no change in tumour thrombus (FU: 6, 24, 49 and 58 weeks), and in three patients, the thrombus' Mayo classification increased (first progressive measurement at 26, 38 and 52 weeks, follow-up until RECIST progression: 34, 38 and 84 weeks, respectively).

Secondary endpoint. Tumours for which follow-up was available experienced a 0.7 cm median reduction in length at the nadir (Inter-Quartile Range: -1.2 to -0.2cm; see fig. 3.13a). In participants whose VTT responded to treatment according to the Mayo/Neves' classification, the thrombus length decreased by a median of 1.4 cm, while non-responders experienced

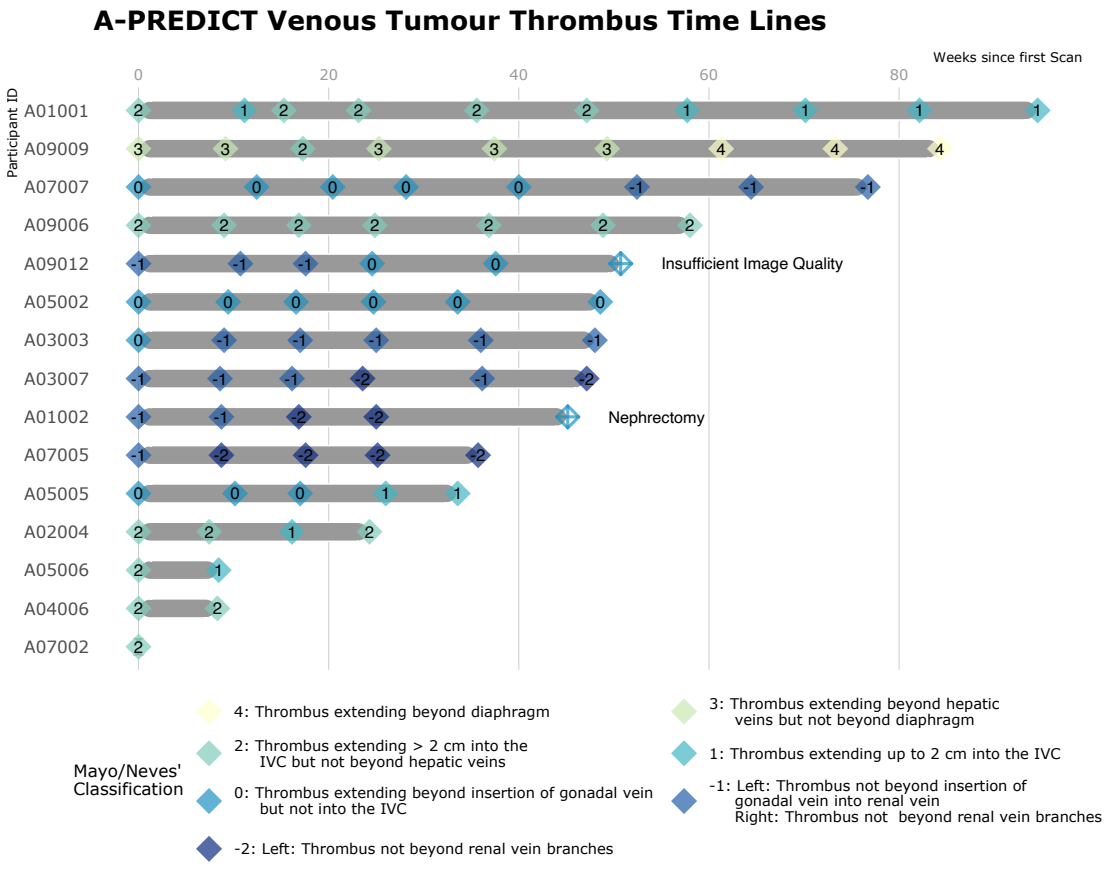


Fig. 3.12 Venous tumour thrombus time lines.

a median decrease by 0.5 cm at the nadir. Participants who achieved a reduction in VTT length exceeding one centimetre did so within six months of starting therapy. In total, 8/15 participants never experienced a decline in VTT length $\geqslant$1 cm.

In six participants, the largest change in thrombus extent was an increase in length (fig. 3.13b). Of those, two patients saw an initial reduction in VTT length $\leqslant$1 cm followed by progression, while one patient experienced an initial increase in tumour length followed by shrinkage.

Figure 3.14 summarises the evolution of the VTT length for individual participants.

Exploratory endpoints. After four months of axitinib treatment, 5/15 participants in the ITT cohort (33.3%; 95% CI 11.8% to 61.6%) presented with a reduced Mayo/Neve's Score, corresponding to 5/12 (41.7%; 95% CI 15.2% to 72.3%) evaluable participants.

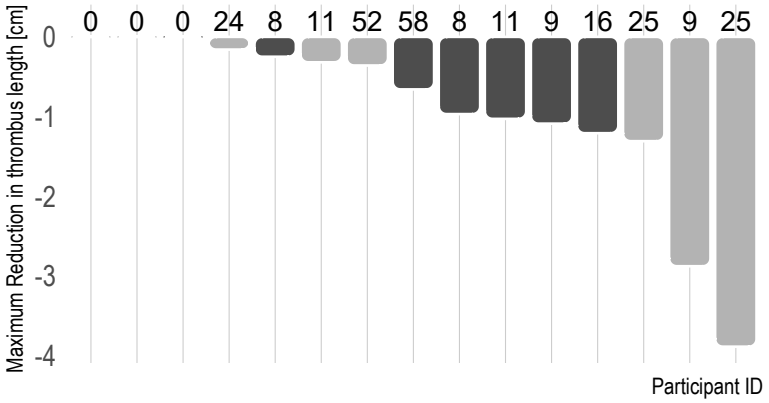
The primary endpoint of the A-PREDICT trial is the proportion of participants who are alive and progression-free after 25 weeks. In the VTT cohort, 4/15 patients (26.7%; 95% CI 7.8% to 55.1%) responded with the tumour thrombus. This corresponded to 4/12 of the evaluable participants (33.3%; 95% CI 9.9% to 65.1%). However, 8/15 patients (53.3%; 95% CI 26.6% to 78.7%) showed a VTT response at any point up to this date.

After one year, 2/15 participants responded with the tumour thrombus (13.3%; 95% CI 1.7% to 40.5%) (evaluable participants: 2/6 (33.3%; 95% CI 4.3% to 77.7%)).

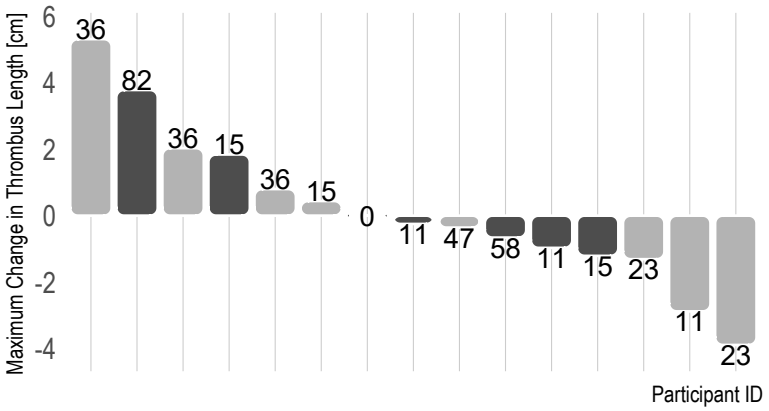
Survival analysis showed a minimal, non-significant progression-free survival advantage for patients with no VTT (0.7 months, log-rank test: $P = 0.16$). No progression-free survival difference was observed between participants with responding and non-responding VTT after 9 and 25 weeks (fig. 3.15).

3.3.4 Discussion

The NAXIVA and A-PREDICT trials demonstrated that axitinib can elicit a significant reduction in tumour thrombus length and Mayo/Neves' level in patients with ccRCC. Within NAXIVA, most responses had occurred by the first follow-up time point three weeks after starting axitinib. Similarly, most patients in A-PREDICT had responded after eight weeks of treatment. This is consistent with previous evidence from primary ccRCC, which showed that most of the reduction in tumour size achievable within 12 weeks of neoadjuvant axitinib occurs in the first seven weeks of treatment [376]. Furthermore, the percentage reduction of the primary tumour diameter is comparable to the loss in VTT length. A rapid length reduction is advantageous for chemoreduction of tumour thrombi. The shrinkage in tumour thrombus length was consistent between weeks three and nine in the NAXIVA cohort.



(a) Maximum absolute reduction in VTT length.



(b) Largest absolute change in VTT length.

Fig. 3.13 Waterfall plots for the maximum absolute reduction in tumour thrombus length (a) and largest absolute change in VTT length (b). Participants in whom the largest absolute change was an increase in length may have achieved a more minor reduction in tumour size at another time point. Participants who achieved their best absolute and relative reductions in the tumour thrombus length in week 0 showed no treatment response in the VTT. Relative changes were not calculated because they are not meaningful with the definition of the tumour thrombus length in the renal vein. Light grey: renal vein thrombus, dark grey: inferior vena cava thrombus.

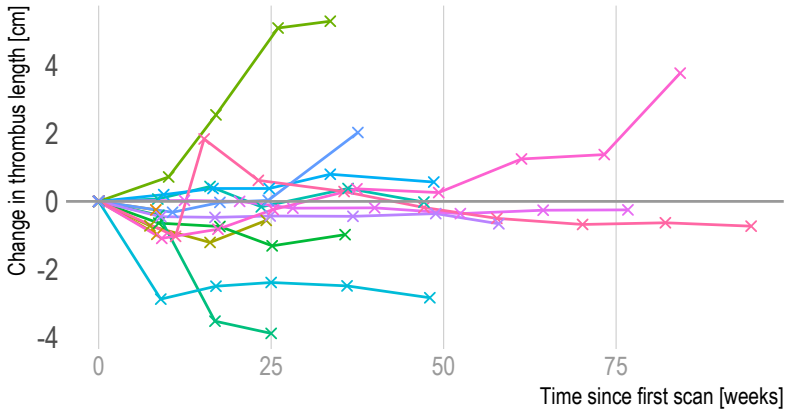
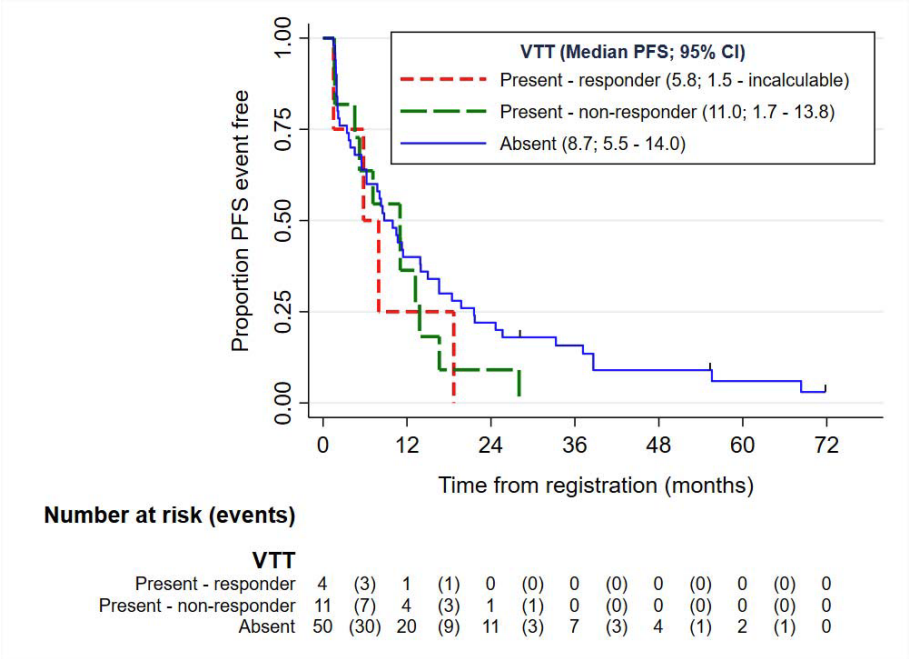


Fig. 3.14 Evolution of VTT length in individual participants.

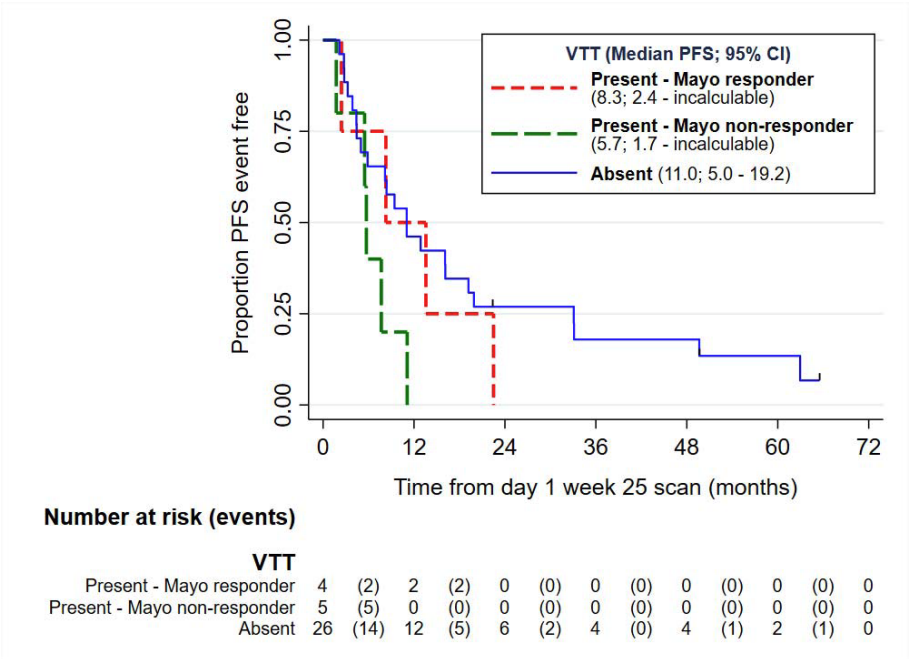
Similarly, most responding participants in the A-PREDICT trial maintained their VTT response over multiple follow-up time points. This facilitates arranging the surgery in the wide temporal window around the time of the optimal response.

Importantly, participants in NAXIVA with IVC thrombi, requiring more extensive surgery, were more likely to respond and show a superior reduction in thrombus length (18 mm after nine weeks) than participants with thrombi limited to the renal vein (1 mm after nine weeks). The A-PREDICT cohort showed an identical effect. However, due to the lower number of participants, it was not statistically significant. In patients with mRCC, the response in the tumour thrombus was not predictive of progression-free survival at a systemic level.

As reported previously, neoadjuvant axitinib in the NAXIVA trial led to a change in the surgical approach in 5/17 (29%) participants with IVC thrombi and a reduction in the surgical extent compared to the initial plan [377]. Two patients became eligible for a minimally invasive approach and the planned incision was reduced in one patient. Overall, 7/17 participants (41%) derived a surgical benefit from neoadjuvant axitinib. No participant required higher risk surgery as a result of the treatment. In the most comprehensive published retrospective case series of 25 patients with VTT undergoing neoadjuvant treatment, 25% of the patients receiving sunitinib experienced a thrombus level change. If patients receiving bevacizumab or temsirolimus are included, the response rate falls to 12% [367]. In a second retrospective study of 19 patients with VTT receiving 12 to 30 weeks of neoadjuvant sunitinib, eight patients (42%) achieved a reduction in the thrombus level [369]. Finally, one retrospective case series of 14 patients found a reduction in thrombus level in 7% [368]. In



(a) PFS by response after 9 weeks of axitinib.



(b) PFS by response after 25 weeks of axitinib.

Fig. 3.15 Kaplan Meier curves of the progression-free survival separated by participants without VTT, responding VTT and non-responding VTT at 9 and 25 weeks. Figure courtesy Dr Ben Jenkins.

summary, axitinib appeared to achieve an equivalent or better response rate compared to sunitinib and did so earlier after the initiation of the treatment.

Four patients in NAXIVA did not undergo surgery, two of them experienced rapid progression in their metastatic disease, and two were deemed unfit for surgery. The rapid metastatic progression within three weeks of starting axitinib allowed these patients to avoid surgery which would have been unlikely to benefit them. Hence, a brief window of neoadjuvant treatment may reveal particularly aggressive tumour phenotypes where higher risk surgery may do more harm than good. Concerns about fitness for surgery were identified but not communicated by the recruiting site before enrollment in two other patients. These were unrelated to the trial treatment. Conversely, one participant in A-PREDICT, whose disease was initially deemed unresectable at recruitment, underwent nephrectomy after 25 weeks of treatment.

This work has several limitations: First, the number of participants is limited as NAXIVA was powered as a phase II trial, and the analysis of the tumour thrombi was an exploratory endpoint of A-PREDICT. Therefore further research and larger phase III trials will be required to confirm the findings and provide appropriate evidence for implementing neoadjuvant treatment of RCC with VTT. Recent signals of new TKI-IO combinations in the metastatic setting (e.g.lenvantanib + pembrolizumab, axitinib + pembrolizumab, and nivolumab + cabozantinib [91, 378, 379]) offer exciting new options for future studies in the neoadjuvant setting. Second, deviations from the imaging protocol occurred, and protocol parameters changed within patients. While length measurements remained unaffected, the impact on the comparability of the DWI and ceMRI data was detrimental. These data would have been most valuable to elucidate the evolution of the ADC in ccRCC treated with TKI between weeks 3 and 9. This time window was not covered by any of the other studies discussed in this thesis. Third, participants in NAXIVA underwent nephrectomy which ranged from curative to cytoreductive. However, none of the patients continued on axitinib post-surgically. As a result, tissue is available for translational analyses, but no information on the long term response in other tumour deposits is available. However, the A-PREDICT trial in patients with more advanced RCC could afford some of this information.

In contrast, NAXIVA was the first trial investigating the ability of neoadjuvant TKI to reduce the extent of VTT in RCC in a prospective and controlled manner. A second independent cohort of patients was available to confirm the findings and demonstrate their transferability from MRI to CT with the A-PREDICT trial. Strikingly, the T_2w sequence was most helpful in determining the extent of the VTT and the ceMRI rarely provided

supplementary insight. Consequently, a simpler and shorter MRI protocol may suffice to assess tumour thrombi.

In conclusion, the NAXIVA trial demonstrated that one in three patients experience a reduction in the VTT Mayo/Neves' level in the first eight weeks of axitinib treatment. The independent A-PREDICT trial confirmed these findings. Furthermore, two in five patients benefited from a less morbid surgical approach than initially anticipated. Therefore, neoadjuvant treatment may have a role to play in patients with RCC and VTT.

Chapter 4

Design of a 3D-printed tumour mould for the co-registration of imaging and tissue samples to achieve multi-regional data integration

This work has been published:
Crispin-Ortuzar M*, Gehrung M*, Ursprung S*, *et. al.*, Stewart GD$^{\Delta}$, Sala E$^{\Delta}$, Markowetz F$^{\Delta}$. Three-Dimensional Printed Molds for Image-Guided Surgical Biopsies: An Open Source Computational Platform. JCO Clinical Cancer Informatics. 2020; 4, 736-748.
*Shared first authorship. $^{\Delta}$Shared last authorship.

Acknowledgements:
The work presented in this chapter received substantial support from Dr Mireia Crispin-Ortuzar and Dr Marcel Gehrung who developped the code for the automated generation of 3D-printed tumour moulds and the calculation of MR-habitat maps jointly. Dr Gehrung was instrumental in setting up the 3D printer and providing training on its use.
Dr Andrew Priest developed the ^{1}H-MRI acquisition protocol and Dr Andrew Gill, Dr Gaspar Delso and Dr Andrew Priest developed code for the reconstruction, motion correction and registration of the multiparametric MRI data.
Stephan Ursprung has conceived the idea for this project in collaboration with Professor Evis Sala, defined the aims of the study and advised on the development of the mould. He obtained ethical approval, recruited participants, performed the image segmentation and analysis, and collected the data for validating the tumour mould design. He conducted the validation jointly with Dr Crispin-Ortuzar.

4.1 Introduction

Multi-omics research strategies employ statistical and machine-learning approaches to derive biological insight from the joint analysis of multiple large scale data streams [380]. The desire to apply several data sources to improving the understanding of cancer biology is rooted in the realisation that multiple factors contribute to the development and progression of malignant tumours [381]. Most commonly, multi-omics approaches interrogate genomic, transcriptomic, proteomic, metabolomic, clinical and imaging data from the whole tumour. However, ccRCC is characterised by a high degree of genetic and metabolic intratumoral heterogeneity [2, 382]. This heterogeneity determines the ability of tumours to progress metastatically and has prognostic implications [150, 148]. Tumours also differ in the speed at which they acquire the capability to metastasise. Cytoreductive nephrectomy, reducing the

genetic diversity, may confer a survival benefit in oligometastatic and slowly progressive RCC [150]. While most genomic analyses, particularly in clinical medicine, derive from a single tumour biopsy [383], four to eight biopsies would be required in ccRCC to detect $\geqslant$75% of the sub-clonal driver mutations [148]. (Section 1.3.1 discusses intratumoral heterogeneity in greater detail.) Multi-regional tumour profiling will become increasingly relevant with the identification of actionable targets for precision oncology.

One challenge in adapting multi-omics analysis pipelines to considering intratumoral heterogeneity is the need for spatially resolved datasets. While imaging intrinsically maps information to the spatial domain, molecular datasets rarely take this into account. In contrast, molecular data streams are easily spatially registered amongst each other if the analyses are carried out on the same tissue sample. Achieving spatial awareness and registration at the same time requires linking imaging data to tissue samples. The registration of radiological imaging and tissue samples is necessary in two fundamentally different contexts: a tissue biopsy where the structure of interest remains in its anatomic context and tissue sampling after surgical resection of the tissue. Beer *et al.* describe the use of image-fusion for the ultrasound-guided biopsy of tumour regions defined on CT [384]. However, renal tumours undergo biopsy infrequently (section 1.2.1). Therefore, a method for spatially registered tissue sampling from the operative specimen is preferable. Shah *et al.* were the first to report using patient-specific 3D-printed tumour moulds derived from MRI for image-guided tissue sampling after prostatectomy [385]. Dwivendi *et al.* first applied the concept to partial tumour nephrectomies [386]. However, the information on intratumoral genetic heterogeneity is arguably more likely to guide future treatment strategies in higher-stage tumours. These undergo radical nephrectomy, exhibit greater genetic diversity and recur more frequently.

The application of a 3D-printed tumour mould to sampling a radical nephrectomy specimen differs from previous applications in prostate, ovarian and renal cancer [385, 387, 386] because the organ is not ordinarily cut in the imaging plane for pathological assessment. Specifically, the tumour is bisected coronally to expose the interface between the tumour and the renal hilum [388, 389]. Furthermore, the kidney, tumour and perirenal adipose tissue are usually resected en bloc. As a result, structures at the surface of the tumour and kidney are not available for orientation, resulting in a less well-defined specimen surface.

This chapter describes the development of a computational framework to produce 3D-printed, patient-specific tumour moulds for multi-parametric MRI-guided tissue sampling from radical nephrectomy specimens. This is the first attempt at designing tumour moulds which re-orient radiological images to adapt to the preferred pathological cutting direction. The code is accessible open-source and automates the mould design, taking the segmentation

as an input and producing a printable 3D model (https://github.com/markowetzlab/cutter). The joint development of the mould with pathologists ensured interoperability with standard clinical pathways and allows pathologists to customise the mould design to the individual case. Finally, we assessed the accuracy of the spatial registration afforded by the final mould design.

4.2 Methods

4.2.1 Study participants and ethics

Study participants were identified at the urological multidisciplinary team meeting at Addenbrooke's hospital and invited to participate during their next clinical appointment. All participants provided written informed consent to the Molecular Imaging and Spectroscopy with Stable Isotopes in Oncology and Neurology (MISSION) study (Approval: Cambridge South Research Ethics Committee, 15/EE/0378). Detailed inclusion and exclusion criteria are provided below (section 5.2.1). The mould design was developed and validated in consecutive participants of the same physiological study. Participants entered the trial between May 2018 and January 2020.

4.2.2 MRI acquisition and data pre-processing

MRI scans for the tumour mould design were acquired on a 3T MRI (Discovery MR750, GE Healthcare, Waukesha WI). The protocol was identical to the MISSION study described below (Section 5.2.3 and table D.1 on page 280 summarise the acquisition parameters). In brief, multiparametric ^{1}H-MRI data containing morphological, IVIM-type DWI, BOLD and DCE information were anonymised and exported from PACS. Affine motion correction, implemented in MATLAB (MathWorks, Natick, MA) employing the Advanced Normalization Tools in the Insight Segmentation and Registration Toolkit (ANTs/ITK), was applied to achieve registration between DWI b-values, T1-mapping flip angles and DCE contrast phases, respectively. Subsequently, MIStar (Apollo Medical Imaging Technology, Melbourne, Australia) generated DCE parameter maps. In-house MATLAB code served to calculate the diffusivity D_0 and perfusion fraction f_p from IVIM-type DWI as well as R_2^* and T_2^* from BOLD-MRI. For details on image processing please see section 5.2.4.

Before the segmentation, the morphological sequences and parameter maps were resampled to match the FoV and resolution of the T_1w LavaFlex sequence and registered rigidly using ITK-SNAP (version 3.6, University of Pennsylvania) [390] and in-house MATLAB

code. The tumour, adjacent normal kidney, Gerota's fascia, upper pole, lower pole, hilum and conspicuous points on the tumour surface were outlined manually on the T_1w LavaFlex sequence using OsiriX (Version 10.0.0, Pixmeo, Switzerland). Registered T_2w HyperCUBE and DCE-MRI sequences informed the segmentation and 3D multi-planar reconstructions allowed identifying the long axis of the kidney and subsequently the upper and lower pole. A genitourinary radiologist with 15 years of experience reviewed all segmentations. ROIs were exported to comma-separated value (.csv) files which encoded the vertices of all ROIs for each slice individually.

4.2.3 3D modelling and tumour mould generation

Summary. The automated tumour mould design commenced with a generic 3D block. The patient-specific shape of the nephrectomy specimen was segmented, re-oriented, smoothed, transformed to a 3D object and subtracted from this mould to create a negative in which the specimen could rest. Then, slits to cut the tumour were subtracted from the block, and a comb to guide the pathologist's knife during the tissue processing was added. The code is freely available on https://github.com/markowetzlab/cutter.

ROI processing. The ROIs defined either tissue volumes (tumour, normal kidney and Gerota's fascia) or landmarks (upper pole, lower pole, hilum and surface markers). The volumes served to define the shape of the tumour mould cavity. Meanwhile, the landmarks facilitated the orientation of the specimen within the mould as desired by the pathologist. Initially, the centroid of each tissue volume was calculated from the mean of all x, y and z coordinates of all voxels contained in the volume mask. A single x, y and z coordinate defined the landmarks. The upper pole, lower pole and hilum are indispensable for the design of the tumour mould. The upper and lower poles define the long axis of the kidney, which can be rotated to achieve transverse or longitudinal cross-sections. After orienting the long axis correctly, the position of the hilum defines the rotation around the long axis, which is required to ensure the bisection of the tumour in the correct plane.

Meanwhile, the surface markers, pointing to easily identifiable structures, help to orient the specimen correctly in the mould but are not essential to the design process. Two centimetre wide openings added to the bottom of the mould where these surface markers are supposed to rest, allow visual and tactile verification of the correct orientation of the specimen. An additional opening was positioned under the renal hilum as its preservation is critical.

MR Image re-slicing. For the rotation of the segmented tumour volume, the images were first resampled to 1 mm isotropic resolution using nearest neighbour interpolation in the Computing Environment for Radiological Research (CERR) in MATLAB [391]. Vectors connecting the centroids define 3D rotations according to eqs. (4.1) to (4.3).

$$v_{\mathrm{L}} = 0.5 \times \left(v_{\mathrm{hilum}} + v_{\mathrm{tumour\ contact}}\right) \tag{4.1}$$

$$v_{\mathrm{LC}} = v_{\mathrm{C}_0} - v_{\mathrm{L}} \tag{4.2}$$

$$v_{\mathrm{ploes}} = v_{\mathrm{upper}} - v_{\mathrm{lower}} \tag{4.3}$$

v_{C_0} designates the vector v from the origin of the coordinate system to the centroid C_0, v_{upper} and v_{lower} describe the vector from the origin to the upper and lower pole. The first rotation aligns the vector through the poles with the y-axis of the mould. The second rotation transforms the segmentation around the y-axis and aligns the x-z projection of v_{LC}, the vector through the centroid of the specimen and the hilum, with the z-axis. The rotated tumour volume and landmark points were output as a 3D integer matrix.

Once the 3D-rotation to fit the nephrectomy specimen into the correct position in the mould has been defined, MR-images had to be re-sliced in the orientation corresponding to the cross-section expected when cutting along the tumour mould. The same 3D rotation was, therefore, applied to the images and re-oriented DICOM images were sliced in the x-z-plane with a spacing of 10 mm, corresponding to the distance between two gaps in the final mould. The position of the images along the y-axis was selected to co-localise with the pathological tissue cross-sections.

Mould generation. The 3D matrix was converted to a mesh in Python with the marching cubes algorithm [392]. A mesh consists of triangular facets forming flat surfaces of the future mould and vertices forming the corner points of a facet. Next, the mesh was cleaned, removing close vertices, duplicate faces and vertices and facets from edges where more than two facets meet. The correct definition of inside and outside of the segmented volume required inversion of the normal vectors. Facet reduction to a maximum of 5000 using quadric edge collapse decimation reduced the complexity of the model and the computational expense in its calculation.

The raw mesh showed considerable surface irregularities, which required surface smoothing: First, the resolution of the MRI was inferior compared to the 3D printer (voxels of $1.8 \times 1.8 \times 4$ mm; slice spacing of 2 mm vs. a printing resolution of 0.2 mm layer height and

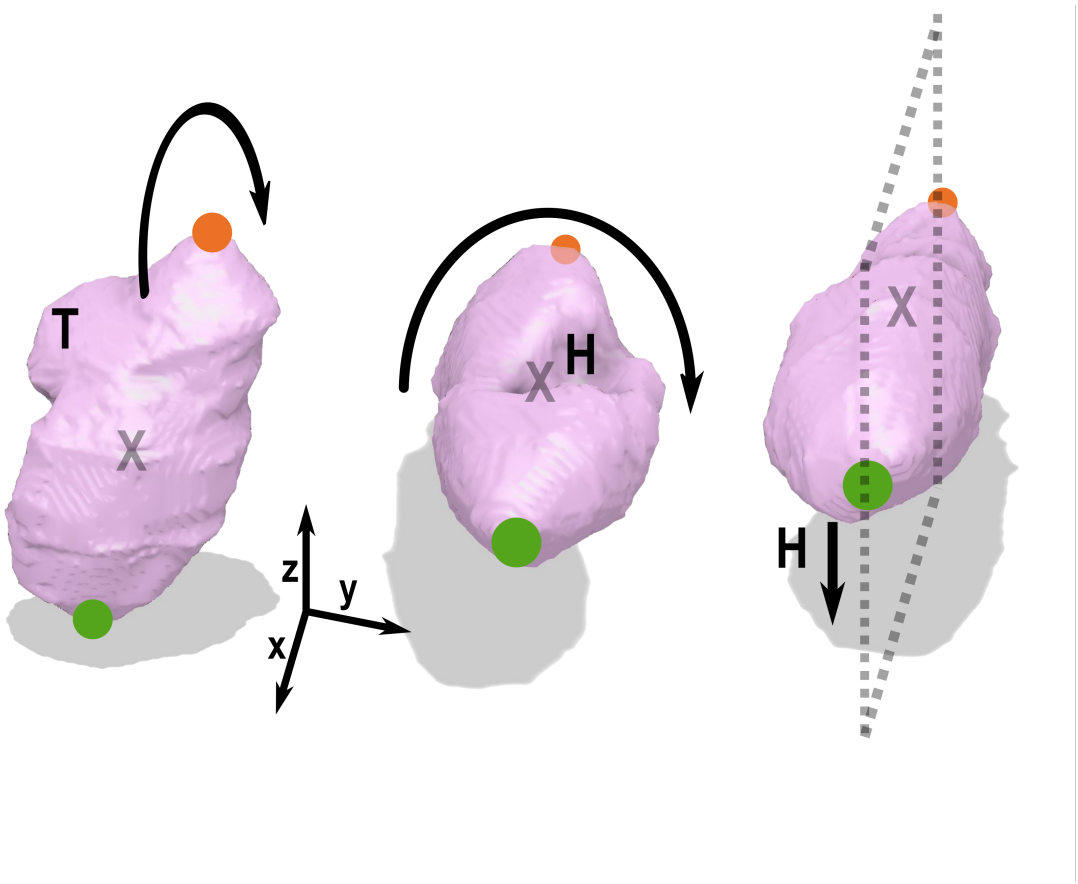


Fig. 4.1 Two sequential rotations align the plane of imaging with the plane of the pathological tissue preparation. First, the long axis of the kidney, defined by the vector through the lower (green) and upper (orange) pole of the specimen is aligned parallel to the x-axis. Second, a rotation around the long axis positions the renal hilum (H) facing downwards. X: centroid of the segmentation.

xy-step size of 0.01 mm in-plane). Second, the manual segmentation introduced edges to the tissue outline. Therefore, a 9×9×9 mm Gaussian filtration kernel with a standard deviation of 3 mm smoothened the mesh as a first step, followed by another mesh cleaning step. A final smoothing step using Taubin smoothing removed final irregularities before holes in the mould's surface were filled in.

As described above, the processed mesh was subtracted from the 3D block, and slices and a vertical comb were added to the mould. The position of each gap in the mould and comb co-localised with a re-oriented image. Markings on the mould indicated which image a gap corresponded to. The subsequent step introduced 2 cm openings in the bottom of the mould underneath the landmarks of the identifiable surface structures and the hilum. Code was written in Python and interfaced with Meshlab [393] and OpenSCAD. Subsequently, the

3D model was sliced in PrusaSlicer (Prusa Research, Prague, Czech Republic). The mould was extruded in polylactic acid (PLA, RS Components, Corby, United Kingdom) on a Prusa i3 MK3s printer at a layer height of 0.2 mm.

Tumour habitat definition and biopsy. Tumour habitats are subregions of a lesion, defined in an unsupervised way based on their unique imaging appearance. Such clustering may reveal the composition of the tumour from biologically distinct tissues. Multiple textures, multiple MR sequences or multiple imaging modalities can define tumour habitats [394].

Multiparametric MRI data defined the habitats in this patient cohort. The selected image contrasts represented a range of tissue properties, aiming to generate biologically meaningful habitats. The T_1w LavaFlex sequence was used as a reference because the structures were segmented in its geometry. Additionally, spatially registered T_2w images, T_1, K^{trans}, D_0, f_p and R_2^* maps provided physiological information on tumour perfusion, diffusivity and oxygenation. Section 5.2.4, 128 page describes the image processing pipeline in detail.

Voxel intensities from each co-registered series and the x-, y-, and z-coordinates were combined through k-means clustering to obtain habitats. The coordinates were included in the clustering algorithm to achieve spatial cohesion of the habitats. Their optimum number was defined so that no habitat had a surface area of less than 3 cm^2, ensuring that habitats could be sampled and would not be "missed" with the 1 cm thick tissue slices.

Assessment of imaging-tissue registration accuracy. The spatial registration accuracy between tissue cross-sections obtained with the 3D-printed tumour mould and the predicted cross-section from the re-oriented MR images was investigated. First, a macrograph of the tissue cross-section was acquired from the fresh specimen when a medical photographer was available and from the formalin-fixed tissue when no photographer was available. Photography before fixation whenever possible avoided shrinking artefacts. Tumours up to 5.5 cm in diameter, smaller than the average lesion included in this study, lose 5% in diameter after one day of fixation in 10% formalin solution, which would have introduced bias [395]. Additionally, distortion may arise if different parts of the tumour shrink to different extents. The renal tumour and adjacent normal-appearing kidney were outlined on the macrographs and overlaid with the segmentation of the corresponding, re-oriented MR image. Manual rigid registration was applied to achieve the best overlap between the two segmentations. The Sørensen-Dice similarity coefficient (DSC, eq. (2.2), page 48) quantified the registration accuracy between image and tissue cross-sections. The DSC approaches 1 for perfect agreement between the segmentations and 0 for non-overlapping segmentations.

4.3 Results

The mould design was improved iteratively in the first patients recruited into the MISSION study between May 2018 and October 2018. Subsequently, the final mould design described above was validated in nine patients recruited between November 2018 and January 2020. Three patients in the validation cohort had to be excluded. One patient experienced disease progression before the planned imaging session and was no longer eligible, one patient was diagnosed with a phaeochromocytoma and was no longer eligible at the time of surgery, and the macrograph was of insufficient quality in one patient. Therefore, six patients were available for the validation of the 3D-printed tumour mould design. Table 4.1 summarises patient characteristics and section 4.3 visualises the mould and habitat generation for one patient.

Table 4.1 Characteristics of patients in the mould design development and validation cohorts

Participant feature	Development	Validation
Number of patients	5	6
Age at consent [median: range; years]	67: 59–71	57.5: 47–81
Sex (m / f)	5 / 0	5 / 1
Histology	4 clear cell RCC 1 chromophobe RCC	5 clear cell RCC 1 Liposarcoma
Stage at diagnosis (RCC only)	3 pT3a pNX/0 cM0 1 pT3a pNX/0 cM1 1 pT3b pN0 cM1	1 pT1b pNX cM0 4 pT3a pNX/0 cM0
Time from imaging to surgery [median: range; days]	14:7–26	16.5: 8–29
Tumour volume [median: range; cm^3]	892: 29–1433	126: 89–1342

4.3.1 Iterative mould design process

The initial iterations generated the mould from the outline of the renal tumour and adjacent normal renal tissue. This approach was highly accurate for patients with little perirenal fat but failed to accommodate larger specimens. The inclusion of a uniform margin for the adipose tissue could have been derived from the existing tumour and kidney segmentation but did not deliver the desired generalisability. Eventually, the additional segmentation of Gerota's fascia, the surgical resection margin, was required to reach a functional mould design. However, with the inclusion of the perirenal fat, the lateral surface was no longer suitable to rest the specimen on. The uncertainty in the amount and compressibility of the

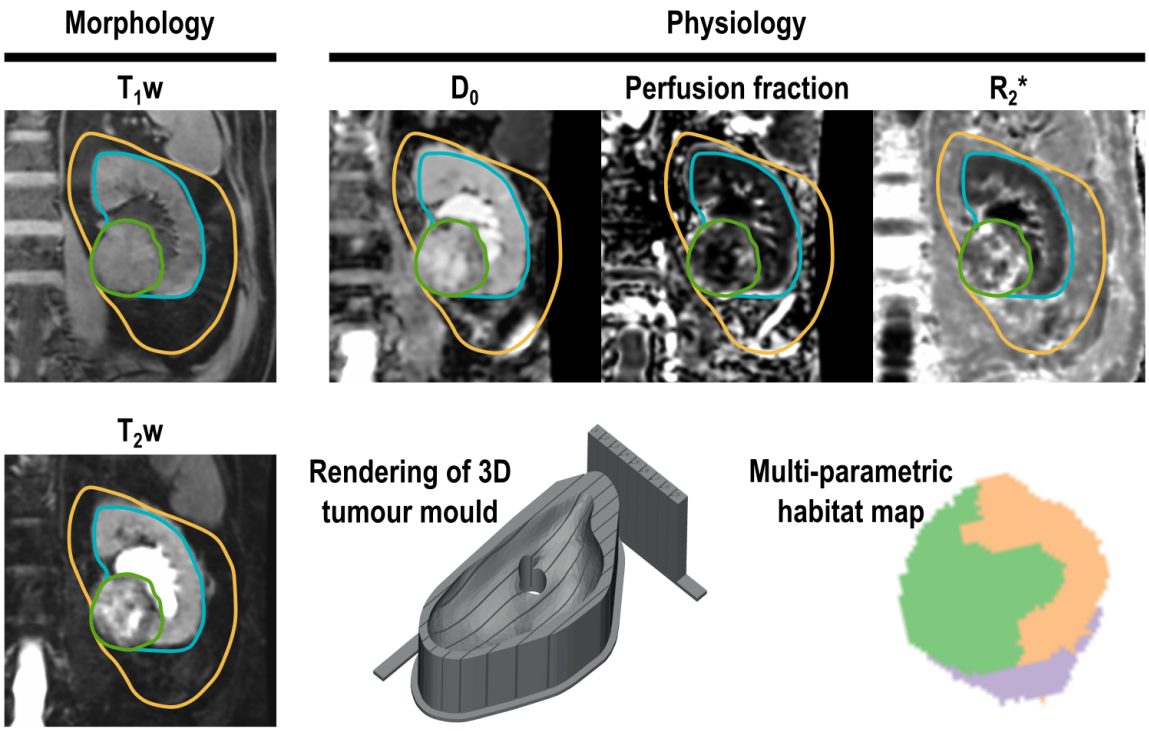


Fig. 4.2 Tumour habitats are generated from multiple co-registered MR sequences capturing morphological and physiological properties of the tumour (green). The normal-appearing kidney (blue) and perirenal adipose tissue (yellow) were outlined manually on the T_1w MRI to generate tumour moulds. Openings in the base of the mould enable visual and tactile verification of the correct positioning of the specimen. K-means clustering defines tumour habitats with unique combinations of morphological and physiological phenotypes on multiple co-registered MRI sequences.

adipose tissue would have resulted in decreased precision in the positioning. Therefore, specimens were placed in the mould with the medial surface facing down and openings were introduced to the mould to verify the correct positioning. They also ensured that the hilum remained intact for pathological diagnosis.

The production of tumour moulds is possible within 48 hours. However, typically two weeks elapsed between imaging and surgery. The manual segmentation of the tumour, normal kidney, Gerota's fascia and landmarks took between 2 and 4 hours, depending on the tumour size. The automated generation of the 3D model, image re-orientation and habitat clustering required less than five minutes. A computational biologist, radiologist and pathologist then verified the mould to ensure its suitability. The pathologist chose to cut the tumour axially in five patients and coronally in one. The time required to print the mould was heavily dependent on its size and could be up to one day for larger tumours and patients with more perirenal adipose tissue.

Clear cell RCC regularly contains cystic components, leading to tissue deformation after the first section. Therefore, the first tumour cross-section had to include as many of the habitats as possible at a reasonable proportion and be sufficiently removed from the hilum.

4.3.2 Validation of the registration accuracy

Initially, the quality of the registration between imaging and tissue sampling was appraised visually. Figure 4.3 shows an example of a tumour with internal complexity where localised haemorrhage and macroscopic fat visible on the MRI can be found at the corresponding location of the tissue cross-section.

The quality of the registration was assessed quantitatively from tissue macrographs and the corresponding segmentation using the DSC. In the validation cohort, the mould achieved a good registration accuracy in the tumour (median: 0.92, range: 0.86 – 0.96). Normal renal tissue was visible on the cross-sections of four of the six patients. The more irregularly shaped normal kidney achieved a median DSC of 0.74 (0.70 – 0.76)

Figure 4.5 shows how tumour habitats can identify different components of a tumour. The three habitats correctly identified a cystic tumour component, necrotic/fibrotic tumour and the viable tumour. The imaging appearance of these regions corresponded with the pathological findings: The cystic tumour region showed no enhancement on the DCE MRI, a high T_2w-signal intensity and high diffusivity. Intermediate diffusivity, low-grade enhancement and a high R_2^* characterised the fibrotic/necrotic region. The viable tumour was highly enhancing, diffusion restricted and had an intermediate R_2^* signal.

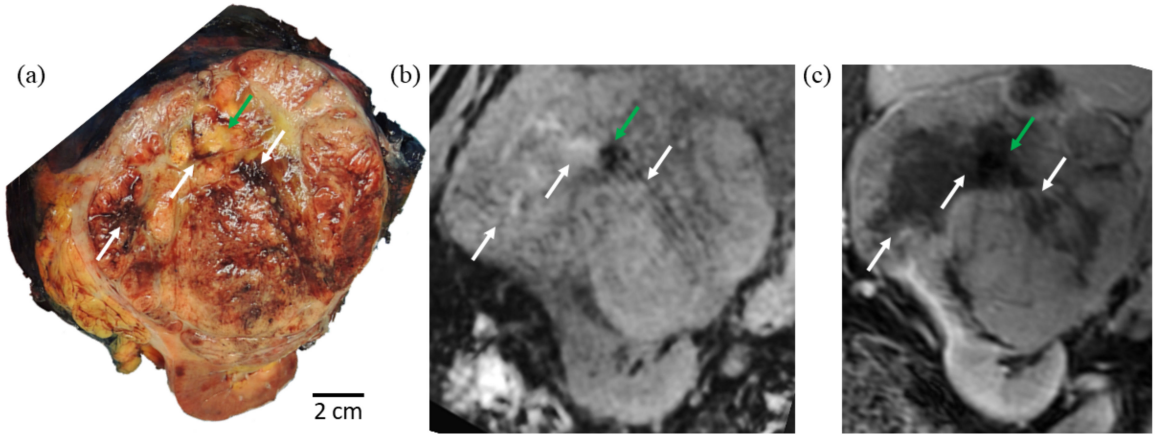


Fig. 4.3 Cross-section of a surgical specimen co-registered to MRI data using a 3D-printed tumour mould. Macrograph (a) with the resection surface inked black, T1-weighted sequence (b) and contrast-enhanced sequence in the arterial phase (c) show a coronal cross-section of the organ (coronal-oblique to the body). At the inferior part of the specimen remaining normal kidney is visible. Haemorrhagic necrosis (white arrows) and intratumoral macroscopic fat (blue arrow) in corresponding positions.

4.4 Discussion

Here we presented an approach to the automated generation of a 3D-printed tumour mould for precision tissue sampling in renal cancer. The design achieved a high registration accuracy between MR-imaging and pathological tissue cross-sections of the tumours. The novelty of this approach is severalfold: Conventionally, the shape of tumour moulds accommodates a structure of interest, for example, a tumour [386] or an organ [385]. In contrast, the design introduced here accounted for the adipose tissue covering the structure of interest, which was a pre-requite for its application to radical nephrectomy specimens. As the exact amount of resected adipose tissue was unpredictable, specimens were placed on the surface with the least distance between the predicted resection margin and the kidney and tumour. This ensured that no fat had to be removed from the specimen, which could have hindered the histological assessment of the resection margin. Second, existing mould approaches adopted the direction of the acquisition of radiological images for tissue slicing. However, our method re-oriented the imaging data and created a tumour mould that respected existing guidelines for pathological tissue preparation. Therefore, seamless integration of precision tissue sampling for research into clinical pathways is possible.

Commandeur *et al.* have taken a complementary approach, maintaining both the conventional orientation of imaging and pathological slicing. Instead, they registered scans of the whole mount histology of the prostate with MR images. This approach achieved registration

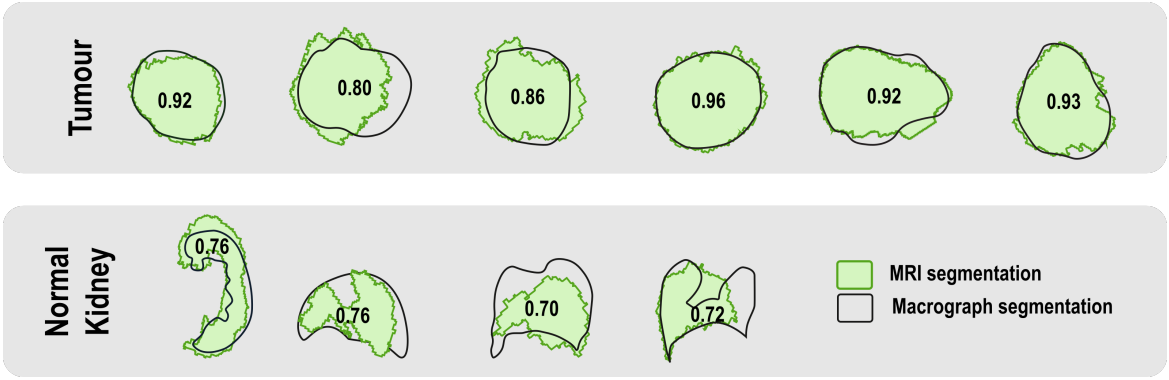


Fig. 4.4 Overlay and Dice similarity coefficients of the segmentation on the T_1w MRI sequence (green) and the macrograph of the tissue slice (black). Adapted from: Crispin-Ortuzar M*, Gehrung M*, Ursprung S*, *et. al.* Three-Dimensional Printed Molds for Image-Guided Surgical Biopsies: An Open Source Computational Platform. JCO Clinical Cancer Informatics. 2020; 4, 736-748.

errors of less than 5 mm but could not aid tissue sampling as the registration occurred *a posteriori* [396]. The pipeline described here relies on identifying landmark structures on imaging that define the rotations required to bring the specimen into the correct location. The rotations to achieve the desired plane for coronal or axial slicing, depending on the case, were achieved in all patients. The method described here applies to organs that are too big for whole-mount histology.

Designing a mould that can integrate with other studies was one of our aims. As a result, the MISSION renal trial adopted the mould presented here to avieve the multi-regional correlation between imaging and tissue biomarkers. The comparison included hyperpolarised [1-^{13}C]-pyruvate MRI parameters, multi-parametric ^{1}H-MRI, immunohistochemistry, liquid chromatography-mass spectrometry (LC-MS), in-situ techniques like RNAscope, RNA and DNA sequencing. Chapter 5 reports the initial results of this study. Subsequently, the WIRE window-of-opportunity clinical trial platform in renal cancer combined MR-ultrasound fusion habitat-guided biopsies (similar to [384]) in treatment-naive patients with mould-guided tissue sampling after the window-of-opportunity treatment and nephrectomy. Chapter 6 summarises the trial protocol and initial results of the WIRE trial.

The mould design had the adaptability to clinical needs and other diseases in mind. For example, the orientation of the specimen in the mould is adaptable to the pathologist's requirements. A first derivative of the mould was used to obtain tissue samples from ovarian cancer deposits after a delayed primary surgery. This application of the mould complements the ultrasound-guided tissue sampling described by Beer *et al.* [384]. Weigelt *et al.* have

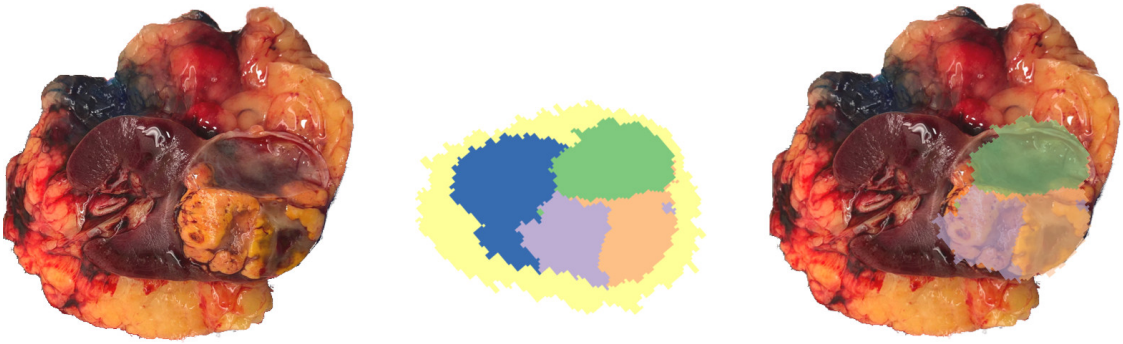


Fig. 4.5 K-means clustering derived habitats align with areas of macroscopically different tissue components of a ccRCC. Cystic tumour component in green, necrotic/fibrotic tumour in orange and viable tumour in purple. The normal kidney and perirenal adipose tissue are displayed in blue and yellow, respectively.

reported a similar technique in patients with ovarian cancer undergoing primary surgery where tumour habitats were able to track genetic aberrations across tumour sites [387].

The work presented here has some limitations. First, manual segmentation is a time-consuming process. Semi-automatic and automatic segmentation tools are available to accelerate the delineation of the tumour. However, segmenting Gerota's fascia required coordination with surgeons about their intended resection margin and is rarely performed. Therefore, no segmentation models are available. Second, the design and printing of a tumour mould required at least 48 hours between the imaging and surgery. Typically, two weeks elapsed in the patients included in this study. While changes in the tumour physiology and architecture could occur, these are likely minor in a cancer type with a tumour doubling time of 18 – 22 months [397, 398].

In summary, we have developed a methodology to register tissue samples and imaging in renal cancer treated with radical nephrectomy. The 3D-printed tumour mould facilitates tissue sampling from pre-defined locations such as those from tumour habitats. Furthermore, it will enable multi-omics data analyses across spatial scales between imaging, metabolomics, transcriptomics and genomics. Furthermore, generating the mould is highly automated and respects existing clinical pathways. In the future, joint analysis of imaging- and tissue-biomarkers may open opportunities for personalised cancer care.

Chapter 5

Hyperpolarised [1-^{13}C]-pyruvate MRI is sensitive to tumour aggressiveness and adverse outcome in clear cell renal cell carcinoma: the MISSION renal study

Acknowledgements:
The work presented in this chapter received substantial support from Dr Mary McLean who developed the MR sequences used for the hyperpolarised ^{13}C MRI acquisition in collaboration with Dr Rolf Schulte. Dr McLean, Dr Arnold Benjamin and Dr James Grist conceived the ^{13}C reconstruction algorithms.
Dr Andrew Priest designed the ^{1}H-MRI acquisition protocol and Dr Andrew Gill, Dr Gaspar Delso and Dr Andrew Priest developed code for the reconstruction, motion correction and registration of the multiparametric MRI data.
Stephan Ursprung has conceived the idea for this project in collaboration with Professor Ferdia Gallagher. Stephan Ursprung defined the research question and developed the research methodology in collaboration with Dr McLean and Dr Priest. He obtained the ethical approval for this sub-study of the MISSION ovary study, recruited the participants, segmented the tumours, conducted the image and data analysis, performed the statistical analysis, visualised and interpreted the results.

5.1 Introduction

The increasingly incidental detection of small renal masses poses a challenge for clinical management which needs to avoid overtreatment of indolent tumours while achieving a favourable oncological outcome for aggressive cancers. However, neither current CT/MRI imaging strategies nor tumour biopsies can inform this decision reliably [5]. Furthermore, cancers recur locally or with distant metastasis in one in five to one in two patients after surgery with curative intent [81, 399]. The tumour WHO/ISUP nuclear grade is the leading risk factor for progression, recurrence and death [59, 400]. However, owing to intratumoral heterogeneity, tumour biopsies frequently underestimate the tumour grade [63]. Consequently, a non-invasive imaging technique that is sensitive to the tumour grade, can distinguish benign and malignant renal tumours and can interrogate the entire tumour volume, thereby overcoming the challenge of intratumoral heterogeneity, could support clinical decision making and improve patient stratification.

Imaging metabolism in RCC. The dysregulation of metabolic pathways including glycolysis (section 1.3.3, page 20) in renal cancer suggests a potential role for metabolic imaging [163]. The almost ubiquitous inactivation of VHL through mutation or promotor hypermethylation stabilises HIF-1α which leads to increased glycolytic activity in ccRCC [134, 137]. However, a deregulated cellular energy metabolism is by no means unique to ccRCC and is recognised as a hallmark of cancer which is necessary to sustain continued proliferation [168]. The shift towards aerobic glycolysis, first described by Otto Warburg, is paralleled by the upregulation of *GLUT1* to compensate for the lower energetic yield of glycolysis compared to oxidative phosphorylation [401]. The widespread clinical use of ^{18}F-fluorodeoxyglucose (a glucose analogue) positron emission tomography (^{18}F-FDG-PET) is a testament to the clinical value of metabolic imaging [402]. ^{18}F-FDG-PET quantifies the first step of the glucose metabolism with exquisite sensitivity at picomolar tracer concentrations while CT and MR contrast agents require millimolar concentrations [403]. However, ^{18}F-FDG-PET is not without limitations. Upregulated glucose uptake is not specific to cancer cells and occurs in inflammation and physiologically in some organs including in the brain, liver, bowel, proliferating endometrium and ovarian follicles [404]. Furthermore, the PET signal cannot differentiate its molecular origin, whether it is the tracer or one of its metabolites. Applications of ^{18}F-FDG-PET in ccRCC are limited to research owing to an intense background signal from the urinary tracer excretion [405].

Magnetic resonance spectroscopy can interrogate proton (^{1}H), phosphorus (^{31}P), carbon (^{13}C) or fluorine (^{19}F) spectra, determine relative and estimate absolute metabolite concentrations with decreasing signal intensity as the nuclei become increasingly rare and the gyromagnetic ratio γ decreases [406]. However, due to the lower abundance of these molecules compared to water and the low sensitivity of MR at clinical field strengths, MRS suffers from a low resolution and long acquisition times even when the molecule of interest is present at millimolar concentrations [407, 194]. Nonetheless, the technology has found some limited clinical applications in metabolic and mitochondrial diseases affecting the brain, brain tumours, and prostate cancer [408]. Importantly, the low temporal resolution limits MRS to measuring steady-state concentrations. Respiratory motion is a major challenge in MRS of renal tumours. However, limited data suggest that the choline signal may be higher in malignant compared to benign tumours and the presence of a lipid peak in a renal tumour is suggestive of ccRCC [197, 198].

Hyperpolarised ^{13}C MRI is an appealing tool for clinical metabolic imaging. Hyperpolarised ^{13}C MRI is an emerging technology based on carbon rather than hydrogen imaging. Non-hyperpolarised carbon NMR can be employed to probe the structure of molecules in organic chemistry, similar to proton NMR. However, the lower gyromagnetic ratio of carbon compared to hydrogen and the low natural abundance of the MR-active ^{13}C isotope of 1.1% result in a sensitivity which is approximately four orders of magnitude lower than that of proton NMR [409]. Dynamic nuclear polarisation (DNP) is a method for overcoming this reduced sensitivity, allowing the rapid *in vivo* imaging of changing concentrations of metabolites at improved spatial resolution. DNP aligns the nuclear spins of isotopically enriched molecules to an external magnetic field beyond the thermal equilibrium dictated by the Boltzmann distribution [410]. A mixture of [1-^{13}C]pyruvate, a gadolinium chelate and a free radical is cooled to about 1 Kelvin and exposed to a strong magnetic field. Microwave irradiation of the sample results in the transfer of the polarisation from electrons of the radical to ^{13}C nuclei [411, 412]. Gadolinium at low concentrations increases the attainable level of nuclear polarisation [413]. Rapid dissolution in a heated buffer solution preserves most of the signal and brings the pH into a physiological range. After removal of the free radical, the solution can be injected into patients. In 2003, this technique was first used to hyperpolarise [^{13}C]urea and image its spatial distribution in a rat [414, 415]. Since then, hyperpolarised ^{13}C pyruvate has been successfully translated into clinical use. and imaged for the first time in patients with prostate cancer [416]. In contrast to ^{1}H or non-hyperpolarised ^{13}C MRS, the

exogenic hyperpolarisation allows the quantification of metabolic fluxes as the background signal is minimal [417].

Pyruvate occurs physiologically and occupies the crossroad of several metabolic pathways (Figure 1.1, page 24). FDG-PET probes tissue perfusion, glucose transport into the cell and its phosphorylation by hexokinase. As such, it is only an indirect marker of tissue glycolysis. Pyruvate is located further downstream in the glycolytic pathway and may provide complementary information on the deregulation of glucose utilisation in cancer. The *VHL* inactivation and accumulation of the transcription factor HIF-1α leads to an increased expression of *GLUT1* and *MCT-1* as well as the activation of pyruvate dehydrogenase kinase 1 (PDK). In summary, these lead to increased cellular uptake of glucose and pyruvate, shunting of pyruvate into lactic acid fermentation and export of lactate [418]. The resulting elevated lactate concentration is found in many types of cancer, was shown to correlate with tumour grade and is exploitable in hyperpolarised [1-^{13}C] pyruvate imaging thanks to the rapid exchange of the ^{13}C-label between pyruvate and lactate [419, 180]. In addition to its function as a potential imaging biomarker, the increased lactate is essential for multiple aspects of carcinogenesis. It stimulates angiogenesis, promotes cell migration and metastasis, and inhibits the immune response to the cancer [420].

Challenges in hyperpolarised ^{13}C MRI. Hyperpolarised MRI comes with challenges, the most noteworthy of which are the temporary nature of the hyperpolarised state and its irreversible destruction during image acquisition. The longitudinal relaxation time (T_1) of hyperpolarised pyruvate depends on the external magnetic field and is 59.8 seconds at 0.75T; after one T_1 37% of the polarisation is lost [421]. However, the transfer through a low field environment before injection and the presence of paramagnetic iron in the blood shorten T_1 *in vivo*. Consequently, *in vivo* T_1 is estimated to be around 30 seconds [422]. The radio-frequency excitation that creates the MR signal in the transverse plane further reduces the longitudinal magnetisation. Compared to proton imaging, it does not recover in hyperpolarised ^{13}C imaging. The desired SNR, number of time points and frequency of acquisitions determine the optimal flip angle. Larger flip angles yield a higher signal intensity and SNR during the initial time points of the dynamic acquisition but exhaust the polarisation quickly, leading to a lower signal intensity towards the end of the time course. Technical advances allow making efficient use of the longitudinal magnetisation: Variable flip angle schemes improve the preservation of the magnetization for later acquisition time points and spectral-spatial pulses excite metabolites selectively. Selective excitation of metabolites allows preserving the signal of more abundant molecules through lower flip

angles while acquiring the maximum signal from molecules at lower abundance through higher flip angles [423–425]. The transient nature of the hyperpolarised signal achievable through conventional DNP, the costly infrastructure, and the required technical expertise hinder a widespread application of the technology. However, efforts to achieve longer-term stability of the hyperpolarised state as a pre-requisite for centralised production and less technologically demanding production methods like parahydrogen induced polarisation are underway [377, 426].

Hyperpolarised ^{13}C-MRI of the kidney. Hyperpolarised MRI has been used to study benign and malignant renal conditions. The first dynamic spectra and spectroscopic images after hyperpolarised [1-^{13}C] pyruvate injection, published in 2007, identified the presence of pyruvate, lactate, alanine, and bicarbonate in rats [427]. Furthermore, the lactate-to-pyruvate ratio increases and the conversion of dehydroascorbate (DHA) to vitamin C decreases in renal ischaemia-reperfusion injury in mice [428]. Similarly, ureteral obstruction leads to an increase in the lactate-to-pyruvate ratio [429]. The induction of diabetes in murine models leads to increased pyruvate-to-lactate conversion and shortening of the HP-^{13}C-urea T_2 in response to tissue hypoxia, observable with hyperpolarised MRI [430, 431]. Hyperpolarised MRI of rodent kidneys also played a role in the development of parallel multi-nuclear imaging techniques as well as the introduction of molecules to interrogate the RedOx state and tissue pH *in vivo* [432–437]. *In vitro*, more aggressive cancer cell lines export lactate more efficiently, leading to decreased intracellular and increased extracellular lactate concentrations despite increased lactate production [201]. This was was confirmed in *in vitro* studies on patient-derived tumour tissue [438]. A reduction in pyruvate-ot-lactate conversion is measurable 24 hours after the initiation of mTOR inhibition in murine xenografts of ccRCC [439]. Intratumoral heterogeneity of the HP-^{13}C-pyruvate-to-lactate conversion was demonstrated in patients in 2019 [199] and a case series in 2021 suggested that the lactate-to-pyruvate ratio is associated with tumour aggressiveness [200].

Monocarboxylate transporter 1 as a determinant of hyperpolarised ^{13}C-pyruvate to lactate conversion. Hyperpolarised [1-^{13}C]-pyruvate MRI interrogates multiple reactions at the end of the glycolytic pathway. Lactic acid fermentation catalysed by lactate dehydrogenase (LDH) is of most interest to oncological applications. However, the expression of cell membrane transporters of pyruvate and lactate, monocarboxylate transporters 1 and 4, influence ^{13}C-lactate signal formation in addition to LDH(A) activity (fig. 1.1, page 24). MCTs are involved in cellular energy metabolism through the transport of pyruvate, lactate and

ketone bodies across the plasma membrane, facilitating the maintenance of the intracellular pH [440]. Findings from human breast and prostate cancer, as well as pre-clinical studies, have identified MCT1 mediated pyruvate uptake as a rate-limiting factor for ^{13}C-lactate formation [202, 205, 441]. However, other pre-clinical studies have also described correlations with MCT4 and LDHA expression [201, 442]. This chapter assesses the properties of HP-^{13}C-MRI as a non-invasive measure of tumour aggressiveness in ccRCC and correlates this with changes in the expression of MCT1 for the first time in patients with RCC. We subsequently validated these changes in gene expression and patient outcome data from The Cancer Genome Atlas Kidney Renal Clear Cell Carcinoma (TCGA-KIRC) database [443]. Additionally, tissue metabolic analyses served to elucidate correlations with the imaging metabolism of HP-^{13}C-pyruvate.

5.2 Methods

5.2.1 Study participants and ethics

Study participants were identified at the urological multidisciplinary team meeting at Addenbrooke's hospital and invited to participate during their next clinical appointment. All participants provided written informed consent to the Molecular Imaging and Spectroscopy with Stable Isotopes in Oncology and Neurology(MISSION) study (Approval: Cambridge South Research Ethics Committee, 15/EE/0378) and the Evaluation of biomarkers in urological diseases bio-repository study (Approval: Cambridge South Research Ethics Committee, 03/018). Patients $>$18 years with a clinical suspicion of renal cell cancer who were candidates for surgery and had an Eastern Cooperative Oncology Group (ECOG) performance status $\leqslant$1 were eligible to participate in the study. Patients with metabolic disorders including poorly controlled diabetes, allergy to gadolinium-based contrast agents, pregnancy or contra-indications for MRI scanning were excluded. Specific to this trial, the MR coil setup was not able to accommodate patients with an anteroposterior abdominal diameter $>$31 cm. Equally, patients with tumours $<$4 cm in diameter were excluded because the resolution of the HP-^{13}C-MRI would cause a significant partial volume effect.

5.2.2 Preparation of the hyperpolarised [1-^{13}C]pyruvate

Kits for the production of fluid paths for the sterile preparation of hyperpolarised [1-^{13}C]pyruvate were manufactured by GE Healthcare (Waukesha WI, USA), filled at the radio-pharmacy at Addenbrooke's hospital and stored at -20 °C. Each path contained 1.47 g

Table 5.1 Reference quality control parameters for hyperpolarised [1-^{13}C]pyruvate preparation

Quality control parameter	Reference value
pH*	6.5 – 8.2
Pyruvate concentration [mmol/L]*	210 – 280
EPA concentration [µmol/L]*	$\leqslant$3
Temperature [°C]	32 – 36
Dissolved volume [mL]	$\geqslant$40
Polarisation [%]	$\geqslant$15

* Critical parameters for which reference values must be respected. EPA: Electron paramagnetic agent.

of [1-^{13}C]pyruvic acid (Sigma Aldrich, St Louis, Missouri, USA) and 15 mmol L^{-1} of an electron paramagnetic agent (EPA, AH111501, Syncom, Groningen, Netherlands). After defrosting, kits were electromagnetically irradiated at a frequency of 139 GHz in the SPINlab clinical hyperpolariser (Research Circle Technology, Niskayuna, NY) for at least three hours at 5 T and 0.8 K. Immediately before injection, the sample was dissolved in 38 mL of superheated water. A filter reduced the concentration of the electron paramagnetic agent before buffering of the solution in 19 mL sterile water and 17.5 ± 0.5 g NaOH/Tris/EDTA (2.4%, 4.03% and 0.033% w/v, Royal Free Hospital, London). The solution passed through an automated quality control module to verify the suitability for use in patients (table 5.1) After a final filtration step (0.2 µL; ZenPure, Manassas, VA, USA), 0.4 mL/kg body weight were injected into a venous catheter in the antecubital fossa at 5 mL s^{-1}, followed by a 25 mL saline chaser (MedRad Spectris Solaris EP MR Injection System, Warrendale, Pennsylvania, USA).

5.2.3 MRI acquisition

The MR acquisition consisted of two parts with differing coil setups. Initially, the hyperpolarised ^{13}C-MRI was acquired followed by repositioning of the patient and acquisition of the ^{1}H-MRI. The ^{13}C-MRI had to be acquired first because even low levels of gadolinium in the blood or tumour would shorten the T_1 of HP-^{13}C-pyruvate and result in a signal loss. Gadoteric acid has an elimination half-life of 1.62 h and an *in vitro* experiment showed complete HP-^{13}C-pyruvate signal suppression at concentrations three hours after the injection of gadolinium-based contrast-agent and taking the distribution volume and clearance into account. All scans were acquired on a clinical 3 T MR system (MR750, GE Healthcare, Waukesha WI, USA). The MRI acquisition schedule is visualised in fig. 5.1.

Hyperpolarised [1-^{13}C]pyruvate MRI

The dedicated ^{13}C-tuned transmit and receive hardware consisted of a clamshell coil for transmission [444] integral to an MRI patient table and an 8-channel array coil (Rapid Biomedical, Rimpar, Germany) for signal reception. The receive coil was centred over the tumour. Patients were either imaged with an Iterative Decomposition with Shifted Echo times and Least Squares Estimation (IDEAL) spiral chemical shift imaging (CSI) sequence [434] or a spectral-spatial pulse sequence. In either case, up to five axial slices of 3 cm thickness with a 5 mm gap were placed through the tumour. A 34×34 cm^2 FOV and acquisition matrix of 40×40 was acquired every 4 s for 64 s (repetition time (TR)=0.5 s) starting 12 s after the injection of the HP-^{13}C-pyruvate.

For the IDEAL spiral acquisition, radiofrequency pulses with a nominal flip angle (α) of 15° excited the ^{13}C-pyruvate and its metabolites simultaneously. Meanwhile, the spectral-spatial sequence excited ^{13}C-pyruvate with a α =10° and ^{13}C-lactate, ^{13}C-alanine and ^{13}C-pyruvate-hydrate with α =40° with a 22.4 ms pulse with flyback gradients, each followed by a single-shot spiral readout (TR=2 s) [445].

The patients were asked to hold their breath during the ^{13}C-MRI acquisition for as long as it was comfortable (approximately 20 s) and then breathe shallowly. This aimed to reduce breathing motion artefacts during the initial phases which contribute most to the signal. Images were upsampled from 17×17 mm^2 to 5×5 mm^2 (128×128) in-plane.

A water image of a T_1w 3D Dixon sequence was acquired immediately before the acquisition of the HP-^{13}C-MRI using the body coil built into the scanner as transmit and receive coil. This allowed tumour segmentation on anatomical images and facilitated the correlation of data from the ^{13} and ^{1}H MRI.

Proton MRI

Patients were re-positioned into a 32-channel cardiac array coil (GE Healthcare, Waukesha WI, USA) on the same MRI scanner immediately following the ^{13}C-MRI. The ^{1}H-MRI protocol described below and summarised in table D.1 on page 280 was acquired for all participants except one who presented with a tumour thrombus in the inferior vena cava and was imaged with the protocol described for the NAXIVA trial (table C.2 in section C.4). In two participants, the number of acquired slices was increased to cover the full tumour, leading to slightly longer breath-holds.

A 3D Dixon sequence (GE implementation: LAVA-Flex) was used to obtain coronal T_1w water, fat, in- and out-of-phase images: TE: 1.1/2.2 ms; TR: 3.8 ms; flip angle 10°;

FoV 40×40 cm^2; slice thickness 4 mm; acquisition (reconstruction) matrix 224×224×40 (256×256×80); receiver bandwidth ±143 kHz; parallel imaging (ARC) factor 1.5; acquisition time 17 s.

A 3D fast spin-echo sequence with inner volume excitation (GE implementation: FOCUS/HyperCUBE) including fat suppression was used to acquire coronal T2-weighted (T_2w) images: TE: 100 ms; TR: one respiratory cycle; FoV: 40×36 cm^2; slice thickness/slice spacing 4 mm; acquisition (reconstruction) matrix 256×224×40 (256×56×80); parallel imaging (ARC) factor 2; echo train length 120; echo-spacing 5 ms; receiver bandwidth ±62.5 kHz; acquisition time 31 breaths (3 min). The coverage was selected identically to the T_1w LAVA-Flex sequence.

A multi-echo gradient-echo (GRE) sequence and 12 echo times from 2.3 and 36.2 ms (3.1 ms echo spacing) was used for coronal R_2^* mapping: TR 110 ms; flip angle 30°; FoV 40×40 cm^2; slice thickness 4 mm; acquisition matrix 256×224; receiver bandwidth ±62.5 kHz; parallel imaging (ASSET) factor 1.5. Two slices were acquired in each 17-second breath-hold, and multiple breath-holds were used to cover the entire tumour. The sagittal extent of the FoV was reduced to the tumour only to achieve a faster acquisition. However, slices remained centred identically to the T_1w LAVA-Flex sequence.

DWI was acquired in coronal orientation using a respiratory navigator-triggered dual spin-echo echo-planar imaging (EPI) sequence and b-values of 0, 10, 20, 30, 100, 300, 500, 700 and 900 s/mm^2: echo time (TE) 80 ms; TR 1 respiratory cycle; FoV 28.8×28.8 cm^2; slice thickness 4 mm; acquisition matrix 96×96; 2 averages for b-values <100 s/mm^2 and 4 averages for higher b-values; receiver bandwidth ±111 kHz; parallel imaging (ASSET: Array coil Spatial Sensitivity Encoding) factor 2; 3 directions averaged to form trace-weighted image; acquisition time 91 breaths (10 min). Saturation bands were used to reduce signals from outside the volume of interest and the slice coverage was identical to the R_2^* mapping above.

T_1-mapping was required to derive contrast concentrations during the DCE-MRI. Therefore, multiple-flip-angle T_1w images (flip angles 2°, 3°, 5°, 8°, 14°); FoV 40×40 cm^2; slice thickness 4 mm; acquisition matrix 196×172×48 (reconstruction matrix: 256×256×96); receiver bandwidth ±43 kHz; parallel imaging (ARC: Autocalibrating Reconstruction for Cartesian imaging) factors 1.2×1.0; scan time 16 s (breath-hold) per flip angle.

B_1 mapping aims to record the spatial inhomogeneity of the radiofrequency transmit (B_1) field and was performed using the Bloch-Siegert method: TE 13 ms; TR 31 ms; nominal flip angle 20°; acquisition matrix 128×128; FoV 40×40 cm^2; slice thickness 12 cm; receiver

bandwidth ±15.6 kHz. The DCE-MRI was processed using the extended Tofts model with a model arterial input function [335, 446].

A Dixon imaging (LAVA-Flex) sequence in coronal orientation was used for the DCE-MRI: TE 1.1/2.2 ms; TR 3.8 ms; flip angle 18°; FoV 40×40 cm^2; slice thickness 4 mm; acquisition (reconstruction) matrix 196×172×48 (256×256×96); receiver bandwidth ±43 kHz; ARC factors 2×1.5; temporal resolution 13 s (consisting of 7 s to acquire the images and 6 s pause for the patient to breathe); and 44 dynamic phases. In each case, the total scan time was approximately 9.5 min. Gd-DOTA (0.1 mmol/kg, Dotarem, Guerbet, Paris, France) was administered intravenously during the dynamic series at 2.5 mL/s followed by a 25 mL saline flush injected at the same rate. The injection commenced 39 s after the start of the scan.

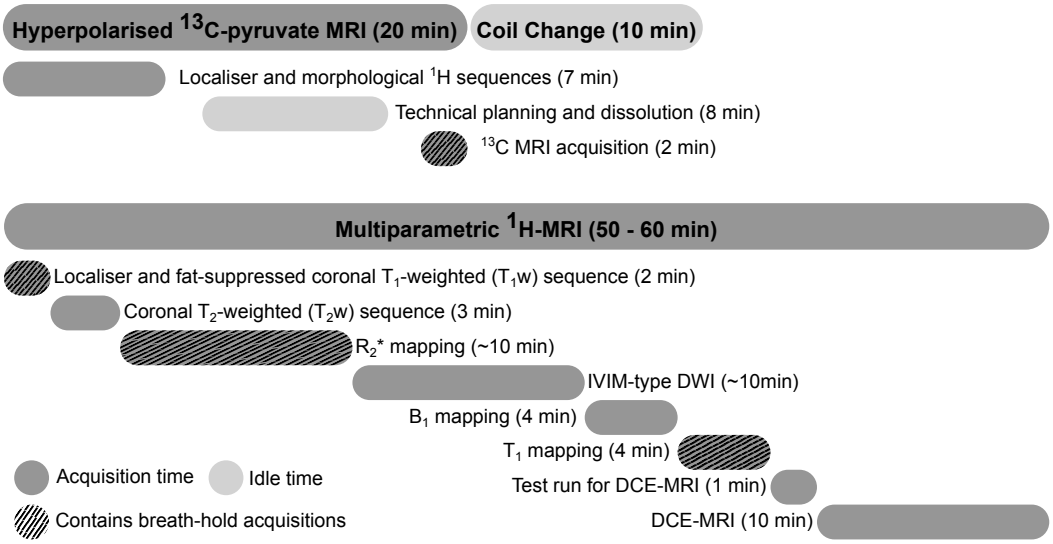


Fig. 5.1 MRI acquisition schedule for the MISSION study. The scanner time is indicated in dark grey, the idle time during the coil changeover is indicated in light grey. MRI sequences which involve breath-holds are striped. Breath-holds did not exceed 17 seconds.

Instead of a DCE-MRI, a dynamic multi-phase contrast-enhanced scan was acquired in the patient with the IVC thrombus. Angiographic triggering was used for the first phase to achieve optimal enhancement of the blood pool. In total, eight contrast phases were acquired over six minutes. Resulting from the low temporal resolution of this data, no quantitative analysis was attempted.

5.2.4 Data analysis

^{13}C MRI data analysis

The imaging data obtained from the ^{13}C-MRI were reconstructed with custom MATLAB code (Version 2018A, The Mathworks, Natick, MA). A singular value decomposition approach as described by Chen *et al.* was used to combine the complex data from all eight channels of the receiver array coil into a single magnitude image [447]. For comparison, a simpler sum-of-squares method of combining coils and a sum-of-squares method weighted by the signal divided by the square of the noise (S/N^2) were also investigated. The weighting of coils by the signal-to-noise ratio eliminated coils contributing mainly noise and very little signal. Metabolite ratios integrated over time were derived after summing the complex data over all time points before the coil combination as this minimised noise propagation.

The T_1w LAVA-Flex sequence acquired with the body coil was used in the manual segmentation of structures of interest. Regions of interest including the renal tumour, contralateral normal kidney, erector spinae muscle, aorta and, if visible on the HP-^{13}C-MRI, the pancreas were segmented and segmentations propagated to the metabolic maps. The ROI of the aorta served to correct the registration manually if the ^{1}H and ^{13}C images were misaligned. The aorta was used for this because of its well-defined ^{13}C-pyruvate signal. Polygonal ROIs were drawn on all slices containing a structure on their entire thickness using OsiriX 10.0 (Pixemo SARL, Switzerland). For the assessment of signal intensities at the location of individual tumour tissue samples (section 5.2.5), cylindrical ROIs of 3 cm diameter and 3 cm height were drawn at the location of the tissue sample. Voxel intensity values for each ROI were exported from OsiriX in the JavaScript Object Notation (JSON) format.

The tumour SNR was calculated for each metabolite according to eq. (5.1).

$$\mathrm{SNR}_{\mathrm{metabolite}} = \frac{\mathrm{mean}(\mathrm{SI}_{\mathrm{ROI}}) - \mathrm{mean}(\mathrm{SI}_{\mathrm{noise}})}{\sqrt{2}\,\mathrm{sd}(\mathrm{SI}_{\mathrm{noise}})} \tag{5.1}$$

Where sd designates the standard deviation and SI the signal intensity. The mean and standard deviation of the noise were derived from an ROI placed in the extracorporal region on the metabolite maps. The extremities of the FoV were avoided owing to the presence of spiral artefacts. The factor of $\sqrt{2}$ accounts for the narrowed Rayleigh distribution of magnitude noise, with an approximate adjustment for the use of multiple receivers. Equations (5.2) and (5.3), where n is the number of channels and C the correction factor, defines the correction

factor which is 1.426 for eight coils [448]. Metabolite SNRs were calculated from the ^{13}C-pyruvate, lactate, alanine and summed carbon ($\overset{Def}{=\!=}$ $\mathrm{SI_{Pyruvat}} + \mathrm{SI_{Lactate}}$) maps.

$$C = \frac{1}{\sqrt{2n - \beta n^2}} \tag{5.2}$$

$$where: \beta = \sqrt{\frac{\pi}{2}} \frac{2n-1!!}{2^{n-1}(n-1)!} \tag{5.3}$$

Only voxels with a summed carbon SNR $\geqslant$5 were included in the analysis: a mask of the intersection of the manual segmentations with the voxels with a summed carbon SNR $\geqslant$5 was constructed. The exclusion of voxels with a low SNR removed data points that did not allow a reliable estimation of the metabolic rate constants.

The exchange of the HP-^{13}C-label between pyruvate and lactate was quantified through the apparent reaction rate constant k_{PL} using a two-site exchange model with a frequency-domain approach and linear least-squares fitting. The ^{13}C-lactate signal loss through the back-conversion of lactate-to-pyruvate (k_{LP}) and the spin-lattice relaxation of the hyperpolarised signal ($\mathrm{T_1}$) were accounted for in the combined effective relaxation term $\mathrm{T_{1eff}}$ [449]. The estimation of the apparent reaction rate constant k_{PL} requires pyruvate and lactate signals with a high temporal resolution and a sufficient signal-to-noise ratio. In contrast, time-summed metabolite signals have an increased signal-to-noise ratio at the expense of the temporal resolution. The ratio of the time-summed lactate to pyruvate signal combines the pyruvate delivery, conversion to lactate, back conversion of lactate to pyruvate and longitudinal relaxation $\mathrm{T_1}$ in a single parameter and has no direct physiological interpretation, in contrast to the k_{PL}. Normalised lactate-to-pyruvate ratio maps (Lac/Pyr) (eq. (5.4)) were obtained after by dividing the lactate signal by the sum of the pyruvate and lactate signal, which prevented the noise in the lactate signal from inflating Lac/Pyr in regions with no pyruvate signal.

$$\mathrm{Lac/Pyr_{normalised}} = \frac{^{13}\mathrm{C\ Lactate}}{^{13}\mathrm{C\ Pyruvate} + ^{13}\mathrm{C\ Lactate}} \tag{5.4}$$

Proton image processing

The data pipeline for ^{1}H-MRI processing allowed completing the initial data analysis within 24 hours. The IVIM-type DWI data underwent affine motion correction applied across all b-values with custom and proprietary software implemented in MATLAB and developed

by GE (GE Healthcare, Chicago IL, USA). Subsequently, in-house code calculated D_0, f_p and ADC parameter maps. Voxelwise D_0 and f_p derived from a non-linear fit of the multi-b-value diffusion information. A standard desktop workstation (Intel Core i7-6700 3.4 GHz (quad-core), 16 GB RAM DDR4 2133 MHz) completed the motion correction in 30 minutes. Meanwhile, the non-linear fitting of the diffusion signal would take between four and six hours depending on the number of slices on a standard workstation and was, therefore, run on a high-performance workstation (Intel Xeon Gold 6148 2.40GHz (40-core), 791 GB RAM) which took thirty minutes.

The multi-TE data underwent voxelwise fitting to a mono-exponential decay using the non-linear Levenberg–Marquardt algorithm and a log-linear approximation to compute the R_2^* values. The relevant in-house developed MATLAB software ran directly on the MRI scanner.

GE provided custom and proprietary MATLAB code for the spatial registration of DCE-MRI contrast phases and multi-flip angle T_1-mapping data. Affine registration across the 44 DCE phases took approximately two hours on a standard workstation while the registration of the T_1-mapping lasted for 20 minutes. The registration employed the Advanced Normalization Tools within the Insight Segmentation and Registration Toolkit (ANTs/ITK) (Version 2.1.0 for Windows). Subsequently, the DCE data were processed in MIStar (Version 3.2.63, Apollo Medical Imaging Technology, Melbourne, Australia) on a dedicated workstation (Intel i7-3970X 3.5 GHz (6-core), 32 GB RAM DDR3 1600 MHz). Initially, B_1 corrected T_1-maps were calculated from the multi-flip angle T_1-mapping data. Subsequently, these T_1-maps were used to derive gadolinium concentrations from the DCE-MRI. Gadolinium concentrations were fitted using the extended Tofts model [335]. The arterial input function derived from the Fritz-Hansen blood-sampled gadolinium concentration curve [446] appended by the Weinmann wash-out. A measured AIF placed in the renal artery on the side of the renal tumour determined the delay of the model AIF. The haematocrit was fixed at 0.4, and the tissue T_1 derived from the T_1 map. The only constraint set on the model was that the extracellular extravascular space v_e should be $<100\%$. The DCE-MRI processing required approximately two hours to complete and maps of the transfer constant K^{trans}, extracellular extravascular volume fraction v_e, vascular plasma volume fraction v_p, rate constant k_{ep}, area under the initial 90 seconds of the contrast concentration curve ($iAUC_{90}$), goodness of fit and T_1 were exported for analysis.

The T_2w sequence and parametric maps were re-sampled to the geometry of the T_1w LAVA-Flex sequence using linear interpolation and registered using ITK-SNAP (version 3.6, University of Pennsylvania) [390] for manual, rigid registration. Three ROIs encompassing

the renal tumour, the tumour including the adjacent normal-appearing kidney as well as Gerota's fascia were drawn in OsiriX on the T_1w LAVA-Flex sequence with reference to the other series and the T_2w HyperCUBE sequence in particular. The T_1w LAVA-Flex sequence was selected for segmentation due to its superior spatial resolution which was advantageous for the subsequent production of the custom 3D-printed tumour mould. Additionally, cylindrical ROIs equivalent to the ones described above were drawn around the location of tissue samples. ROIs were propagated to the other series, voxelwise intensity information exported in the JSON format as described above and masks excluding voxels with poor goodness of fit (R^2 <0.75) were constructed to analyse the DCE-MRI parameters.

5.2.5 Image-guided tissue sampling through 3D-printed patient-specific tumour moulds and tissue processing

Tumour moulds were created as described previously (chapter 4, page 103 and [450]). Briefly, the outlines of the tumour, normal kidney and Gerota's fascia were re-oriented using in-house MATLAB code and the mould generated using in-house code implemented in Python, interfacing with Meshlab and OpenSCAD to achieve smoothing of the segmentation for printing, subtraction from a mould template, creation of openings in the mould to facilitate the orientation of the specimen and addition of a knife-guide to ensure the sampling accuracy. PrusaSlicer (Version 2.3.0+, Prusa Research, Prague, Czech Republic) sliced the 3D model output as a STereoLithography (.stl) file and a Prusa i3 MK3S printer produced the mould from polylactic acid (PLA) filament. Furthermore, tumour habitats were calculated as described previously [450].

After the radical nephrectomy, the specimen was transferred to the tumour mould and transported to the pathology suite for sampling. Axial tumour slices were selected to allow taking three to nine tissue samples from each tumour and covering all habitats. Locations of the samples were recorded on habitat maps which were reconstructed to represent the anatomy corresponding to the selected slice. Samples were collected with a 6 mm tissue punch and half of each tissue sample was formalin-fixed and paraffin-embedded (FFPE) while the remaining two quarters were snap frozen individually.

One frozen sample from each biopsy was transferred to the Sanger Institute (Hixton, UK) for whole-exome sequencing and transcriptomic analysis. The other frozen sample was lysed for LC-MS. After preparation of sample extracts from approximately 30 mg of snap-frozen tumour tissue in 25 µL/mg wet tissue weight of 50% methanol/30% acetonitrile/20% water extraction solution using Precellys tubes prefilled with ceramic beads and a Precellys 24

homogenizer (Stretton Scientific), samples were randomized to avoid bias due to machine drifts and processed blindly. LC-MS analysis of sample extracts was performed on a Q Exactive mass spectrometer coupled to a Dionex UltiMate 3000 Rapid Separation liquid chromatography system (ThermoFisher Scientific, Milton Keynes UK). The setup of the LC-MS system and the analysis of acquired spectra have been described previously [451]. Tissue pyruvate and lactate levels were measured to validate MR-based metabolic measurements biologically.

Sections of 4 μm thickness were cut from the FFPE tissue blocks and stained with hematoxylin and eosin. A uro-pathologist reviewed all samples, excluding samples with $<75\%$ viable tumour tissue from further analysis. Tumours were graded with the WHO/ISUP grading system for ccRCC [59] and, additionally, the pathologist pseudo-graded each tissue sample. Immunohistochemical (IHC) staining of enzymes involved in the cellular uptake of pyruvate and secretion of lactate was achieved with Leica's Polymer Refine Detection System (DS9800) in combination with their Bond automated system (Leica Biosystems Newcastle Ltd). Sections were incubated in Tris EDTA for 20 min for antigen retrieval. Endogenous peroxidase activity was quenched using 3-4% (v/v) hydrogen peroxide. The sections were incubated with the primary antibodies HPA003324 directed against MCT1 and HPA021451 against MCT4 (Atlas Antibodies, Bromma, Sweden) diluted 1:500. The sections were then incubated in Anti-rabbit Poly-HRP-IgG polymer (<25 μg/mL) containing 10% (v/v) animal serum in Tris-buffered saline (0.09% ProClin 950, Sigma-Aldrich, Gillingham, United Kingdom). The complex was visualized using 66 mmol/L 3,3'-Diaminobenzidine tetrahydrochloride hydrate in a stabilizer solution and $\leqslant 0.1\%$ (v/v) hydrogen peroxide. Leica DAB Enhancer was added to enhance staining. Cell nuclei were counter-stained with $<0.1\%$ hematoxylin.

Tissue slices adjacent to the ones used for IHC were stained with RNAscope. RNAscope is an in-situ hybridisation technique for the detection of mRNA on formalin-fixed or fresh frozen tissue. It is available as a chromogenic and fluorescent assay. While the chromogenic assay is suitable for up to two targets at a time, up to four RNA targets plus the nuclei can be stained in a multiplex fluorescent assay. The assay uses two paired Z oligonucleotide probes which hybridise to their immediately adjacent targets on the mRNA strand, which affords a high specificity. Each probe contains an 18 – 25 base pair (bp) region which is complementary to the target, a spacer and a 14 bp tail region. The probes are available in different channels, defined by the tail region, specific to the pre-amplifier which only binds paired probes. Twenty channel-specific amplifiers then bind each pre-amplifier and each amplifier has twenty binding sites for fluorescent labels. Finally, an assay contains

twenty probe pairs described above which bind to the same mRNA molecule within a 1kb region, which leads to a theoretical maximum signal amplification of 8000-fold [452]. Tissue was processed on the Bond automated system (Leica Biosystems Newcastle Ltd.). Tissue slices were deparaffinised and dehydrated. After heat-induced antigen retrieval and protease treatment, slides were incubated with the target probe, pre-amplifier, amplifier and the fluorophore in direct succession. Target probes for LDHA, LDHB and PDHA1 and multiplex reagent kits were acquired from ACD Bio-Techne (Minneapolis, MN, USA). The Opal™570, 690 and 620 fluorophores (Akoya Biosciences, Marlborough, MA, USA) bound the probe complex on LDHA, LDHB and PDHA1 mRNA, respectively. Positive controls staining for housekeeping genes and negative controls staining the bacterial gene *dapB* to assess the background signal were run in parallel [453].

All image and tissue analysis was carried out blinded to the clinical outcome.

5.2.6 IHC analysis

Slides were scanned with a Leica Aperio AT2 at 20x magnification and immunohistochemical staining was quantified with the HALO software (Version: 2.2.1870.15, Indica Labs, Albuquerque NM, USA) and its module for area quantification (Version: 2.1.11). Tissue areas with weak, moderate and strong staining were identified automatically and divided by the total tissue area to obtain the relative areas of weakly, moderately, strongly and total MCT1 and MCT4 positive tissue. The boundaries for optical densities representing weak, moderate and strong staining were identified automatically as >0.2164, >0.3274 and >0.4938, respectively. Biopsy staining was summarised at the patient level to account for spatially heterogeneous MCT1/MCT4 expression through the median percentage of stained tissue (weakly, moderately and strongly). Additionally, the cell density in each biopsy sample was quantified using HALO by dividing the number of cells counted in each biopsy by the tissue area and the median used as the patient level summary metric.

RNAscope slides were digitised on a Zeiss AxioScan Z1 at 40x and images processed in HALO using the FISH and area quantification modules. First, the number of cells in each slide was derived from the count of the DAPI stained nuclei. Secondly, the number of mRNA transcripts of each of the genes, one from each of the three channels, was quantified and the number of RNA copies per cell calculated [454].

5.2.7 TCGA-KIRC extension dataset

The Cancer Genome Atlas ccRCC dataset (TCGA-KIRC) served to validate some of the findings from the prospective MISSION renal study in a large, external database and extend the analyses to involve long-term outcome information. Normalised RNA-Seq by Expectation Maximisation (RSEM) gene expression data were retrieved from the Broad Institute, Cambridge, USA [443] and z-transformed before analysis. The Kruskal-Wallis test identified differentially expressed genes among ccRCC tumour grades and pairwise Wilcoxon rank-sum tests with the Benjamini-Hochberg correction for multiple testing identified pairs of tumour grades with different gene expression. Univariate and multivariate Cox regression models established clinical and genetic variables predictive of survival and Kaplan-Meier curves visualised univariate associations with survival at the pre-specified cut-off of the 85^{th} percentile. The 85^{th} percentile was chosen to represent the distinction between patients with a high expression and patients with a moderate to low expression of a gene while maintaining sufficient statistics for analysis. Furthermore, it is consistent with previous analyses of TCGA RNA data in breast cancer in our lab [455].

5.2.8 Statistical analysis

Statistical analyses were performed in R (version 4.0.0, R foundation for statistical computing, Vienna, Austria) and the dplyr (version 1.0.2), ggplot2 (3.3.3), GGally (2.1.0) survival (3.2-7), survminer (0.4.8) packages used for data processing, figure design and survival analysis. The mean and standard deviation summarised (approximatively) normally distributed continuous variables. Where appropriate, the Shapiro-Wilk test assessed the normal distribution of a sample. Median and range represented skewed data. Unpaired samples were compared with the Wilcoxon rank-sum test, while the Wilcoxon signed-rank test was used for paired samples. Finally, the Spearman correlation was employed for skewed data. In contrast, (approximately) normally distributed data were analysed with Student's t-test and Pearson's correlation. Biopsy-level data were analysed with linear mixed models and two-way ordinal analysis of variance (ANOVA) to account for dependencies of samples obtained from the same patient.

5.3 Results

5.3.1 Participants

Thirty-five patients fulfilled the inclusion criteria for the MISSION study and were approached for participation. Seventeen patients agreed to participate. Sixteen underwent multiparametric ^{1}H-MRI, and nine underwent HP-^{13}C-MRI successfully. One participant was imaged with an alternative ^{1}H-MRI protocol. The ^{13}C-pyruvate-MRI failed in two participants owing to the non-availability of fluid paths and in three patients to the non-conformance of the QC parameters (table 5.1). Furthermore, one participant did not meet the size requirement of the ^{13}C-tuned clamshell coil, and one participant with rapid disease progression was no longer eligible on the day of the planned imaging. Finally, during one of the injections, the SPINlab suffered a technical failure (fig. 5.2). Table 5.2 summarises participant characteristics. HP-^{13}C-pyruvate and multiparametric ^{1}H imaging took place shortly before the nephrectomy in participants with malignant tumours (median 12 days pre-surgery; interquartile range: 10 – 21 days). The diagnosis of malignant tumours was confirmed histologically on the nephrectomy specimen, where six of eight participants showed clear cell histology. One participant, suspected to harbour an RCC pre-operatively, was diagnosed with a sarcoma not otherwise specified (NOS) that was re-classified as a dedifferentiated renal liposarcoma after detection of an MDM2 amplification during the review by a reference centre. A second tumour, diagnosed as a renal tumour on imaging, appearing confluent with the normal kidney and encapsulating the renal vessels, was diagnosed as a phaeochromocytoma on postoperative histology. Finally, one participant underwent a multi-core renal mass biopsy where they were diagnosed with a renal oncocytoma. This participant allowed comparing the metabolic activity of benign and malignant renal tumours. Tissue samples for research were available for all six participants with a ccRCC and from the dedifferentiated sarcoma. The clinical diagnosis required all tissue samples from the participant with the phaeochromocytoma, and no tissue was available for research.

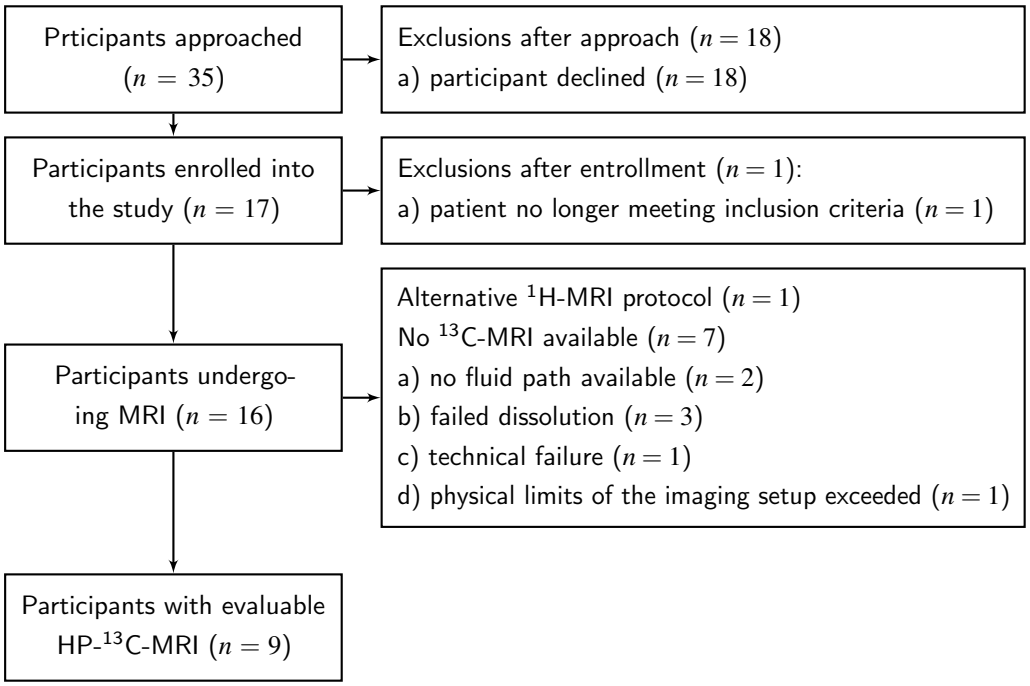


Fig. 5.2 MISSION: CONSORT diagram

5.3.2 Hyperpolarised ^{13}C-MRI

A visual representation of the HP-^{13}C-MRI results including the ^{13}C-pyruvate and lactate signal summed across the acquisition time course, the apparent reaction rate constant and the Lac/Pyr ratio can be found in fig. 5.3. The spectra represent time-summed information of the entire slice containing the largest tumour cross-section. These spectra are not suitable to derive metabolic information from the tumour alone but serve to illustrate which metabolites were detected and where their resonance frequencies fall within the ^{13}C spectrum. Table D.2 on page 279 summarises the results of the automated quality control module of the SPINlab hyperpolariser.

Similar [1-^{13}C]pyruvate SNRs were measured in ccRCC and the normal kidney (median: 26.7 (16.8 – 61.3) and 30.1 (14.8 – 64.7, respectively). Similarly, the [1-^{13}C]lactate SNR of ccRCC and normal kidney (median: 5.7 (1.9 – 9.6) and 3.4 (2.0 – 9.3; $P = 0.51$), respectively) were comparable. The dedifferentiated liposarcoma and pheochromocytoma showed a pyruvate (31.9 and 34.0) and lactate SNR (10.5 and 6.0) in a similar range to the ccRCCs. Meanwhile, the oncocytoma showed pyruvate (12.3) and lactate SNRs (1.0) which were lower than in any other tumour.

Table 5.2 Characteristics of MISSION patients with evaluable ^{13}C-MRI

Participant feature	
Number of patients	9
Age at consent [years]	median: 59.5
	SD: 8.7
Sex (m / f)	8 / 1
Histology	6 clear cell RCC
	1 Phaeochromocytoma
	1 Dedifferentiated liposarcoma
	1 Oncocytoma
Stage at diagnosis (RCC only)	2 pT1b pNX cM0
	3 pT3a pN0 cM0
	1 pT3b pNX cM1
WHO/ISUP grade at surgery	1 Grade 2
(RCC only)	2 Grade 3
	3 Grade 4
Metastatic deposits	1 Lung
Patient weight [kg]	median: 90.1
	IQR: 13.5
Time from imaging to	median: 12
surgery [days]	IQR: 11
Laterality	5 left / 4 right
Plasma glucose [mmol/L]	median: 5.0
	IQR: 0.3

IQR: interquartile range

The metabolic activity measured with the median Lac/Pyr and k_{PL} was comparable between ccRCCs and the normal kidney. The Lac/Pyr reached 0.13 (range 0.10 – 0.42) in primary renal cancers and 0.14 (0.12 – 0.22; $P = 0.40$) in the normal kidney, respectively. The liposarcoma, phaeochromocytoma and oncocytoma showed Lac/Pyr of 0.35, 0.17 and 0.14, corresponding to the range observed for ccRCC. The median k_{PL} did not differentiate ccRCC 0.0065 (range 0.0024 - 0.0151) and the normal kidney either 0.0043 (0.0028-0.0076). However, the liposarcoma showed a higher apparent metabolic rate constant than any other tumour (0.0152) and the oncocytoma a lower value than any other tumour (0.0022). The k_{PL} of the phaeochromocytoma was similar to ccRCC (0.0086). The Lac/Pyr and k_{PL} were strongly correlated (r = 0.88, P = 0.03, fig. 5.4a). Meanwhile, the lactate SNR was largely independent from the pyruvate signal, suggesting that the perfusion and pyruvate delivery were not the dominant determinant of the ^{13}C-lactate signal (fig. 5.4b).

The WHO/ISUP tumour grade in ccRCC ($n = 6$) was correlated psoitively with the k_{PL} ($\rho = 0.92$, $P = 0.009$; fig. 5.4). At the individual biopsy level ($n = 44$), the k_{PL} remained

correlated with the tumour grade even when taking into account that imaging measurements from a single patient are not independent ($P = 0.03$). Allowing for patient dependence in the statistical analysis is crucial since measurements were clustered according to patients. Lac/Pyr significantly increased in higher-grade tumour samples ($P = 0.005$). However, at the tumour level, the relationship failed to reach statistical significance.

None of the metabolic parameters, pyruvate SNR, lactate SNR, k_{PL}, and Lac/Pyr was associated with tumour volume.

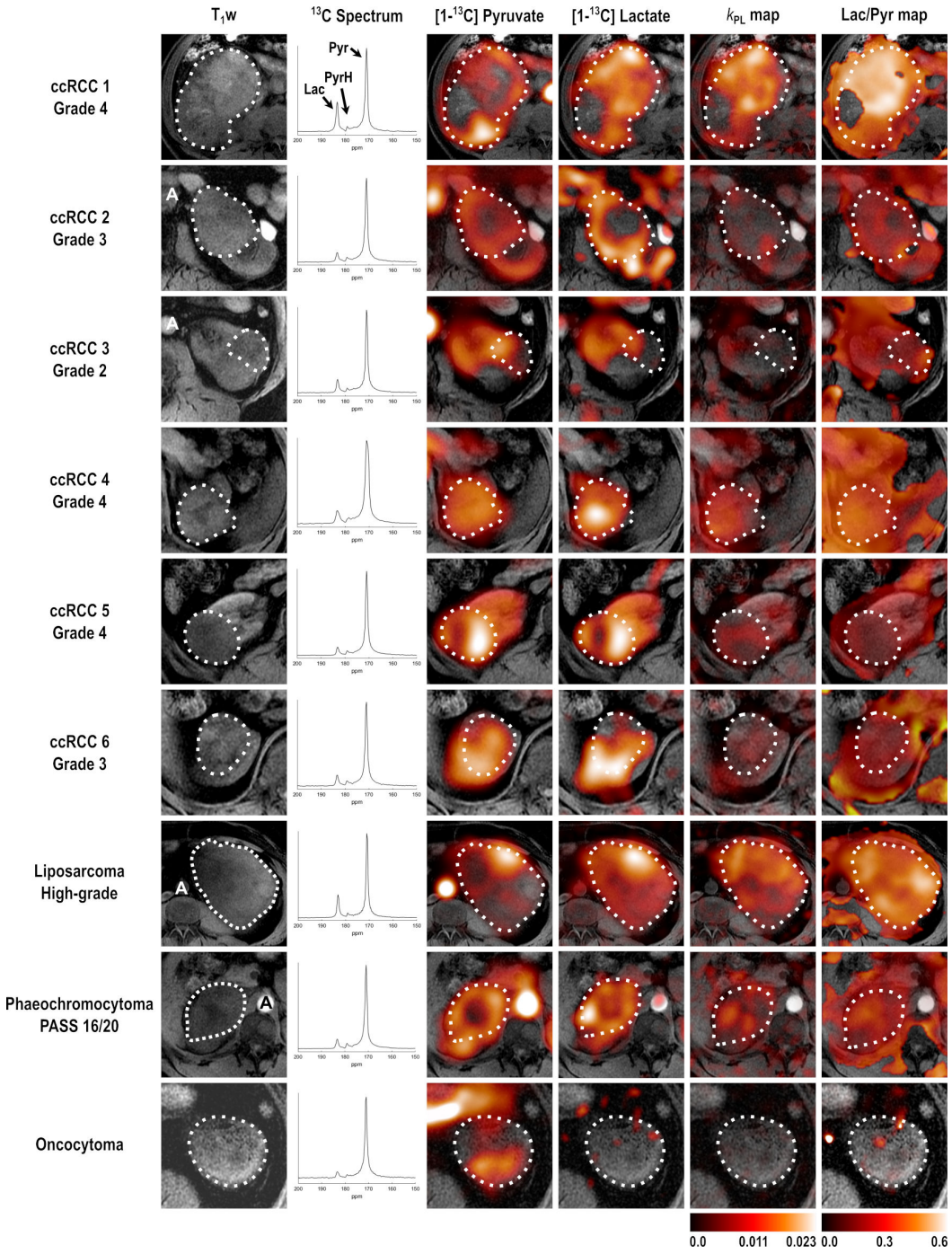
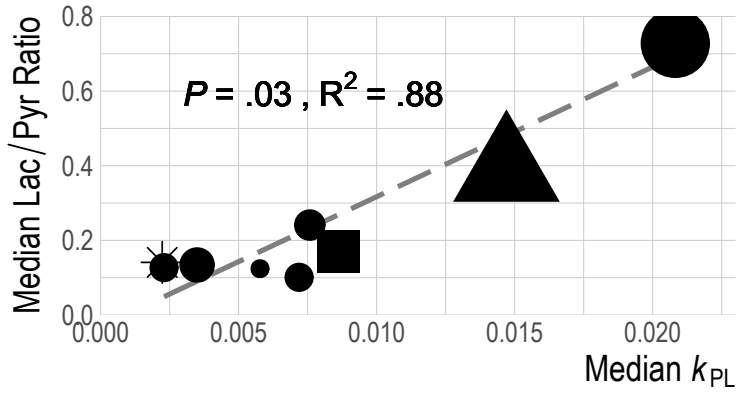
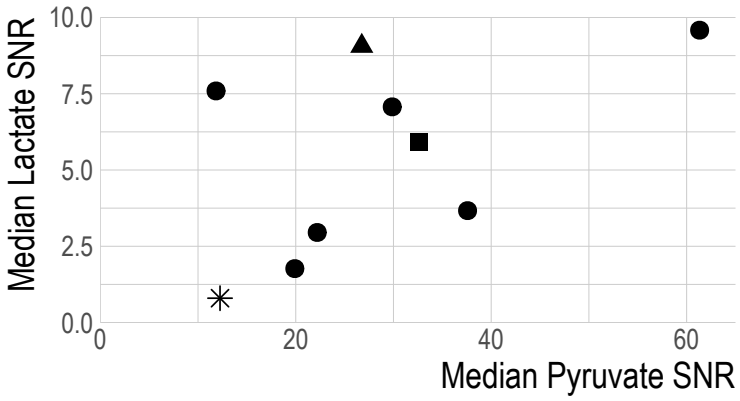


Fig. 5.3 Summary figure presenting the spatial heterogeneity of the ^{13}C-pyruvate and lactate signal as well as the apparent reaction rate constant k_{PL} and lac/pyr of all tumours in the MISSION cohort. Axial T_1w images for anatomical reference. The ^{13}C spectra represent the slice encompassing the largest tumour cross-section in a non-localised manner. Signals were summed across all timepoints. The tumour segmentation is displayed in white. A: Aorta, Lac: Lactate, Pyr: Pyruvate, PyrH: Pyruvate hydrate. PASS scores (Thompson, The American Journal of Surgical Pathology: May 2002) of $\leqslant$4 are associated with an increased risk of clinically aggressive behaviour in pheochromocytoma.

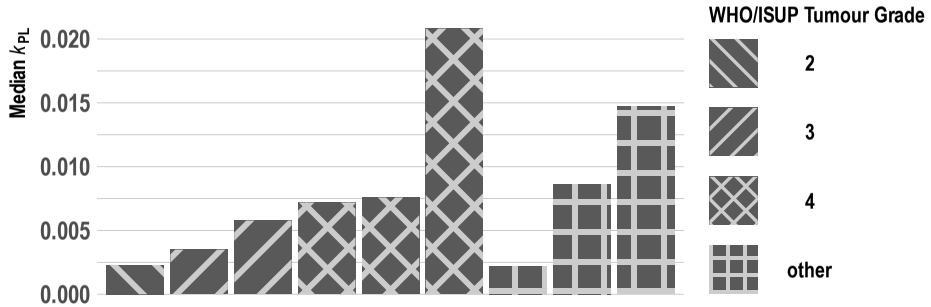


(a)

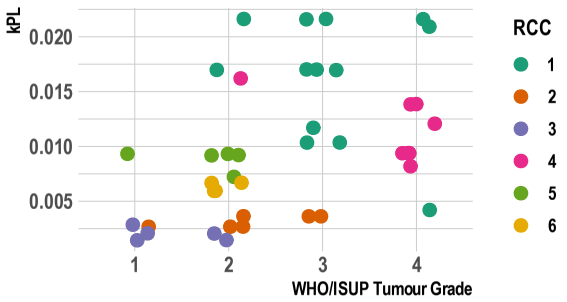


(b)

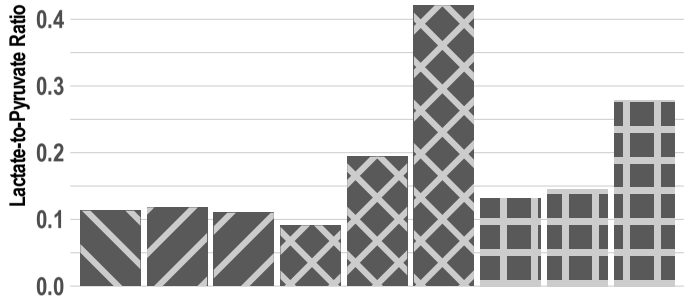
Fig. 5.4 (a) Correlation of the lactate-to-pyruvate ratio and k_{PL}. The correlation was independent of the tumour volume which ranged from 53 to 1350 cm^3 (53 – 908 cm^3 for ccRCC). (b) No correlation between the lactate and pyruvate SNR was observed for tumours in the MISSION cohort ($R^2 = 0.17$, $P = 0.14$), indicating that pyruvate delivery is not the rate limiting step in lactate formation. ccRCC: circle, liposarcoma: triangle, phaeochromocytoma: square, oncocytoma: asterisk



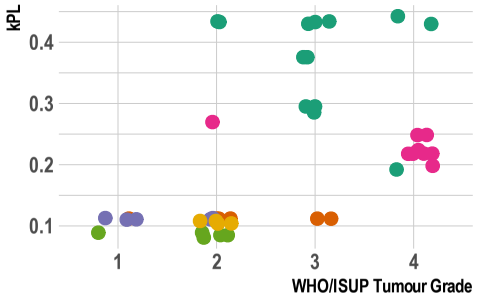
(a)



(b)



(c)



(d)

Fig. 5.5 Correlation of the k_{PL} with tumour grade at the tumour (a) and biopsy (b) level. Corresponding information for the lactate-to-pyruvate ratio (c and d). Non-ccRCC tumours are listed as 'other' as WHO/ISUP grades are not defined for these tumours.

5.3.3 ^{1}H MRI

Owing to the relatively small number of patients and high dimensionality of the dataset, the following analysis should be understood as exploratory. It helps define more specific research questions for future studies and serve as a baseline for power calculations. Figure 5.6 shows an example of the ^{1}H-MRI image contrasts acquired in the MISSION study. The perfusion fraction, representing the relative capillary volume, was negatively associated with the WHO/ISUP tumour grade (fig. 5.7a). However, this relationship was not maintained at the biopsy level (fig. 5.7b).

None of the DCE-MRI parameters (K^{trans}, v_e, v_p or the delay in contrast arrival), the diffusivity D_0 or R_2^* were associated with the tumour grade. Furthermore, the tumour grade was unrelated to the tumour volume (fig. 5.7d). However, the Lac/Pyr negatively correlated with D_0 ($P = 0.003$ and $r = -0.98$; figs. 5.7e and 5.7f). Meanwhile, the k_{PL} was not correlated with any of the ^{1}H-MRI parameters and showed only a trend for correlation with the perfusion fraction ($P = 0.30$, $r = -0.58$; fig. 5.7c).

5.3.4 Immunohistochemistry

The expression of MCT1, quantified through the positive pixel count on immunohistochemistry, stained ccRCC and normal renal tissue similarly strongly (4.0±6.0% and 6.2±3.4%, mean ± S.D.; $P = 0.17$, $n = 6$). However, ccRCC overexpressed MCT4 compared to normal renal tissue (37±10.5% vs 10.2±6.0%, mean ± S.D.; $P = 0.009$). MCT1 showed a WHO/ISUP grade-dependent expression with the weakest staining in grade 2 tumours and the strongest staining in grade 4 tumours (fig. 5.8). MCT1 expression in the dedifferentiated liposarcoma was comparable to ccRCC (5.4% vs 6.2%) while the MCT4 expression was lower (16% vs 37%).

The pyruvate-to-lactate conversion in ccRCC measured non-invasively correlated strongly with the expression of MCT1. The median k_{PL} ($r = 0.89$, $P = 0.016$) and Lac/Pyr ($r = 0.94$, $P = 0.005$) both correlated positively with the membrane transporter expression (fig. 5.9). When comparing individual biopsies and accounting for the dependence of ^{13}C-MRI measurements on patients, k_{PL} and MCT1 retained a borderline significant correlation ($P = 0.052$). Neither the k_{PL} nor the lactate to pyruvate ratio were correlated with the cell density of the tumour biopsies ($P = 0.95$ and 0.4, respectively). Additionally, the diffusivity D_0 was not correlated with the cell density in tumour biopsies ($P = 0.50$) in this small number of patients. In contrast, a meta-analysis of studies assessing the correlation between cell density and the ADC found a significant negative correlation ($\rho = -0.53$) [456]. However, only two studies

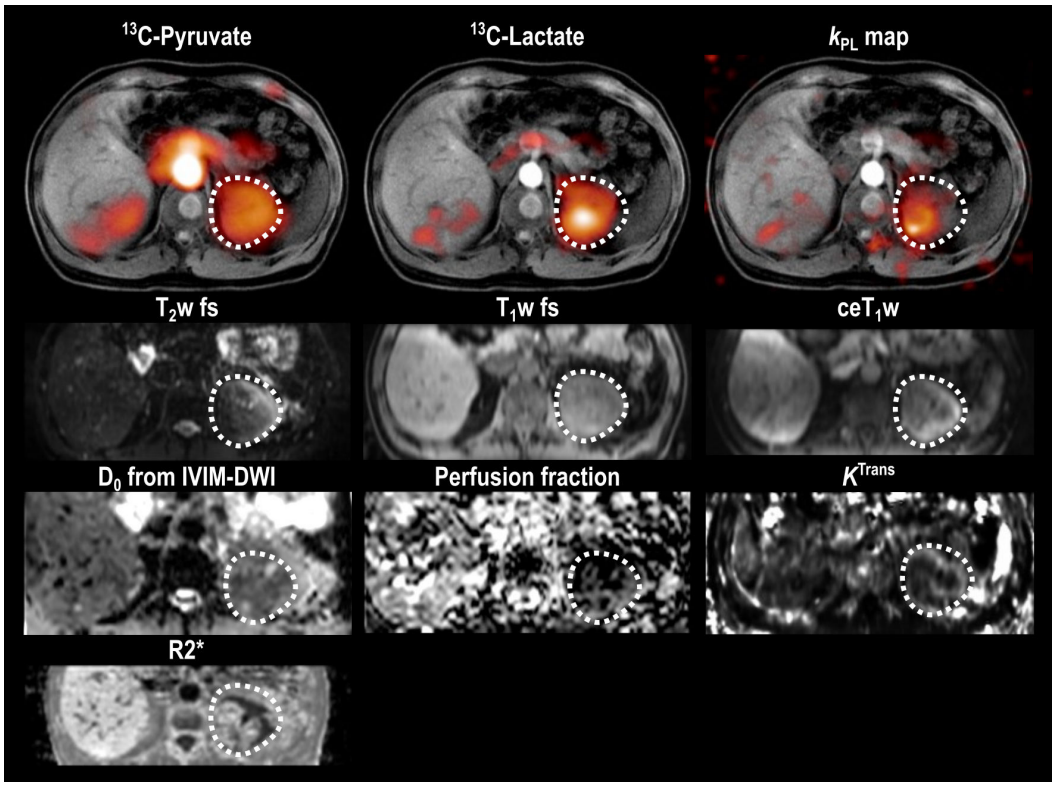


Fig. 5.6 Summary of the ^{13}C and ^{1}H MRI sequences acquired in a participant with a grade 4 ccRCC. T_2w with fat saturation, T_1w with fat saturation, contrast-enhanced T_1w, summed ^{13}C-pyruvate and ^{13}C-lactate images display arbitrary units. The k_{PL} is displayed in the range between 0 and 0.02. The D0 map in the range of 0.0 to 3.0×10^{-3}mm/s^2. Perfusion fraction ranging from 0 – 60%, K^{trans} from 0.0 – 0.8 min^{-1} and R_2^* from 0 to 100 Hz. The ^{1}H-MRI was acquired coronally and reformatted in the axial plane to allow direct comparison with the ^{13}C-MRI. The tumour was segmented in white.

with 63 patients in total addressed this question. The MCT4 expression was independent from the metabolic activity.

5.3.5 RNAscope

Despite multiple attempts by our histopathological core lab, no antibody against LDHA has demonstrated sufficient specificity against control cell lines to be used routinely. Since immunohistochemistry for this enzyme was not available, we developed a spatial transcriptomic approach to address this problem. RNAscope staining provided insight into gene rather than protein expression levels. Importantly, RNA expression is not a direct surrogate of enzymatic activity which was measured in the HP-^{13}C-MRI. Post-transcriptional regulation, post-translational modifications and protein degradation will modulate the concentration and

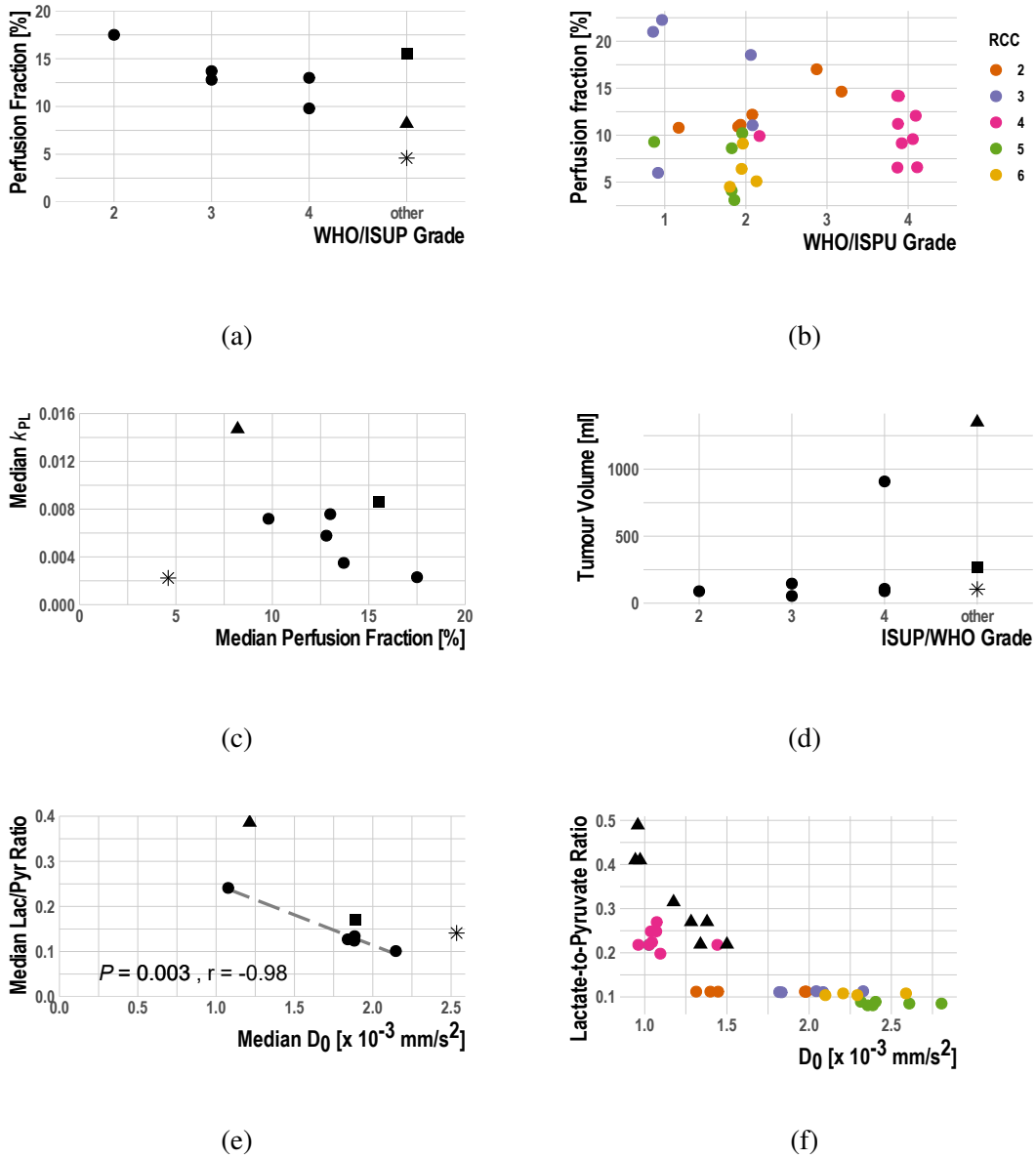


Fig. 5.7 The tumour grade was negatively associated with the perfusion fraction at the tumour (a) but not at the single biopsy level (b). The k_{PL} was only weakly related with the perfusion fraction (c). The tumour grade was independent of the tumour volume (d). Finally, the Lac/Pyr ratio correlated negatively with the diffusivity D_0 (e and f). ccRCC: dot, liposarcoma: triangle, phaeochromocytoma: square, oncocytoma: asterisk

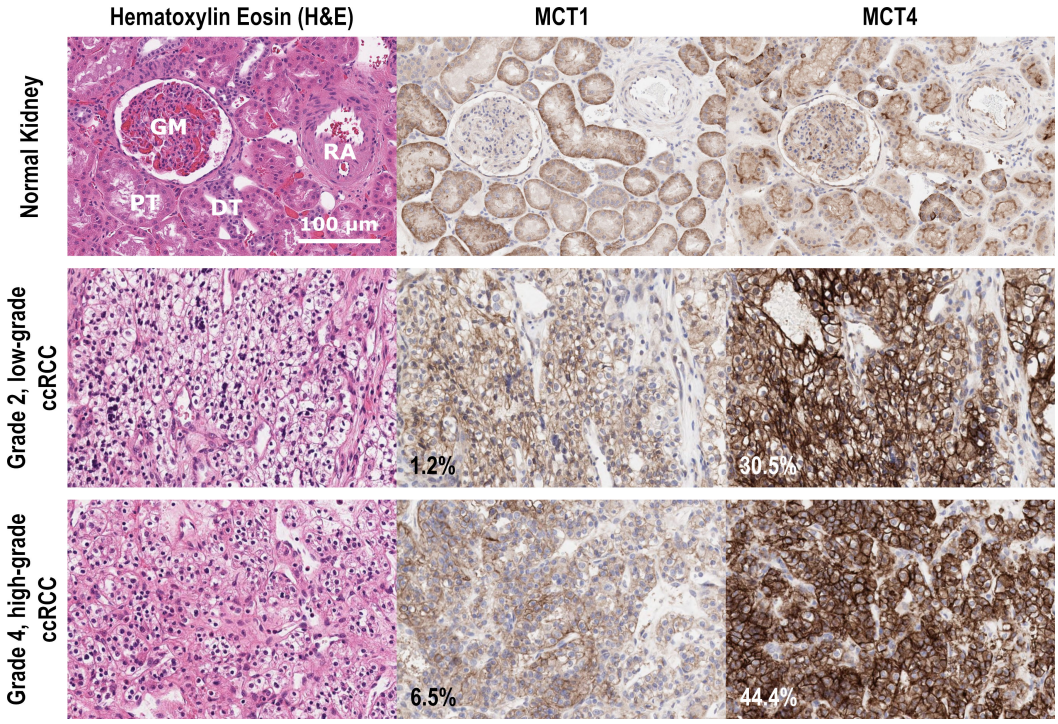


Fig. 5.8 Expression of MCT1 and MCT4 in low- and high-grade ccRCC and comparison with the adjacent normal kidney. DT: Distal tubule, GM: Glomerulus, PT: Proximal tubule, RA: Interlobular artery.

activity of proteins independently from RNA expression. Methods such as western blotting could have measured protein concentrations, a more closely related surrogate of metabolic activity than RNA expression if sufficient amounts of tissue had been available. In contrast, RNAscope is able to resolve expression levels by cell type.

One to four biopsies from each ccRCC in this cohort were stained with probes for LDHA, LDHB and PDHA1. PDHA1 is the alpha subunit of the E1 complex which is part of pyruvate dehydrogenase. Positive and negative controls demonstrated little to no background signal and the correct staining of housekeeping genes. Figure 5.10 shows an example of RNAscope images. The histologically and metabolically most aggressive tumour had been archived for two years and undergone oxidative degradation of the RNA targets. Consequently, only minimal amounts of RNA of any of the genes investigated were detectable in any biopsy from this tumour (n = 4). Therefore, these tumour samples were excluded from the following quantitative analysis. However, they are included in the figures as a visual reference.

The median copy number of LDHA mRNA expressed in each cell was 0.94 (IQR: 0.69 – 1.33). Meanwhile, LDHB and PDHA1 showed copy numbers per cell of 12.5 (6.27 – 22.6) and 3.16 (2.86 – 3.78), respectively. The LDHA mRNA copy number was lower in the

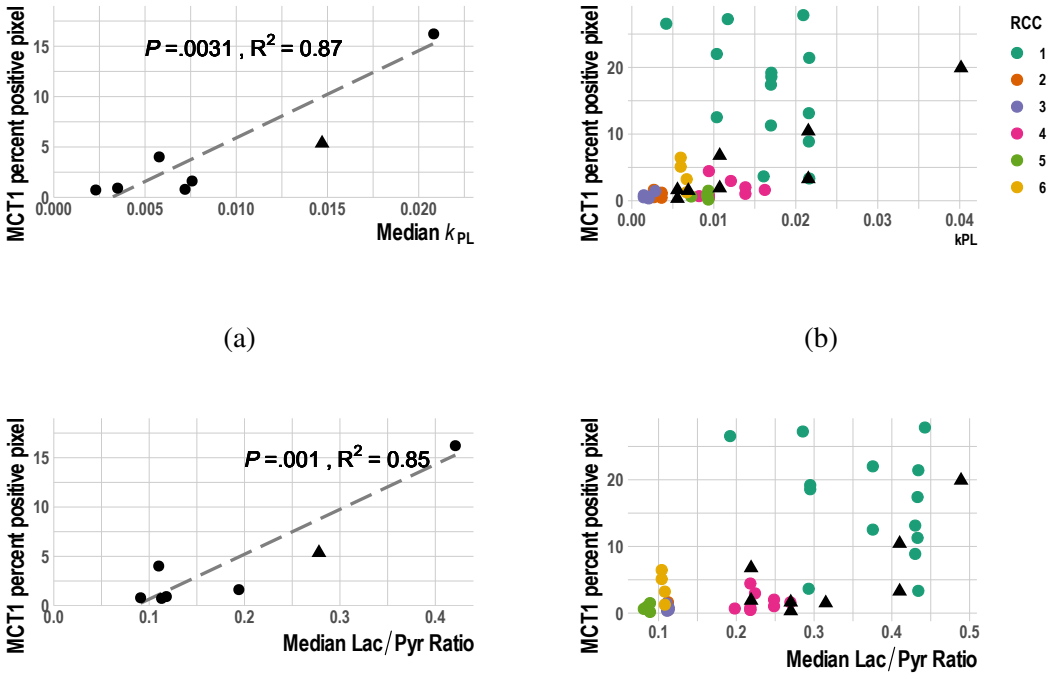


Fig. 5.9 The lactate signal on hyperpolarized ^{13}C-MRI and lactate concentrations from LC-MS in RCC are correlated and spatially heterogeneous (a). Comparison of lactate measurements with LC-MS and MRI in tissue obtained from perfused (grey) and ischaemic (black) tissue. Normal renal tissue (triangles) showed consistently lower lactate levels on imaging and LC-MS compared to the respective tumour samples. Tumour samples show a correlation between increasing ^{13}C-lactate SNR with increasing tissue lactate concentrations. For image-based measurements, the SNR was determined over a circular 3cm ROI centred on the biopsy site, ensuring robust measurements. The correlation between the apparent rate constant k_{PL} and tissue lactate concentrations was not significant(b).

normal renal tissue compared to any of the tumour samples (0.025). In contrast, the normal kidney showed a higher expression of LDHB and PHDA1 than most tumour samples (25.9 and 4.25, respectively).

High-grade tumours showed an increased RNA expression of LHDA compared to low-grade tumours (fig. 5.11a). However, this trend could not be observed at the single biopsy level. In contrast, LDHB was unrelated to the tumour grade (fig. 5.11b). The k_{PL} trended towards a positive associated with the LDHA and LDHB expression at the tumour and biopsy level (figs. 5.11d, 5.11e, 5.11g and 5.11h).

No correlation was found between the expression of LDHA, LDHB, the lactate-to-pyruvate ratio and the lactate SNR. The expression of PDHA1 was independent of the tumour grade and imaging parameters.

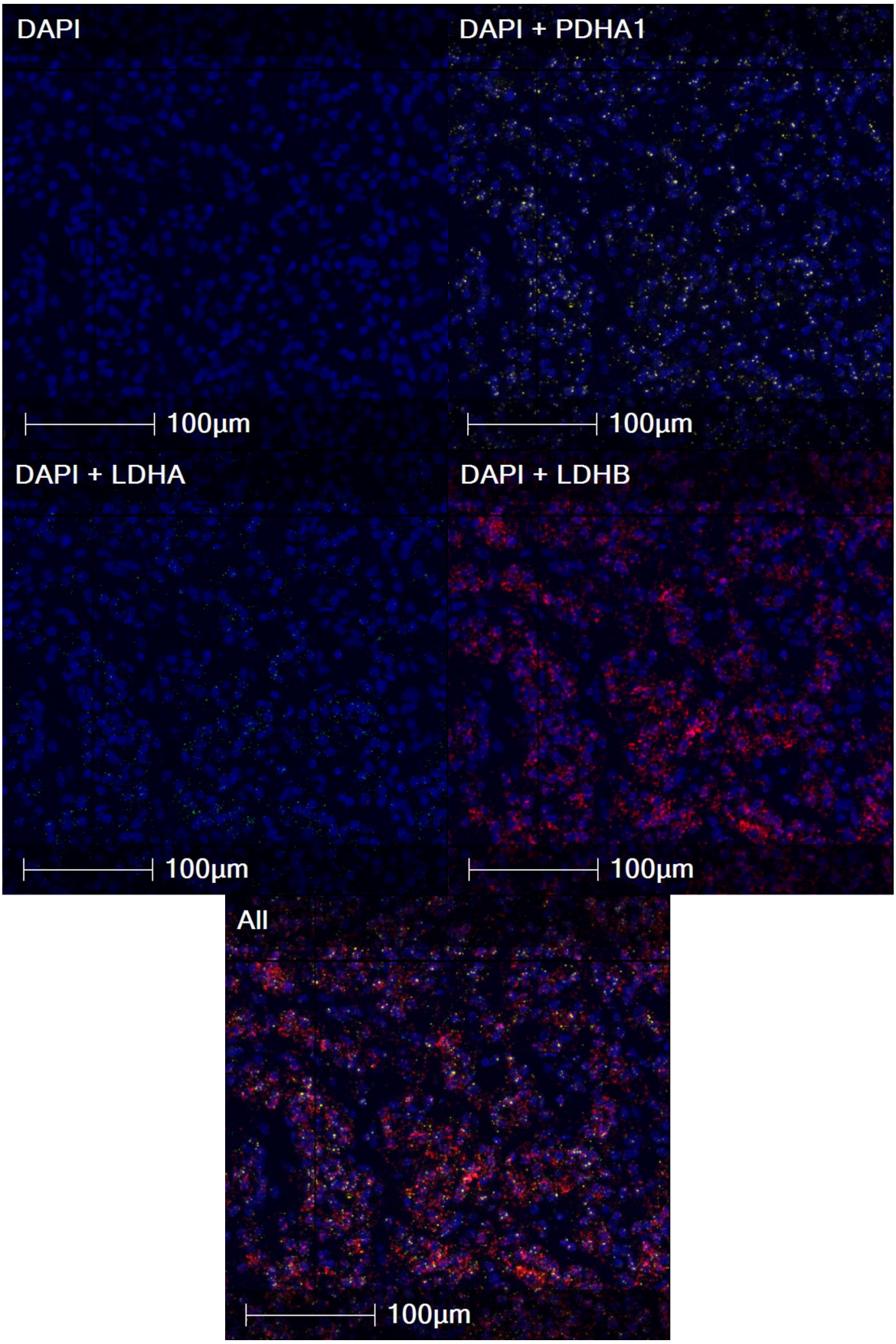


Fig. 5.10 Example of RNAscope images. Cellular nuclei were stained with DAPI. Probes for LDHA, LDHB and PDHA1 were fluorescent at distinct wavelengths and are pseudocoloured in green (LDHA), yellow (PDHA1) and red (LDHB).

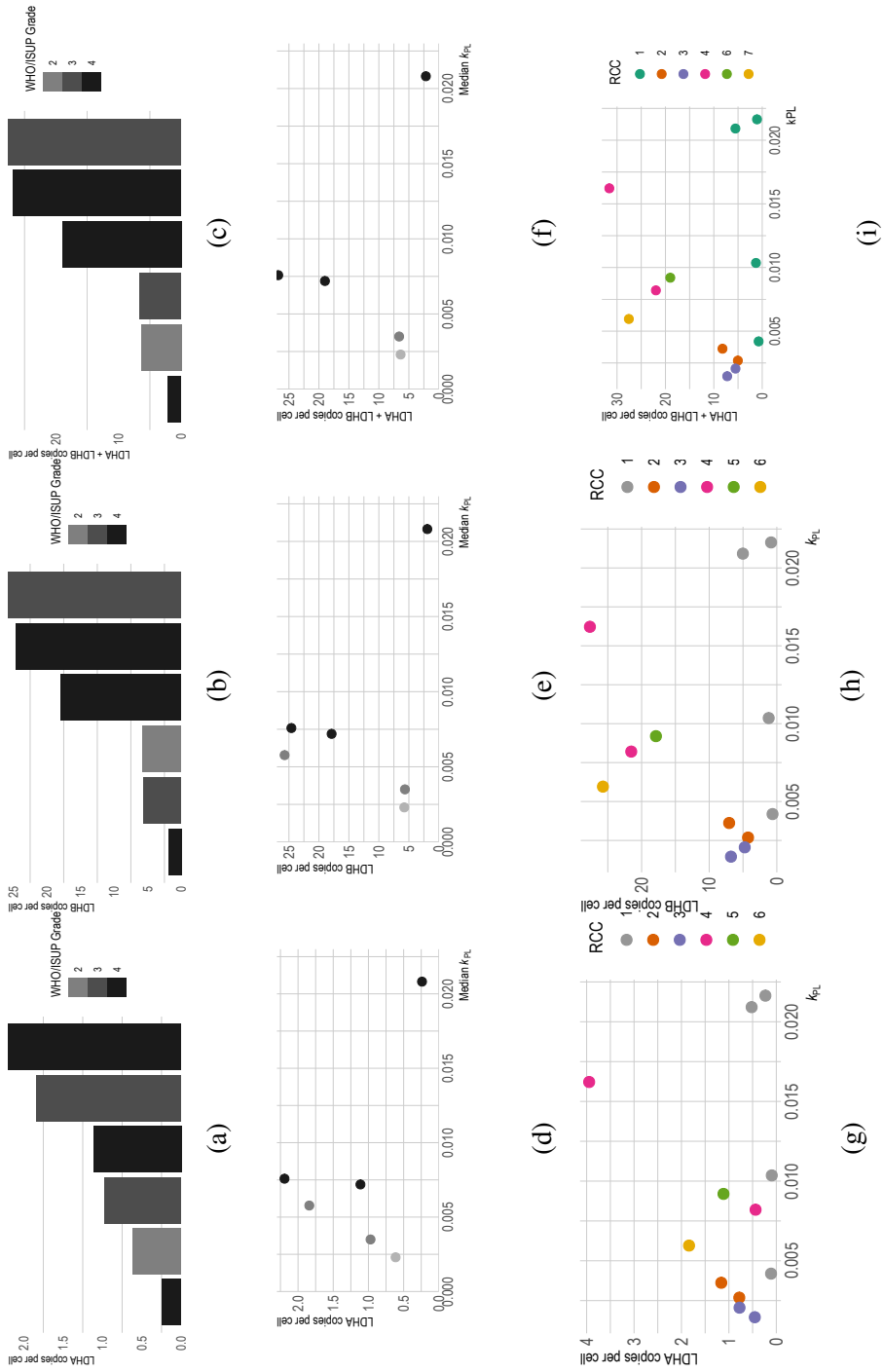


Fig. 5.11 Association of LDHA and LDHB and total LDHA RNA copy number per cell from RNAscope with WHO/ISUP tumour grade (a, b and c). Correlation of LDHA and LDHB and total LDH copy number per cell with the k_{PL} from hyperpolarised ^{13}C-pyruvate MRI at the tumour (d, e and f) and biopsy (g, h and i) level. The tissue samples which were excluded owing to oxidative degradation of the RNA targets are displayed in grey.

5.3.6 TCGA-KIRC RNA expression analysis

The TCGA-KIRC database contains RNA expression data and clinical parameters for 526 patients with ccRCC. The initial analysis of MCT1 and MCT4 expression relative to tumour grade revealed a grade-dependent expression of the two proteins ($P < 0.001$ and 0.008, respectively, figs. 5.12a and 5.12b). Pairwise comparisons demonstrated significant overexpression of MCT1 and MCT 4 in grade 4 compared to grade 2 and 3 tumours, even after correction for multiple comparisons. The number of grade 1 tumours was most likely too low (13 patients) to reach statistical significance. The expression of LDHA, a third molecule involved in the generation of the ^{13}C-lactate signal, was non-significantly related to tumour grade ($P = 0.06$, fig. 5.12c). Finally, the variability in each of the three genes explained only a small proportion of the variability in others (R^2_{adj} MCT1/MCT4: 0.03, MCT1/LDHA: 0.16, MCT4/LDHA: 0.22) despite them being highly correlated ($P < 0.001$, fig. 5.12d).

Explorative survival analysis showed an overall and progression-free survival advantage for patients with an MCT1 expression below the 85th percentile. Similarly, MCT4 overexpression trended towards an association with inferior overall survival and LDHA overexpression was associated with shorter progression-free survival (fig. 5.13). A univariate Cox regression confirmed these findings. A multivariate Cox regression including clinical variables (age, sex, lymph node and distant metastasis, tumour size and grade) showed that only MCT1 expression remained independently predictive of overall survival (table 5.3). The expression of none of the genes was predictive of recurrence-free survival (table 5.4). The MCT1 expression was the second strongest predictor of overall survival after the presence of distant metastases. Similarly, the presence of distant metastases was the strongest predictor of recurrence-free survival followed by lymph node metastases.

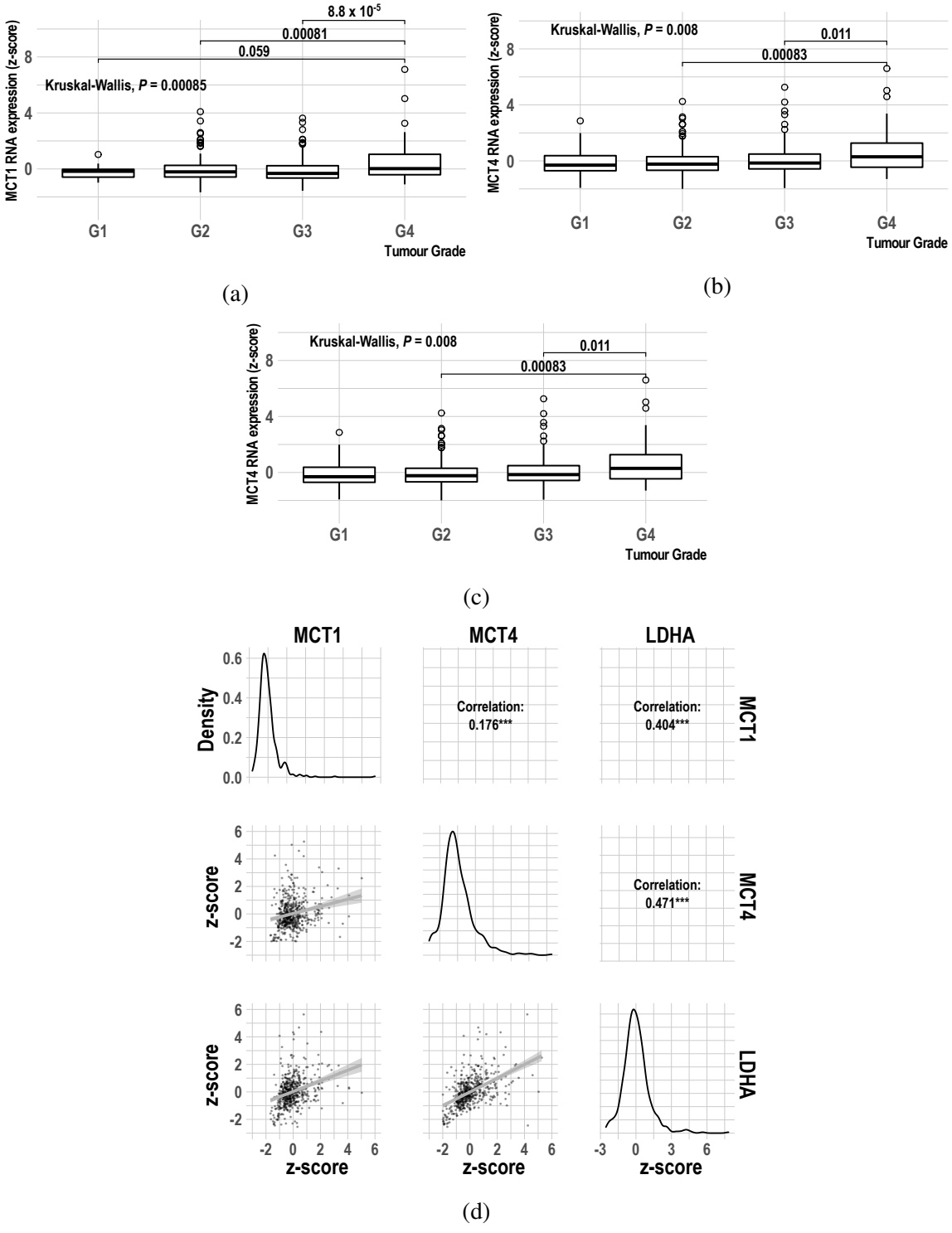


Fig. 5.12 Association of the expression levels of MCT1 (a), MCT4 (b) and LDHA (c) with the WHO/ISUP tumour grade in the TCGA-KRIC dataset. Cross-correlation of the MCT1/4 and LDHA gene expression. ***: P < 0.001. Grey lines and sector show the regression line and the 95%-confidence interval.

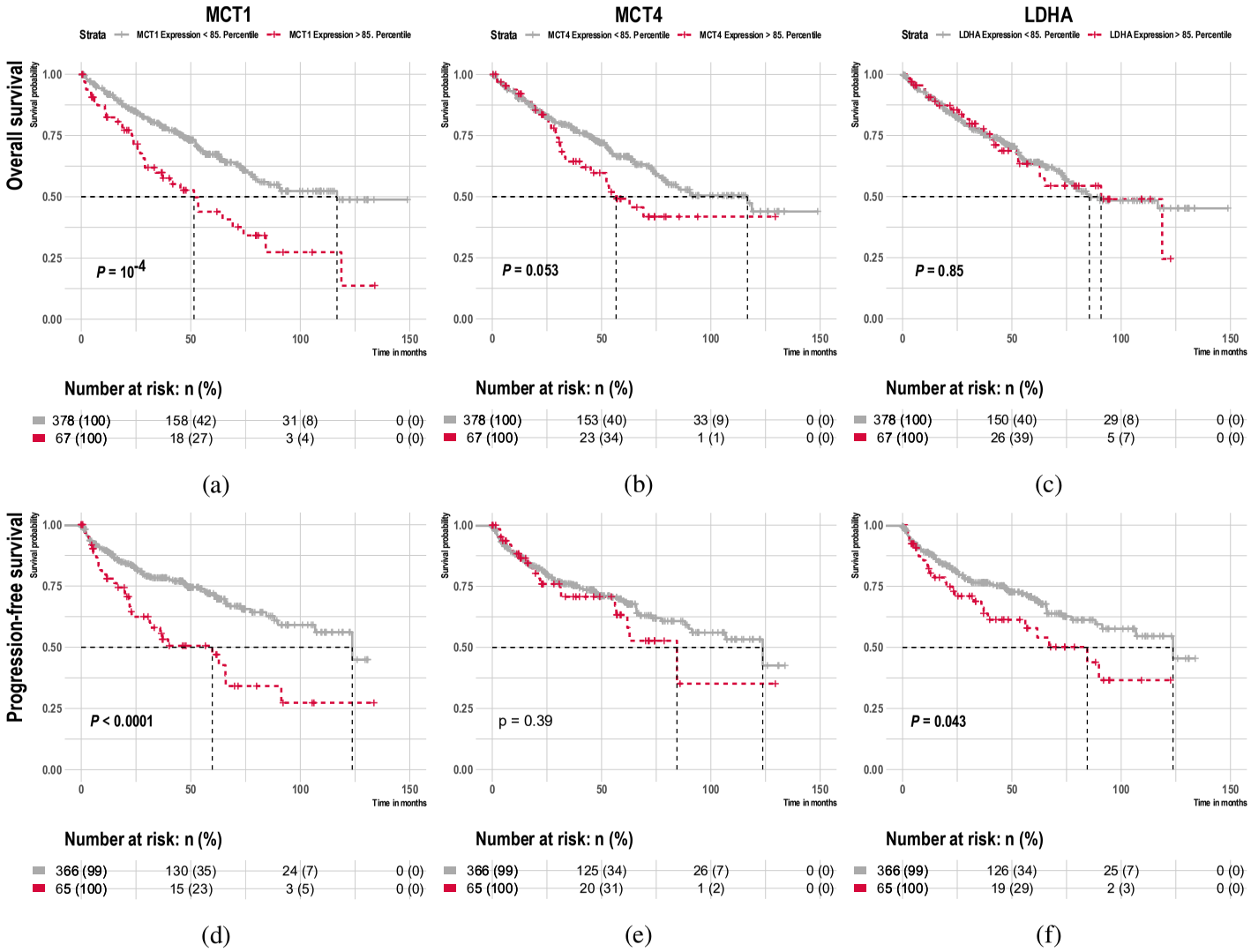


Fig. 5.13 Kaplan-Meier Curves for the association of overall and progression-free survival with MCT1 (a and d), MCT4 (b and e) and LDHA (c and f) expression in the TCGA KIRC dataset.

Table 5.3 Multivariate logistic regression for the prediction of overall survival

Covariates	Coefficient	Standard error	P value	HR	95% Confidence Interval	
					Lower	Upper
Age	**0.034**	**0.011**	**0.002**	**1.034**	**1.012**	**1.058**
Female Sex	-0.363	0.262	0.17	0.696	0.416	1.162
LN	0.457	0.441	0.30	1.579	0.665	3.748
Metastasis	**1.157**	**0.281**	**<0.001**	**3.179**	**1.834**	**5.510**
Size	0.175	0.192	0.36	1.191	0.818	1.734
Grade						
Grade 1	-1.547	3141	0.99	0.001	0	∞
Grade 2	-0.525	0.369	0.15	0.591	0.287	1.218
Grade 3	-0.225	0.335	0.50	0.799	0.414	1.540
MCT1	**0.369**	**0.105**	**0.010**	**1.309**	**1.065**	**1.609**
MCT4	0.070	0.119	0.55	1.073	0.850	1.355
LDHA	-0.0269	0.131	0.082	0.797	0.617	1.029

HR: Hazard ratio, LN: Lymph node metastasis

Table 5.4 Multivariate logistic regression for the prediction of progression-free survival

Covariates	Coefficient	Standard error	P value	HR	95% Confidence Interval	
					Lower	Upper
Age	0.020	0.012	0.10	1.020	0.996	1.044
Female Sex	-0.134	0.280	0.63	0.875	0.505	1.514
Lymph nodes	**1.115**	**0.452**	**0.013**	**3.048**	**1.256**	**7.399**
Metastasis	**1.760**	**0.285**	**<0.001**	**5.812**	**3.323**	**10.166**
Size	0.140	0.217	0.52	1.150	0.752	1.759
Grade						
Grade 1	0.759	1.105	0.49	2.135	0.245	18.606
Grade 2	**-0.809**	**0.390**	**0.038**	**0.445**	**0.208**	**0.957**
Grade 3	-0.326	0.366	0.37	0.722	0.352	1.481
MCT1	0.164	0.105	0.11	1.178	0.960	1.45
MCT4	0.024	0.132	0.85	1.024	0.790	1.327
LDHA	-0.0124	0.113	0.91	0.988	0.792	1.232

HR: Hazard ratio, LN: Lymph node metastasis

5.3.7 Correlation of ^{13}C-MRI and tissue metabolite concentrations

LC-MS measurements of the tissue lactate concentration were available for participant 1 in the MISSION renal cohort while the samples from the remaining participants were still being processed at the time of the writing of this thesis.

The parametric maps demonstrated at least two metabolically distinct tumour regions (fig. 5.3). The two avidly enhancing tumour regions on contrast-enhanced MRI, located antero-medially and postero-medially, demonstrated a high pyruvate signal but were distinguished based on differences in metabolism (k_{PL}: 0.030 s^{-1} and 0.013 s^{-1}; table 5.5). A region of haemorrhagic necrosis showed minimal pyruvate and lactate signals.

Table 5.5 Spectroscopic measurements and kinetic modelling of the pyruvate to lactate conversion of the tumour and normal kidney. (Perfused tissue samples)

Parameter	Normal kidney	Tumour (median)	Highly metabolically active tumour	Lowly metabolically active tumour
Lac/Pyr	0.19	0.42	0.46	0.25
k_{PL} [s^{-1}]	0.007	0.021	0.030	0.013
Pyr SNR	14.1	11.8	13.1	26.0
Lac SNR	2.01	7.59	11.1	7.3
ADC [$\times 10^3$ mm^2/s]	2.54	1.42	1.47	1.38
Relative enhancement in arterial phase	+230%	+97%	+70%	+109%
Pyr (LCMS) [ng/mg wet tissue weight]	8.10	3.88	4.92	2.85
Lac (LCMS) [ng/mg wet tissue weight]	642	1270	1410	1130

ADC: apparent diffusion coefficient, IQR: Interquartile range, Lac: lactate, LC-MS: liquid chromatography-mass spectrometry, Pyr: pyruvate, SNR: signal-to-noise ratio. Image measurements represent the median of the region of interest.

Measurements of the apparent diffusion co-efficient showed restricted water diffusion in the tumour (ADC: 1.43×10^{-3} mm^2/s) compared to normal kidney (2.54×10^{-3} mm^2/s). The metabolically distinct tumour habitats showed very similar ADCs (1.47×10^{-3} mm^2/s vs 1.38×10^{-3} mm^2/s) and relative gadolinium contrast-enhancement on the portal venous (109% vs 94%) and delayed phases (99% vs 101%; percentage increase in signal compared to pre-enhancement). However, the most avidly enhancing areas on arterial phase gadolinium-enhanced imaging also showed the highest ^{13}C-pyruvate signal on ^{13}C-MRI. Conversely, ^{13}C-lactate signal was higher in regions of low arterial gadolinium-enhancement, compared to regions with avid enhancement (relative enhancement of 52% and 109%).

Tissue pyruvate and lactate levels were assessed using liquid chromatography mass-spectrometry to validate the hyperpolarised ^{13}C-MRI measurements. Lactate concentrations in the perfused tumour were on average 200% higher compared to the normal kidney (range: 180-220%). The lactate SNR corresponded with the tissue lactate concentrations measured

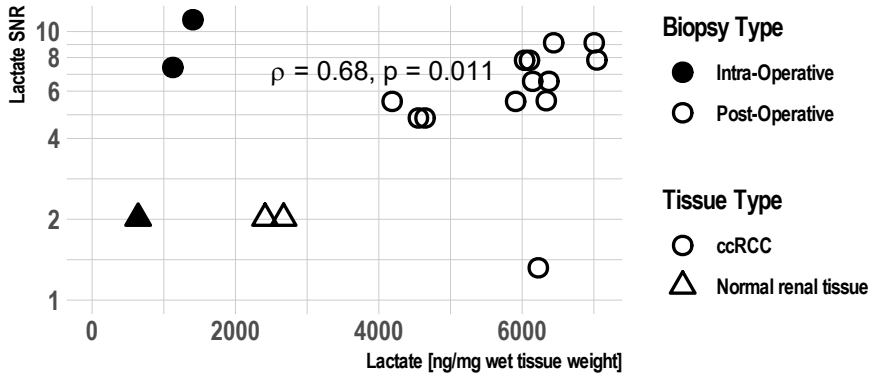
using LC-MS (fig. 5.14). Biopsies obtained after nephrectomy showed signs of ischaemia and energy depletion, indicated by increased lactate and succinate concentrations and depletion of ATP, alpha-ketoglutarate and citrate [457] (table D.3, page 282). The average lactate concentration in the normal renal tissue and tumour increased by 97% and 130%, respectively, when comparing ischaemic to perfused samples. However, the relative metabolic differences between samples persisted, revealing increased lactate in the tumour compared to normal kidney and heterogeneity in lactate concentrations across the tumour (fig. 5.14a). Spearman's correlation revealed a strong, positive correlation between lactate signal intensities and lactate concentrations ($\rho = 0.73$, $P = 0.003$). No significant association between the metabolic rate constant k_{PL} and the tissue lactate concentration was found in this patient (fig. 5.14b).

5.3.8 Repeatability and comparison of IDEAL-spiral and spectral-spatial pulse sequences

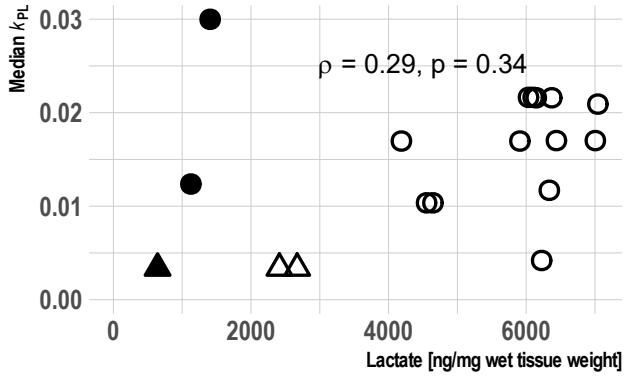
Spectral-spatial pulse sequences (SpSp) may utilise the polarisation signal more efficiently. Exciting a much narrower range of resonance frequencies than the IDEAL spiral sequence, it excites an individual metabolite with each pulse. This allows higher flip-angles for lower abundance metabolites while preserving the hyperpolarised state of more abundant pyruvate. The participant with the 6.4×5.7×6.1 cm oncocytoma diagnosed by multi-core needle biopsy was imaged twice with hyperpolarised ^{13}C-pyruvate MRI 10 days apart, once using the spectral-spatial and once the IDEAL spiral pulse sequence. The quality control parameters are summarised in table 5.6.

While the ^{13}C-pyruvate signal-to-noise ratio of the oncocytoma was comparable in both acquisitions with a median of 12.3 in the IDEAL spiral and 10.3 in SpSp, the lactate SNR was lower at 0.8 using the IDEAL spiral pulse sequence compared to 4.9 with the SpSp sequence. The k_{PL} was measured three times higher at 0.0022 in the IDEAL spiral compared to 0.0008 in the SpSp acquisition. The lactate-to-pyruvate ratio was measured at 0.13 and 0.10.

The contralateral normal kidney showed a higher pyruvate SNR of 40.3 with the IDEAL spiral sequence compared to 13.5 with SpSp. However, the lactate SNR remained lower at 2.6 compared to 8.6, respectively. The k_{PL} reached 0.0024 and 0.0016. Meanwhile, the lactate-to-pyruvate ratio was 0.07 and 0.12. No alanine was measurable in the normal kidney with either method. However, an alanine signal was measurable in the erector spinae muscle with the SpSp sequence while the IDEAL spiral sequence could not measure any metabolite signal. The median SNR of pyruvate, lactate and alanine were 3.6, 9.8 and 10.6, respectively.



(a)



(b)

Fig. 5.14 The lactate signal on hyperpolarized ^{13}C-MRI and lactate concentrations from LC-MS in RCC are correlated and spatially heterogeneous (a). Comparison of lactate measurements with LC-MS and MRI in tissue obtained from perfused (grey) and ischaemic (black) tissue. Normal renal tissue (triangles) showed consistently lower lactate levels on imaging and LC-MS compared to the respective tumour samples. Tumour samples show a correlation between increasing ^{13}C-lactate signal-to-noise ratio (SNR) with increasing tissue lactate concentrations. For image-based measurements, the SNR was determined over a circular 3 cm ROI centred on the biopsy site, ensuring robust measurements. The correlation between the apparent rate constant k_{PL} and tissue lactate concentrations was not significant (b).

The resulting metabolic rate constants and metabolite ratios were: k_{PL}: 0.0044, Lac/Pyr: 0.27, k_{PA}: 0.0034, Ala/Pyr: 0.24 (fig. 5.15).

Table 5.6 Quality control parameters for the comparison of the IDEAL spiral and spectral-spatial sequence

Quality control parameter	SpSp	IDEAL spiral
Polarisation [%]	21.3	10.5
pH	7.9	7.7
EPA concentration [µmol/L]	0.1	0.3
Temperature [°C]	34.1	33.3
Pyruvate concentration [mmol/L]	265	264
Injected volume [mL]	36.8	36.8
Time dissolution – injection [s]	63	56

EPA: Electron paramagnetic agent, SpSp: spectral-spatial.

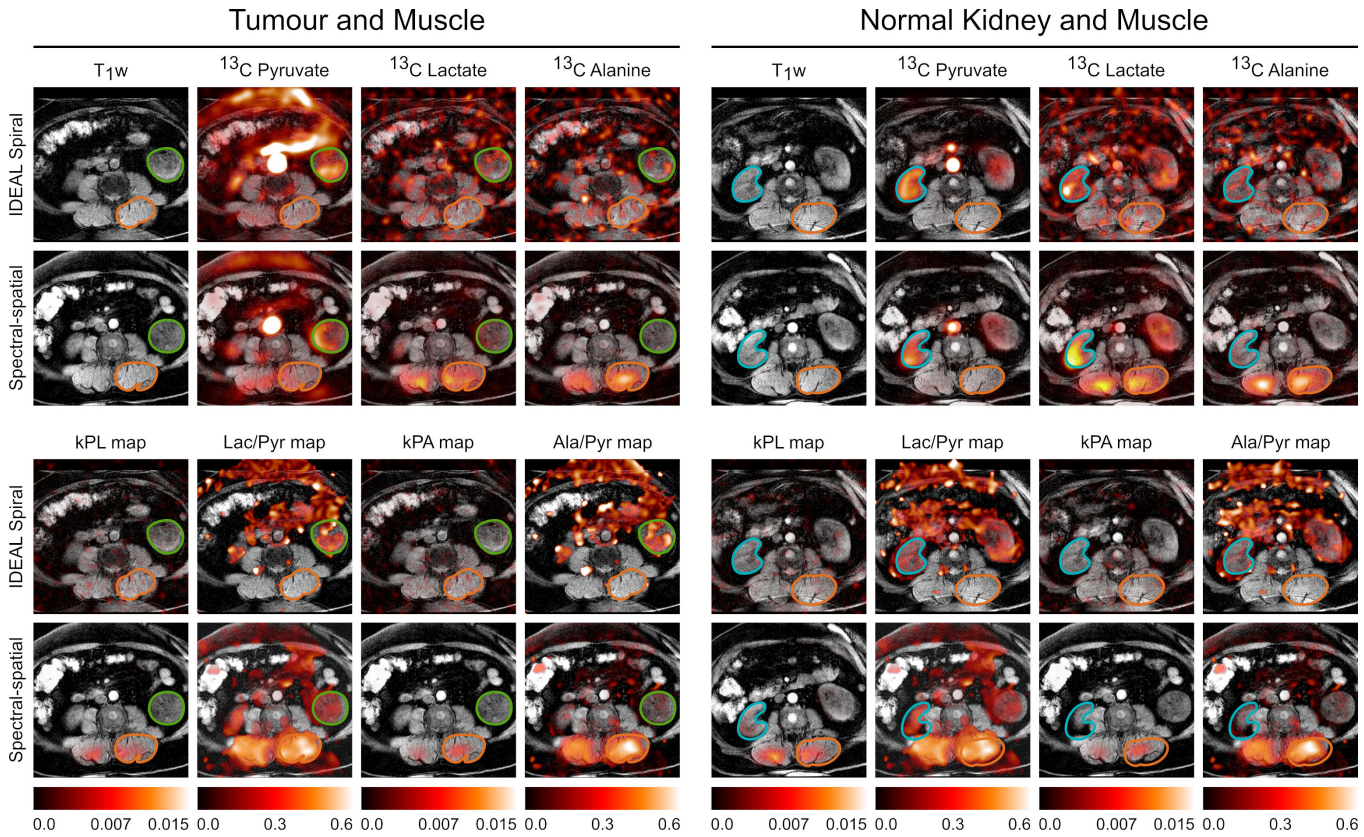


Fig. 5.15 Patient scanned with the IDEAL spiral and spectral-spatial pulse sequence. The renal tumour, contralateral normal kidney and erector spinae muscle are outlined in green, blue and orange, respectively. The pyruvate signal (arbitrary units) has been scaled differently between the slice containing the normal kidney and the tumour to facilitate the visual perception of the heterogeneity within a region of interest. In the IDEAL spiral acquisition, the tumour containing slice was multiplied four-fold and in the spectral-spatial acquisition 2.5-fold compared to the corresponding normal kidney slices. All other metabolite signals are scaled identically.

The ^{13}C-pyruvate MRI was repeated after 45 minutes in two patients with grade 4 ccRCC to determine the repeatability of metabolic measurements. Unfortunately, the second fluid path achieved a very low polarisation of 5.9% in one participant, compared to 26.4% during the first dissolution. This resulted in a decrease in the ^{13}C-pyruvate SNR from 61 to 21 in the tumour and 65 to 14 in the normal kidney. Meanwhile, the ^{13}C-lactate SNR was no longer measurable above the noise level. In contrast, polarisation rates were similar in both instances for the second patient (29.7% and 33%). Similar ^{13}C-pyruvate SNRs were measured in the tumour and normal kidney during both scans (29 vs 33 and 31 vs 31, respectively). Corresponding ^{13}C-lactate measurements were also broadly comparable (8.0 vs 5.7 and 3.0 and 3.8, respectively). While only a moderate reduction in the apparent metabolic rate constant k_{PL} was observed in the normal kidney (-14.9%), the tumour k_{PL} reduced substantially (-54%). The Lac/Pyr ratio fell by 5.8% and 37% in the normal kidney and tumour, respectively.

5.4 Discussion

The MISSION renal study provided preliminary evidence that measurement of the tumour pyruvate metabolism using hyperpolarised ^{13}C-MRI provided a non-invasive readout of aggressive tumour histology and MCT1 expression. Increased pyruvate-to-lactate ^{13}C-label exchange quantified with the apparent rate constant k_{PL} and the lactate-to-pyruvate ratio correlated with a higher WHO/ISUP tumour grade. This correlation was observed at the tumour level and confirmed in the analysis of the multi-regional biopsies. Each biopsy was graded individually and associated with corresponding imaging measurements through a personalised 3D-printed tumour mould, thereby increasing the statistical power of the analyses. Nonetheless, these findings require confirmation in appropriately powered studies. Hyperpolarised ^{13}C-pyruvate-MRI has demonstrated its ability to measure tumour aggressiveness previously. Patients with more aggressive grade 3 and triple-negative breast cancer show a higher Lac/Pyr than patients with grade 2 tumours [202]. *In vitro*, *in vivo* and clinical experience with the technique suggests increased lactate labelling in prostate cancers with higher Gleason grade [205, 204, 206]. Initial data in patients with ccRCC also suggest a trend towards higher Lac/Pyr in high-grade tumours [200] which is supported by *in vitro* experiments demonstrating an increased lactate production in more aggressive RCC cell lines [201]. However, MISSION is the first study to demonstrate an association of tumour grade with k_{PL}, a modelled parameter of pyruvate-to-lactate conversion with a direct physiological interpretation. Furthermore, it was the first study to investigate benign

and malignant differential diagnoses of RCC, showing that HP-^{13}C-pyruvate MRI may aid the differentiation of benign and malignant tumours of the kidney. Finally, extensive tissue analyses provided insight into the biological mechanism underpinning the ^{13}C-lactate signal formation in ccRCC.

HP-^{13}C-pyruvate MRI provides complementary information to ^{1}H-MRI. Participants in MISSION underwent complementary, multiparametric morphological and physiological ^{1}H-MRI, which was critical to understand the incremental value of HP-^{13}C-pyruvate MRI and investigate potential confounding factors. First, we showed that the k_{PL} was the best predictor of tumour aggressiveness, followed by the perfusion fraction from IVIM-type DWI. The perfusion fraction and k_{PL} were not significantly correlated at the tumour and biopsy levels, suggesting that the metabolic activity was not determined by differences in tumour microvascular volume. Measuring the perfusion fraction has the advantage of not requiring contrast application. Tumours with a higher diffusion restriction (a lower D_0) showed a higher Lac/Pyr, suggesting that the increased metabolic activity in these tumours may be owing to an increased number of cells present in the tumour rather than differences in cellular metabolism. However, the absolute cell density obtained from the histological specimens failed to substantiate this hypothesis. Tumour volume may influence tumour metabolism as larger tumours are more likely to experience central hypoxia. Hypoxia may drive the increase in the lactate signal and Lac/Pyr with tumour size in breast cancer for example [202]. However, no similar correlation occurred in ccRCC, which may be explained by the two to three orders of magnitude higher tumour volumes. In summary, HP-^{13}C-MRI appears to provide biologically relevant information which is complementary to multiparametric ^{1}H-MRI.

The expression of the monocarboxylate transporter 1 determines the rate of pyruvate to lactate conversion measurable with HP-^{13}C-MRI. MCT1, a membrane transporter primarily involved in the cellular uptake of pyruvate [458] correlates with the metabolic activity in breast and prostate cancer [202, 205]. Rao *et. al.* have suggested MCT1 expression as the rate-limiting step in the conversion of pyruvate to lactate in multiple *in vitro* and *in vivo* cancer models of pancreatic and breast cancer [441]. In these cancer models, the Lac/Pyr was independent of the LDHA expression, even though LDHA expression may correlate with MCT1 expression in some tumours and, therefore, appear to correlate with Lac/Pyr. In the MISSION renal study, both the k_{PL} and Lac/Pyr correlated with the MCT1 expression at the tumour and single biopsy level, providing evidence that MCT1 may be the

rate-limiting step in renal cancer also. The observation of MCT1 overexpression among grade 4 ccRCC in the TCGA-KIRC database supports the generalisability and potential clinical significance of our findings. Furthermore, the correlation of MCT1 expression with overall survival reported previously [459] points towards a potential prognostic significance of HP-^{13}C-pyruvate MRI. Independent from the TCGA dataset, MCT1 overexpression measured on immunohistochemistry was associated with a more aggressive tumour phenotype and adverse survival outcomes in two cohorts of patients with ccRCC [460, 461]. The increasing glycolytic activity of higher-grade ccRCC results in intracellular acidification, leading to MCT1 and MCT4 overexpression. Meanwhile, MCT1/4 mediated H^+/lactate symport supports the stabilisation of the intracellular pH [460]. While MCT4 expression was unrelated to tumour grade in the MISSION renal cohort, stronger MCT4 staining was observed in tumours compared to the normal kidney, which is in agreement with previous findings [201, 462]. Lactate export through MCT4, in particular, is not only required to maintain the glycolytic activity but also favours tumour progression through increased cell motility and invasion [463, 464]. Neither k_{PL} nor Lac/Pyr differentiated between ccRCC and the normal kidney. Simultaneously, the two tissue types showed a comparable MCT1 expression, which is in keeping with Rao *et al.*'s postulation that MCT1 is the rate-limiting step in HP-^{13}C-pyruvate MRI [441]. However, molecular drivers are likely different between tumour types and Rao *et al.* have not investigated RCC. In contrast to these findings, Sriram *et al.* measured the highest ^{13}C-lactate concentrations in tissue slices of benign renal tumours followed by ccRCC and the normal kidney [438] and attributed this to the more rapid lactate export in ccRCC. The experimental design of *in vitro* compared to clinical imaging may explain this discrepancy.

The RNAscope analysis revealed a trend towards a relative overexpression of LDHA mRNA in patients with higher-grade tumours. A non-significant trend towards this correlation was also observed in the TCGA KIRC RNAseq dataset. Furthermore, the LDHA and LDHB expression trended towards a positive correlation with the apparent rate constant of the hyperpolarised ^{13}C-pyruvate to lactate conversion k_{PL}. However, no statistical quantification of this association was possible owing to the limited number of RNAscope samples. In contrast, no correlation was observed between the LDHA expression and the lactate-to-pyruvate ratio which is consistent with previous findings in breast cancer [202]. The tissue samples of one patient showed very low staining which was most likely due to the older age of the tissue slides. Oxidation and hydrolysis contribute to the relatively rapid degradation of targets for RNA in-situ hybridisation techniques like RNAscope. Reduced staining occurred in tissue that was stored as unstained slices as well as paraffin-embedded blocks. The

reduction in staining was significant after one year and greater than one order of magnitude after 3 – 7 years [465]. RNA targets are best preserved if tissue is stored at -20°C.

The pyruvate metabolism of the normal kidney and correlation with HP-^{13}C-MRI. The normal kidney showed notably high pyruvate to lactate conversion on ^{13}C-pyruvate imaging, which was comparable to all but the most metabolically active tumours. The renal pyruvate/lactate metabolism differs between the well-oxygenated cortex and the relatively hypoxic medulla. Proximal renal tubules perform gluconeogenesis as part of the Cori cycle [466]. Meanwhile, the distal tubule is highly glycolytically active [467]. Renal gluconeogenesis was long thought to serve the kidneys' energy requirements. However, it is the second most important contributor to systemic gluconeogenesis with 40% in a fasted state [468]. Lactate is the main substrate for renal gluconeogenesis accounting for 50% of its anabolic activity [469]. The renal glomerulus filtrates lactate freely which is then taken up by sodium-coupled monocarboxylate transporters SMCT1 and SMCT2 through sodium symport along the sodium concentration gradient [470]. Simultaneously, proximal renal tubule cells express MCT1 and MCT2 on their basolateral membrane which are thought to serve either for transport of lactate into the blood or the uptake of pyruvate and lactate from the circulation for gluconeogenesis [468]. Information on the localisation of MCT4 in the human kidney is scarce, however, The Human Protein Atlas (http://www.proteinatlas.org/) describes a medium expression in tubule cells and its RNA data suggests that the highest expression is present in the collecting duct followed by the proximal tubule. The atlas also describes the highest expression of MCT1 in the proximal tubule. No data on immunohistochemical staining for SMCT1 is available, however, RNA gene expression data support its localisation to the proximal tubule. Lactate dehydrogenase exists in two isoforms, LDHA forming the M (muscle) subunits and LDHB forming the H (heart) sub-units of LDH. LDHA shows intermediate immunohistochemical staining in the renal tubule while staining is absent from the glomerulus. RNA expression is highest in the collecting duct. Meanwhile, immunohistochemistry reveals strong staining for LDHB in the tubule while the glomerulus shows intermediate staining [471]. It is highly expressed in the proximal tubule and the thick ascending limb of the loop of Henle [472]. This agrees with the preferential conversion of pyruvate to lactate by LDHA and lactate to pyruvate by LDHB, respectively. The presence of LDHB in the well-oxygenated parts of the tubule supports renal gluconeogenesis. A challenge with the supra-physiological concentrations of pyruvate which are present in HP-^{13}C-MRI may shift the equilibrium of these reactions towards the conversion of pyruvate to lactate which was observed in this study. Unfortunately, the spatial resolution of the

^{13}C-MRI does not allow differentiating the origin of metabolic signals within the kidney. The underexpression of LDHA mRNA and overexpression of LDHB in the normal kidney relative to tumour tissue in our study would suggest that tumours have a higher propensity to convert pyruvate to lactate. Nonetheless, the two structures were metabolically comparable. In light of the comparable expression of MCT1, this agrees with the report by Rao *et al.* that MCT1 is the rate-limiting step in hyperpolarised ^{13}C-lactate formation [441].

Choice of the optimal pulse sequence for HP-^{13}C-MRI. Most data for this study have been acquired with the IDEAL spiral pulse sequence. This sequence has the advantage of being relatively insensitive to the miscalibration of the resonance frequencies as an entire spectrum is acquired and the metabolite frequencies are selected based on the ^{13}C-pyruvate peak. Additionally, the sequence is less sensitive to the miscalibration or spatial variation of flip angles. On the other hand, spectral-spatial pulses promise a more efficient utilisation of the hyperpolarised signal as metabolites can be excited individually and with variable flip angles. However, this comes at the expense of requiring the resonance frequency to be precisely defined (passband width 80 Hz). Spectral-spatial pulse sequences can also be programmed to acquire more abundant metabolites with a higher spatial resolution or metabolites with a very low signal with a more coarse matrix. One patient in the MISSION cohort was imaged with both pulse sequences. Owing to the higher flip angle at the resonance frequency of ^{13}C-lactate, the signal to noise ratio in the tumour could be increased six-fold. The metabolic activity in the tumour was too low for a reasonable comparison of the sequences. However, the k_{PL} and Lac/Pyr varied by 30–40%, indicating that measurements are not directly comparable between the sequences.

Further research will improve the understanding of the biological foundations of HP-^{13}C-MRI signals. Initial results comparing tissue lactate concentrations with the ^{13}C-lactate signal on MRI were promising and showed that the two correlate. This is in agreement with the very limited existing data [199]. Further elucidation of the molecular biology underpinning the ^{13}C-lactate signal formation will be possible as results from LC-MS, DNA and RNA sequencing analyses become available.

Future works will have to answer questions which this study was unable to elucidate. First, despite involving the largest cohort of suspected RCC, larger studies will be required to determine the generalisability of these findings. However, the multi-regional tissue and image analysis using the custom 3D-printed tumour mould is an asset to this study and strengthened its conclusions. Additionally, this was the first study of HP-^{13}C-MRI to attempt

to overcome the metabolic and genetic intra-tumour heterogeneity, which is particularly notable in RCC [2, 382]. The integration of imaging, histological, metabolomic, genomic and transcriptomic data will require more work. Initial data in breast cancer outlined a potential role for HP-^{13}C-pyruvate-MRI in the detection of very early treatment response [203], which the WIRE window-of-opportunity trial in renal cancer will investigate. The definition of the reproducibility and repeatability of the HP-^{13}C-MRI parameters will be critical in establishing thresholds for treatment response. Two patients initially thought to harbour RCC were diagnosed with liposarcoma and phaeochromocytoma postoperatively. While reducing the sample size of ccRCC in our cohort, this offered the first opportunity to study their metabolism with HP-^{13}C-MRI and provided the first evidence that HP-^{13}C-MRI may not aid the differential diagnosis in this case. Finally, patients had to be repositioned between the acquisition of the HP-^{13}C-MRI and the multiparametric ^{1}H-MRI which may have led to a mis-registration between the scans and is a disadvantage compared to studies investigating glucose metabolism through FDG-PET on integrated PET-MRI systems. However, dual-tuned coils which are able to receive the MR-signals from ^{1}H and ^{13}C nuclei will solve this problem in the future.

In conclusion, the k_{PL} derived from HP-^{13}C-MRI provides a non-invasive measure of tumour aggressiveness in ccRCC. Its correlation with the expression of MCT1, the membrane transporter mediating pyruvate uptake, provided evidence of the molecular biological processes underpinning the HP-^{13}C-lactate signal formation. As an independent predictor of survival in patients with ccRCC, it highlights potential prognostic implications of non-invasive metabolic imaging. In the future, HP-^{13}C-MRI may aid treatment response monitoring and personalisation of cancer care. This could include intensifying treatments in small but metabolically aggressive renal tumours or identifying treatment resistance before morphological progression.

Chapter 6

Physiological MRI as a measure of treatment activity in novel RCC treatments

Part of this work has been submitted for publication:
Ursprung S *et al.* The WIRE study a phase II, multi-arm, multi-centre, non-randomised window-of-opportunity clinical trial platform using a Bayesian adaptive design for proof-of-mechanism of novel treatment strategies in operable renal cell cancer.

Acknowledgements:
The study presented in this chapter was conceived by Prof Grant Stewart. Prof Stewart and Dr Sarah Welsh developed the study protocol.
Stephan Ursprung has defined the imaging endpoints in collaboration with Professor Ferdia Gallager and Professor Evis Sala. He developed the relevant methods under their oversight. Stephan Ursprung planned and oversaw the image acquisition, performed the segmentation, analysed the images, performed the statistical analysis, visualised and interpreted the results.

6.1 Introduction

Surgery is the preferred treatment option in patients where curative treatment is possible and, consequently, 56% of patients with RCC undergo surgery as part of their treatment

plan [65, 473]. While neoadjuvant or adjuvant systemic therapy is not routinely indicated in RCC undergoing surgery with curative intent, a growing number of surgical, interventional, targeted drug and immune-checkpoint inhibitor treatments are available for the 30% of patients developing a recurrence [474–476].

The rational design of drug trials becomes increasingly difficult as the number of novel agents, combination treatments and possible standard of care comparators increases. Meanwhile, the evaluation of treatments in randomised controlled trials is costly. Estimates of research and development costs for bringing anti-cancer drugs to market vary considerably, ranging from 314 million USD to 2.1 billion USD. Two recent analyses estimated costs to reach 757 million USD and 1.35 billion USD, respectively [477, 478]. Clinical trials are one contributor to the total cost of bringing a drug to market. Later phase clinical trials are associated with increased cost as the participant number is the major cost driver [479, 480]. Phase-III oncological trials are estimated to cost 19 million USD as a median [481]. These costs underestimate the true cost of drug trials since they do not take trials into account which eventually result in no approval of a drug. Therefore, evidence for prioritising promising therapies for later phase trials is required. Window-of-opportunity trials take place during the time in which a patient would normally wait for their standard of care treatment [482]. In patients with resectable renal cancer, this is the time between diagnosis and nephrectomy. In Canada, the United Kingdom and the United States, waiting times of one month and more are typical [483–485]. Window-of-opportunity trials utilise this time to investigate the pharmacodynamic effects of agents in this treatment naïve cohort. Window-of-opportunity trials can support an extensive translational agenda as patients eventually undergo surgery where significant amounts of tissue become available and no previous treatments have introduced bias [482]. One of the primary risks of window-of-opportunity trials is the potential for delays in the standard of care treatment resulting from trial-related adverse events. This is seen as a reason why, historically, window-of-opportunity trials were slow to accrue patients [486, 487].

The integrity of the DNA is essential for cell survival. Conversely, DNA instability is a hallmark of cancer and enables its development and progression [168]. Multiple, interlinked enzymatic pathways ensure that DNA damage resulting from the interaction of endogenous reactive oxygen species, exogenous carcinogens or radiation with DNA or during replication is repaired [488]. These DNA damage repair (DDR) pathways include homologous recombination, non-homologous end joining, base excision repair, nucleotide excision repair and mismatch repair [489]. Many cancer types are deficient in one of these pathways and rely on the redundancy of DDR pathways. Prominent examples are breast and

ovarian cancer in patients with a somatic or germline BRCA-1 or BRCA-2 mutation [490]. Inactivation of BRCA-1/2 inhibits homologous recombination repair. Instead, cells rely on base excision repair to correct DNA single-strand breaks. DDR inhibitors (DDRi) exploit this reliance on a single pathway [489]. Poly (ADP-ribose) polymerase-1 and 2 (PARP) bind to DNA single-strand breaks and recruit the enzymes required for its repair by attaching poly (ADP-ribose) to itself and other molecules in the repair complex (PARylation) [491]. PARP inhibitors (PARPi) bind either the active site of PARP or bind the PARP-DNA complex and prevent its activation. This leads to the accumulation of DNA double-strand breaks in cells deficient in homologous recombination [490]. Persistent DNA double-strand breaks result in cell cycle arrest during mitosis and eventually apoptosis. The coincidence of a germline or somatic mutation leading to the loss of one DNA single-strand repair mechanism and the pharmacological inhibition of a second pathway is called synthetic lethality.

The WIndow-of-opportunity clinical trial platform for evaluation of novel treatment strategies in REnal cell cancer (WIRE) is lead by Professor Grant Stewart and Dr Sarah Welsh and was developed in collaboration with the Department of Radiology. The trial aims to exploit the tumour-specific genomic instability, which is in contrast to chemo- and radiotherapy that result in damage to normal tissue owing to their unspecific nature [492]. DDRi and immune checkpoint inhibitors have been successfully applied in the treatment of multiple cancers [493–495]. However, somatic mutations in genes of the DNA damage response pathways are uncommon in renal cancer. BRCA-1/2 mutations are not associated with an increased risk of RCC [496, 497]. DDRi may, nonetheless, have a role to play in the treatment of RCC. In contrast to synthetic lethality described above, contextual synthetic lethality describes the phenomenon where two pharmacological agents interact synergistically to elicit an anti-tumour response [498]. Loss of VHL (85% of ccRCC) and SETD2 mutations (12% of ccRCC) are more common than BRCA mutations in renal cancer and result in replication stress in RCC. Replication stress was measured by the appearance of γH2AX foci and tumour suppressor p53-binding protein 1 foci, markers of DNA double-strand breaks [499, 500]. VEGF-directed TKI also result in replication stress, suggesting that their combination with DDRi may be an effective therapeutic strategy. Olaparib, a PARPi, confers prolonged PFS in BRCA-mutated ovarian cancer as demonstrated in two recent placebo-controlled phase-III clinical trials (19.1 vs. 5.5 months) [501, 502]. Additionally, patients with human epidermal growth factor 2 (HER2)-negative breast cancer harbouring a germline BRCA-1/2 mutation who were treated with olaparib achieved a longer PFS of 7.0 compared to 4.2 months for patients treated with either capecitabine, eribulin or vinorelbine in a phase-III clinical trial [494]. Olaparib has also demonstrated efficacy in prostate cancer,

achieving a longer OS (17.3 vs. 14.0 months) compared to the combination of enzalutamide or abiraterone acetate and prednisone [503]. None of the PARP inhibitors has been evaluated in renal cell carcinoma.

WIRE will also investigate the activity of olaparib in combination with cediranib, a potent VEGF-A and VEGFR-2 inhibitor in renal cancer. Cediranib elicits a rapid reduction in tumour vascular volume and therefore blood flow and is hypothesised to increase replication stress in ccRCC. It has shown activity against ccRCC in a placebo-controlled, randomised, double-blinded phase-II clinical trial with a significant reduction in tumour size after 12 weeks. [504]. Furthermore, DCE-MRI is a sensitive marker of response to cediranib as early as 7 days after the start of treatment [505]. The combination of cediranib and olaparib resulted in prolonged PFS in ovarian cancer compared to olaparib alone (17.7 vs. 9.0 months) in a phase-II clinical trial [506].

Immune checkpoint inhibitors are a mainstay of current first-line treatments in ccRCC. Durvalumab, a PD-L1 inhibitor, with approval for the treatment of non-small cell lung cancer achieved a longer OS in patients after chemoradiotherapy in a placebo-controlled phase-III trial [507]. Owing to their existing role in the treatment of ccRCC [508, 90] and the proposed synergistic activity of immune checkpoint inhibitors and DDRi, the combination of durvalumab and olaparib is a promising therapeutic strategy in ccRCC. The proposed efficacy of the combination treatment is hypothesised to result from the increased expression of PD-L1 and activation of the stimulator of interferon genes (STING) pathway following PARP inhibition observed *in vivo* [509, 510]. These immunomodulatory effects are independent of the BRCA-1/2 expression status and the combination treatment of a PARPi with an immune checkpoint inhibitor showed greater efficacy than either agent alone in *in vivo* models of ovarian and breast cancer [511].

This chapter presents the initial imaging results of the WIRE clinical trial platform. The trial comprises five arms and will evaluate the engagement of the DDRi (olaparib) alone and in combination with the TKI cediranib and immune checkpoint inhibitor durvalumab with their putative pharmacological targets in patients with renal cell cancer. Future treatment arms can include new systemic anti-cancer agents, interventional or radiation therapy.

6.2 Methods

6.2.1 Participants

The WIRE trial is a multi-arm, multi-centre, non-randomised phase-II window-of-opportunity trial aiming to achieve proof of mechanism for cediranib, olaparib and durvalumab single-agent as well as cediranib + olaparib and olaparib + durvalumab combination treatments in patients with clear cell renal cell carcinoma. The trial was registered on clinicaltrials.gov (NCT03741426). Participants will be recruited from three hospitals in the UK: Cambridge University Hospitals, Oxford University Hospitals and the Queen Elizabeth University Hospital Glasgow.

Inclusion and exclusion cirteria

Patients included in the WIRE trial must:

- be $\geqslant$18 years of age
- have biopsy-proven, surgically resectable, cT1b+, cN0-1, cM0-1 ccRCC
- reach Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1
- have one or more measurable lesions as per RECIST 1.1 [114]
- have sufficient organ and bone marrow function to ensure the safety of the treatment
- use contraception (two forms of highly active contraception in women of childbearing potential) until three months after the trial

Patients may not:

- have cT1a, N0, M0 ccRCC
- have brain or leptomeningeal metastasis
- have spinal cord compression unless definitely treated and stable for 28 days
- have a bodyweight $\leqslant$30 kg
- have had any other invasive malignancy in the last two years
- have undergone prior treatment with PARP, tyrosine kinase, immune checkpoint or mammalian target of rapamycin (mTOR) inhibitors

- have had cardio- or cerebrovascular complications in the last 12 months
- participate in a concurrent interventional clinical trial
- be immunosuppressed or harbour an active infection
- be unable to undergo MRI scanning
- be pregnant

Reference values for adequate organ and bone marrow function, as well as in- and exclusion criteria specific to individual treatment arms, are detailed in the appendix (section E.1).

Power calculation

The trial uses a Bayesian adaptive design to select the optimum number of participants in each of the five treatment arms in a dynamic way. Information from the initial patients in each arm is used to determine the recruitment target. Figure 6.1 shows the intermediary targets and the decision boundaries for single- and combination-agent arms. When a recruitment target has been reached in one arm, a subsequent arm will open while the interim analysis is conducted. If according to the decision boundaries an arm is likely to achieve efficacy after recruiting further patients, it is immediately re-opened following the interim analysis. Participants will be allocated to an arm at screening and may not cross to another arm or enter a reserve pool if they are ineligible for the current arm. In total, no more than 76 evaluable patients will be recruited into the first five arms of the WIRE trial.

The maximum number of evaluable patients to be recruited into monotherapy arms was determined to be 12. With the null hypothesis that no more than 20% will respond to the treatment (H0: $p \leqslant 0.2$) and the alternative hypothesis that $> 20\%$ will respond (H1: $p > 0.2$), the overall type I error rate (α) is 20% and the power (β) is 92% when $p=0.5$.

The maximum number of evaluable patients to be recruited into combination treatment arms was determined to be 20. With the null hypothesis that no more than 30% will respond to the treatment (H0: $p \leqslant 0.3$) and the alternative hypothesis that $>30\%$ will respond (H1: $p > 0.3$), the overall type I error rate (α)is 10% and the power (β) is 93% when $p=0.6$.

6.2.2 Ethics

Ethical approval for this trial was obtained from the London-Harrow research ethics committee (NRES: 19/LO/1462, IRAS project ID: 252392). Patients provided written informed consent.

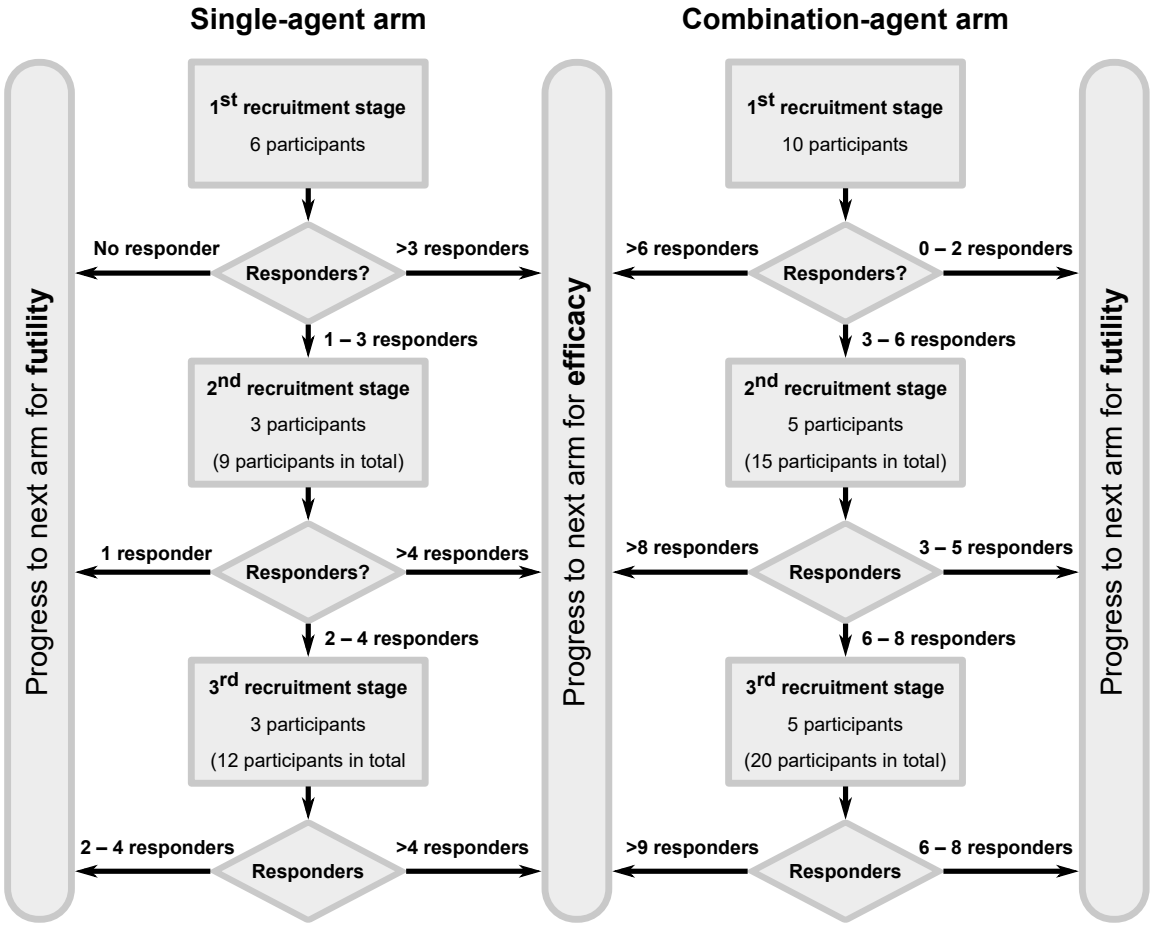


Fig. 6.1 Recruitment for each treatment arm of WIRE occurs in up to three stages. While an interim analysis is ongoing, recruitment continues into a subsequent arm. Depending on the number of responders at each interim analysis, an arm may close for efficacy, close for futility, or progress to the next recruitment stage.

6.2.3 Objectives and outcomes

Primary

This window-of-opportunity trial aims to demonstrate the engagement of novel single-agent and combination-agent treatments proposed for ccRCC with their putative cellular pathways. Cediranib, the TKI, is expected to result in changes in vascular permeability and a reduction in the median K^{trans} by 30% following at least 14 days of treatment is defined as indicative of treatment response. This threshold for the repeatability of K^{trans} measurements was updated after the NeoSun trial to reflect more recent findings. Test-re-test analyses for DCE-MRI in RCC showed coefficients of variation (CV) for K^{trans} of 15.6% between subsequent measurements taken 48 to 72 hours apart in the same patient [265]. This is in keeping with observations in pancreatic adenocarcinoma, glioblastoma, and prostate cancer, where CV reached 15.6%, 13.9% and 12%, respectively. For the present study, changes in K^{trans} by $1.96 \cdot 15.6\% = 30\%$, representing the upper limit of the 95% confidence interval, were accepted as representing a biological effect rather than test-retest variability. K^{trans} is a sensitive and early marker of treatment response in VEGFR-directed TKI therapy [332, 512]. The rationale for this endpoint is derived from the NeoSun trial described in section 3.2. *In vivo* studies identified a reduction in the migration of vascular endothelial cells in a mouse model which results in impaired angiogenesis [513]. Therefore, K^{trans} will serve as the primary endpoint for the cediranib, olaparib, and cediranib + olaparib arms.

Durvalumab, similar to other immune checkpoint inhibitors, is expected to elicit an increase in tumour infiltrating CD8+ T-lymphocytes. An increase in tumour infiltrating CD8+ T-lymphocytes is associated with better treatment response in melanoma [514, 515]. An increase in tumour infiltrating CD8+ T-cells by 30% will serve as the primary endpoint for the durvalumab containing arms.

Secondary

Secondarily, the changes in tumour size and systemic response were assessed with the Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1 [114]. Arms in which K^{trans} and the CD8+ T-cell count are not primary endpoints will evaluate these as secondary endpoints.

Exploratory

Exploratory objectives will aim to define the molecular biological response to the IMPs in blood, urine, and tissue samples. Additionally, the relationship of the primary endpoints with genetic mutations in genes related to the pathogenesis of ccRCC will be investigated. Finally, the ability of automated histological image analysis and hyperpolarised [1-^{13}C]pyruvate MRI imaging to detect biological treatment response will be explored.

6.2.4 Trial procedures

Following informed consent, patients will undergo multiparametric MRI including DCE-MRI and, if possible hyperpolarised [1-^{13}C]-pyruvate MRI as part of the screening procedures. The proton and ^{13}C MRI protocol and data analysis pathway will be identical to the one described for the MISSION renal project (section 5.2.2, page 122). After the screening MRI, patients will undergo multi-regional US-guided tumour biopsies to verify the histologic subtype of the tumour before definite inclusion in the trial. The US-guided biopsy will target tumour habitats defined by their unique imaging phenotype on multi-parametric proton MRI (section 4.2.3, page 110). Six biopsies will be taken from each tumour, two of which will be formalin-fixed, two snap-frozen and two fresh. No more than 28 days after any of the screening examinations, participants will commence treatment in one of the five arms in WIRE (table 6.1). They will receive a minimum of 14 days and a maximum of 28 days of treatment (fig. 6.2). Participants will undergo a repeat multiparametric proton and hyperpolarised [1-^{13}C]-pyruvate MRI up to 72 hours before nephrectomy. The surgery is scheduled between 17 and 28 days after the first day of treatment. In patients undergoing open surgery, fresh-frozen core biopsies will be obtained from the perfused tumour immediately before ligation of the supplying vessel(s). Imaging habitats and a 3D-printed tumour mould will be derived from the pre-surgical imaging as described previously (section 4.2, page 106) to achieve spatial registration between the imaging data and tissue samples. Multi-regional formalin-fixed, snap-frozen, and fresh tissue samples will be collected after nephrectomy. During the screening examination, immediately before surgery and during the follow-up, blood and urine samples will be collected for exploratory analyses.

Dose adjustments and participant replacement

Dose adjustments are permissible in case of patient toxicity. Reduced dose levels are summarised in table 6.1. Both dugs will be dose-reduced in combination agent arms unless the toxicity can be ascribed to a single agent. Participants will be discontinued if toxicities

Table 6.1 Initial drug regimens in the WIRE trial

Arm	Drug (combination)	Regimen	Dose reduction(s)
I	Cediranib	20mg od until 1½ days before surgery	reduction to 15 mg
II	Cediranib + Olaparib	20 mg od + 300 mg b.d.	as for single agents
III	Olaparib	300 mg b.d. until day of surgery	1st reduction to 250 mg b.d. 2nd reduction to 200 mg b.d.
IV	Durvalumab	1500 mg i.v.	NA
V	Olaparib + Durvalumab	300 mg b.d. + 1500 mg i.v.	as for single agents

b.d.: twice daily, od: once daily, i.v. intravenous, NA: not applicable. Oral dosing unless stated otherwise.

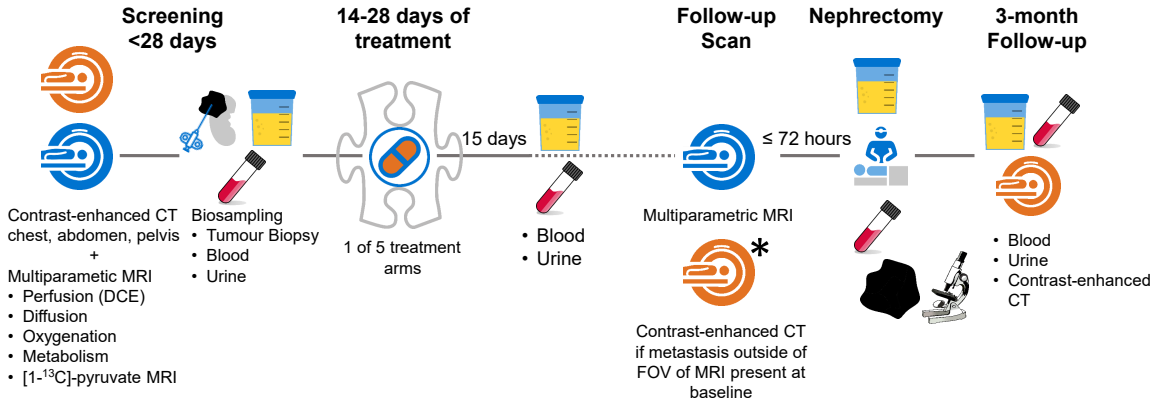


Fig. 6.2 WIRE participant pathway. CT: Computed tomography, DCE: Dynamic contrast enhanced MRI, FOV: Field-of-view, MRI: Magnetic resonance imaging.

persist. Furthermore, the treatment will be discontinued if patients progress during treatment, become pregnant, lose weight to $\leqslant$30 kg, or on patient request. Participants who have received less than 14 days of treatment or a single infusion of durvalumab will be replaced. Furthermore, patients who had to interrupt their treatment for more than seven consecutive days or were unable to take their treatment on more than 4 days in the week leading up to the follow-up imaging, were unable to undergo follow-up imaging or undergo surgery more than 28 days after starting treatment will be substituted.

6.2.5 Image and tissue analysis

Assessment of the primary endpoint

Image data from all participating centres will be anonymised locally and analysed centrally. The methodology for the analysis of the DCE-MRI has been described in detail in section 5.2.2, page 122. In brief, affine registration will be applied to the T_1-mapping and dynamic contrast-enhanced sequences. Kinetic modelling will be undertaken in the MiStar software (Apollo Medical, Melbourne). The B_1 and T_1 map will be employed to calculate maps of the gadolinium concentration. The kinetic rate constants K^{trans}, k_{ep}, v_e and the goodness of fit (R^2) will be modelled using the extended Tofts model [335]. Owing to the temporal resolution of the DCE-MRI of 12 s, the Fritz-Hansen blood-sampled curve with values after the first pass extrapolated from the Weinmann function and scaled accordingly, served as the arterial input function [446, 516]. The parameter maps will be resampled in the same space as the T_1w Lava-Flex sequence. After rigid registration, the 3-dimensional ROI encompassing the entire tumour will be propagated to the parametric maps. Voxels with a goodness of fit $\leqslant$0.75 will be excluded from the calculation of the median tumour K^{trans}, k_{ep} and v_e.

The number of CD8+ T-lymphocytes per unit area will be determined centrally. Formalin-fixed paraffin-embedded tissue blocks will be transferred to Cambridge. After slicing, the tissue is incubated in a hydrogen peroxide solution to block endogenous peroxidase which would otherwise cause high, non-specific background staining. Subsequently, the tissue will be heated to retrieve the antigenic sites for primary anti-body binding which methylene bridges may have obscured after formalin fixation. A primary mouse monoclonal anti-human CD8 antibody (DAKO Cytomation, Glostrup, Denmark, 1:100 dilution) will be used. Slides will be digitised and two board-certified uropathologists will quantify the CD8+ T-cells in ten random fields of view per tumour sample, expressed as the median number of cells/mm^2 before and after treatment.

Translational investigations

The WIRE platform will support a broad translational agenda covering imaging and molecular research to explore future markers of treatment response. The MRI protocol includes physiological and metabolic imaging sequences providing spatially resolved information on tumour properties including perfusion, cellularity, blood oxygenation and carbohydrate metabolism. Existing data provide a rationale that all of these could be sensitive to TKI, PARPi, and immune-checkpoint inhibitor treatment.

Tissue samples will support an equally broad range of translational research: DNA, RNA, proteomic and metabolic analyses from corresponding, co-registered multi-regional tissue samples will serve to understand the molecular basis of intratumoral imaging heterogeneity. In addition, changes in DNA, RNA, ctDNA and cfDNA before and after treatment will be investigated for molecular biomarkers of treatment response and to confirm the proposed drug mechanism of action. The representation of intratumoral genetic heterogeneity in multi-regional tissue samples and liquid biopsies will be compared. Finally, the trial aims to generate patient-derived xenografts and cell lines as models for future research.

6.2.6 Statistical analysis

Interim analyses are scheduled after every recruitment target and determine if an arm has reached efficacy or futility thresholds or should progress towards the next recruitment target. The Bayesian analysis will use a Beta (0.3, 0.7) prior for the biological response rate which will be updated with the observed response rate using a Beta-Binomial conjugate distribution which in turn determines the Beta posterior distribution p.

If the probability for rejection of H0 in the monotherapy arms is above 98%, the arm will be declared efficacious and no further recruitment will occur. If the probability of attaining statistical significance at a one-sided level of $p<0.1$ after recruiting all 12 patients is $\leqslant 2\%$, the arm will be stopped for futility. The results of an arm will be considered positive if the recruitment stops early owing to efficacy or if the final p-value calculated after all 12 participants have completed the arm is $\leqslant 0.1$, assuming a binomial distribution with $p=0.2$.

If the probability for rejection of H0 in the combination treatment arms is above 98%, the arm will be declared efficacious and no further recruitment will occur. If the probability of attaining statistical significance at a one-sided level of $p<0.05$ after recruiting all 20 patients is $\leqslant 2\%$, the arm will be stopped for futility. The results of an arm will be considered positive if the recruitment stops early owing to efficacy or if the final p-value calculated after all 20 participants have completed the arm is $\leqslant 0.05$, assuming a binomial distribution with $p=0.3$.

Statistical analyses and figure design were performed in the R environment for statistical computing (Vienna, Austria), Version 4.0.4. The ggplot2 package version 3.3.3 was employed for figure design.

6.3 Results

This section describes the initial results of the imaging aspects of the WIRE trial. Thirteen patients were screened for inclusion until the 13th May 2021. Eight patients participated in the trial, six of whom received window-of-opportunity treatment with cediranib and two with the combination of cediranib and olaparib. Figure 6.3 summarises reasons for exclusion.

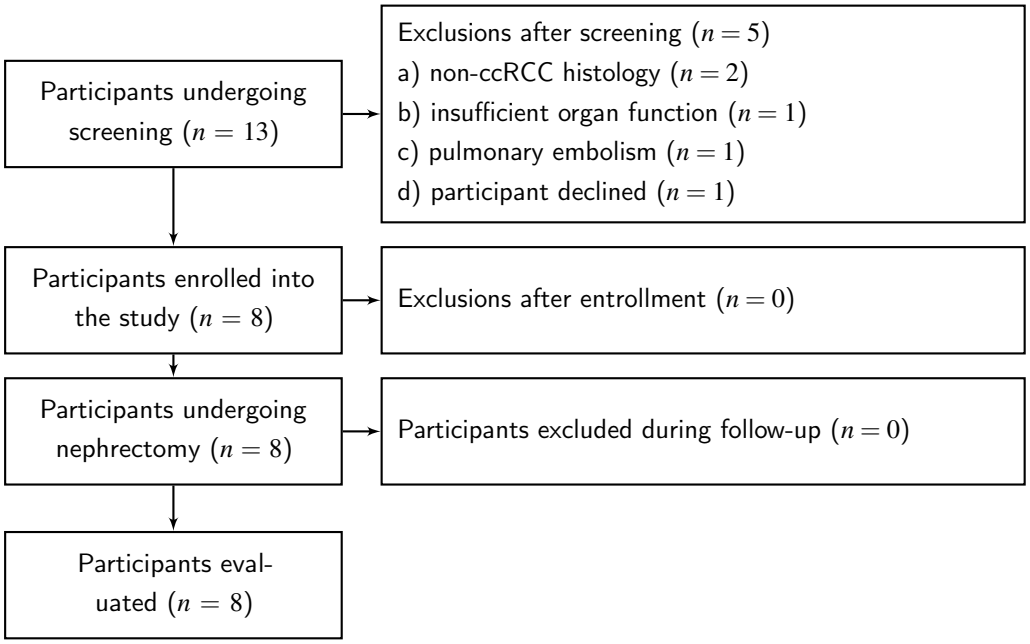


Fig. 6.3 WIRE: CONSORT Diagram

6.3.1 Multiparametric ^{1}H MRI

All participants underwent multiparametric ^{1}H MRI successfully at baseline and after at least 14 days of trial treatment. Participant demographics are summarised in table 6.2. The patients in the cediranib arm presented with a median baseline tumour volume of 113 cm^3 (range: 49 to 307 cm^3). The participants received a median of 24 days of treatment before undergoing the follow-up MRI (range: 17 to 25). After treatment, the volume had fallen significantly to 101 cm^3 (IQR: 41 to 194 cm^3) equating to a relative volume change of -14% (range: -37% to -5%; *P*=0.031). The patients undergoing combination treatment with cediranib and olaparib experienced a change in tumour volume from 400 to 404 cm^3 (+1%) after 25 days of treatment and from 694 to 646 cm^3 (-7

Table 6.2 Characteristics of the participants in the WIRE trial

Characteristic	Distribution
Sex [female/male]	1/7
Age at diagnosis [years, median (min – max)]	48 (62 – 71)
WHO/ISUP grade at biopsy	Grade 2: 5 Grade 3: 2 Grade 4: 1
WHO/ISUP grade at surgery	Grade 3: 5 Grade 4: 2
TNM-stage at surgery	ypT3a, pNX, cM0: 3 ypT3a, pN0, cM0: 1 ypT3a, pN1, cM0: 1 ypT3a, pNX, cM1: 1 ypT3b, pNX, cM1: 1
Trial treatment	Cediranib: 6 Cediranib + olaparib: 2

Dynamic contrast-enhanced MRI

In the cediranib arm, all but one participant experienced a reduction in the contrast transfer constant K^{trans}. The median K^{trans} fell from 0.44 min^{-1} (range: 0.25 to 0.78 min^{-1}) to 0.26 min^{-1} (range: 0.08 to 0.46 min^{-1}). The median relative change was -47% (range: -80% to +80%). Four participants met the pre-determined threshold of a reduction in K^{trans} by 30%. According to the decision boundaries laid out in fig. 6.1, the null hypothesis that cediranib leads to a reduction in no more than 1 in 5 patients was rejected and the arm was closed for efficacy after the first recruitment stage. One patient experienced a reduction in K^{trans} by 12% which failed to reach the limit set for unequivocal response and K^{trans} increased by 80% in one patient (fig. 6.4a). The first participant in the cediranib + olaparib arm achieved a 37% reduction from 0.21 to 0.14 min^{-1}, meeting the criteria for achieving the primary endpoint. Meanwhile the second patient responded with a reduction in K^{trans} by 12% from 0.15 to 0.13 min^{-1} which was below the threshold for tumour respone.

The area under the contrast-enhancement curve during the initial 90 seconds (iAUC_{90}), a surrogate marker for K^{trans} relying on fewer assumptions about contrast agent kinetics, confirmed the findings of the K^{trans} analysis. The iAUC_{90} fell by 44% (range: -59 to +31%) from 0.63 mmol/L min (range: 0.52 to 0.86 mmol/L min) to 0.34 mmol/L min (range: 0.30

to 0.75, fig. 6.4b). In agreement with the K^{trans} assessment, the only patient who experienced an increase in K^{trans} was also the only patient with and increased iAUC_{90} at the follow-up time point (+28%). Furthermore, the iAUC_{90} decreased least in the patient in the first arm with the least reduction in K^{trans} (-20%) and the non-responding patient in the second arm (-18%). The reduction in iAUC_{90} reached or exceeded 30% in all other patients including the first patient receiving cediranib and olaparib combination treatment.

Following treatment with the TKI cediranib, the vascular plasma volume fraction v_p decreased from a median of 14.2% (range: 7 to 40%) to 7.8% (range: 1 to 36%). The significant loss in plasma volume ranged from 2.3 to 13.6% (median: 3.8%, $P = 0.031$, fig. 6.4c). While all patients showed a reduction in v_p, the volume fraction of the extravascular extracellular space v_e remained unchanged before and after treatment at a median of 37% and 38%, respectively ($P = 1.00$, fig. 6.4d). Neither parameter was able to differentiate the response groups identified by K^{trans}. The response in the responding patient from the second arm agreed with these findings and was in keeping with their response status, showing a reduction in v_p by 35% from 25% to 16%. In contrast, the non-responding patient in the second arm was the only patient with an increase in v_e by 13% from 4.5% to 5.1%.

Finally, the delay in contrast arrival decreased from 9 s (range: 1 to 13 s) to 3 s (range 2 to 14 s). While most patients showed a decreased delay during the follow-up MRI (median: -71%), the delay increased by 41% and 150% in the two non-responding lesions (fig. 6.4e). The delay increased from 14.0 to 15.2 seconds (+8%) in the patient in the second arm of the WIRE trial. They were the only responder with an increasing delay. However, relative to the standard deviation of 5 s of the delay within the tumour, the increase between the baseline and follow-up MRI was minor. This patient showed the highest standard deviation in contrast delay reaching a median of 2 s across patients. The non-responding patient in the second arm experienced an increase in the delay by 99% which is in keeping with the response status.

Intravoxel incoherent motion-type diffusion-weighted imaging

Tumour diffusivity D_0 from IVIM-type DWI was reduced from a median of $1.96\times10^{-3}\text{mm}^2\text{s}^{-1}$ (range: 1.52 to $2.84\times10^{-3}\text{mm}^2\text{s}^{-1}$) to $1.82\times10^{-3}\text{mm}^2\text{s}^{-1}$ (range: 1.49 to $2.32\times10^{-3}\text{mm}^2\text{s}^{-1}$) following cediranib treatment. All responding and the poorly responding lesion according to K^{trans} showed a reduction in D_0 -14% (range: -2 to -18%) while the progressive lesion showed an increase (+4%). However, none of the changes in D_0 reached the threshold of a change $>20\%$ defined as indicative of true change which is unlikely to be the result of test-retest variability as used in the NeoSun trial (section 3.2.2 on page 68). The perfusion fraction f_p,

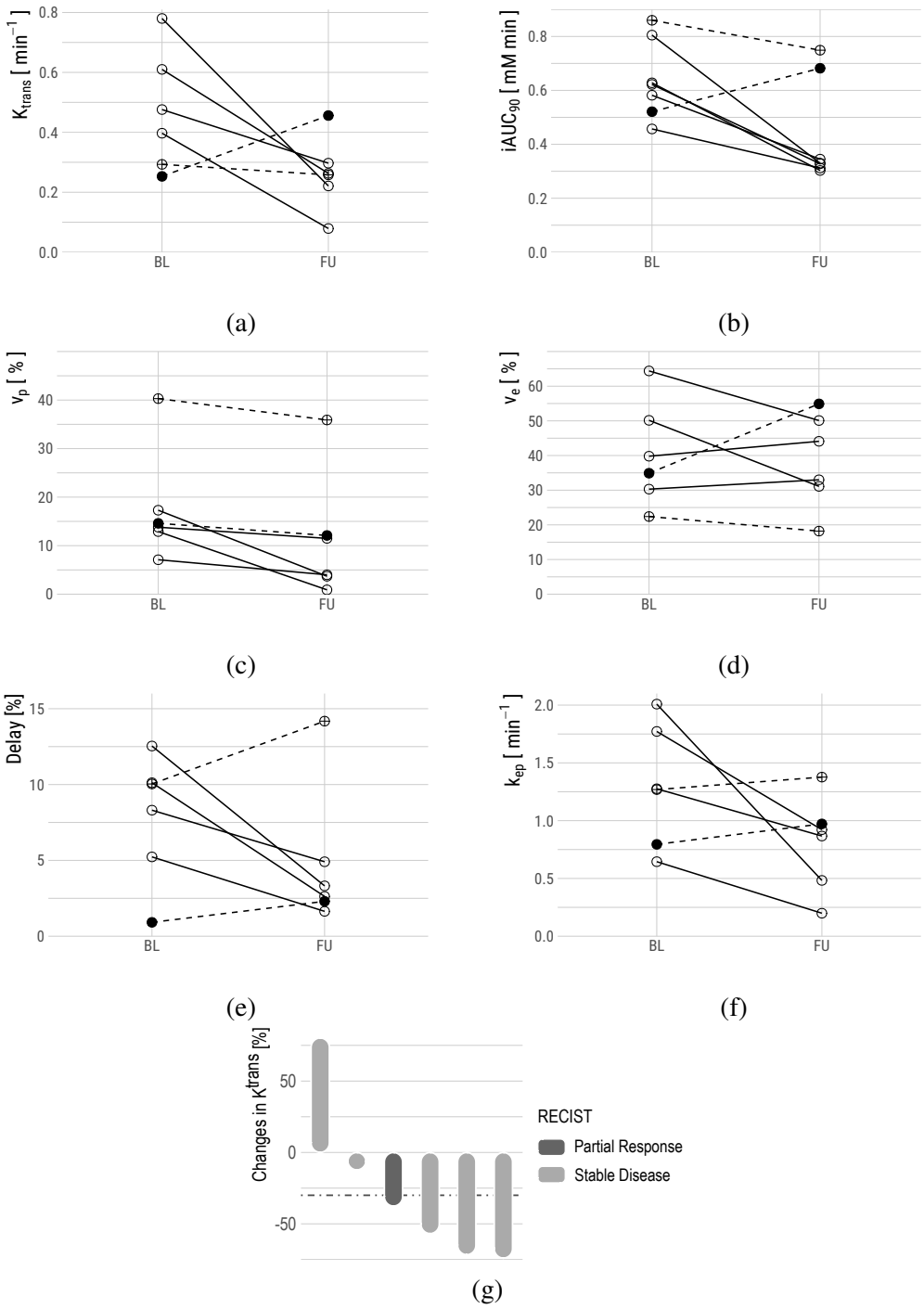
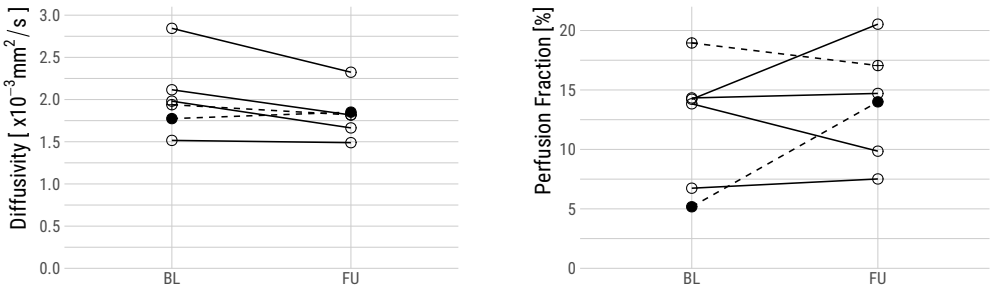


Fig. 6.4 Change in parameters from dynamic contrast-enhanced MRI following at least 14-days of cediranib treatment in the first arm of the WIRE trial. Dashed lines represent non-responders according to the primary endpoint. The closed circle represents the patient with increasing K^{trans}. The waterfall plot (g) shows the reduction in K^{trans} relative to the cutoff of 30% (dotted line)

a measure of microvascular volume which can be differentiated from the extravascular and intracellular space based on the more rapid motion of water molecules along the vasculature, remained unchanged after cediranib treatment $P = 0.56$. The microvascular volume was calculated at 11.9% (range: 2.1 to 19.7%) and 13.6% (5.2 to 16.7%) at the baseline and follow-up MRI, respectively. No pattern differentiating responders and non-responders was apparent.

D_0 fell by 14.5% in the first patient receiving cediranib + olaparib from 1.21×10^{-3}mm^{2}s^{-1} to 1.00×10^{-3}mm^{2}s^{-1} comparable to the findings in patients receiving cediranib alone. In agreement with the non-responding lesions in the first arm, the diffusivity in the patient who achieved a reduction in $K^{\text{trans}} < 30\%$ with cedrianib + olaparib remained unchanged (1.60 and 1.66). The f_p increased from 7.8% to 9.7% in the responder and decreased from 11.6% to 8.2% in the non-responder.



(a) (b)

Fig. 6.5 Change in parameters from intravoxel incoherent motion-type diffusion-weighted ^{1}H-MRI following at least 14-days of cediranib treatment. Dashed lines represent non-responders according to the primary endpoint. The closed circle represents the patient with increasing K^{trans}.

Blood oxygenation level-dependent MRI

R_2^*, which is sensitive to tissue oxygenation, increased significantly from 25.5 Hz (range: 8 to 37 Hz) to 37 Hz (range: 26 to 44 Hz) following cediranib treatment ($P = 0.031$). The increase in R_2^* was more pronounced in responders (median: 61%, range: 8% to 313%) compared to non-responders (+4% and +11%). An increasing R_2^* can be the result of tissue hypoxia and increased deoxyhaemoglobin concentrations. In agreement with their response status, the participants undergoing treatment with cediranib and olaparib showed an increase in R_2^* by 33% from 33 Hz to 47 Hz and 12% (24 Hz to 27 Hz).

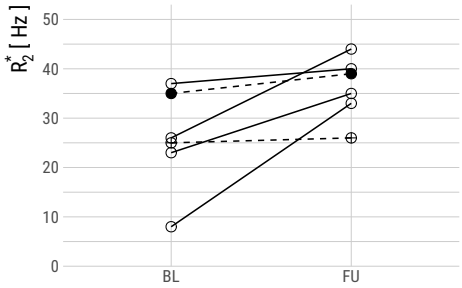


Fig. 6.6 Change R_2^* relaxation on ^{1}H-MRI following at least 14-days of cediranib treatment. Dashed lines represent non-responders according to the primary endpoint. The closed circle represents the patient with increasing K^{trans}.

6.4 Discussion

This chapter presented a platform trial for a phase-II clinical study in ccRCC including five single agent and combinational treatment arms and showed that K^{trans} is sensitive to early physiological changes during cediranib treatment.

The number of participants in each treatment arm is optimised according to Bayesian statistics to recruit as few participants as possible to demonstrate futility or effectiveness. In the case of the cediranib arm where four of six patients were found to have responded with a reduction in $K^{\text{trans}} > 30\%$, this resulted in the closure of the arm after only half of the potential full recruitment target had been reached. This demonstrated how an adaptive trial design can accelerate these early phase clinical trials.

The early reduction in K^{trans} with a median of 47% is consistent with previous findings where sunitinib, a less potent TKI compared to cediranib, caused a median reduction in K^{trans} by 28% after 12 days (section 3.2, page 65), just half of the duration of the median treatment in the first arm of WIRE. Identical findings were reported in mRCC treated with sunitinib for 14 days [517]. Others have reported >30% reduction in K^{trans} in 55% and 65% of patients with metastatic, castration-resistant prostate cancer after one day and one month of cediranib treatment and 46% in a range of advanced cancers after 28 days, albeit treated with a higher daily dose of cediranib [518, 505]. Murine models of colorectal cancer and glioma show reductions in K^{trans} by 33% after two days [519]. In metastatic RCC, 3 to 12 weeks of sorafenib treatment induces a decline in K^{trans} by a median of 60% [326]. Furthermore, a dose-dependent reduction in K^{trans} in response to sorafenib was observed at 28 days (-14% at 200 mg b.d. and -24% at 400 mg b.d.) [520]. Finally, pazopanib treatment in mRCC leads to a reduction in K^{trans} by 49% after one week [521].

We found a good agreement between the reduction in K^{trans} and iAUC_{90} in this study. A moderate correlation between the two parameters has been described previously, e.g. in rectal cancer and glioma [522, 523]. Non-parametric analyses of DCE-MRI data have some advantages, including their simple calculation, derivation directly from the signal time curve and freedom from modelling assumptions. While this avoids the non-linear conversion of signal intensities to contrast concentrations [335], it comes at the expense of reproducibility as these parameters depend on acquisition parameters [318]. The area under the initial contrast-enhancement curve has been proposed as a middle ground, which is reasonably reproducible but suffers from a complex relationship with the underlying physiology determined by perfusion, intravascular volume, vessel permeability and tissue contrast retention [524]. As a result of its complex relationship with K^{trans} and v_e, iAUC_{90} should not be seen as a surrogate marker for either [525].

Other physiological changes observed after the treatment with cediranib were consistent with the proposed mechanism of action of the TKI. The reduction in v_p agrees with previous findings describing a mean 23% and 34% reduction in v_p in response to 200 mg b.d. and 400 mg b.d. sorafenib at 28 days [520]. These findings were similar in range to our findings with a median reduction in v_p by 30% (mean 43%). This is also consistent with a reduction in micro-vessel density measured in *in vivo* models of ccRCC where treatment with cediranib resulted in a decrease in vascular density measured with CD31 (Platelet endothelial cell adhesion molecule [PECAM-1]) immunohistochemistry [526, 527]. In this study, we found no significant change in v_e after approximately three weeks of cediranib treatment, which is in agreement with observations in metastatic RCC treated with 28 days of sorafenib [520]. However, others have found a significant decrease (mean 23%) in v_e after six weeks in RCC in response to sorafenib [528]. The R_2^* value, which is influenced by blood and tissue oxygenation, increased in all patients following treatment. R_2^* is influenced by the vascular volume and geometry in addition to oxygenation. The results are in keeping with higher relative concentrations of deoxyhaemoglobin and tissue hypoxia as a result of VEGFR blockade. Importantly, R_2^* is no reliable measure of absolute tissue oxygen concentration but can serve as a qualitative parameter of changes in blood oxygenation [529].

Similar to the findings in the NeoSun trial, we again observed a reduction in water diffusivity (D_0) within the tumour following treatment. The reduction in diffusivity was greater in responding compared to non-responding tumours. As discussed for the NeoSun trial (section 3.2.4 on page 74), change in tumour diffusivity may be temporally variable with an initial decrease owing to oedematous changes followed by an increase in diffusivity as tumour cell death ensues and another decrease as tumour tissue is either replaced by

fibrous tissue or develops treatment resistance [530]. The temporally varying changes in D_0 complicate the use of this otherwise advantageous imaging biomarker of treatment response. DWI avoids the cost of contrast agents and the risks and uncertainty surrounding gadolinium deposition in the brain [531]. However, the rate of change of DWI was inferior to DCE parameters in this patient sample.

The median tumour volume in the WIRE trial was considerably smaller than in NeoSun. Additionally, the relative reduction in tumour volume was smaller with 14% compared to 25%, despite the longer treatment duration. The significance of this difference at such an early time point is unclear and may be due to the competing processes of impaired vascularisation leading to cancer cell death and oedematous transformation.

The analyses presented in this chapter have several limitations. Recruitment for the WIRE trial has commenced less than one year ago. As a result, the trial remains early in its roll-out and the other trial centres have not yet initiated the trial at their site. Consequently, only few patients were available for analysis and bio-samples had not yet been processed. Unfortunately, hyperpolarised ^{13}C-MRI before and after treatment was only available for one patient. In the remaining participants, no fluid paths were available or the image acquisition failed owing to technical defects. Finally, the tumour biopsy revealed alternative histologies in two of the patients who underwent hyperpolarised ^{13}C-MRI. Going forward, it will be important to determine the comparability of metabolic measurements between the IDEAL spiral and SpSp pulse sequences.

The window-of-opportunity design itself has limitations. First, patients may have early-stage disease and are exposed to the study compound for a limited time only before their standard of care nephrectomy. Consequently, this study setting cannot inform on the survival benefits of the investigational treatment. Similarly, tumour size as the most common marker of treatment response in adjuvant trials is not sensitive enough for this very early treatment response. Therefore, alternative biomarkers which are clinically less well-established are required to quantify the engagement of the compounds with their biological target. The uncertainty about the response range to new treatments is higher for these endpoints, but the Bayesian framework allows adjusting for these unknown factors. The trial also risks delaying standard-of-care treatment in case of significant adverse events. After a careful screening to a standard above the clinical routine and close monitoring for the early identification of toxicity, none of the first eight patients experienced a delay in their nephrectomy surgery. The Bayesian adaptive trial design mentioned above is another step in risk mitigation as it avoids larger than necessary recruitment targets. In contrast, the window-of-opportunity design has the advantage of studying treatment-naive patients without genetic alterations induced by

previous lines of treatment and providing access to large amounts of tissue for translational analyses. WIRE is expandable and future treatment arms may also include other treatment modalities such as radiotherapy or interventional procedures. Finally, the treatment may only start to show its physiological effects, measured as the trial endpoint, in the time frame of the trial treatment. Therefore, its effect could be underestimated.

In summary, the WIRE trial prioritises promising anti-cancer treatments in ccRCC for later phase-II and III clinical trials, which are required to investigate survival benefits in larger patient cohorts. The first arm of the trial investigating cediranib has shown statistically significant biological effects in just six patients and confirmed that K^{trans} is a sensitive marker of early response to TKI. Additionally, WIRE supports the identification of new biomarkers of treatment response through imaging and bio-sampling. These may serve as endpoints in future arms of the trial.

Chapter 7

Conclusions and future directions

7.1 Summary of findings

Applications of artificial intelligence to imaging RCC. The use of machine learning methods, classical or deep-learning-based, in imaging research has increased rapidly and shows promise for automating repetitive tasks and answering clinical questions beyond subjective image assessment. However, a widening gap separates advances in radiomic research from clinical impact. A review of published AI algorithms in renal cancer revealed methodological shortcomings in much of the literature which will hinder the further development and validation of these algorithms. Overfitting high-dimensional radiomic models is a risk requiring tight control. Nonetheless, only half of the studies used feature reduction or controlled for multiple comparisons. Furthermore, validation of algorithms on unseen data was uncommon and only one of 57 algorithms was validated on data from another institution. Additionally, prospective evidence is missing. So far, no cost-effectiveness analyses have been conducted and no code has been shared to allow independent reproduction of the findings. Together, these factors make the clinical translation of the algorithms unnecessarily challenging. A meta-analysis of studies employing radiomics for the differentiation of RCC and AML without visible fat showed evidence of publication bias and an over-representation of positive findings. Prospective studies and the registration of study protocols could minimise this risk.

Imaging response to TKI in ccRCC using ^{1}H-MRI and CT. An increasing number of single-agent and combinational treatments is available for patients with advanced renal cancer. TKI and immuno-oncology agents are the two principal classes of drugs used in the systemic treatment of ccRCC. However, the assessment of treatment response in the clinical routine has largely relied on the single concept of tumour size. We have shown

that a significant reduction in solid tumour size and a reduction in tumour perfusion and oxygenation related physiological MRI parameters are measurable 12 days after the start of sunitinib treatment. The restriction in water diffusion varies with time after the start of TKI treatment and may predict progression-free survival at 12 days. An increase in diffusion restriction was associated with partial or complete response while patients with an early reduction in diffusion restriction experienced stable or progressive disease.

Patients with ccRCC and venous tumour thrombi are challenging to treat owing to the morbid but potentially curative surgical resection and the poor survival in untreated patients. Reducing the extent of venous involvement through pre-surgical chemoreduction is appealing but underexplored. Axitinib, a potent TKI, elicited a rapid reduction in the IVC thrombus length by a median of 1.8 cm at 9 weeks. Most of the reduction in length was realised within three weeks, opening a wide time window for the surgery. Fortunately, the reduction was greatest in patients with the most advanced tumour thrombus and 27% achieved a decrease in tumour thrombus levels and 41% derived a surgical benefit. No patient required more invasive surgery as a result of the neoadjuvant treatment. In an independent cohort of patients with VTT who were ineligible for surgery, a similar proportion of patients achieved a reduction in VTT level after nine weeks (4/15).

Registration of imaging and tissue samples in RCC using 3D-printed tumour moulds. Renal cancer is genetically, metabolically, histologically, phenotypically and functionally heterogeneous. The registration of tissue samples and imaging data is one requirement for an integrative understanding of this intratumoral heterogeneity. The 3D-printed tumour mould proposed here was the first to extend this approach to radical nephrectomy specimens and achieved accurate orientation, slicing and sampling from specimens that are encapsulated in adipose tissue. Furthermore, it is the first mould design which resamples the imaging data to respect the conventions of pathological workup. The final mould design achieved a median dice similarity coefficient between the observed cross-section and the one predicted from the MRI of 0.92 and the internal structure of the MRI and tissue macrograph corresponded well. The composition of tumours from necrotic/cystic, fibrotic and viable components was successfully predicted with imaging habitats from multiparametric ^{1}H-MRI.

Hyperpolarised ^{13}C-pyruvate MRI in renal tumours. Metabolic alterations are a central feature of ccRCC and a direct consequence of the almost ubiquitous inactivation of VHL. Hyperpolarised ^{13}C-pyruvate MRI investigates metabolism at the crossroads of glycolysis, lactic acid fermentation, amino acid synthesis and the citric acid cycle non-invasively. Corre-

lations of metabolic imaging features with tumour aggressiveness have been documented in cancer of the breast, prostate and kidney. The apparent metabolic rate constant of ^{13}C label exchange between pyruvate and lactate was correlated with the tumour grade and the expression of MCT1, the membrane transporter which is responsible for the pyruvate uptake. In the TCGA-KIRC dataset, MCT1 expression was shown to be increased in higher-grade tumours and a negative predictor of overall survival. The association between tumour grade, MCT1 and k_{PL} was confirmed in multi-regional biopsy samples. Hyperpolarised ^{13}C-pyruvate MRI provided complementary information to ^{1}H-MRI and was not simply a measure of tissue cellularity.

Early assessment of treatment effects using physiological MRI. Window-of-opportunity trials facilitate the understanding of the biological activity of novel anti-cancer treatments in the time between the diagnosis and definitive treatment of newly diagnosed patients. The WIRE window-of-opportunity trial in renal cancer combines a Bayesian adaptive trial design to minimise the number of patients needed to attain statistical significance with an extensive translational imaging and tissue agenda. In the first arm of the trial, a median of 24 days of cediranib, a VEGF TKI, resulted in a significant decrease in the volume transfer constant K^{trans} on DCE-MRI. The diffusivity decreased in all responding patients according to the K^{trans} threshold which is in agreement with observations in the NeoSun trial. The changes in the other physiological parameters agreed with these findings. The second arm of the WIRE trial combined cediranib and olaparib, a PARP inhibitor. One of two patients achieved a K^{trans} response.

7.2 Future Directions

This thesis has investigated imaging techniques that can support precision medicine in the treatment of renal cancer. Radiomics and alternative response measurements can be implemented with existing imaging protocols. Physiological MRI, while requiring protocol changes, is available with the existing hardware in most places. Finally, hyperpolarised ^{13}C-MRI requires significant investment in production and image acquisition hardware as well as staff training. Such a multi-layered approach may give rise to some solutions which are broadly applicable while others are reserved for specific indications at sub-speciality centres. However, above all, advances in renal cancer imaging are needed to realise the full potential of current treatment options and provide objective evidence of treatment benefits from new drugs.

There is increasing evidence that genomics, transcriptomics, circulating tumour DNA or imaging alone will not afford the breadth and depth of information required to guide increasingly complex treatment decisions. Integration of multiple datastreams will be required to harness the advantages of each one. DNA and RNA sequencing, metabolomics and proteomics on tissue samples may phenotype only a small fraction of the tumour deeply. Additionally, tissue sampling is an invasive procedure. In contrast, if ctDNA is detectable, it could provide genetic information on the entire tumour burden in a spatially unresolved, minimally invasive manner. Finally, imaging interrogates large volumes of tumour at a macroscopic scale efficiently and non-invasively. Future research will have to define the most appropriate modalities for the diagnosis, treatment stratification, post-surgical follow-up, re-staging, treatment response monitoring and outcome prediction. This thesis proposes a 3D-printed tumour mould approach for the spatially resolved integration of molecular and imaging data after nephrectomy to address some of these challenges. The flexibility in the mould design minimises disruptions to existing workflows and can hopefully be translated to other tumour types in addition to renal and ovarian cancer. Another area of great potential is image-guided multi-regional tissue biopsies. Sampling intratumoral heterogeneity with biopsies is particularly challenging because of the narrow risk-benefit ratio between multiple biopsies and the potential value of additional information. The fusion of multiple imaging modalities will allow targeting imaging habitats, however, more accurate image registration algorithms and detection of the exact biopsy location could increase the precision of the approach. Pressing questions which the integration of multiple data streams could solve are the identification of tumours not requiring treatment because of their very slow progression, patients who are most at risk of recurrence and may benefit from adjuvant treatment, and patients with slowly progressive systemic disease who may benefit from cytoreductive nephrectomy or radiotherapy to single metastases.

The research discussed in this thesis has focused on ^{1}H- and ^{13}C-MR and CT imaging. Hyperpolarised ^{13}C-pyruvate MRI has been successfully translated into clinical research. Future research will define its potential as a clinical imaging modality. The non-invasive differentiation of tumour aggressiveness is a potential application for this technique, which requires further investigation. The early detection of treatment response is another potentially promising application of the technique which will be investigated alongside the WIRE trial. This thesis joins research in other organs, particularly the breast and prostate, which has shown promising results also. Simultaneously, further biological validation will be undertaken on tissue and plasma samples from the WIRE trial, and reproducibility of the results will be determined at a second trial site. The hope for the future is that hp ^{13}C-pyruvate

MRI will become a quantitative imaging technique that can be used to interrogate tumour metabolism repeatedly and without exposing the patient to ionising radiation.

However, advances in other imaging modalities also have the potential to benefit patients with RCC. The investigation of a small molecule ^{18}F-labeled radiotracer targeting CAIX for the detection of ccRCC metastases, Na^{18}F-PET for bone metastases and ^{11}C-choline PET for the identification of RCC lesions in acquired cystic kidney disease are examples of initiatives in nuclear medicine research [50, 532, 533]. Other promising avenues of imaging research in RCC are ^{23}Na- and ^{2}H-MRI. Sodium MRI measures the tightly controlled balance between the intracellular and extracellular sodium concentrations. As a result of the tumour hypoxia, the intracellular sodium concentration is hypothesised to increase in cancer. The feasibility of deuterium MRI in humans has been demonstrated in brain studies of [6,6'-$^{2}H_2$]glucose metabolism which imaged conversion to glutamine/glutamate and lactate. Deuterium MRI is technologically somewhat simpler compared to hyperpolarised MRI but the lower signal and resolution, as well as longer acquisition times, are a challenge [534].

Most renal cell cancers are detected incidentally and the increasing use of cross-sectional imaging techniques has led to a decrease in stage at diagnosis over time. However, a minority of patients still presents with advanced, prognostically poor renal cancer. Screening is indicated for significant health risks with a lead time and effective treatment options. However, imaging has so far lacked sensitivity and specificity for early diagnoses in particular. Recent advances in the analysis of the epigenome of circulating tumour DNA coupled with machine learning have shown promising results and classifiers correctly identified 97.1% of plasma samples from patients with RCC based on their DNA methylation pattern [162]. The combination of a sensitive and relatively specific blood-based biomarker with an equally sensitive imaging approach would be required to capitalise on the earlier detection of renal malignancies and initiate curative surgical or interventional treatment.

Finally, advanced image analysis, which harnesses information already present in images but not accessible to human interpretation, will complement the development of new imaging modalities. Texture-based and DL models have the potential to revolutionise the non-invasive diagnosis of renal tumours. These techniques could be of value in the determination of histological subtypes and the biological aggressiveness of renal lesions and guide the selection of treatments. Furthermore, they may provide a more sensitive measure of treatment response. However, for artificial intelligence to realise its full potential, higher quality research and adapted regulatory processes will be critical.

In conclusion, radiological imaging in its many facets has the potential to be a cornerstone of the rollout of precision medicine to patients with renal cancer. Future research will define

its precise role in the screening, diagnosis, treatment response assessment and follow-up of patients with renal masses. Hopefully, the joint evolution of treatment approaches, diagnostic imaging and molecular biomarkers will achieve better outcomes for patients with early and advanced RCC.

References

[1] J. Ferlay, M. Ervik, F. Lam, et al. Global Cancer Observatory: Cancer Today. https://gco.iarc.fr/today/home [accessed: 23.03.2020].

[2] M. Gerlinger, A. J. Rowan, S. Horswell, et al. Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. *New England Journal of Medicine*, 366(10):883–892, mar 2012.

[3] G. Low, G. Huang, W. Fu, et al. Review of renal cell carcinoma and its common subtypes in radiology. *World Journal of Radiology*, 8(5):484, may 2016.

[4] B. Escudier, C. Porta, M. Schmidinger, et al. Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology*, 30(5):706–720, feb 2019.

[5] B. Ljungberg, K. Bensalah, A. Bex, et al. Guidelines on Renal Cell Carcinoma. *European Association of Urology Guidelines Office, Arnhem, The Netherlands.*, 2015.

[6] B. Musaddaq, T. Musaddaq, A. Gupta, et al. Renal Cell Carcinoma: The Evolving Role of Imaging in the 21st Century. *Seminars in Ultrasound, CT and MRI*, 41(4):344–350, aug 2020.

[7] P. Li, A. Znaor, I. Holcatova, et al. Regional geographic variations in kidney cancer incidence rates in European countries. *European Urology*, 67(6):1134–1141, jun 2015.

[8] J. Ferlay, M. Colombet, and F. Bray. Cancer Incidence in Five Continents. http://ci5.iarc.fr [accessed: 23.03.2020], 2018.

[9] Cancer Research UK. Kidney Cancer Statistics. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/kidney-cancer#heading-Zero [accessed: 20.03.2020], 2017.

[10] S. Q. Wong, J. M. Raleigh, J. Callahan, et al. Circulating Tumor DNA Analysis and Functional Imaging Provide Complementary Approaches for Comprehensive Disease Monitoring in Metastatic Melanoma. *JCO Precision Oncology*, 1:1–14, apr 2017.

[11] E. Gumundsson, H. Hellborg, S. Lundstam, et al. Metastatic potential in renal cell carcinomas <7 cm: Swedish kidney cancer quality register data. *European Urology*, 60(5):975–982, nov 2011.

[12] H. D. Patel, M. Gupta, G. A. Joice, et al. Clinical Stage Migration and Survival for Renal Cell Carcinoma in the United States. *European Urology Oncology*, 2(4):343–348, jul 2019.

[13] R. Smith-Bindman, D. L. Miglioretti, E. Johnson, et al. Use of diagnostic imaging studies and associated radiation exposure for patients enrolled in large integrated health care systems, 1996-2010. *JAMA*, 307(22):2400–2409, jun 2012.

[14] *Cancer survival in England - adults diagnosed 2013-2017*. Office for National Statistics, 2018.

[15] Cancer survival in England - adults diagnosed 2003-2007. Technical report, 2008.

[16] S. H. Rossi, T. Klatte, J. Usher-Smith, and G. D. Stewart. Epidemiology and screening for renal cancer. *World Journal of Urology*, 36(9):1341–1353, sep 2018.

[17] C. L. Callahan, J. N. Hofmann, D. A. Corley, et al. Obesity and renal cell carcinoma risk by histologic subtype: A nested case-control study and meta-analysis. *Cancer Epidemiology*, 56:31–37, oct 2018.

[18] A. G. Renehan, M. Tyson, M. Egger, et al. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *The Lancet*, 371(9612):569–578, feb 2008.

[19] E. E. Calle and R. Kaaks. Overweight, obesity and cancer: Epidemiological evidence and proposed mechanisms. *Nature Reviews Cancer*, 4(8):579–591, aug 2004.

[20] S. Weikert, H. Boeing, T. Pischon, et al. Blood pressure and risk of renal cell carcinoma in the European prospective investigation into cancer and nutrition. *American Journal of Epidemiology*, 167(4):438–446, feb 2008.

[21] S. Christakoudi, A. Kakourou, G. Markozannes, et al. Blood pressure and risk of cancer in the European Prospective Investigation into Cancer and Nutrition. *International Journal of Cancer*, jul 2019.

[22] J. D. Hunt, O. L. Van Der Hel, G. P. McMillan, et al. Renal cell carcinoma in relation to cigarette smoking: Meta-analysis of 24 studies. *International Journal of Cancer*, 114(1):101–108, mar 2005.

[23] H. Fajkovic, S. F. Shariat, T. Klatte, et al. Impact of smoking status on survival after cytoreductive nephrectomy for metastatic renal cell carcinoma. *World Journal of Urology*, 34(10):1411–1419, oct 2016.

[24] N. Kroeger, H. Li, G. De Velasco, et al. Active Smoking Is Associated With Worse Prognosis in Metastatic Renal Cell Carcinoma Patients Treated With Targeted Therapies. *Clinical Genitourinary Cancer*, 17(1):65–71, feb 2019.

[25] J. H. Stewart, C. M. Vajdic, M. T. van Leeuwen, et al. The pattern of excess cancer in dialysis and transplantation. *Nephrology Dialysis Transplantation*, 24(10):3225–31, oct 2009.

[26] A. Znaor, J. Lortet-Tieulent, M. Laversanne, et al. International Variations and Trends in Renal Cell Carcinoma Incidence and Mortality. *European Urology*, 67(3):519–530, mar 2015.

[27] C. T. Lee, J. Katz, P. A. Fearn, and P. Russo. Mode of presentation of renal cell carcinoma provides prognostic information. *Urologic Oncology: Seminars and Original Investigations*, 7(4):135–140, jul 2002.

[28] E. Sacco, F. Pinto, F. Sasso, et al. Paraneoplastic syndromes in patients with urological malignancies, jul 2009.

[29] Z. I. Carrim and J. T. Murchison. The prevalence of simple renal and hepatic cysts detected by spiral computed tomography. *Clinical Radiology*, 58(8):626–9, aug 2003.

[30] S. Sevcenco, C. Spick, T. H. Helbich, et al. Malignancy rates and diagnostic performance of the Bosniak classification for the diagnosis of cystic renal lesions in computed tomography - a systematic review and meta-analysis. *European radiology*, 27(6):2239–2247, jun 2017.

[31] S. G. Silverman, G. M. Israel, B. R. Herts, and J. P. Richie. Management of the incidental renal mass. *Radiology*, 249(1):16–31, oct 2008.

[32] I. G. Schoots, K. Zaccai, M. G. Hunink, and P. C. Verhagen. Bosniak Classification for Complex Renal Cysts Reevaluated: A Systematic Review. *Journal of Urology*, 198(1):12–21, jul 2017.

[33] S. G. Silverman, I. Pedrosa, J. H. Ellis, et al. Bosniak classification of cystic renal masses, version 2019: An update proposal and needs assessment. *Radiology*, 292(2):475–488, jun 2019.

[34] B. Escudier, C. Porta, M. Schmidinger, et al. Renal cell carcinoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Annals of Oncology*, 27(January):v58–v68, jan 2016.

[35] F. U. Kay and I. Pedrosa. Imaging of Solid Renal Masses. *Urologic Clinics of North America*, 45(3):311–330, aug 2018.

[36] Z. J. Wang, P. Nikolaidis, G. Khatri, et al. ACR Apropriateness Criteria: Indeterminate Renal Masses. https://acsearch.acr.org/docs/69367/Narrative/ [accessed: 14.12.2020], 2020.

[37] C. Vogel, B. Ziegelmüller, B. Ljungberg, et al. Imaging in Suspected Renal-Cell Carcinoma: Systematic Review. *Clinical Genitourinary Cancer*, 17(2):e345–e355, apr 2019.

[38] U. G. Mueller-Lisse and U. L. Mueller-Lisse. Imaging of advanced renal cell carcinoma. *World Journal of Urology*, 28(3):253–261, jun 2010.

[39] R. Vikram, M. D. Beland, M. D. Blaufox, et al. ACR Appropriateness Criteria Renal Cell Carcinoma Staging. *Journal of the American College of Radiology*, 13(5):518–525, may 2016.

[40] P. T. Johnson, K. M. Horton, and E. K. Fishman. How Not to Miss or Mischaracterize a Renal Cell Carcinoma: Protocols, Pearls, and Pitfalls. *American Journal of Roentgenology*, 194(4):W307–W315, apr 2010.

[41] M. Jinzaki, J. D. McTavish, K. H. Zou, et al. Evaluation of Small (smaller than 3 cm) Renal Masses with MDCT: Benefits of Thin Overlapping Reconstructions. *American Journal of Roentgenology*, 183(1):223–228, jul 2004.

[42] D. H. Szolar, F. Kammerhuber, S. Altziebler, et al. Multiphasic helical CT of the kidney: increased conspicuity for detection and characterization of small (< 3-cm) renal masses. *Radiology*, 202(1):211–7, jan 1997.

[43] Z. Jane Wang, P. S. Leader Matthew Davenport, C.-C. G. Stuart Silverman, et al. MRI renal mass protocol v1.0 Society of Abdominal Radiology Disease Focused Panel on Renal Cell Carcinoma. *Society of Abdominal Radiology Disease Focused Panel on Renal Cell Carcinoma*, 2017.

[44] M. Bianchi, M. Sun, C. Jeldres, et al. Distribution of metastatic sites in renal cell carcinoma: a population-based analysis. *Annals of Oncology*, 23(4):973–980, apr 2012.

[45] R. J. Motzer, E. Jonasch, S. Boyle, et al. NCCN Guidelines Insight: Kideny Cancer, Version 1.2021. *Journal of the National Comprehensive Cancer Network*, 18(9), sep 2020.

[46] A. Larcher, P. Dell'Oglio, N. Fossati, et al. When to perform preoperative chest computed tomography for renal cancer staging. *BJU International*, 120(4):490–496, oct 2017.

[47] H. Wu, R. Yen, Y. Shen, et al. Comparing whole body 18F-2-deoxyglucose positron emission tomography and technetium-99m methylene diphosphate bone scan to detect bone metastases in patients with renal cell carcinomas - A preliminary report. *Journal of Cancer Research and Clinical Oncology*, 128(9):503–506, sep 2002.

[48] R. Kumar, P. Sharma, S. Karunanithi, et al. 18F-Fluoride PET/CT for detection of bone metastasis in patients with renal cell carcinoma: A pilot study. *Nuclear Medicine Communications*, 35(12):1247–1253, dec 2014.

[49] S. A. Sohaib, G. Cook, S. D. Allen, et al. Comparison of whole-body MRI and bone scintigraphy in the detection of bone metastases in renal cancer. *The British Journal of Radiology*, 82(980):632–9, aug 2009.

[50] E. L. Gerety, E. M. Lawrence, J. Wason, et al. Prospective study evaluating the relative sensitivity of 18F-NaF PET/CT for detecting skeletal metastases from renal cell carcinoma in comparison to multidetector CT and 99mTc-MDP bone scintigraphy, using an adaptive trial design. *Annals of Oncology*, 26(10):2113–2118, oct 2015.

[51] R. E. Bechtold and R. J. Zagoria. Imaging approach to staging of renal cell carcinoma. *Urologic Clinics of North America*, 24(3):507–522, aug 1997.

[52] M. Ernest Marshall, T. Pearson, W. Simpson, et al. Low incidence of asymptomatic brainmetastases in patients with renal cell carcinoma. *Urology*, 36(4):300–302, oct 1990.

[53] R. Kotecha, A. Redzematovic, R. J. Motzer, and M. H. Voss. Rates of occult brain metastases in patients with advanced renal cell carcinoma: A cohort study from patients treated across 22 clinical trials. *Journal of Clinical Oncology*, 37(7):673–673, mar 2019.

[54] M. Yanagawa, K. Arima, K. Yamakawa, et al. Prediction of residual total renal function before nephron-sparing surgery using99mTc-DMSA renal scintigraphy. *Clinical and Experimental Nephrology*, 2(3):245–248, sep 1998.

[55] H. Moch, A. L. Cubilla, P. A. Humphrey, et al. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs—Part A: Renal, Penile, and Testicular Tumours. *European Urology*, 70(1):93–105, jul 2016.

[56] S. R. Prasad, P. A. Humphrey, J. R. Catena, et al. Common and Uncommon Histologic Subtypes of Renal Cell Carcinoma: Imaging Spectrum with Pathologic Correlation. *RadioGraphics*, 26(6):1795–1806, nov 2006.

[57] V. F. Muglia and A. Prando. Renal cell carcinoma: histological classification and correlation with imaging findings. *Radiologia brasileira*, 48(3):166–74, may 2015.

[58] L. Lipworth, A. K. Morgans, T. L. Edwards, et al. Renal cell cancer histological subtype distribution differs by race and sex. *BJU International*, 117(2):260–265, feb 2016.

[59] B. Delahunt, J. C. Cheville, G. Martignoni, et al. The International Society of Urological Pathology (ISUP) grading system for renal cell carcinoma and other prognostic parameters. *American Journal of Surgical Pathology*, 37(10):1490–1504, oct 2013.

[60] U. Capitanio, V. Cloutier, L. Zini, et al. A critical assessment of the prognostic value of clear cell, papillary and chromophobe histological subtypes in renal cell carcinoma: A population-based study. *BJU International*, 103(11):1496–1500, jun 2009.

[61] B. C. Leibovich, C. M. Lohse, P. L. Crispen, et al. Histological Subtype is an Independent Predictor of Outcome for Patients With Renal Cell Carcinoma. *Journal of Urology*, 183(4):1309–1316, apr 2010.

[62] K. A. Keegan, C. W. Schupp, K. Chamie, et al. Histopathology of surgically treated renal cell carcinoma: Survival differences by subtype and stage. *Journal of Urology*, 188(2):391–397, aug 2012.

[63] L. Marconi, S. Dabestani, T. B. Lam, et al. Systematic review and meta-analysis of diagnostic accuracy of percutaneous renal tumour biopsy. *European Urology*, 69(4):660–673, apr 2016.

[64] P. O. Richard, L. T. Lavallée, F. Pouliot, et al. Is routine use of renal tumor biopsy associated with lower rates of benign histology following nephrectomy for small renal masses? *The Journal of Urology*, 200(4):731–736, oct 2018.

[65] S. Campbell, R. G. Uzzo, M. E. Allaf, et al. Renal Mass and Localized Renal Cancer: AUA Guideline. *The Journal of Urology*, 198(3):520–529, sep 2017.

[66] A. Antonelli, A. Mari, N. Longo, et al. Role of Clinical and Surgical Factors for the Prediction of Immediate, Early and Late Functional Results, and its Relationship with Cardiovascular Outcome after Partial Nephrectomy: Results from the Prospective Multicenter RECORd 1 Project. *The Journal of Urology*, 199(4):927–932, apr 2018.

[67] T. Takagi, T. Kondo, K. Omae, et al. Comparison of progression to end-stage renal disease requiring dialysis after partial or radical nephrectomy for renal cell carcinoma in patients with severe chronic kidney disease. *International Urology and Nephrology*, 48(9):1421–1427, sep 2016.

[68] W. T. Lowrance, D. S. Yee, C. Savage, et al. Complications After Radical and Partial Nephrectomy as a Function of Age. *Journal of Urology*, 183(5):1725–1730, may 2010.

[69] H. Van Poppel, L. Da Pozzo, W. Albrecht, et al. A prospective, randomised EORTC intergroup phase 3 study comparing the oncologic outcome of elective nephron-sparing surgery and radical nephrectomy for low-stage renal cell carcinoma. *European Urology*, 59(4):543–52, apr 2011.

[70] B. Gershman, R. H. Thompson, S. A. Boorjian, et al. Radical Versus Partial Nephrectomy for cT1 Renal Cell Carcinoma. *European Urology*, 74(6):825–832, dec 2018.

[71] E. L. Wood, M. Adibi, W. Qiao, et al. Local Tumor Bed Recurrence Following Partial Nephrectomy in Patients with Small Renal Masses. *The Journal of Urology*, 199(2):393–400, feb 2018.

[72] M. C. Mir, I. Derweesh, F. Porpiglia, et al. Partial Nephrectomy Versus Radical Nephrectomy for Clinical T1b and T2 Renal Tumors: A Systematic Review and Meta-analysis of Comparative Studies. *European Urology*, 71(4):606–617, apr 2017.

[73] A. Kutikov and R. G. Uzzo. The R.E.N.A.L. Nephrometry Score: A Comprehensive Standardized System for Quantitating Renal Tumor Size, Location and Depth. *The Journal of Urology*, 182(3):844–853, sep 2009.

[74] V. Ficarra, G. Novara, S. Secco, et al. Preoperative Aspects and Dimensions Used for an Anatomical (PADUA) Classification of Renal Tumours in Patients who are Candidates for Nephron-Sparing Surgery. *European Urology*, 56(5):786–793, nov 2009.

[75] J. Simhan, M. C. Smaldone, K. J. Tsai, et al. Objective Measures of Renal Mass Anatomic Complexity Predict Rates of Major Complications Following Partial Nephrectomy. *European Urology*, 60(4):724–730, oct 2011.

[76] D. Canter, A. Kutikov, B. Manley, et al. Utility of the R.E.N.A.L. nephrometry scoring system in objectifying treatment decision-making of the enhancing renal mass. *Urology*, 78(5):1089–94, nov 2011.

[77] M. Spaliviero, B. Y. Poon, C. A. Karlo, et al. An Arterial Based Complexity (ABC) Scoring System to Assess the Morbidity Profile of Partial Nephrectomy. *European Urology*, 69(1):72–79, jan 2016.

[78] P. F. Hsieh, Y. D. Wang, C. P. Huang, et al. A Mathematical Method to Calculate Tumor Contact Surface Area: An Effective Parameter to Predict Renal Function after Partial Nephrectomy. *Journal of Urology*, 196(1):33–40, jul 2016.

[79] M. N. Simmons, C. B. Ching, M. K. Samplaski, et al. Kidney Tumor Location Measurement Using the C Index Method. *Journal of Urology*, 183(5):1708–1713, may 2010.

[80] J. R. Andrews, T. Atwell, G. Schmit, et al. Oncologic Outcomes Following Partial Nephrectomy and Percutaneous Ablation for cT1 Renal Masses(Figure presented.). *European Urology*, 76(2):244–251, aug 2019.

[81] N. K. Janzen, H. L. Kim, R. A. Figlin, and A. S. Belldegrun. Surveillance after radical or partial nephrectomy for localized renal cell carcinoma and management of recurrent disease. *The Urologic clinics of North America*, 30(4):843–52, nov 2003.

[82] B. C. Leibovich, M. L. Blute, J. C. Cheville, et al. Prediction of progression after radical nephrectomy for patients with clear cell renal cell carcinoma. *Cancer*, 97(7):1663–1671, apr 2003.

[83] W. H. da Costa, R. R. Moniz, I. W. da Cunha, et al. Impact of renal vein invasion and fat invasion in pT3a renal cell carcinoma. *BJU International*, 109(4):544–548, feb 2012.

[84] B. Y. Zhang, R. H. Thompson, C. M. Lohse, et al. A novel prognostic model for patients with sarcomatoid renal cell carcinoma. *BJU International*, 115(3):405–411, mar 2015.

[85] C. Lebacle, K. Bensalah, J.-C. Bernhard, et al. Evaluation of axitinib to downstage cT2a renal tumours and allow partial nephrectomy: a phase II study. *BJU International*, 123(5):804–810, may 2019.

[86] J. L. Silberstein, F. Millard, R. Mehrazin, et al. Feasibility and efficacy of neoadjuvant sunitinib before nephron-sparing surgery. *BJU International*, 106(9):1270–1276, nov 2010.

[87] B. I. Rini, E. R. Plimack, T. Takagi, et al. A Phase II Study of Pazopanib in Patients with Localized Renal Cell Carcinoma to Optimize Preservation of Renal Parenchyma. *Journal of Urology*, 194(2):297–303, aug 2015.

[88] R. J. Motzer, T. E. Hutson, P. Tomczak, et al. Sunitinib versus Interferon Alfa in Metastatic Renal-Cell Carcinoma. *New England Journal of Medicine*, 356(2):115–124, jan 2007.

[89] C. N. Sternberg, I. D. Davis, J. Mardiak, et al. Pazopanib in Locally Advanced or Metastatic Renal Cell Carcinoma: Results of a Randomized Phase III Trial. *Journal of Clinical Oncology*, 28(6):1061–1068, feb 2010.

[90] R. J. Motzer, N. M. Tannir, D. F. McDermott, et al. Nivolumab plus Ipilimumab versus Sunitinib in Advanced Renal-Cell Carcinoma. *New England Journal of Medicine*, 378(14):1277–1290, apr 2018.

[91] B. I. Rini, E. R. Plimack, V. Stus, et al. Pembrolizumab plus Axitinib versus Sunitinib for Advanced Renal-Cell Carcinoma. *New England Journal of Medicine*, 380(12):1116–1127, mar 2019.

[92] R. J. Motzer, T. E. Hutson, P. Tomczak, et al. Overall survival and updated results for sunitinib compared with interferon alfa in patients with metastatic renal cell carcinoma. *Journal of Clinical Oncology*, 27(22):3584–3590, aug 2009.

[93] C. N. Sternberg, R. E. Hawkins, J. Wagstaff, et al. A randomised, double-blind phase III study of pazopanib in patients with advanced and/or metastatic renal cell carcinoma: Final overall survival results and safety update. *European Journal of Cancer*, 49(6):1287–1296, apr 2013.

[94] B. I. Rini, T. B. Dorff, P. Elson, et al. Active surveillance in metastatic renal-cell carcinoma: a prospective, phase 2 trial. *The Lancet Oncology*, 17(9):1317–1324, sep 2016.

[95] T. Powles, L. Albiges, A. Bex, et al. eUpdate – Renal Cell Carcinoma Treatment Recommendations. https://www.esmo.org/guidelines/genitourinary-cancers/renal-cell-carcinoma/eupdate-renal-cell-carcinoma-treatment-recommendations [accessed: 29.04.2020], feb 2020.

[96] D. Y. Heng, W. Xie, M. M. Regan, et al. Prognostic Factors for Overall Survival in Patients With Metastatic Renal Cell Carcinoma Treated With Vascular Endothelial Growth Factor–Targeted Agents: Results From a Large, Multicenter Study. *Journal of Clinical Oncology*, 27(34):5794–5799, dec 2009.

[97] L. Khoja, M. O. Butler, S. P. Kang, et al. Pembrolizumab. *Journal for ImmunoTherapy of Cancer*, 3(1), aug 2015.

[98] S. Konishi, S. Hatakeyama, T. Tanaka, et al. Comparison of axitinib and sunitinib as first-line therapies for metastatic renal cell carcinoma: a real-world multicenter analysis. *Medical oncology (Northwood, London, England)*, 36(1):6, nov 2018.

[99] C. S. Abhinand, R. Raju, S. J. Soumya, et al. VEGF-A/VEGFR2 signaling network in endothelial cells relevant to angiogenesis. *Journal of Cell Communication and Signaling*, 10(4):347–354, dec 2016.

[100] A. Tarhini, E. Lo, and D. R. Minor. Releasing the brake on the immune system: Ipilimumab in melanoma and other tumors. *Cancer Biotherapy and Radiopharmaceuticals*, 25(6):601–613, dec 2010.

[101] R. M. Wong, R. R. Scotland, R. L. Lau, et al. Programmed death-1 blockade enhances expansion and functional capacity of human melanoma antigen-specific CTLs. *International Immunology*, 19(10):1223–1234, oct 2007.

[102] S. Yip, C. Wells, R. B. Moreira, et al. Real world experience of immuno-oncology agents in metastatic renal cell carcinoma: Results from the IMDC. *Journal of Clinical Oncology*, 35(6):492–492, feb 2017.

[103] D. A. Karnofsky. *The clinical evaluation of chemotherapeutic agents in cancer*. Columbia University Press, 1949.

[104] D. Y. Heng, W. Xie, M. M. Regan, et al. External validation and comparison with other models of the International Metastatic Renal-Cell Carcinoma Database Consortium prognostic model: A population-based study. *The Lancet Oncology*, 14(2):141–148, feb 2013.

[105] A. Méjean, A. Ravaud, S. Thezenas, et al. Sunitinib Alone or after Nephrectomy in Metastatic Renal-Cell Carcinoma. *New England Journal of Medicine*, 379(5):417–427, aug 2018.

[106] A. Bex, P. Mulders, M. Jewett, et al. Comparison of Immediate vs Deferred Cytoreductive Nephrectomy in Patients with Synchronous Metastatic Renal Cell Carcinoma Receiving Sunitinib: The SURTIME Randomized Clinical Trial. *JAMA Oncology*, 5(2):164–170, feb 2019.

[107] R. C. Flanigan, G. Mickisch, R. Sylvester, et al. Cytoreductive Nephrectomy in Patients With Metastatic Renal Cancer: A Combined Analysis. *The Journal of Urology*, 171(3):1071–1076, mar 2004.

[108] S. Dabestani, L. Marconi, F. Hofmann, et al. Local treatments for metastases of renal cell carcinoma: A systematic review. *The Lancet Oncology*, 15(12):e549–e561, nov 2014.

[109] A. L. Alt, S. A. Boorjian, C. M. Lohse, et al. Survival after complete surgical resection of multiple metastases from renal cell carcinoma. *Cancer*, 117(13):2873–2882, jul 2011.

[110] H. Ikushima, K. Tokuuye, M. Sumi, et al. Fractionated stereotactic radiotherapy of brain metastases from renal cell carcinoma1. *International Journal of Radiation Oncology, Biology, Physics*, 48(5):1389–1393, dec 2000.

[111] G. K. Hunter, E. H. Balagamwala, S. A. Koyfman, et al. The efficacy of external beam radiotherapy and stereotactic body radiotherapy for painful spinal metastases from renal cell carcinoma. *Practical Radiation Oncology*, 2(4), oct 2012.

[112] N. J. Maxwell, N. Saleem Amer, E. Rogers, et al. Renal artery embolisation in the palliative treatment of renal carcinoma. *The British Journal of Radiology*, 80(950):96–102, feb 2007.

[113] W. Kassouf, L. L. Monteiro, D. E. Drachenberg, et al. Canadian Urological Association guideline for followup of patients after treatment of non-metastatic renal cell carcinoma. *Canadian Urological Association Journal*, 12(8):231–238, aug 2018.

[114] E. A. Eisenhauer, P. Therasse, J. Bogaerts, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *European Journal of Cancer*, 45(2):228–247, jan 2009.

[115] T. Tirkes, M. A. Hollar, M. Tann, et al. Response criteria in oncologic imaging: Review of traditional and new criteria. *RadioGraphics*, 33(5):1323–1341, sep 2013.

[116] C. K. Kuhl, Y. Alparslan, J. Schmoee, et al. Validity of RECIST version 1.1 for response assessment in metastatic cancer: A prospective, multireader study. *Radiology*, 290(3):349–356, mar 2019.

[117] S. R. Prasad, K. S. Jhaveri, S. Saini, et al. CT tumor measurement for therapeutic response assessment: Comparison of unidimensional, bidimensional, and volumetric techniques - Initial observations. *Radiology*, 225(2):416–419, nov 2002.

[118] T. Frauenfelder, M. Tutic, W. Weder, et al. Volumetry: An alternative to assess therapy response for malignant pleural mesothelioma? *European Respiratory Journal*, 38(1):162–168, jul 2011.

[119] L. Fournier, A. Bellucci, Y. Vano, et al. Imaging Response of Antiangiogenic and Immune-Oncology Drugs in Metastatic Renal Cell Carcinoma (mRCC): Current Status and Future Challenges. *Kidney Cancer*, 1(2):107–114, nov 2017.

[120] B. P. Brufau, C. S. Cerqueda, L. B. Villalba, et al. Metastatic Renal Cell Carcinoma: Radiologic Findings and Assessment of Response to Targeted Antiangiogenic Therapy by Using Multidetector CT. *RadioGraphics*, 33(6):1691–1716, oct 2013.

[121] F. Soria, A. I. Beleni, D. D'Andrea, et al. Pseudoprogression and hyperprogression during immune checkpoint inhibitor therapy for urothelial and kidney cancer. *World Journal of Urology*, 36(11):1703–1709, jan 2018.

[122] H. J. Park, K. W. Kim, J. Pyo, et al. Incidence of Pseudoprogression during Immune Checkpoint Inhibitor Therapy for Solid Tumors: A Systematic Review and Meta-Analysis. *Radiology*, 297(1):87–96, oct 2020.

[123] S. Champiat, L. Dercle, S. Ammari, et al. Hyperprogressive disease is a new pattern of progression in cancer patients treated by anti-PD-1/PD-L1. *Clinical Cancer Research*, 23(8):1920–1928, apr 2017.

[124] J. Y. Kim, K. H. Lee, J. Kang, et al. Hyperprogressive disease during anti-PD-1 (PDCD1) / PD-L1 (CD274) therapy: A systematic review and meta-analysis. *Cancers*, 11(11), nov 2019.

[125] J. J. Adashek, I. M. Subbiah, I. Matos, et al. Hyperprogression and Immunotherapy: Fact, Fiction, or Alternative Fact? *Trends in Cancer*, 6(3):181–191, mar 2020.

[126] J. D. Wolchok, A. Hoos, S. O'Day, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: Immune-related response criteria. *Clinical Cancer Research*, 15(23):7412–7420, dec 2009.

[127] O. Bohnsack, A. Hoos, and K. Ludajic. Adaptation and modification of the immune related response criteria (IRRC): IrRECIST. *Journal of Clinical Oncology*, 32(15), may 2014.

[128] L. Seymour, J. Bogaerts, A. Perrone, et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *The Lancet Oncology*, 18(3):143–152, mar 2017.

[129] R. L. Wahl, H. Jacene, Y. Kasamon, and M. A. Lodge. From RECIST to PERCIST: Evolving considerations for PET response criteria in solid tumors. *Journal of Nuclear Medicine*, 50(SUPPL. 1), may 2009.

[130] L. Goldfarb, B. Duchemann, K. Chouahnia, et al. Monitoring anti-PD-1-based immunotherapy in non-small cell lung cancer with FDG PET: introduction of iPERCIST. *EJNMMI Research*, 9(1), jan 2019.

[131] F. Latif, K. Tory, J. Gnarra, et al. Identification of the von Hippel-Lindau disease tumor suppressor gene. *Science*, 260(5112):1317–20, may 1993.

[132] L. Schmidt, F.-M. Duh, F. Chen, et al. Germline and somatic mutations in the tyrosine kinase domain of the MET proto-oncogene in papillary renal carcinomas. *Nature Genetics*, 16(1):68–73, may 1997.

[133] M.-H. Wei, O. Toure, G. M. Glenn, et al. Novel mutations in FH and expansion of the spectrum of phenotypes expressed in families with hereditary leiomyomatosis and renal cell cancer. *Journal of Medical Genetics*, 43(1):18–27, jan 2006.

[134] M. L. Nickerson, E. Jaeger, Y. Shi, et al. Improved Identification of von Hippel-Lindau Gene Alterations in Clear Cell Renal Tumors. *Clinical Cancer Research*, 14(15):4726–4734, aug 2008.

[135] L. E. Moore, M. L. Nickerson, P. Brennan, et al. Von Hippel-Lindau (VHL) Inactivation in Sporadic Clear Cell Renal Cancer: Associations with Germline VHL Polymorphisms and Etiologic Risk Factors. *PLoS Genetics*, 7(10):e1002312, oct 2011.

[136] A. Franovic, L. Gunaratnam, K. Smith, et al. Translational up-regulation of the EGFR by tumor hypoxia provides a nonmutational explanation for its overexpression in human cancer. *Proceedings of the National Academy of Sciences*, 104(32):13092–13097, aug 2007.

[137] W. G. Kaelin Jr. The von Hippel–Lindau tumour suppressor protein: O2 sensing and cancer. *Nature Reviews Cancer*, 8(11):865–873, nov 2008.

[138] L. Albiges, J. Guegan, A. Le Formal, et al. MET is a potential target across all papillary renal cell carcinomas: result from a large molecular study of pRCC with CGH array and matching gene expression array. *Clinical Cancer research*, 20(13):3411–21, jul 2014.

[139] I. P. M. Tomlinson, N. A. Alam, A. J. Rowan, et al. Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer. *Nature Genetics*, 30(4):406–410, apr 2002.

[140] R. L. Grubb, M. E. Franks, J. Toro, et al. Hereditary Leiomyomatosis and Renal Cell Cancer: A Syndrome Associated With an Aggressive Form of Inherited Renal Cancer. *The Journal of Urology*, 177(6):2074–2080, jun 2007.

[141] P. R. Benusiglio, S. Giraud, S. Deveaux, et al. Renal cell tumour characteristics in patients with the Birt-Hogg-Dubé cancer susceptibility syndrome: a retrospective, multicentre study. *Orphanet Journal of Rare Diseases*, 9(1):163, dec 2014.

[142] A. Nagy, D. Zoubakov, Z. Stupar, and G. Kovacs. Lack of mutation of the folliculin gene in sporadic chromophobe renal cell carcinoma and renal oncocytoma. *International Journal of Cancer*, 109(3):472–475, apr 2004.

[143] C. P. Freier, A. Stiasny, C. Kuhn, et al. Immunohistochemical Evaluation of the Role of p53 Mutation in Cervical Cancer: Ser-20 p53-Mutant Correlates with Better Prognosis. *Anticancer Research*, 36(6):3131–7, jun 2016.

[144] M. Kunizaki, T. Sawai, H. Takeshita, et al. Clinical Value of Serum p53 Antibody in the Diagnosis and Prognosis of Colorectal Cancer. *Anticancer Research*, 36(8):4171–5, aug 2016.

[145] J. Dimas-González, V. Maldonado-Lagunas, J. Díaz-Chávez, et al. Overexpression of p53 protein is a marker of poor prognosis in Mexican women with breast cancer. *Oncology Reports*, 37(5):3026–3036, may 2017.

[146] Z. Wang, S. Peng, N. Jiang, et al. Prognostic and clinicopathological value of p53 expression in renal cell carcinoma: a meta-analysis. *Oncotarget*, 8(60):102361–102370, nov 2017.

[147] R. Fisher, L. Pusztai, and C. Swanton. Cancer heterogeneity: implications for targeted therapeutics. *British Journal of Cancer*, 108(3):479–485, feb 2013.

[148] S. Turajlic, H. Xu, K. Litchfield, et al. Deterministic Evolutionary Trajectories Influence Primary Tumor Growth: TRACERx Renal. *Cell*, 173(0):595–610, apr 2018.

[149] I. Varela, P. Tarpey, K. Raine, et al. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. *Nature*, 469(7331):539–542, jan 2011.

[150] S. Turajlic, H. Xu, K. Litchfield, et al. Tracking Cancer Evolution Reveals Constrained Routes to Metastases: TRACERx Renal. *Cell*, 173:581–594, 2018.

[151] K. E. Maturen, H. V. Nghiem, E. M. Caoili, et al. Renal mass core biopsy: Accuracy and impact on clinical management. *American Journal of Roentgenology*, 188(2), feb 2007.

[152] A. Volpe, A. Finelli, I. S. Gill, et al. Rationale for Percutaneous Biopsy and Histologic Characterisation of Renal Tumours. *European Urology*, 62(3):491–504, sep 2012.

[153] A. Torres Muñoz, R. Valdez-Ortiz, C. González-Parra, et al. Percutaneous renal biopsy of native kidneys: efficiency, safety and risk factors associated with major complications. *Archives of medical science : AMS*, 7(5):823–31, oct 2011.

[154] B. A. Shannon, R. J. Cohen, H. de Bruto, and R. J. Davies. The Value of Preoperative Needle Core Biopsy for Diagnosing Benign Lesions Among Small, Incidentally Detected Renal Masses. *The Journal of Urology*, 180(4):1257–1261, oct 2008.

[155] M. J. Leveridge, A. Finelli, J. R. Kachura, et al. Outcomes of Small Renal Mass Needle Core Biopsy, Nondiagnostic Percutaneous Biopsy, and the Role of Repeat Biopsy. *European Urology*, 60(3):578–584, sep 2011.

[156] M. De Martino, T. Klatte, A. Haitel, and M. Marberger. Serum cell-free DNA in renal cell carcinoma: A diagnostic and prognostic marker. *Cancer*, 118(1):82–90, jan 2012.

[157] J. Wan, L. Zhu, Z. Jiang, and K. Cheng. Monitoring of Plasma Cell-Free DNA in Predicting Postoperative Recurrence of Clear Cell Renal Cell Carcinoma. *Urologia Internationalis*, 91(3):273–278, jul 2013.

[158] C. G. Smith, T. Moser, F. Mouliere, et al. Comprehensive characterization of cell-free tumor DNA in plasma and urine of patients with renal tumors. *Genome Medicine*, 12(1), feb 2020.

[159] C. Bettegowda, M. Sausen, R. J. Leary, et al. Detection of Circulating Tumor DNA in Early- and Late-Stage Human Malignancies. *Science Translational Medicine*, 6(224):224ra24–224ra24, feb 2014.

[160] C. Corrò, T. Hejhal, C. Poyet, et al. Detecting circulating tumor DNA in renal cancer: An open challenge. *Experimental and Molecular Pathology*, 102(2):255–261, apr 2017.

[161] M. C. Maia, P. G. Bergerot, N. Dizman, et al. Association of Circulating Tumor DNA (ctDNA) Detection in Metastatic Renal Cell Carcinoma (mRCC) with Tumor Burden. *Kidney Cancer*, 1(1):65–70, jul 2017.

[162] P. V. Nuzzo, J. E. Berchuck, K. Korthauer, et al. Detection of renal cell carcinoma using plasma and urine cell-free DNA methylomes. *Nature Medicine*, 26(7):1041–1043, jul 2020.

[163] W. M. Linehan, R. Srinivasan, and L. S. Schmidt. The genetic basis of kidney cancer: a metabolic disease. *Nature Reviews Urology*, 7(5):277–285, may 2010.

[164] P. H. Maxwell, M. S. Wlesener, G. W. Chang, et al. The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature*, 399(6733):271–275, may 1999.

[165] M. Ivan, K. Kondo, H. Yang, et al. HIFα targeted for VHL-mediated destruction by proline hydroxylation: Implications for O2 sensing. *Science*, 292(5516):464–468, apr 2001.

[166] K. Kondo, W. Y. Kim, M. Lechpammer, and W. G. Kaelin. Inhibition of HIF2α is sufficient to suppress pVHL-defective tumor growth. *PLoS Biology*, 1(3), dec 2003.

[167] A. Nagao, M. Kobayashi, S. Koyasu, et al. HIF-1-dependent reprogramming of glucose metabolic pathway of cancer cells and its therapeutic significance. *International Journal of Molecular Sciences*, 20(2), jan 2019.

[168] D. Hanahan and R. A. Weinberg. Hallmarks of cancer: the next generation. *Cell*, 144(5):646–74, mar 2011.

[169] K. Fischer, P. Hoffmann, S. Voelkl, et al. Inhibitory effect of tumor cell-derived lactic acid on human T cells. *Blood*, 109(9):3812–3819, may 2007.

[170] V. Estrella, T. Chen, M. Lloyd, et al. Acidity generated by the tumor microenvironment drives local invasion. *Cancer Research*, 73(5):1524–1535, mar 2013.

[171] M. V. Liberti and J. W. Locasale. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends in Biochemical Sciences*, 41(3):211–218, mar 2016.

[172] O. R. Colegio. Lactic acid polarizes macrophages to a tumor-promoting state. *OncoImmunology*, 5(3), mar 2016.

[173] B. Z. Qian, J. Li, H. Zhang, et al. CCL2 recruits inflammatory monocytes to facilitate breast-tumour metastasis. *Nature*, 475(7355):222–225, jul 2011.

[174] R. Noy and J. W. Pollard. Tumor-Associated Macrophages: From Mechanisms to Therapy. *Immunity*, 41(1):49–61, jul 2014.

[175] N. Pandey, V. Lanke, and P. K. Vinod. Network-based metabolic characterization of renal cell carcinoma. *Scientific Reports*, 10(1):1–13, dec 2020.

[176] A. Lameirinhas, V. Miranda-Gonçalves, R. Henrique, and C. Jerónimo. The complex interplay between metabolic reprogramming and epigenetic alterations in renal cell carcinoma. *Genes*, 10(4), apr 2019.

[177] K. Trpkov, O. Hes, A. Agaimy, et al. Fumarate hydratase-deficient renal cell carcinoma is strongly correlated with fumarate hydratase mutation and hereditary leiomyomatosis and renal cell carcinoma syndrome. *American Journal of Surgical Pathology*, 40(7):865–875, jul 2016.

[178] J. S. Isaacs, J. J. Yun, D. R. Mole, et al. HIF overexpression correlates with biallelic loss of fumarate hydratase in renal cancer: Novel role of fumarate in regulation of HIF stability. *Cancer Cell*, 8(2):143–153, aug 2005.

[179] A. A. Hakimi, E. Reznik, C.-H. Lee, et al. An Integrated Metabolic Atlas of Clear Cell Renal Cell Carcinoma. *Cancer Cell*, 29:104–116, 2016.

[180] L. Jing, J. M. Guigonis, D. Borchiellini, et al. LC-MS based metabolomic profiling for renal cell carcinoma histologic subtypes. *Scientific Reports*, 9(1):1–10, dec 2019.

[181] X. Liu, M. Zhang, X. Liu, et al. Urine Metabolomics for Renal Cell Carcinoma (RCC) Prediction: Tryptophan Metabolism as an Important Pathway in RCC. *Frontiers in Oncology*, 9:663, jul 2019.

[182] X. Liu, M. Zhang, X. Cheng, et al. LC-MS-Based Plasma Metabolomics and Lipidomics Analyses for Differential Diagnosis of Bladder Cancer and Renal Cell Carcinoma. *Frontiers in Oncology*, 10:717, may 2020.

[183] F. Gatto, I. Nookaew, and J. Nielsen. Chromosome 3p loss of heterozygosity is associated with a unique metabolic network in clear cell renal carcinoma. *Proceedings of the National Academy of Sciences of the United States of America*, 111(9):E866–75, mar 2014.

[184] M. R. Raspollini, F. Castiglione, G. Martignoni, et al. Unlike in clear cell renal cell carcinoma, KRAS is not mutated in multilocular cystic clear cell renal cell neoplasm of low potential. *Virchows Archiv*, 467(6):687–693, dec 2015.

[185] T. Mochizuki, E. Tsukamoto, Y. Kuge, et al. FDG Uptake and Glucose Transporter Subtype Expressions in Experimental Tumor and Inflammation Models. *Soc Nuclear Med*, 42(10):1551–1555, 2001.

[186] B. M. Gallagher, J. S. Fowler, N. I. Gutterson, et al. Metabolic Trapping as a Principle of Radiopharmaceutical Design: Some Factors Responsible for the Biodistribution of [18F] 2-Deoxy-2-Fluoro-D-Glucose. *Journal of Nuclear Medicine*, 19(10), 1978.

[187] A. Mahajan and G. Cook. Clinical applications of PET/CT in oncology. In *Basic Science of PET Imaging*, pages 429–450. Springer International Publishing, jan 2016.

[188] Y. Liu. The place of FDG PET/CT in renal cell carcinoma: Value and limitations. *Frontiers in Oncology*, 7(SEP):201, sep 2016.

[189] M. Takahashi, H. Kume, K. Koyama, et al. Preoperative evaluation of renal cell carcinoma by using 18F-FDG PET/CT. *Clinical Nuclear Medicine*, 40(12):936–940, 2015.

[190] E. H. Dibble, S. Kravets, S. Cheng, et al. Utility of FDG-PET/CT in Patients with Advanced Renal Cell Carcinoma with Osseous Metastases: Comparison with CT and 99mTc-MDP Bone Scan in a Prospective Clinical Trial. *Kidney Cancer*, 3(4):241–251, dec 2019.

[191] D. W. McRobbie, E. A. Moore, and M. J. Graves. *MRI from Picture to Proton*. Cambridge University Press, Cambridge, 2017.

[192] I. I. Rabi, J. R. Zacharias, S. Millman, and P. Kusch. A New Method of Measuring Nuclear Magnetic Moment. *Physical Review*, 53(4):318–318, feb 1938.

[193] K. G. de la Cruz-López, L. J. Castro-Muñoz, D. O. Reyes-Hernández, et al. Lactate in the Regulation of Tumor Microenvironment and Therapeutic Approaches. *Frontiers in Oncology*, 9, nov 2019.

[194] T. Martín Noguerol, J. Sánchez-González, J. P. Martínez Barbero, et al. Clinical Imaging of Tumor Metabolism with ^{1}H Magnetic Resonance Spectroscopy. *Magnetic Resonance Imaging Clinics of North America*, 24(1):57–86, feb 2016.

[195] G. Öz, J. R. Alger, P. B. Barker, et al. Clinical proton MR spectroscopy in central nervous system disorders. *Radiology*, 270(3):658–679, mar 2014.

[196] N. Mottet, J. Bellmunt, M. Bolla, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. *European Urology*, 71(4):618–629, apr 2017.

[197] S. Sevcenco, M. Krssak, D. Javor, et al. Diagnosis of renal tumors by in vivo proton magnetic resonance spectroscopy. *World Journal of Urology*, 33(1):17–23, mar 2014.

[198] H. A. Ali, M. J. Couch, R. Menezes, et al. Predictive Value of In Vivo MR Spectroscopy With Semilocalization by Adiabatic Selective Refocusing in Differentiating Clear Cell Renal Cell Carcinoma From Other Subtypes. *American Journal of Roentgenology*, 214(4):817–824, apr 2020.

[199] M. Tran, A. Latifoltojar, J. B. Neves, et al. First-in-human in vivo non-invasive assessment of intra-tumoral metabolic heterogeneity in renal cell carcinoma. *BJR case reports*, 5(3):20190003, sep 2019.

[200] S. Tang, M. V. Meng, J. B. Slater, et al. Metabolic imaging with hyperpolarized 13C pyruvate magnetic resonance imaging in patients with renal tumors—Initial experience. *Cancer*, 127(15):2693–2704, aug 2021.

[201] K. R. Keshari, R. Sriram, B. L. Koelsch, et al. Hyperpolarized ^{13}C-pyruvate magnetic resonance reveals rapid lactate export in metastatic renal cell carcinomas. *Cancer Research*, 73(2):529–538, jan 2013.

[202] F. A. Gallagher, R. Woitek, M. A. McLean, et al. Imaging breast cancer using hyperpolarized carbon-13 MRI. *Proceedings of the National Academy of Sciences of the United States of America*, 117(4):2092–2098, jan 2020.

[203] R. Woitek, M. A. McLean, A. B. Gill, et al. Hyperpolarized ^{13}C MRI of Tumor Metabolism Demonstrates Early Metabolic Response to Neoadjuvant Chemotherapy in Breast Cancer. *Radiology: Imaging Cancer*, 2(4):e200017, jul 2020.

[204] M. J. Albers, R. Bok, A. P. Chen, et al. Hyperpolarized ^{13}C lactate, pyruvate, and alanine: Noninvasive biomarkers for prostate cancer detection and grading. *Cancer Research*, 68(20):8607–8615, oct 2008.

[205] K. L. Granlund, S. S. Tee, H. A. Vargas, et al. Hyperpolarized MRI of Human Prostate Cancer Reveals Increased Lactate with Tumor Grade Driven by Monocarboxylate Transporter 1. *Cell Metabolism*, 31(1):105–114.e3, jan 2020.

[206] R. Sriram, M. Van Criekinge, J. De Los Santos, et al. Elevated tumor lactate and efflux in high-grade prostate cancer demonstrated by hyperpolarized ^{13}C magnetic resonance spectroscopy of prostate tissue slice cultures. *Cancers*, 12(3):537, mar 2020.

[207] V. Z. Miloushev, K. L. Granlund, R. Boltyanskiy, et al. Metabolic Imaging of the Human Brain with Hyperpolarized ^{13}C Pyruvate Demonstrates ^{13}C Lactate Production in Brain Tumor Patients. *Cancer Research*, page canres.0221.2018, may 2018.

[208] J. T. Grist, M. A. McLean, F. Riemer, et al. Quantifying normal human brain metabolism using hyperpolarized [1–^{13}C]pyruvate and magnetic resonance imaging. *NeuroImage*, 189:171–179, apr 2019.

[209] O. J. Rider, A. Apps, J. J. Miller, et al. Noninvasive In Vivo Assessment of Cardiac Metabolism in the Healthy and Diabetic Human Heart Using Hyperpolarized ^{13}C MRI. *Circulation Research*, 126(6):725–736, mar 2020.

[210] D. C. Sullivan, N. A. Obuchowski, L. G. Kessler, et al. Metrology Standards for Quantitative Imaging Biomarkers. *Radiology*, 277(3):813–825, dec 2015.

[211] G. Castellano, L. Bonilha, L. M. Li, and F. Cendes. Texture analysis of medical images. *Clinical Radiology*, 59(12):1061–9, dec 2004.

[212] R. J. Gillies, P. E. Kinahan, and H. Hricak. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*, 278(2):563–577, feb 2016.

[213] N. Hindman, L. Ngo, E. M. Genega, et al. Angiomyolipoma with Minimal Fat: Can It Be Differentiated from Clear Cell Renal Cell Carcinoma by Using Standard MR Techniques? *Radiology*, 265(2):468–477, nov 2012.

[214] S. Livens. Wavelets for texture analysis, an overview. In *6th International Conference on Image Processing and its Applications*, volume 1997, pages 581–585. IEE, 1997.

[215] N. Aggarwal and R. K. Agrawal. First and Second Order Statistics Features for Classification of Magnetic Resonance Brain Images. *Journal of Signal and Information Processing*, 03(02):146–153, may 2012.

[216] R. M. Haralick, I. Dinstein, K. Shanmugam, and I. Dinstein. Textural Features for Image Classification. SMC-3(6):610–621, nov 1973.

[217] C. Nikias and J. Mendel. Signal processing with higher-order spectra. *IEEE Signal Processing Magazine*, 10(3):10–37, jul 1993.

[218] F. Orlhac, C. Nioche, and I. Buvat. Technical Appendix Local Image Features Extraction LIFEx. 2016.

[219] K. Sasaguri, N. Takahashi, D. Gomez-Cardona, et al. Small (< 4 cm) renal mass: Differentiation of oncocytoma from renal cell carcinoma on biphasic contrast-enhanced CT. *American Journal of Roentgenology*, 205(5):999–1007, nov 2015.

[220] N. Takahashi, S. Leng, K. Kitajima, et al. Small (< 4 cm) renal masses: Differentiation of angiomyolipoma without visible fat from renal cell carcinoma using unenhanced and contrast-enhanced CT. *American Journal of Roentgenology*, 205(6):1194–1202, dec 2015.

[221] Z. Feng, P. Rong, P. Cao, et al. Machine learning-based quantitative texture analysis of CT images of small renal masses: Differentiation of angiomyolipoma without visible fat from renal cell carcinoma. *European Radiology*, 28(4):1625–1633, apr 2018.

[222] U. N. Hoang, S. Mojdeh Mirmomen, O. Meirelles, et al. Assessment of multiphasic contrast-enhanced MR textures in differentiating small renal mass subtypes. *Abdominal Radiology*, 43(12):3400–3409, dec 2018.

[223] S. P. Raman, Y. Chen, J. L. Schroeder, et al. CT texture analysis of renal masses: Pilot study using random forest classification for prediction of pathology. *Academic Radiology*, 21(12):1587–1596, dec 2014.

[224] H. S. Yu, J. Scalera, M. Khalid, et al. Texture analysis as a radiomic marker for differentiating renal tumors. *Abdominal Radiology*, 42(10):2470–2478, oct 2017.

[225] M. G. Lubner, N. Stabo, E. J. Abel, et al. CT textural analysis of large primary renal cell carcinomas: Pretreatment tumor heterogeneity correlates with histologic findings and clinical outcomes. *American Journal of Roentgenology*, 207(1):96–105, jul 2016.

[226] M. A. Haider, A. Vosough, F. Khalvati, et al. CT texture analysis: A potential tool for prediction of survival in patients with metastatic clear cell carcinoma treated with sunitinib. *Cancer Imaging*, 17(1):1–9, dec 2017.

[227] J. J. Ko, W. Xie, N. Kroeger, et al. The International Metastatic Renal Cell Carcinoma Database Consortium model as a prognostic tool in patients with metastatic renal cell carcinoma previously treated with first-line targeted therapy: a population-based study. *The Lancet Oncology*, 16(3):293–300, mar 2015.

[228] V. Goh, P. Nathan, J. K. Juttla, et al. Assessment of Response to Tyrosine Kinase Inhibitors in Metastatic Renal Cell Cancer: CT Texture as a Predictive Biomarker. *Radiology*, 261(1):165–171, oct 2011.

[229] R. Berenguer, M. d. R. Pastor-Juan, J. Canales-Vázquez, et al. Radiomics of CT Features May Be Nonreproducible and Redundant: Influence of CT Acquisition Parameters. *Radiology*, 288(2):172361, apr 2018.

[230] M. Shafiq-ul Hassan, G. G. Zhang, K. Latifi, et al. Intrinsic dependencies of CT radiomics features on voxel size and number of gray levels. *Medical Physiccs*, 44(3):1050–1062, mar 2017.

[231] K. Y. A. M. C. M. O. Kiryu;, H. Akai, D. Mackin, et al. Precision of quantitative computed tomography texture analysis using image filtering. *Medicine*, 96(21):e6993, may 2017.

[232] J. P. B. O'Connor, E. O. Aboagye, J. E. Adams, et al. Imaging biomarker roadmap for cancer studies. *Nature Reviews Clinical Oncology*, 14(3):169–186, mar 2017.

[233] L. P. Clarke, R. J. Nordstrom, H. Zhang, et al. The Quantitative Imaging Network: NCI's Historical Perspective and Planned Goals. *Translational Oncology*, 7(1):1–4, feb 2014.

[234] L. K. Shankar. The clinical evaluation of novel imaging methods for cancer management. *Nature Reviews Clinical Oncology*, 9(12):738–744, dec 2012.

[235] E. S. o. R. (ESR). ESR statement on the stepwise development of imaging biomarkers. *Insights into Imaging*, 4(2):147–152, apr 2013.

[236] P. Lambin, R. T. Leijenaar, T. M. Deist, et al. Radiomics: The bridge between medical imaging and personalized medicine. *Nature Reviews Clinical Oncology*, 14(12):749–762, dec 2017.

[237] A. Zwanenburg, M. Vallières, M. A. Abdalah, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology*, 295(2):328–338, mar 2020.

[238] M. G. Lubner, A. D. Smith, K. Sandrasegaran, et al. CT Texture Analysis: Definitions, Applications, Biologic Correlates, and Challenges. *RadioGraphics*, 37(5):1483–1503, sep 2017.

[239] K. A. Miles. How to use CT texture analysis for prognostication of non-small cell lung cancer. *Cancer Imaging*, 16(1):10, dec 2016.

[240] S. Ursprung, L. Beer, A. Bruining, et al. Radiomics of computed tomography and magnetic resonance imaging in renal cell carcinoma—a systematic review and meta-analysis. *European Radiology*, pages 1–9, feb 2020.

[241] S. Sanduleanu, H. C. Woodruff, E. E. C. de Jong, et al. Tracking tumor biology with radiomics: A systematic review utilizing a radiomics quality score. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*, 127(3):349–360, jun 2018.

[242] J. E. Park, H. S. Kim, D. Kim, et al. A systematic review reporting quality of radiomics research in neuro-oncology: Toward clinical utility and quality improvement using high-dimensional imaging features. *BMC Cancer*, 20(1), jan 2020.

[243] J. E. Park, D. Kim, H. S. Kim, et al. Quality of science and reporting of radiomics in oncologic studies: room for improvement according to radiomics quality score and TRIPOD statement. *European Radiology*, 30(1):523–536, jan 2020.

[244] B. Kocak, A. H. Yardimci, C. T. Bektas, et al. Textural differences between renal cell carcinoma subtypes: Machine learning-based quantitative computed tomography texture analysis with independent external validation. *European Journal of Radiology*, 107:149–157, oct 2018.

[245] N. Bharwani, M. E. Miquel, T. Powles, et al. Diffusion-weighted and multiphase contrast-enhanced MRI as surrogate markers of response to neoadjuvant sunitinib in metastatic renal cell carcinoma. *British Journal of Cancer*, 110(3):616–624, feb 2014.

[246] M. D. F. McInnes, D. Moher, B. D. Thombs, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies. *JAMA*, 319(4):388, jan 2018.

[247] P. F. Whiting, A. W. Rutjes, M. E. Westwood, et al. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Annals of Internal Medicine*, 155(8):529, oct 2011.

[248] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016.

[249] T. K. Koo and M. Y. Li. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2):155–163, jun 2016.

[250] D. Marasini, P. Quatto, and E. Ripamonti. Assessing the inter-rater agreement for ordinal data through weighted indexes. *Statistical Methods in Medical Research*, 25(6):2611–2633, dec 2016.

[251] A. R. Feinstein and D. V. Cicchetti. High agreement but low Kappa: I. the problems of two paradoxes. *Journal of Clinical Epidemiology*, 43(6):543–549, jan 1990.

[252] M. Maclure and W. C. Willett. Misinterpretation and Misuse of the Kappa Statistic. *American Journal of Epidemiology*, 126(2):161–169, aug 1987.

[253] P. Quatto and E. Ripamonti. *A Modification of Fleiss' Kappa in Case of Nominal and Ordinal Variables*. R package version 2.0.1, 2014.

[254] W. J. Conover. *Practical Nonparametric Statistics*. Wiley, New York, 3rd ed edition, 1999.

[255] J. P. T. Higgins and S. G. Thompson. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11):1539–1558, jun 2002.

[256] H.-y. Y. Wang, Z.-h. H. Su, X. Xu, et al. Dynamic Contrast-enhanced MR Imaging in Renal Cell Carcinoma: Reproducibility of Histogram Analysis on Pharmacokinetic Parameters. *Scientific Reports*, 6, jul 2016.

[257] L. Treanor, R. A. Frank, L. A. Cherpak, et al. Publication bias in diagnostic imaging: conference abstracts with positive conclusions are more likely to be published. *European Radiology*, 30(5):2964–2972, may 2020.

[258] P. Peduzzi, J. Concato, E. Kemper, et al. A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology*, 49(12):1373–1379, dec 1996.

[259] B. K. Park. Renal Angiomyolipoma Based on New Classification: How to Differentiate It From Renal Cell Carcinoma. *American Journal of Roentgenology*, pages 1–7, jan 2019.

[260] M. Shafiq-ul hassan, G. G. Zhang, D. C. Hunt, et al. Accounting for reconstruction kernel- induced variability in CT radiomic features using noise power spectra Accounting for reconstruction kernel-induced variability in CT radiomic features using noise power spectra. *Journal of Medical Imaging*, 5(1):1, dec 2017.

[261] B. Zhao, Y. Tan, W. Y. Tsai, et al. Exploring Variability in CT Characterization of Tumors: A Preliminary Phantom Study. *Translational Oncology*, 7(1):88–93, feb 2014.

[262] F. Deckers, B. De Foer, F. Van Mieghem, et al. Apparent diffusion coefficient measurements as very early predictive markers of response to chemotherapy in hepatic metastasis: A preliminary investigation of reproducibility and diagnostic value. *Journal of Magnetic Resonance Imaging*, 40(2):448–456, aug 2014.

[263] L. Heijmen, E. E. Ter Voert, I. D. Nagtegaal, et al. Diffusion-weighted MR imaging in liver metastases of colorectal cancer: Reproducibility and biological validation. *European Radiology*, 23(3):748–756, mar 2013.

[264] R. Klaassen, O. J. Gurney-Champion, J. W. Wilmink, et al. Repeatability and correlations of dynamic contrast enhanced and T2* MRI in patients with advanced pancreatic ductal adenocarcinoma. *Magnetic Resonance Imaging*, 50:1–9, jul 2018.

[265] H. Wang, Z. Su, H. Ye, et al. Reproducibility of Dynamic Contrast-Enhanced MRI in Renal Cell Carcinoma. *Medicine*, 94(37):e1529, sep 2015.

[266] A. L. Beers, Y.-F. Yen, K. E. Emblem, et al. Repeatability of ktrans derived from DCE-MRI in newly diagnosed glioblastoma across multiple baseline images and processing methods. In *Proc. Intl. Soc. Mag. Reson. Med.*, page 4172, 2017.

[267] E. M. Klawer, P. J. van Houdt, F. F. Simonis, et al. Improved repeatability of dynamic contrast-enhanced MRI using the complex MRI signal to derive arterial input functions: a test-retest study in prostate cancer patients. *Magnetic Resonance in Medicine*, 81(5):3358–3369, may 2019.

[268] L. Fournier, L. Costaridou, L. Bidaut, et al. Incorporating radiomics into clinical trials: expert consensus on considerations for data-driven compared to biologically driven quantitative biomarkers. *European Radiology*, pages 1–12, jan 2021.

[269] C. Y. Chen, T. H. Tsai, T. S. Jaw, et al. Diagnostic performance of split-bolus portal venous phase dual-energy CT urography in patients with hematuria. *American Journal of Roentgenology*, 206(5):1013–1022, may 2016.

[270] N. Morrison, S. Bryden, and A. F. Costa. Split vs. Single Bolus CT Urography: Comparison of Scan Time, Image Quality and Radiation Dose. *Tomography*, 7(2):210–218, may 2021.

[271] N. Heller, N. Sathianathen, A. Kalapara, et al. The KiTS19 Challenge Data: 300 Kidney Tumor Cases with Clinical Context, CT Semantic Segmentations, and Surgical Outcomes. ArXiv ID: 1904.00445v2. 2020.

[272] E. P. Le, L. Rundo, J. M. Tarkin, et al. Assessing robustness of carotid artery CT angiography radiomics in the identification of culprit lesions in cerebrovascular events. *Scientific Reports*, 11(1):3499, dec 2021.

[273] X. Li, P. S. Morgan, J. Ashburner, et al. The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *Journal of Neuroscience Methods*, 264:47–56, may 2016.

[274] R. M. Haralick. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5):786–804, 1979.

[275] C. Sun and W. G. Wee. Neighboring gray level dependence matrix for texture classification. *Computer Vision, Graphics and Image Processing*, 23(3):341–352, sep 1983.

[276] M. M. Galloway. Texture analysis using gray level run lengths. *Computer Graphics and Image Processing*, 4(2):172–179, jun 1975.

[277] G. Thibault, J. Angulo, and F. Meyer. Advanced statistical matrices for texture characterization: Application to cell classification. *IEEE Transactions on Biomedical Engineering*, 61(3):630–637, mar 2014.

[278] M. Amadasun and R. King. Texural Features Corresponding to Texural Properties. *IEEE Transactions on Systems, Man and Cybernetics*, 19(5):1264–1274, sep 1989.

[279] R. Brun and F. Rademakers. ROOT - An object oriented data analysis framework. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 389(1-2):81–86, apr 1997.

[280] M. Frelaut, P. du Rusquec, A. de Moura, et al. Pseudoprogression and Hyperprogression as New Forms of Response to Immunotherapy. *BioDrugs : clinical immunotherapeutics, biopharmaceuticals and gene therapy*, 34(4):463–476, aug 2020.

[281] H. C. Thoeny and B. D. Ross. Predicting and monitoring cancer treatment response with diffusion-weighted MRI. *Journal of Magnetic Resonance Imaging*, 32(1):2–16, jul 2010.

[282] S. K. Kang, A. Zhang, P. V. Pandharipande, et al. DWI for Renal Mass Characterization: Systematic Review and Meta-Analysis of Diagnostic Test Performance. *American Journal of Roentgenology*, 205(2):317–324, aug 2015.

[283] E. O. Stejskal. Use of Spin Echoes in a Pulsed Magnetic-Field Gradient to Study Anisotropic, Restricted Diffusion and Flow. *The Journal of Chemical Physics*, 43(10):3597–3603, nov 1965.

[284] D. G. Taylor and M. C. Bushell. The spatial mapping of translational diffusion coefficients by the NMR imaging technique. *Physics in Medicine and Biology*, 30(4):345–349, apr 1985.

[285] O. Dietrich, A. Biffar, A. Baur-Melnyk, and M. F. Reiser. Technical aspects of MR diffusion imaging of the body. *European Journal of Radiology*, 76(3):314–322, dec 2010.

[286] D. Le Bihan, E. Breton, D. Lallemand, et al. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology*, 161(2):401–407, nov 1986.

[287] A. Bizzi, A. Righini, R. Turner, et al. MR of diffusion slowing in global cerebral ischemia. *AJNR. American journal of neuroradiology*, 14(6):1347–54, apr 2001.

[288] D. Le Bihan. Apparent Diffusion Coefficient and Beyond: What Diffusion MR Imaging Can Tell Us about Tissue Structure. *Radiology*, 268(2):318–322, aug 2013.

[289] C. P. Corona-Villalobos, L. Pan, V. G. Halappa, et al. Agreement and Reproducibility of Apparent Diffusion Coefficient Measurements of Dual-b-Value and Multi-b-Value Diffusion-Weighted Magnetic Resonance Imaging at 1.5 Tesla in Phantom and in Soft Tissues of the Abdomen. *Journal of Computer Assisted Tomography*, 37(1):46–51, jan 2013.

[290] S. Y. Park, C. K. Kim, B. K. Park, and G. Y. Kwon. Comparison of Apparent Diffusion Coefficient Calculation Between Two-Point and Multipoint b Value Analyses in Prostate Cancer and Benign Prostate Tissue at 3 T: Preliminary Experience. *American Journal of Roentgenology*, 203(3):W287–W294, sep 2014.

[291] B. Paudyal, P. Paudyal, Y. Tsushima, et al. The role of the ADC value in the characterisation of renal carcinoma by diffusion-weighted MRI. *The British Journal of Radiology*, 83(988):336–43, apr 2010.

[292] H. Wang, L. Cheng, X. Zhang, et al. Renal Cell Carcinoma: Diffusion-weighted MR Imaging for Subtype Differentiation at 3.0 T. *Radiology*, 257(1):135–143, oct 2010.

[293] B. Taouli, R. K. Thakur, L. Mannelli, et al. Renal Lesions: Characterization with Diffusion-weighted Imaging versus Contrast-enhanced MR Imaging. *Radiology*, 251(2):398–407, may 2009.

[294] K. Sandrasegaran, C. P. Sundaram, R. Ramaswamy, et al. Usefulness of Diffusion-Weighted Imaging in the Evaluation of Renal Masses. *American Journal of Roentgenology*, 194(2):438–445, feb 2010.

[295] A. Yamamoto, T. Tamada, K. Ito, et al. Differentiation of subtypes of renal cell carcinoma: dynamic contrast-enhanced magnetic resonance imaging versus diffusion-weighted magnetic resonance imaging. *Clinical Imaging*, 41:53–58, jan 2017.

[296] M. Tordjman, R. Mali, G. Madelin, et al. Diagnostic test accuracy of ADC values for identification of clear cell renal cell carcinoma: systematic review and meta-analysis. *European Radiology*, 30(7):4023–4038, jul 2020.

[297] A. B. Rosenkrantz, B. E. Niver, E. F. Fitzgerald, et al. Utility of the Apparent Diffusion Coefficient for Distinguishing Clear Cell Renal Cell Carcinoma of Low and High Nuclear Grade. *American Journal of Roentgenology*, 195(5):W344–W351, nov 2010.

[298] A. Goyal, R. Sharma, A. S. Bhalla, et al. Diffusion-weighted MRI in renal cell carcinoma: A surrogate marker for predicting nuclear grade and histological subtype. *Acta Radiologica*, 53(3):349–358, apr 2012.

[299] F. Farhadi, M. Nikpanah, A. K. Paschall, et al. Clear Cell Renal Cell Carcinoma Growth Correlates with Baseline Diffusion-weighted MRI in Von Hippel–Lindau Disease. *Radiology*, 295(3):583–590, jun 2020.

[300] J. H. Jensen, J. A. Helpern, A. Ramani, et al. Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine*, 53(6):1432–1440, jun 2005.

[301] Y. Dai, Q. Yao, G. Wu, et al. Characterization of clear cell renal cell carcinoma with diffusion kurtosis imaging: correlation between diffusion kurtosis parameters and tumor cellularity. *NMR in Biomedicine*, 29(7):873–881, jul 2016.

[302] J. Ye, Q. Xu, S. A. Wang, et al. Quantitative Evaluation of Intravoxel Incoherent Motion and Diffusion Kurtosis Imaging in Assessment of Pathological Grade of Clear Cell Renal Cell Carcinoma. *Academic Radiology*, 27(7):e176–e182, jul 2020.

[303] A. B. Rosenkrantz, A. R. Padhani, T. L. Chenevert, et al. Body diffusion kurtosis imaging: Basic principles, applications, and considerations for clinical practice. *Journal of Magnetic Resonance Imaging*, 42(5):1190–1202, nov 2015.

[304] D. Le Bihan, E. Breton, D. Lallemand, et al. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology*, 168(2):497–505, aug 1988.

[305] M. Iima and D. Le Bihan. Clinical Intravoxel Incoherent Motion and Diffusion MR Imaging: Past, Present, and Future. *Radiology*, 278(1):13–32, jan 2016.

[306] M. Müller, P. Prasad, and R. Edelman. Can the IVIM model be used for renal perfusion imaging? *European Journal of Radiology*, 26(3):297–303, feb 1998.

[307] H. C. Thoeny, F. De Keyzer, R. H. Oyen, and R. R. Peeters. Diffusion-weighted MR Imaging of Kidneys in Healthy Volunteers and Patients with Parenchymal Diseases: Initial Experience. *Radiology*, 235(3):911–917, jun 2005.

[308] H. Chandarana, S. K. Kang, S. Wong, et al. Diffusion-Weighted Intravoxel Incoherent Motion Imaging of Renal Tumors With Histopathologic Correlation. *Investigative Radiology*, 47(12):688–696, dec 2012.

[309] Y. Ding, M. Zeng, S. Rao, et al. Comparison of Biexponential and Monoexponential Model of Diffusion-Weighted Imaging for Distinguishing between Common Renal Cell Carcinoma and Fat Poor Angiomyolipoma. *Korean Journal of Radiology*, 17(6):853–863, nov 2016.

[310] B. G. S. H. B. P. Stoffel;Hersh, E. E. Sigmund, W. C. Huang, et al. Subtype Differentiation of Renal Tumors Using Voxel-Based Histogram Analysis of Intravoxel Incoherent Motion Parameters. *Investigative Radiology*, 50(3):144–152, mar 2015.

[311] K. L. Eales, K. E. R. Hollinshead, and D. A. Tennant. Hypoxia and metabolic adaptation of cancer cells. *Oncogenesis*, 5(1):e190–e190, jan 2016.

[312] N. Lawrentschuk, A. M. Poon, S. S. Foo, et al. Assessing regional hypoxia in human renal tumours using 18F-fluoromisonidazole positron emission tomography. *BJU International*, 96(4):540–546, sep 2005.

[313] D. Hernando, J. H. Kramer, and S. B. Reeder. Multipeak fat-corrected complex R2* relaxometry: theory, optimization, and clinical validation. *Magnetic resonance in medicine*, 70(5):1319–31, nov 2013.

[314] P. V. Prasad, R. R. Edelman, and F. H. Epstein. Noninvasive evaluation of intrarenal oxygenation with BOLD MRI. *Circulation*, 94(12):3271–5, dec 1996.

[315] M. Notohamiprodjo, M. Staehler, N. Steiner, et al. No Title. *Academic Radiology*, 20(6):685–693, jun 2013.

[316] G.-Y. Wu, S.-T. Suo, Q. Lu, et al. The value of blood oxygenation level-dependent (BOLD) MR imaging in differentiation of renal solid mass and grading of renal cell carcinoma (RCC): analysis based on the largest cross-sectional area versus the entire whole tumour. *PLoS one*, 10(4):e0123431, apr 2015.

[317] Y. A. Choi, C. K. Kim, S. Y. Park, et al. Subtype Differentiation of Renal Cell Carcinoma Using Diffusion-Weighted and Blood Oxygenation Level–Dependent MRI. *American Journal of Roentgenology*, 203(1):W78–W84, jul 2014.

[318] J. P. B. O'Connor, P. S. Tofts, K. A. Miles, et al. Dynamic contrast-enhanced imaging techniques: CT and MRI. *The British Journal of Radiology*, 84(2):S112–S120, dec 2011.

[319] F. Khalifa, A. Soliman, A. El-Baz, et al. Models and methods for analyzing DCE-MRI: A review. *Medical Physics*, 41(12):124301, dec 2014.

[320] K. S. S. Lawrence and T.-Y. Lee. An Adiabatic Approximation to the Tissue Homogeneity Model for Water Exchange in the Brain: I Theoretical Derivation. *Journal of Cerebral Blood Flow and Metabolism*, 18(12):1365–1377, dec 1998.

[321] L. E. Kershaw and H.-L. M. Cheng. Temporal resolution and SNR requirements for accurate DCE-MRI data analysis using the AATH model. *Magnetic Resonance in Medicine*, 64(6):1772–1780, dec 2010.

[322] P. S. Tofts. Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. *Journal of magnetic resonance imaging*, 7(1):91–101, jan 1997.

[323] P. S. Tofts and A. G. Kermode. Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. Fundamental concepts. *Magnetic Resonance in Medicine*, 17(2):357–67, feb 1991.

[324] J. C. Miller, H. H. Pien, D. Sahani, et al. Imaging Angiogenesis: Applications and Potential for Drug Development. *JNCI*, 97(3):172–187, feb 2005.

[325] H.-y. Wang, Z.-h. Su, X. Xu, et al. Dynamic Contrast-enhanced MRI in Renal Tumors: Common Subtype Differentiation using Pharmacokinetics. *Scientific Reports*, 7(1):3117, dec 2017.

[326] K. T. Flaherty, M. A. Rosen, D. F. Heitjan, et al. Pilot study of DCE-MRI to predict progression-free survival with sorafenib therapy in renal cell carcinoma. *Cancer Biology and Therapy*, 7(4):496–501, apr 2008.

[327] R. S. Lanzman, P. M. Robson, M. R. Sun, et al. Arterial Spin-labeling MR Imaging of Renal Masses: Correlation with Histopathologic Findings. *Radiology*, 265(3):799–808, dec 2012.

[328] A. Rodriguez-Vida, T. E. Hutson, J. Bellmunt, and M. H. Strijbos. New treatment options for metastatic renal cell carcinoma. *ESMO Open*, 2(2):e000185, jun 2017.

[329] V. Grünwald, R. R. McKay, K. M. Krajewski, et al. Depth of remission is a prognostic factor for survival in patients with metastatic renal cell carcinoma. *European Urology*, 67(5):952–958, may 2015.

[330] D. Huang, Y. Ding, Y. Li, et al. Sunitinib acts primarily on tumor endothelium rather than tumor cells to inhibit the growth of renal cell carcinoma. *Cancer Research*, 70(3):1053–1062, feb 2010.

[331] J.-V. Gaustad, T. G. Simonsen, M. N. Leinaas, and E. K. Rofstad. Sunitinib treatment does not improve blood supply but induces hypoxia in human melanoma xenografts. *BMC Cancer*, 12(1):388, dec 2012.

[332] S. Welsh, K. Fife, A. Matakidou, et al. A phase II clinical study evaluating the efficacy and safety of neoadjuvant and adjuvant sunitinib in previously untreated patients with metastatic renal cell carcinoma. *Journal of Clinical Oncology*, 35(15):16087–16087, may 2017.

[333] National Cancer Institute. Common Terminology Criteria for Adverse Events (CTCAE). Available at: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm [Accessed: 28.01.2021].

[334] G. J. M. Parker, C. Roberts, A. Macdonald, et al. Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced MRI. *Magnetic Resonance in Medicine*, 56(5):993–1000, nov 2006.

[335] P. S. Tofts, G. Brix, D. L. Buckley, et al. Estimating kinetic parameters from dynamic contrast-enhanced T1- weighted MRI of a diffusable tracer: Standardized quantities and symbols. *Journal of Magnetic Resonance Imaging*, 10(3):223–232, sep 1999.

[336] T. Heye, E. M. Merkle, C. S. Reiner, et al. Reproducibility of Dynamic Contrast-enhanced MR Imaging. Part II. Comparison of Intra- and Interobserver Variability with Manual Region of Interest Placement versus Semiautomatic Lesion Segmentation and Histogram Analysis. *Radiology*, 266(3):812–821, mar 2013.

[337] S. J. Welsh, N. Thompson, A. Warren, et al. Dynamic biomarker and imaging changes from a phase II study of pre- and post-surgical sunitinib. *BJU international*, nov 2021.

[338] M. Shibuya. Vascular Endothelial Growth Factor (VEGF) and Its Receptor (VEGFR) Signaling in Angiogenesis: A Crucial Target for Anti- and Pro-Angiogenic Therapies. *Genes and Cancer*, 2(12):1097–1105, dec 2011.

[339] N. Lassau, S. Koscielny, L. Albiges, et al. Metastatic Renal Cell Carcinoma Treated with Sunitinib: Early Evaluation of Treatment Response Using Dynamic Contrast-Enhanced Ultrasonography. *Clinical Cancer Research*, 16(4):1216–1225, feb 2010.

[340] R. Williams, J. M. Hudson, B. A. Lloyd, et al. Dynamic microbubble contrast-enhanced US to measure tumor response to targeted therapy: A proposed clinical protocol with results from renal cell carcinoma patients receiving antiangiogenic therapy. *Radiology*, 260(2):581–590, aug 2011.

[341] N. Lassau, J. Bonastre, M. Kind, et al. Validation of dynamic contrast-enhanced ultrasound in predicting outcomes of antiangiogenic therapy for solid tumors: The French multicenter support for innovative and expensive techniques study. *Investigative Radiology*, 49(12):794–800, dec 2014.

[342] S. M. Crusz, Y. Z. Tang, S. J. Sarker, et al. Heterogeneous response and progression patterns reveal phenotypic heterogeneity of tyrosine kinase inhibitor response in metastatic renal cell carcinoma. *BMC Medicine*, 14(1):185, nov 2016.

[343] I. Desar, E. ter Voert, T. Hambrock, et al. Functional MRI techniques demonstrate early vascular changes in renal cell cancer patients treated with sunitinib: a pilot study. *Cancer Imaging*, 11(1):259, jan 2011.

[344] Q. Yin, S.-C. C. Hung, L. Wang, et al. Associations between Tumor Vascularity, Vascular Endothelial Growth Factor Expression and PET/MRI Radiomic Signatures in Primary Clear-Cell–Renal-Cell-Carcinoma: Proof-of-Concept Study. *Scientific Reports*, 7(March):43356, mar 2017.

[345] V. F. Marshall, R. G. Middleton, G. R. Holswade, and E. I. Goldsmith. Surgery for renal cell carcinoma in the vena cava. *The Journal of Urology*, 103(4):414–420, apr 1970.

[346] R. J. Neves and H. Zincke. Surgical Treatment of Renal Cancer with Vena Cava Extension. *British Journal of Urology*, 59(5):390–395, mar 1987.

[347] A. Moinzadeh and J. A. Libertino. Prognostic significance of tumor thrombus level in patients with renal cell carcinoma and venous tumor thrombus extension. Is all T3b the same? *Journal of Urology*, 171(2 I):598–601, 2004.

[348] B. Wagner, J. J. Patard, A. Méjean, et al. Prognostic Value of Renal Vein and Inferior Vena Cava Involvement in Renal Cell Carcinoma. *European Urology*, 55(2):452–460, feb 2009.

[349] A. C. Reese, J. M. Whitson, and M. V. Meng. Natural history of untreated renal cell carcinoma with venous tumor thrombus. *Urologic Oncology: Seminars and Original Investigations*, 31(7):1305–1309, oct 2013.

[350] J. M. Whitson, A. C. Reese, and M. V. Meng. Factors associated with surgery in patients with renal cell carcinoma and venous tumor thrombus. *BJU International*, 107(5):729–734, mar 2011.

[351] O. Rodriguez Faba, E. Linares, D. Tilki, et al. Impact of Microscopic Wall Invasion of the Renal Vein or Inferior Vena Cava on Cancer-specific Survival in Patients with Renal Cell Carcinoma and Tumor Thrombus: A Multi-institutional Analysis from the International Renal Cell Carcinoma-Venous Thrombus Co. *European Urology Focus*, 4(3):435–441, apr 2018.

[352] A. C. Novick and S. B. Pontes. *Stewart's operative urology: Vol1: The Kidneys, Adrenal Glands and retroperitonium*. Williams & Wilkins, Philadelphia, 1989.

[353] G. Ciancio, S. P. Shirodkar, M. S. Soloway, et al. Renal Carcinoma With Supradiaphragmatic Tumor Thrombus: Avoiding Sternotomy and Cardiopulmonary Bypass. *Annals of Thoracic Surgery*, 89(2):505–510, feb 2010.

[354] A. Mandhani, N. Patidar, P. Aga, et al. A new classification of inferior vena cava thrombus in renal cell carcinoma could define the need for cardiopulmonary or venovenous bypass. *Indian Journal of Urology*, 31(4):327–332, oct 2015.

[355] B. C. Leibovich, C. M. Lohse, J. C. Cheville, et al. Renal Cell Carcinoma with Inferior Vena Cava Extension: Can Classification Be Optimized to Predict Perioperative Outcomes? *Kidney Cancer*, 4(2):111–115, may 2020.

[356] E. H. Lambert, P. M. Pierorazio, A. Shabsigh, et al. Prognostic Risk Stratification and Clinical Outcomes in Patients Undergoing Surgical Treatment for Renal Cell Carcinoma with Vascular Tumor Thrombus. *Urology*, 69(6):1054–1058, jun 2007.

[357] S. A. Boorjian, S. Sengupta, and M. L. Blute. Renal cell carcinoma: Vena caval involvement. *BJU International*, 99(5):1239–1244, may 2007.

[358] E. J. Abel, R. H. Thompson, V. Margulis, et al. Perioperative outcomes following surgical resection of renal cell carcinoma with inferior vena cava thrombus extending above the hepatic veins: A contemporary multicenter experience. *European Urology*, 66(3):584–592, sep 2014.

[359] P. Toren, R. Abouassaly, N. Timilshina, et al. Results of a national population-based study of outcomes of surgery for renal tumors associated with inferior vena cava thrombus. *Urology*, 82(3):572–578, sep 2013.

[360] M. Blute, B. Leibovich, C. Lohse, et al. The Mayo Clinic experience with surgical management, complications and outcome for patients with renal cell carcinoma and venous tumour thrombus. *BJU international*, 94(1):33–41, jul 2004.

[361] C. Gratzke, M. Seitz, F. Bayrle, et al. Quality of life and perioperative outcomes after retroperitoneoscopic radical nephrectomy (RN), open RN and nephron-sparing surgery in patients with renal cell carcinoma. *BJU International*, 104(4):470–475, aug 2009.

[362] R. C. Hsu, T. Salika, J. Maw, et al. Influence of hospital volume on nephrectomy mortality and complications: A systematic review and meta-analysis stratified by surgical type. *BMJ Open*, 7(9):e016833, sep 2017.

[363] C. J. Wallis, G. Bjarnason, J. Byrne, et al. Morbidity and Mortality of Radical Nephrectomy for Patients With Disseminated Cancer: An Analysis of the National Surgical Quality Improvement Program Database. *Urology*, 95:95–102, sep 2016.

[364] S. P. Psutka, S. A. Boorjian, R. H. Thompson, et al. Clinical and radiographic predictors of the need for inferior vena cava resection during nephrectomy for patients with renal cell carcinoma and caval tumour thrombus. *BJU International*, 116(3):388–396, sep 2015.

[365] L. C. Adams, B. Ralla, Y. N. Y. Bender, et al. Renal cell carcinoma with venous extension: Prediction of inferior vena cava wall invasion by MRI. *Cancer Imaging*, 18(1), may 2018.

[366] S. P. Psutka and B. C. Leibovich. Management of inferior vena cava tumor thrombus in locally advanced renal cell carcinoma. *Therapeutic Advances in Urology*, 7(4):216–229, aug 2015.

[367] N. G. Cost, S. E. Delacroix, J. P. Sleeper, et al. The impact of targeted molecular therapies on the level of renal cell carcinoma vena caval tumor thrombus. *European Urology*, 59(6):912–918, jun 2011.

[368] P. Bigot, T. Fardoun, J. C. Bernhard, et al. Neoadjuvant targeted molecular therapies in patients undergoing nephrectomy and inferior vena cava thrombectomy: Is it useful? *World Journal of Urology*, 32(1):109–114, feb 2014.

[369] C. A. Field, B. H. Cotta, J. Jimenez, et al. Neoadjuvant Sunitinib Decreases Inferior Vena Caval Thrombus Size and Is Associated With Improved Oncologic Outcomes: A Multicenter Comparative Analysis. *Clinical Genitourinary Cancer*, 17(3):505–512, jun 2019.

[370] D. D. Hu-Lowe, H. Y. Zou, M. L. Grazzini, et al. Nonclinical antiangiogenesis and antitumor activities of axitinib (AG-013736), an oral, potent, and selective inhibitor of vascular endothelial growth factor receptor tyrosine kinases 1, 2, 3. *Clinical Cancer Research*, 14(22):7272–7283, nov 2008.

[371] T. Powles. Recent eUpdate to the ESMO Clinical Practice Guidelines on renal cell carcinoma on cabozantinib and nivolumab for first-line clear cell renal cancer: Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up 1. *Annals of Oncology*, 32(3):422–423, mar 2021.

[372] B. Escudier and M. Gore. Axitinib for the management of metastatic renal cell carcinoma. *Drugs in R and D*, 11(2):113–126, jun 2011.

[373] O. Rixe, R. M. Bukowski, M. D. Michaelson, et al. Axitinib treatment in patients with cytokine-refractory metastatic renal-cell cancer: a phase II study. *Lancet Oncology*, 8(11):975–984, nov 2007.

[374] W. Cai, B. Cai, J. Zhou, et al. Comparison of efficacy and safety among axitinib, sunitinib, and sorafenib as neoadjuvant therapy for renal cell carcinoma: A retrospective study, oct 2019.

[375] T. E. Hutson, V. Lesovoy, S. Al-Shukri, et al. Axitinib versus sorafenib as first-line therapy in patients with metastatic renal-cell carcinoma: A randomised open-label phase 3 trial. *The Lancet Oncology*, 14(13):1287–1294, dec 2013.

[376] J. A. Karam, C. E. Devine, D. L. Urbauer, et al. Phase 2 trial of neoadjuvant axitinib in patients with locally advanced nonmetastatic clear cell renal cell carcinoma. *European Urology*, 66(5):874–880, nov 2014.

[377] N. J. Stewart and S. Matsumoto. Biomedical applications of the dynamic nuclear polarization and parahydrogen induced polarization techniques for hyperpolarized ^{13}C MR imaging. *Magnetic Resonance in Medical Sciences*, 20(1):1–17, mar 2021.

[378] R. Motzer, B. Alekseev, S.-Y. Rha, et al. Lenvatinib plus Pembrolizumab or Everolimus for Advanced Renal Cell Carcinoma. *New England Journal of Medicine*, 384(14):1289–1300, feb 2021.

[379] T. K. Choueiri, T. Powles, M. Burotto, et al. Nivolumab plus Cabozantinib versus Sunitinib for Advanced Renal-Cell Carcinoma. *New England Journal of Medicine*, 384(9):829–841, mar 2021.

[380] G. de Anda-Jáuregui and E. Hernández-Lemus. Computational Oncology in the Multi-Omics Era: State of the Art. *Frontiers in Oncology*, 10:423, apr 2020.

[381] W. Du and O. Elemento. Cancer systems biology: Embracing complexity to develop better anticancer therapeutic strategies. *Oncogene*, 34(25):3215–3225, jun 2015.

[382] T. Okegawa, M. Morimoto, S. Nishizawa, et al. Intratumor Heterogeneity in Primary Kidney Cancer Revealed by Metabolic Profiling of Multiple Spatially Separated Samples within Tumors. *EBioMedicine*, 19:31–38, may 2017.

[383] D. L. Longo. Tumor Heterogeneity and Personalized Medicine. *New England Journal of Medicine*, 366(10):956–957, mar 2012.

[384] L. Beer, P. Martin-Gonzalez, M. Delgado-Ortet, et al. Ultrasound-guided targeted biopsies of CT-based radiomic tumour habitats: technical development and initial experience in metastatic ovarian cancer. *European Radiology*, 31(6):3765–3772, jun 2021.

[385] V. Shah, T. Pohida, B. Turkbey, et al. A method for correlating in vivo prostate magnetic resonance imaging and histopathology using individualized magnetic resonance -based molds. *Review of Scientific Instruments*, 80(10), oct 2009.

[386] D. K. Dwivedi, Y. Chatzinoff, Y. Zhang, et al. Development of a Patient-specific Tumor Mold Using Magnetic Resonance Imaging and 3-Dimensional Printing Technology for Targeted Tissue Procurement and Radiomics Analysis of Renal Masses. *Urology*, 112:209–214, feb 2018.

[387] B. Weigelt, H. A. Vargas, P. Selenica, et al. Radiogenomics Analysis of Intratumor Heterogeneity in a Patient With High-Grade Serous Ovarian Cancer. *JCO Precision Oncology*, 3(3):1–9, dec 2019.

[388] A. Y. Warren, H. D. D. Griffiths, and S. Fleming. *Standards and datasets for reporting cancers Dataset for histopathological reporting of adult renal parenchyma neoplasms*. The Royal College of Pathologists, London, 2017.

[389] K. Trpkov, D. J. Grignon, S. M. Bonsib, et al. Handling and staging of renal cell carcinoma: The International Society of Urological Pathology Consensus (ISUP) conference recommendations. *American Journal of Surgical Pathology*, 37(10):1505–1517, oct 2013.

[390] P. A. Yushkevich, J. Piven, H. C. Hazlett, et al. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage*, 31(3):1116–1128, jul 2006.

[391] J. O. Deasy, A. I. Blanco, and V. H. Clark. CERR: A computational environment for radiotherapy research. *Medical Physics*, 30(5):979–985, may 2003.

[392] T. Lewiner, H. Lopes, A. W. Vieira, and G. Tavares. Efficient Implementation of Marching Cubes' Cases with Topological Guarantees. *Journal of Graphics Tools*, 8(2):1–15, jan 2003.

[393] P. Cignoni, M. Callieri, M. Corsini, et al. *MeshLab: an Open-Source Mesh Processing Tool*. The Eurographics Association, Aire-la-Ville, Switzerland, 2008.

[394] S. Napel, W. Mu, B. V. Jardim-Perassi, et al. Quantitative imaging of cancer in the postgenomic era: Radio(geno)mics, deep learning, and habitats. *Cancer*, 124(24):4633–4649, dec 2018.

[395] T. Tran, C. P. Sundaram, C. D. Bahler, et al. Correcting the shrinkage effects of formalin fixation and tissue processing for renal tumors: Toward standardization of pathological reporting of tumor size. *Journal of Cancer*, 6(8):759–766, jul 2015.

[396] F. Commandeur, O. Acosta, A. Simon, et al. Prostate whole-mount histology reconstruction and registration to MRI for correlating in-vivo observations with biological findings. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, volume 11, pages 2399–2402. Institute of Electrical and Electronics Engineers Inc., nov 2015.

[397] J. Y. Lee, C. K. Kim, D. Choi, and B. K. Park. Volume doubling time and growth rate of renal cell carcinoma determined by helical CT: A single-institution experience. *European Radiology*, 18(4):731–737, apr 2008.

[398] S. Ozono, N. Miyao, T. Igarashi, et al. Tumor doubling time of renal cell carcinoma measured by CT: Collaboration of Japanese Society of Renal Cancer. *Japanese Journal of Clinical Oncology*, 34(2):82–85, feb 2004.

[399] N. S. Vasudev, M. Hutchinson, S. Trainor, et al. UK Multicenter Prospective Evaluation of the Leibovich Score in Localized Renal Cell Carcinoma: Performance has Altered Over Time. *Urology*, 136:162–168, feb 2020.

[400] J. Dagher, B. Delahunt, N. Rioux-Leclercq, et al. Clear cell renal cell carcinoma: validation of World Health Organization/International Society of Urological Pathology grading. *Histopathology*, 71(6):918–925, dec 2017.

[401] R. G. Jones and C. B. Thompson. Tumor suppressors and cell metabolism: a recipe for cancer growth. *Genes and Development*, 23(5):537–548, mar 2009.

[402] C. Horiuchi, M. Tsukuda, T. Taguchi, et al. Correlation between FDG-PET findings and GLUT1 expression in salivary gland pleomorphic adenomas. *Annals of Nuclear Medicine*, 22(8):693–698, oct 2008.

[403] A. Kiani, A. Esquevin, N. Lepareur, et al. Main applications of hybrid PET-MRI contrast agents: a review. *Contrast Media and Molecular Imaging*, mar 2015.

[404] N. M. Long and C. S. Smith. Causes and imaging features of false positives and false negatives on FDG-PET/CT in oncologic imaging. *Insights into Imaging*, 2(6):679–698, dec 2011.

[405] T. H. Witney, M. I. Kettunen, S. E. Day, et al. A comparison between radiolabeled fluorodeoxyglucose uptake and hyperpolarized ^{13}C-labeled pyruvate utilization as methods for detecting tumor response to treatment. *Neoplasia*, 11(6):574–82, jun 2009.

[406] R. W. Prost. Magnetic Resonance Spectroscopic Imaging. *Comprehensive Biomedical Physics*, 3:331–345, jan 2014.

[407] S. K. Gujar, S. Maheshwari, I. Björkman-Burtscher, and P. C. Sundgren. Magnetic resonance spectroscopy. *Journal of neuro-ophthalmology*, 25(3):217–26, sep 2005.

[408] R. Faghihi, B. Zeinali-Rafsanjani, M.-A. Mosleh-Shirazi, et al. Magnetic Resonance Spectroscopy and its Clinical Applications: A Review. *Journal of Medical Imaging and Radiation Sciences*, 48(3):233–253, sep 2017.

[409] R. M. Silverstein, F. X. Webster, D. J. Kiemle, and D. L. Bryce. *Spectrometric identification of organic compounds*. John Wiley & Sons, Hoboken, NJ, 2015.

[410] A. Abragam and M. Goldman. Principles of dynamic nuclear polarisation. *Reports on Progress in Physics*, 41(3):395–467, mar 1978.

[411] S. J. Nelson, D. Vigneron, J. Kurhanewicz, et al. DNP-Hyperpolarized ^{13}C Magnetic Resonance Metabolic Imaging for Cancer Applications. *Applied magnetic resonance*, 34(3-4):533–544, mar 2008.

[412] K. Golman, R. in 't Zandt, and M. Thaning. Real-time metabolic imaging. *Proceedings of the National Academy of Sciences*, 103(30):11270–11275, jul 2006.

[413] L. Lumata, M. E. Merritt, C. R. Malloy, et al. Impact of Gd 3+ on DNP of [1-^{13}C]Pyruvate Doped with Trityl OX063, BDPA, or 4-Oxo-TEMPO. *The Journal of Physical Chemistry A*, 116(21):5129–5138, may 2012.

[414] J. H. Ardenkjaer-Larsen, B. Fridlund, A. Gram, et al. Increase in signal-to-noise ratio of more than 10,000 times in liquid-state NMR. *Proceedings of the National Academy of Sciences of the United States of America*, 100(18):10158–63, sep 2003.

[415] K. Golman, J. H. Ardenkjaer-Larsen, J. S. Petersson, et al. Molecular imaging with endogenous substances. *Proceedings of the National Academy of Sciences of the United States of America*, 100(18):10435–9, sep 2003.

[416] S. J. Nelson, J. Kurhanewicz, D. B. Vigneron, et al. Metabolic imaging of patients with prostate cancer using hyperpolarized [1-^{13}C]pyruvate. *Science Translational Medicine*, 5(198):198ra108, aug 2013.

[417] C. J. Daniels, M. A. McLean, R. F. Schulte, et al. A comparison of quantitative methods for clinical imaging with hyperpolarized ^{13}C-pyruvate. *NMR in Biomedicine*, 29(4):387–399, apr 2016.

[418] F. Hirschhaeuser, U. G. A. Sattler, and W. Mueller-Klieser. Lactate: a metabolic key player in cancer. *Cancer Research*, 71(22):6921–5, nov 2011.

[419] H. I. Wettersten, A. A. Hakimi, D. Morin, et al. Grade-dependent metabolic reprogramming in kidney cancer revealed by combined proteomics and metabolomics analysis. *Cancer Research*, 75(12):2541–2552, jun 2015.

[420] I. San-Millán and G. A. Brooks. Reexamining cancer metabolism: lactate production for carcinogenesis could be the purpose and explanation of the Warburg Effect. *Carcinogenesis*, 38(2):bgw127, dec 2016.

[421] N. Chattergoon, F. Martínez-Santiesteban, W. B. Handler, et al. Field dependence of T1 for hyperpolarized [1-^{13}C]pyruvate. *Contrast Media and Molecular Imaging*, 8(1):57–62, jan 2013.

[422] K. M. Brindle, S. E. Bohndiek, F. A. Gallagher, and M. I. Kettunen. Tumor imaging using hyperpolarized ^{13}C magnetic resonance spectroscopy. *Magnetic Resonance in Medicine*, 66(2):505–519, aug 2011.

[423] C. H. Cunningham, A. P. Chen, M. Lustig, et al. Pulse sequence for dynamic volumetric imaging of hyperpolarized metabolic products. *Journal of Magnetic Resonance*, 193(1):139–46, jul 2008.

[424] Y. Xing, G. D. Reed, J. M. Pauly, et al. Optimal variable flip angle schemes for dynamic acquisition of exchanging hyperpolarized substrates. *Journal of Magnetic Resonance*, 234:75–81, sep 2013.

[425] J. Maidens, J. W. Gordon, M. Arcak, and P. E. Z. Larson. Optimizing Flip Angles for Metabolic Rate Estimation in Hyperpolarized Carbon-13 MRI. *IEEE Transactions on Medical Imaging*, 35(11):2403–2412, nov 2016.

[426] A. Comment and A. Chen. Patent: System and method for hyperpolarizing a substance. *United States Patent and Trademark Office*, US 10,942,:20, 2021.

[427] S. Kohler, Y. Yen, J. Wolber, et al. In vivo ^{13}carbon metabolic imaging at 3T with hyperpolarized ^{13}C-1-pyruvate. *Magnetic Resonance in Medicine*, 58(1):65–69, jul 2007.

[428] C. Baligand, H. Qin, A. True-Yasaki, et al. Hyperpolarized ^{13}C magnetic resonance evaluation of renal ischemia reperfusion injury in a murine model. *NMR in Biomedicine*, 30(10):e3765, oct 2017.

[429] D. J. Niles, J. W. Gordon, G. Huang, et al. Evaluation of renal metabolic response to partial ureteral obstruction with hyperpolarized ^{13}C MRI. *NMR in Biomedicine*, 31(1):e3846, jan 2018.

[430] C. Laustsen, J. A. Østergaard, M. H. Lauritzen, et al. Assessment of early diabetic renal changes with hyperpolarized [1-^{13}C]pyruvate. *Diabetes/Metabolism Research and Reviews*, 29(2):125–129, feb 2013.

[431] C. Laustsen, T. Stokholm Nørlinger, D. Christoffer Hansen, et al. Hyperpolarized ^{13}C urea relaxation mechanism reveals renal changes in diabetic nephropathy. *Magnetic Resonance in Medicine*, 75(2):515–518, feb 2016.

[432] K. R. Keshari, J. Kurhanewicz, R. Bok, et al. Hyperpolarized ^{13}C dehydroascorbate as an endogenous redox sensor for in vivo metabolic imaging. *Proceedings of the National Academy of Sciences*, 108(46):1806–18611, sep 2011.

[433] T. Xu, D. Mayer, M. Gu, et al. Quantification of in vivo metabolic kinetics of hyperpolarized pyruvate in rat kidneys using dynamic ^{13}C MRSI. *NMR in Biomedicine*, 24(8):997–1005, oct 2011.

[434] F. Wiesinger, E. Weidl, M. I. Menzel, et al. IDEAL spiral CSI for dynamic metabolic MR imaging of hyperpolarized [1-^{13}C]pyruvate. *Magnetic Resonance in Medicine*, 68(1):8–16, jul 2012.

[435] J. W. Gordon, S. B. Fain, D. J. Niles, et al. Simultaneous imaging of ^{13}C metabolism and ^{1}H structure: technical considerations and potential applications. *NMR in Biomedicine*, 28(5):576–582, may 2015.

[436] R. Schmidt, C. Laustsen, J. N. Dumez, et al. In vivo single-shot ^{13}C spectroscopic imaging of hyperpolarized metabolites by spatiotemporal encoding. *Journal of Magnetic Resonance*, 240:8–15, mar 2014.

[437] D. J. Scholz, M. A. Janich, U. Köllisch, et al. Quantified pH imaging with hyperpolarized ^{13}C-bicarbonate. *Magnetic Resonance in Medicine*, 73(6):2274–2282, jun 2015.

[438] R. Sriram, M. V. Criekinge, J. D. Santos, et al. Non-Invasive Differentiation of Benign Renal Tumors from Clear Cell Renal Cell Carcinomas Using Clinically Translatable Hyperpolarized C Pyruvate Magnetic Resonance. *Tomography*, 2(March):35–42, mar 2016.

[439] Y. Dong, R. Eskandari, C. Ray, et al. Hyperpolarized MRI visualizes Warburg effects and predicts treatment response to mTOR inhibitors in patient-derived CCRCC xenograft models. *Cancer Research*, 79(1):242–250, jan 2019.

[440] M. A. Felmlee, R. S. Jones, V. Rodriguez-Cruz, et al. Monocarboxylate transporters (SLC16): Function, regulation, and role in health and disease. *Pharmacological Reviews*, 72(2):466–485, apr 2020.

[441] Y. Rao, S. Gammon, N. M. Zacharias, et al. Hyperpolarized [1-^{13}C]pyruvate-to-[1-^{13}C]lactate conversion is rate-limited by monocarboxylate transporter-1 in the plasma membrane. *Proceedings of the National Academy of Sciences of the United States of America*, 117(36):22378–22389, sep 2020.

[442] Y. Adler-Levy, A. Nardi-Schreiber, T. Harris, et al. In-cell determination of lactate dehydrogenase activity in a luminal breast cancer model ex vivo investigation of excised xenograft tumor slices using dDNP hyperpolarized [1-^{13}C]pyruvate. *Sensors*, 19(9), may 2019.

[443] C. J. Creighton, M. Morgan, P. H. Gunaratne, et al. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature*, 499(7456):43–49, jul 2013.

[444] J. Tropp, J. M. Lupo, A. Chen, et al. Multi-channel metabolic imaging, with SENSE reconstruction, of hyperpolarized [1-^{13}C] pyruvate in a live rat at 3.0 tesla on a clinical MR scanner. *Journal of Magnetic Resonance*, 208(1):171–177, jan 2011.

[445] R. F. Schulte, J. I. Sperl, E. Weidl, et al. Saturation-recovery metabolic-exchange rate imaging with hyperpolarized [1-^{13}C] pyruvate using spectral-spatial excitation. *Magnetic Resonance in Medicine*, 69(5):1209–1216, may 2013.

[446] T. Fritz-Hansen, E. Rostrup, H. B. Larsson, et al. Measurement of the arterial concentration of Gd-DTPA using MRI: A step toward quantitative perfusion imaging. *Magnetic Resonance in Medicine*, 36(2):225–231, aug 1996.

[447] H. Y. Chen, A. W. Autry, J. R. Brender, et al. Tensor image enhancement and optimal multichannel receiver combination analyses for human hyperpolarized ^{13}C MRSI. *Magnetic Resonance in Medicine*, dec 2020.

[448] G. Gilbert. Measurement of signal-to-noise ratios in sum-of-squares MR images. *Journal of Magnetic Resonance Imaging*, 26(6):1678–1678, dec 2007.

[449] O. Khegai, R. F. Schulte, M. A. Janich, et al. Apparent rate constant mapping using hyperpolarized [1-^{13}C]pyruvate. *NMR in Biomedicine*, 27(10):1256–1265, oct 2014.

[450] M. Crispin-Ortuzar, M. Gehrung, S. Ursprung, et al. Three-Dimensional Printed Molds for Image-Guided Surgical Biopsies: An Open Source Computational Platform. *JCO Clinical Cancer Informatics*, 4(4):736–748, sep 2020.

[451] P. A. Gammage, C. Viscomi, M.-L. Simard, et al. Genome editing in mitochondria corrects a pathogenic mtDNA mutation in vivo. *Nature Medicine*, 24(11):1691–1695, nov 2018.

[452] F. Wang, J. Flanagan, N. Su, et al. RNAscope: A novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *Journal of Molecular Diagnostics*, 14(1):22–29, jan 2012.

[453] J. Jones, H. Zecchini, and S. Nagarajan. Multiplexed Detection and Analysis of Low-Abundance Long Noncoding RNA Using RNAscope™ in Cultured Cells. In *Methods in Molecular Biology*, volume 2148, pages 111–125. Humana Press Inc., 2020.

[454] C. Brodie. Overcoming Autofluorescence (AF) and Tissue Variation in Image Analysis of In Situ Hybridization. In *Methods in Molecular Biology*, volume 2148, pages 19–32. Humana Press Inc., 2020.

[455] R. Woitek, M. A. McLean, S. Ursprung, et al. Hyperpolarized Carbon-13 MRI for Early Response Assessment of Neoadjuvant Chemotherapy in Breast Cancer Patients. *Cancer research*, 81(23):canres.1499.2021, 2021.

[456] A. Surov, H. J. Meyer, and A. Wienke. Correlation between apparent diffusion coefficient (ADC) and cellularity is different in several tumors: a meta-analysis. *Oncotarget*, 8(35):59492, 2017.

[457] E. T. Chouchani, V. R. Pell, E. Gaude, et al. Ischaemic accumulation of succinate controls reperfusion injury through mitochondrial ROS. *Nature*, 515(7527):431–435, nov 2014.

[458] S. K. Parks, J. Chiche, and J. Pouysségur. Disrupting proton dynamics and energy metabolism for cancer therapy. *Nature Reviews Cancer*, 13(9):611–623, sep 2013.

[459] Y. S. Kim, J. W. Choi, J. H. Lee, and Y. S. Kim. Expression of lactate/H+ symporters MCT1 and MCT4 and their chaperone CD147 predicts tumor progression in clear cell renal cell carcinoma: Immunohistochemical and the Cancer Genome Atlas data analyses. *Human Pathology*, 46(1):104–112, jan 2015.

[460] D. Ambrosetti, M. Dufies, B. Dadone, et al. The two glycolytic markers GLUT1 and MCT1 correlate with tumor grade and survival in clear-cell renal cell carcinoma. *PLoS ONE*, 13(2):e0193477, feb 2018.

[461] Y. W. Cao, Y. Liu, Z. Dong, et al. Monocarboxylate transporters MCT1 and MCT4 are independent prognostic biomarkers for the survival of patients with clear cell renal cell carcinoma and those receiving therapy targeting angiogenesis. *Urologic Oncology: Seminars and Original Investigations*, 36(6):311.e15–311.e25, jun 2018.

[462] P. Fisel, S. Kruck, S. Winter, et al. DNA methylation of the SLC16A3 promoter regulates expression of the human lactate transporter MCT4 in renal Cancer with consequences for clinical outcome. *Clinical Cancer Research*, 19(18):5170–5181, sep 2013.

[463] K. S. Dimmer, B. Friedrich, F. Lang, et al. The low-affinity monocarboxylate transporter MCT4 is adapted to the export of lactate in highly glycolytic cells. *Biochemical Journal*, 350(1):219–227, aug 2000.

[464] M. Gerlinger, C. R. Santos, B. Spencer-Dene, et al. Genome-wide RNA interference analysis of renal carcinoma survival regulators identifies MCT4 as a Warburg effect metabolic target. *Journal of Pathology*, 227(2):146–156, jun 2012.

[465] J. A. Baena-Del Valle, Q. Zheng, J. L. Hicks, et al. Rapid Loss of RNA Detection by in Situ Hybridization in Stored Tissue Blocks and Preservation by Cold Storage of Unstained Slides. *American Journal of Clinical Pathology*, 148(5):398–415, nov 2017.

[466] A. Vandewalle, G. Wirthensohn, H. G. Heidrich, and W. G. Guder. Distribution of hexokinase and phosphoenolpyruvate carboxykinase along the rabbit nephron. *American Journal of Physiology - Renal Fluid and Electrolyte Physiology*, 9(6), jun 1981.

[467] B. D. Ross, J. Espinal, and P. Silva. Glucose metabolism in renal tubular function. *Kidney International*, 29(1):54–67, jan 1986.

[468] D. Legouis, A. Faivre, P. E. Cippà, and S. de Seigneux. Renal gluconeogenesis: an underestimated role of the kidney in systemic glucose metabolism. *Nephrology Dialysis Transplantation*, nov 2020.

[469] C. Meyer, M. Stumvoll, J. Dostou, et al. Renal substrate exchange and gluconeogenesis in normal postabsorptive humans. *American Journal of Physiology - Endocrinology and Metabolism*, 282(2 45-2), feb 2002.

[470] H. M. Becker, N. Mohebbi, A. Perna, et al. Localization of members of MCT monocarboxylate transporter family Slc16 in the kidney and regulation during metabolic acidosis. *American Journal of Physiology-Renal Physiology*, 299(1):F141–F154, jul 2010.

[471] M. Uhlén, L. Fagerberg, B. M. Hallström, et al. Tissue-based map of the human proteome. *Science*, 347(6220), jan 2015.

[472] J. W. Lee, C. L. Chou, and M. A. Knepper. Deep sequencing in microdissected renal tubules identifies nephron segment-specific transcriptomes. *Journal of the American Society of Nephrology*, 26(11):2669–2677, nov 2015.

[473] National Cancer Registration and Analysis Service Public Health England. Chemotherapy, Radiotherapy and Surgical Tumour Resections in England 2013-2016. https://www.cancerdata.nhs.uk/treatments [accessed: 10.09.2020], 2020.

[474] ESMO Guidelines Committee. Standard Operating Procedures (SOPs) for Authors and templates for ESMO Clinical Practice Guidelines (CPGs) and ESMO-MCBS Scores. https://www.esmo.org/content/download/77789/1426712/file/ESMO-Clinical-Practice-Guidelines-Standard-Operating-Procedures.pdf , 2020.

[475] B. Ljungberg, L. Albiges, Y. Abu-Ghanem, et al. European Association of Urology Guidelines on Renal Cell Carcinoma: The 2019 Update. *European Urology*, 75(5):799–810, may 2019.

[476] B. I. Rini, S. C. Campbell, and B. Escudier. Renal cell carcinoma. *The Lancet*, 373(9669):1119–1132, mar 2009.

[477] V. Prasad and S. Mailankody. Research and development spending to bring a single cancer drug to market and revenues after approval. *JAMA Internal Medicine*, 177(11):1569–1575, nov 2017.

[478] O. J. Wouters, M. McKee, and J. Luyten. Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009-2018. *JAMA*, 323(9):844–853, mar 2020.

[479] A. Sertkaya, H. H. Wong, A. Jessup, and T. Beleche. Key cost drivers of pharmaceutical clinical trials in the United States. *Clinical Trials*, 13(2):117–126, apr 2016.

[480] T. J. Moore, J. Heyward, G. Anderson, and G. C. Alexander. Variation in the estimated costs of pivotal clinical benefit trials supporting the US approval of new therapeutic agents, 2015-2017: a cross-sectional study. *BMJ Open*, 10:38863, jun 2020.

[481] T. J. Moore, H. Zhang, G. Anderson, and G. C. Alexander. Estimated Costs of Pivotal Trials for Novel Therapeutic Agents Approved by the US Food and Drug Administration, 2015-2016. *JAMA Internal Medicine*, 178(11):1451–1457, nov 2018.

[482] B. Glimelius and M. Lahn. Window-of-opportunity trials to evaluate clinical activity of new molecular entities in oncology. *Annals of Oncology*, 22(8):1717–1725, aug 2011.

[483] K. R. Subramonian, S. Puranik, and G. R. Mufti. How will the two-weeks-wait rule affect delays in management of urological cancers? *Journal of the Royal Society of Medicine*, 96(8):398–399, aug 2003.

[484] A. A. Stec, B. J. Coons, S. S. Chang, et al. Waiting Time From Initial Urological Consultation to Nephrectomy for Renal Cell Carcinoma-Does it Affect Survival? *Journal of Urology*, 179(6):2152–2157, jun 2008.

[485] B. Shiff, R. Breau, P. Patel, et al. Impact of time-to-surgery and surgical delay on oncologic outcomes for renal cell carcinoma. *Journal of Urology*, 201(Supplement 4), apr 2019.

[486] S. Schmitz, F. Duhoux, and J. P. Machiels. Window of opportunity studies: Do they fulfil our expectations? *Cancer Treatment Reviews*, 43:50–57, feb 2016.

[487] J. L. Farlow, A. C. Birkeland, P. L. Swiecicki, et al. Window of opportunity trials in head and neck cancer. *Journal of Cancer Metastasis and Treatment*, 2019, mar 2019.

[488] L. Carrassa and G. Damia. DNA damage response inhibitors: Mechanisms and potential applications in cancer therapy. *Cancer Treatment Reviews*, 60:139–151, nov 2017.

[489] C. J. Lord and A. Ashworth. PARP inhibitors: Synthetic lethality in the clinic. *Science*, 355(6330):1152–1158, mar 2017.

[490] F. Zheng, Y. Zhang, S. Chen, et al. Mechanism and current progress of Poly ADP-ribose polymerase (PARP) inhibitors in the treatment of ovarian cancer. *Biomedicine and Pharmacotherapy*, 123:109661, mar 2020.

[491] H. Wei and X. Yu. Functions of PARylation in DNA Damage Repair Pathways. *Genomics, Proteomics and Bioinformatics*, 14(3):131–139, jun 2016.

[492] Y. Chen, P. Jungsuwadee, M. Vore, et al. Collateral damage in cancer chemotherapy: Oxidative stress in nontargeted tissues. *Molecular Interventions*, 7(3):147–156, jun 2007.

[493] T. Golan, P. Hammel, M. Reni, et al. Maintenance olaparib for germline BRCA-mutated metastatic pancreatic cancer. *New England Journal of Medicine*, 381(4):317–327, jul 2019.

[494] M. Robson, S.-A. Im, E. Senkus, et al. Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. *New England Journal of Medicine*, 377(6):523–533, aug 2017.

[495] J. Ledermann, P. Harter, C. Gourley, et al. Olaparib Maintenance Therapy in Platinum-Sensitive Relapsed Ovarian Cancer. *New England Journal of Medicine*, 366(15):1382–1392, apr 2012.

[496] M. I. Carlo, S. Mukherjee, D. Mandelker, et al. Prevalence of germline mutations in cancer susceptibility genes in patients with advanced renal cell carcinoma. *JAMA Oncology*, 4(9):1228–1235, sep 2018.

[497] P. Sun, Y. Li, X. Chao, et al. Clinical characteristics and prognostic implications of BRCA-associated tumors in males: a pan-tumor survey. *BMC Cancer*, 20(1):994, dec 2020.

[498] S. Nowsheen, T. Cooper, J. A. Stanley, and E. S. Yang. Synthetic Lethal Interactions between EGFR and PARP Inhibition in Human Triple Negative Breast Cancer Cells. *PLoS ONE*, 7(10):e46614, oct 2012.

[499] J. Espana-Agusti, A. Warren, S. K. Chew, et al. Loss of PBRM1 rescues VHL dependent replication stress to promote renal carcinogenesis. *Nature Communications*, 8(1):1–11, dec 2017.

[500] N. Kanu, E. Grönroos, P. Martinez, et al. SETD2 loss-of-function promotes renal cancer branched evolution through replication stress and impaired DNA repair. *Oncogene*, 34(46):5699–5708, nov 2015.

[501] K. Moore, N. Colombo, G. Scambia, et al. Maintenance Olaparib in Patients with Newly Diagnosed Advanced Ovarian Cancer. *New England Journal of Medicine*, 379(26):2495–2505, dec 2018.

[502] E. Pujade-Lauraine, J. A. Ledermann, F. Selle, et al. Olaparib tablets as maintenance therapy in patients with platinum-sensitive, relapsed ovarian cancer and a BRCA1/2 mutation (SOLO2/ENGOT-Ov21): a double-blind, randomised, placebo-controlled, phase 3 trial. *The Lancet Oncology*, 18(9):1274–1284, sep 2017.

[503] M. Hussain, J. Mateo, K. Fizazi, et al. Survival with Olaparib in Metastatic Castration-Resistant Prostate Cancer. *New England Journal of Medicine*, 383(24):2345–2357, dec 2020.

[504] P. Mulders, R. Hawkins, P. Nathan, et al. Cediranib monotherapy in patients with advanced renal cell carcinoma: Results of a randomised phase II study. *European Journal of Cancer*, 48(4):527–537, mar 2012.

[505] C. Messiou, M. Orton, J. E. Ang, et al. Advanced Solid Tumors Treated with Cediranib: Comparison of Dynamic Contrast-enhanced MR Imaging and CT as Markers of Vascular Activity. *Radiology*, 265(2):426–436, nov 2012.

[506] J. F. Liu, W. T. Barry, M. Birrer, et al. Combination cediranib and olaparib versus olaparib alone for women with recurrent platinum-sensitive ovarian cancer: A randomised phase 2 study. *The Lancet Oncology*, 15(11):1207–1214, oct 2014.

[507] S. J. Antonia, A. Villegas, D. Daniel, et al. Overall survival with durvalumab after chemoradiotherapy in stage III NSCLC. *New England Journal of Medicine*, 379(24):2342–2350, dec 2018.

[508] R. J. Motzer, B. Escudier, D. F. McDermott, et al. Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma. *New England Journal of Medicine*, 373(19):1803–1813, nov 2015.

[509] J. Shen, W. Zhao, Z. Ju, et al. PARPI triggers the STING-dependent immune response and enhances the therapeutic efficacy of immune checkpoint blockade independent of BRCANEss. *Cancer Research*, 79(2):311–319, jan 2019.

[510] S. Jiao, W. Xia, H. Yamaguchi, et al. PARP inhibitor upregulates PD-L1 expression and enhances cancer-associated immunosuppression. *Clinical Cancer Research*, 23(14):3711–3720, jul 2017.

[511] L. Ding, H. Kim, Q. Wang, et al. PARP Inhibition Elicits STING-Dependent Antitumor Immunity in Brca1-Deficient Ovarian Cancer. *Cell Reports*, 25(11):2972–2980.e5, dec 2018.

[512] J. M. Lee, J. B. Trepel, P. Choyke, et al. CECs and IL-8 have prognostic and predictive utility in patients with recurrent platinum-sensitive ovarian cancer: Biomarker correlates from the randomized phase-2 trial of olaparib and cediranib compared with olaparib in recurrent platinum-sensitive ovar. *Frontiers in Oncology*, 5, jun 2015.

[513] L. Tentori, P. M. Lacal, A. Muzi, et al. Poly(ADP-ribose) polymerase (PARP) inhibition or PARP-1 gene deletion reduces angiogenesis. *European Journal of Cancer*, 43(14):2124–2133, sep 2007.

[514] P. C. Tumeh, C. L. Harview, J. H. Yearley, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature*, 515(7528):568–571, nov 2014.

[515] A. C. Huang, M. A. Postow, R. J. Orlowski, et al. T-cell invigoration to tumour burden ratio associated with anti-PD-1 response. *Nature*, 545(7652):60–65, may 2017.

[516] H.-J. Weinmann, M. Laniado, and W. Mützel. Pharmacokinetics of GdDTPA/dimeglumine after intravenous injection into healthy volunteers. *Physiological chemistry and physics and medical NMR*, 16(2):167–172, 1984.

[517] J. M. Hudson, C. Bailey, M. Atri, et al. The prognostic and predictive value of vascular response parameters measured by dynamic contrast-enhanced-CT , -MRI and -US in patients with metastatic renal cell carcinoma receiving sunitinib. *European Radiology*, 28:2281–2290, jun 2018.

[518] W. L. Dahut, R. A. Madan, J. J. Karakunnel, et al. Phase II clinical trial of cediranib in patients with metastatic castration-resistant prostate cancer. *BJU International*, 111(8):1269–1280, jun 2013.

[519] D. P. Bradley, J. J. Tessier, T. Lacey, et al. Examining the acute effects of cediranib (RECENTIN, AZD2171) treatment in tumor models: a dynamic contrast-enhanced MRI study using gadopentate. *Magnetic Resonance Imaging*, 27(3):377–384, apr 2009.

[520] O. M. Hahn, C. Yang, M. Medved, et al. Dynamic contrast-enhanced magnetic resonance imaging pharmacodynamic biomarker study of sorafenib in metastatic renal carcinoma. *Journal of Clinical Oncology*, 26(28):4572–4578, oct 2008.

[521] R. F. Sweis, M. Medved, S. Towey, et al. Dynamic Contrast-Enhanced Magnetic Resonance Imaging as a Pharmacodynamic Biomarker for Pazopanib in Metastatic Renal Carcinoma. *Clinical Genitourinary Cancer*, 15(2):1–6, apr 2016.

[522] F. Shen, J. Lu, L. Chen, et al. Diagnostic value of dynamic contrast-enhanced magnetic resonance imaging in rectal cancer and its correlation with tumor differentiation. *Molecular and Clinical Oncology*, 4(4):500–506, apr 2016.

[523] A. Ulyte, V. K. Katsaros, E. Liouta, et al. Prognostic value of preoperative dynamic contrast-enhanced MRI perfusion parameters for high-grade glioma patients. *Neuroradiology*, 58(12):1197–1208, dec 2016.

[524] M. O. Leach, K. M. Brindle, J. L. Evelhoch, et al. The assessment of antiangiogenic and antivascular therapies in early-stage clinical trials using magnetic resonance imaging: Issues and recommendations. *British Journal of Cancer*, 92(9):1599–1610, may 2005.

[525] S. Walker-Samuel, M. O. Leach, and D. J. Collins. Evaluation of response to treatment using DCE-MRI: The relationship between initial area under the gadolinium curve (IAUGC) and quantitative pharmacokinetic analysis. *Physics in Medicine and Biology*, 51(14):3593–3602, jul 2006.

[526] M. Medinger, N. Esser, J. Soltau, et al. Antitumor effect of the vascular-disrupting agent ZD6126 in a murine renal cell carcinoma model. *International Journal of Oncology*, 38(2):455–464, feb 2011.

[527] D. W. Siemann, W. D. Brazelle, and J. M. Jürgensmeier. The Vascular Endothelial Growth Factor Receptor-2 Tyrosine Kinase Inhibitor Cediranib (Recentin; AZD2171) Inhibits Endothelial Cell Function and Growth of Human Renal Tumor Xenografts. *International Journal of Radiation Oncology Biology Physics*, 73(3):897–903, mar 2009.

[528] P. J. O'Dwyer, M. Rosen, M. Gallagher, et al. Pharmacodynamic study of BAY 43-9006 in patients with metastatic renal cell carcinoma. *Journal of Clinical Oncology*, 23(16):3005–3005, jun 2005.

[529] C. Baudelet and B. Gallez. How does blood oxygen level-dependent (BOLD) contrast correlate with oxygen partial pressure (pO2) inside tumors? *Magnetic Resonance in Medicine*, 48(6):980–986, dec 2002.

[530] A. Afaq, A. Andreou, and D. M. Koh. Diffusion-weighted magnetic resonance imaging for tumour response assessment: Why, when and how? *Cancer Imaging*, 10(SPEC. ISS. A):S179, oct 2010.

[531] T. Kanda, K. Ishii, H. Kawaguchi, et al. High signal intensity in the dentate nucleus and globus pallidus on unenhanced T1-weighted MR images: Relationship with increasing cumulative dose of a gadoliniumbased contrast material. *Radiology*, 270(3):834–841, mar 2014.

[532] B. Turkbey, M. L. Lindenberg, S. Adler, et al. PET/CT imaging of renal cell carcinoma with 18F-VM4-037: a phase II pilot study. *Abdominal Radiology*, 41(1):109–118, jan 2016.

[533] K. Kitajima, S. Yamamoto, Y. Kawanaka, et al. Imaging of renal cell carcinoma in patients with acquired cystic disease of the kidney: comparison 11 C-choline and FDG PET/CT with dynamic contrast-enhanced CT. *Japanese Journal of Radiology*, 37(2):165–177, feb 2019.

[534] H. M. De Feyter, K. L. Behar, Z. A. Corbin, et al. Deuterium metabolic imaging (DMI) for MRI-based 3D mapping of metabolism in vivo. *Science Advances*, 4(8):7314–7336, aug 2018.

[535] C. A. Dykewicz. Summary of the guidelines for preventing opportunistic infections among hematopoietic stem cell transplant recipients. *Clinical Infectious Diseases*, 33(2):139–144, jul 2001.

Appendix A

Supplementary tables

A.1 Levels and grades of evidence

The definition of levels and grades of evidence follows the European Society of Medical Oncology (ESMO) [474] and was adapted from the Infectious Diseases Society of America [535].

A.1.1 Levels of evidence

Table A.1 Levels of evidence

Grade	Explanation
I	Evidence from at least one large randomised, controlled trial of good methodological quality or meta-analyses of well-conducted randomised trials without heterogeneity.
II	Small randomised trials or large randomised trials with a suspicion of bias or meta-analyses of such trials or of trials with demonstrated heterogeneity.
III	Prospective cohort studies
IV	Retrospective cohort studies or case-control studies
V	Studies without control group, case reports, expert opinions

A.1.2 Grade of clinical practice recommendations

Table A.2 Grade of practice recommendations

Grade	Explanation
A	Strong evidence for efficacy with a substantial clinical benefit, strongly recommended.
B	Strong or Moderate evidence for efficacy but with a limited clinical benefit, generally recommended.
C	Insufficient evidence for efficacy or benefit does not outweigh the risk or the disadvantages (adverse events, costs, ets.), optional.
D	Moderate evidence against efficacy or for adverse outcome, generally not recommended.
E	Strong evidence against efficacy or for adverse outcome, never recommended.

Appendix B

Radiomics studies: supplementary documentation

B.1 Texture analysis in RCC - Systematic Review and Meta-Analysis

Table B.1 RCC Radiomics: Study details and average RQS rating

Fist author	Year	Journal	PMID	#Pat	Modality	Research question	RQS (%)
Antunes	2016	Transl Oncol	27084432	2	PET/MR	Detecting treatment response to antiangiogenic therapy	8.5
Bektas	2019	Eur Rad	14321084	53	CT	Differentiation of low and high nuclear ccRCC	5.7
Bharwani	2014	Br. J. Cancer	24366299	20	MRI	Treatment-associated texture changes on sunitinib and correlation with overall survival	7.5
Bier	2018	PLoS ONE	29668695	106	CT	Prediction of tumour adherence to perinephric fat	6.5
Boos	2017	AJR	28570097	19	CT	Correlation of CT texture to best response according to RECIST under VEGFR TKI	-4.0
Catalano	2008	Radiology	18413886	50	CT	Differentiation of AMLwvf and ccRCC	-1.5
Chandarana	2012	Radiology	30404863	73	MRI	Differentiation of ccRCC and pRCC	-0.5
Chaudhry	2012	AJR	22268181	64	CT	Differentiation of AMLwvf, ccRCC and pRCC	-1.0
Chen	2017	Abdom Rad	27595574	94	CT	Differentiation of ccRCC and oncocytoma	4.5
Chen	2015	Springerplus	25694862	61	CT	Differentiation of ccRCC and pRCC	4.0
Ding	2018	Eur J Rad	29803385	206	CT	Differentiation of low and high nuclear grade ccRCC	13.5
Doshi	2016	AJR	26901013	37	MRI	Differentiation of Type 1 and Type 2 pRCC	6.0
Feng	2018	Eur Rad	29134348	58	CT	Differentiation of small AMLwvf and RCC subtypes	6.5
Gaing	2015	Invest Radiol	25387050	44	MRI	Differentiation of ccRCC, pRCC, chRCC, cystic	12.0

AMLwvf: Angiomyolipoma without visible fat, ccRCC: clear cell RCC, chRCC: chromophobe RCC, CT: Computed tomography, MRI: Magnetic resonance imaging, NA: Not available, Onc: Oncocytoma, PET: Positron emission tomography, PMID: PubMed identifier, pRCC: papillary RCC, RCC: Renal cell carcinoma, RECIST: Response evaluation criteria in solid tumours, TKI: Tyrosine kinase inhibitor, VEGFR: Vascular endothelial growth factor receptor

Table B.1 Continued: RCC Radiomics: Study details and average RQS rating

Fist author	Year	Journal	PMID	#Pat	Modality	Research question	RQS (%)
						RCC, oncocytoma and AML	
Ghosh	2015	J Med Imag	26839909	78	CT	Prediction of BAP1 mutation status in ccRCC	5.5
Goh	2011	Radiology	21813743	39	CT	Correlation of CT texture at baseline and under TKI with time to progression	-1.7
Haider	2017	Cancer Imag	28114978	40	CT	Overall and progression-free survival under sunitinib	-2.0
Hoang	2018	Abdom Rad	29858935	41	MRI	Differentiation of small ccRCC, pRCC and oncocytoma	4.5
Hodgdon	2015	Radiology	25906183	100	CT	Differentiation of AMLwvf from RCC subtypes	5.0
Huhdanpaa	2015	Abdom Imag	26304585	65	CT	Differentiation of low and high nuclear grade ccRCC	-2.0
Khene	2018	World J Urol	29675631	70	CT	Prediction of tumour adherence to perinephric fat	4.5
Kierans	2014	AJR	25415729	61	MRI	Differentiation of low and high stage ccRCC	5.5
Kim	2008	Radiology	18094264	144	CT	Differentiation of AMLwvf from multiple RCC subtypes	-2.0
Kocak	2018	Eur J Rad	30292260	93	CT	Differentiation of ccRCC, pRCC and chRCC	14.3
Kunapuli	2018	Digit Imag	29980960	150	CT	Differentiation of benign and malignant renal lesions	5.5
Lee	2017	Med Phys	28376281	50	CT	Differentiation of AMLwvf and ccRCC	2.7
Lee	2018	Med Phys	29474742	80	CT	Differentiation of AMLwvf and ccRCC	3.0
Leng	2017	Abdom Rad	28025654	139	CT	Differentiation of AMLwvf from multiple RCC subtypes	0.0
Li	2018	AJR	29812980	92	MRI	Differentiation AMLwvf, ccRCC, pRCC, chRCC, onc	0.0
Li	2019	Acad Rad	30087067	140	MRI	Differentiation of AMLwvf and ccRCC	0.0
Liu	2017	ClinExpMed	NA	44	CT	Differentiation ccRCC, pRCC, chRCC and sarcoma	4.0

AMLwvf: Angiomyolipoma without visible fat, ccRCC: clear cell RCC, chRCC: chromophobe RCC, CT: Computed tomography, MRI: Magnetic resonance imaging, NA: Not available, Onc: Oncocytoma, PET: Positron emission tomography, PMID: PubMed identifier, pRCC: papillary RCC, RCC: Renal cell carcinoma, RECIST: Response evaluation criteria in solid tumours, TKI: Tyrosine kinase inhibitor, VEGFR: Vascular endothelial growth factor receptor

Table B.1 Continued: RCC Radiomics: Study details and average RQS rating

Fist author	Year	Journal	PMID	#Pat	Modality	Research question	RQS (%)
Linguraru	2011	Med Phys	21992388	43	CT	Differentiation of VHL, BHD, HPRC and HLRCC	5.0
Lubner	2016	AJR	27145377	157	CT	Overall survival, time to recurrence, histologic subtype, nuclear grade, sarcomatoid transformation and presence of metastasis	3.5
Mains	2017	Br J Rad	29144161	69	CT	Correlation with overall and progression-free survival on various treatments	7.0
Paschall	2018	Abdom Rad	29520425	55	MRI	Differentiation of ccRCC, pRCC and oncocytoma	-3.0
Raman	2014	Acad Rad	25239842	99	CT	Differentiation of ccRCC, pRCC, oncocytoma and cysts	7.0
Ramesh	2018	ClinDiagRes	NA	188	CT	Differentiation of normal renal tissue, benign and malignant renal masses	-3.0
Reynolds	2018	PLoS ONE	30114235	12	MRI	Delta radiomics: early response to radiotherapy	4.5
Sasaguri	2015	AJR	26496547	166	CT	Differentiation of ccRCC and oncocytoma	5.5
Schieda	2018	AJR	29547054	37	CT	Differentiation of low/high nuclear grade chRCC	6.0
Schieda	2015	AJR	25905936	31	CT	Differentiation of ccRCC and sarcomatoid RCC	5.5
Scrima	2019	Abdom Rad	29804215	249	CT	Differentiation of ccRCC and non-ccRCC, correlation with DNA expression	3.5
Simpfendorfer	2009	AJR	19155407	36	CT	Differentiation of AMLwvf from RCC	0.0
Takahashi	2015	AJR	26587925	153	CT	Differentiation of AMLwvf from ccRCC	5.0
Takahashi	2016	Abdom Rad	27015866	112	CT	Differentiation of AML from ccRCC	-2.5

AMLwvf: Angiomyolipoma without visible fat, ccRCC: clear cell RCC, chRCC: chromophobe RCC, CT: Computed tomography, MRI: Magnetic resonance imaging, NA: Not available, Onc: Oncocytoma, PET: Positron emission tomography, PMID: PubMed identifier, pRCC: papillary RCC, RCC: Renal cell carcinoma, RECIST: Response evaluation criteria in solid tumours, TKI: Tyrosine kinase inhibitor, VEGFR: Vascular endothelial growth factor receptor

Table B.1 Continued: RCC Radiomics: Study details and average RQS rating

Fist author	Year	Journal	PMID	#Pat	Modality	Research question	RQS (%)
Tanaka	2011	Int J Urol	21815937	41	MRI	Differentiation of AMLwvf from RCC	-3.0
Varghese	2018	AJR	30240299	174	CT	Differentiation of benign/malignant fat-poor masses	5.7
Varghese	2018	Br J Rad	29888982	156	CT	Differentiation of benign/malignant fat-poor masses	-1.7
Vendrami	2018	AJR	30240294	41	MRI	Differentiation of type 1 and type 2 pRCC	5.0
Wang	2016	Sci Rep	27380733	21	MRI	Reproducibility of DCE-based histogram parameters	5.0
Xi	2018	Eur Rad	28681074	16	MRI	Differentiation of low and high nuclear grade ccRCC	6.5
Yan	2015	Acad Rad	26031228	48	CT	Differentiation of AMLwvf, ccRCC and pRCC	4.0
Yap	2018	Urology	29305199	150	CT	Differentiation of predominantly solid masses	-1.7
Yin	2017	Sci Rep	28256615	9	PET/MR	Differentiation of ccA and ccB molecular subtypes	12.0
Yin	2018	Clin Rad	29801658	8	PET/MR	Differentiation of ccA and ccB molecular	3.5
Yu	2017	Abdom Rad	28421244	119	CT	Differentiation of ccRCC, pRCC, chRCC and onc	0.0
Zhang	2015	AJR	26204307	46	MRI	Differentiation of low and high nuclear grade ccRCC	-1.0

AMLwvf: Angiomyolipoma without visible fat, ccRCC: clear cell RCC, chRCC: chromophobe RCC, CT: Computed tomography,
MRI: Magnetic resonance imaging, NA: Not available, Onc: Oncocytoma, PET: Positron emission tomography, PMID: PubMed identifier,
pRCC: papillary RCC, RCC: Renal cell carcinoma, RECIST: Response evaluation criteria in solid tumours,
TKI: Tyrosine kinase inhibitor, VEGFR: Vascular endothelial growth factor receptor

Table B.2 RCC Radiomics: QUADAS-2 risk of bias assessment

First author + year	Risk of bias												Applicability								
	Patient selection			Index test			Reference standard			Flow and timing			Patient selection			Index test			Reference standard		
Antunes2016	H	H	-	I	I	-	L	L	-	L	I	-	L	L	-	L	L	-	L	L	-
Bektas2019	H	L	L	I	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
Bharwani2014	H	-	L	L	-	L	L	-	L	L	-	L	L	-	L	H	-	L	L	-	L
Bier2018	L	L	-	L	L	-	L	L	-	L	L	-	L	L	-	L	L	-	L	L	-
Boos2017	L	-	L	H	-	H	L	-	L	I	-	I	L	-	L	H	-	H	L	-	L
Catalano2008	H	-	H	I	-	I	L	-	L	I	-	I	L	-	L	L	-	L	L	-	L
Chandarana2012	H	-	L	H	-	H	L	-	L	L	-	L	L	-	L	L	-	L	L	-	L
Chaudhry2012	H	-	H	L	-	L	L	-	L	I	-	I	L	-	L	L	-	L	L	-	L
Chen2017	H	-	L	H	-	H	L	-	L	I	-	I	L	-	L	L	-	L	L	-	L
Chen2015	H	-	H	H	-	H	L	-	L	I	-	I	L	-	L	L	-	L	L	-	L
Ding2018	H	L	-	L	L	-	L	L	-	I	L	-	L	L	-	L	L	-	L	L	-
Doshi2016	H	L	-	L	L	-	L	L	-	I	I	-	L	L	-	L	L	-	L	L	-
Feng2018	H	H	-	I	H	-	L	L	-	I	L	-	L	L	-	L	L	-	L	L	-
Gaing2015	H	-	H	L	-	L	L	-	L	L	-	L	L	-	L	L	-	L	L	-	L
Ghosh2015	H	L	-	L	L	-	L	L	-	I	I	-	L	L	-	L	L	-	L	L	-
Goh2011	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
Haider2017	L	L	-	L	L	-	L	L	-	L	L	-	L	L	-	L	L	-	L	L	-

The three sub columns correspond to the assessment by each of the three readers. H: High risk of bias / High level of applicability concerns, I: Indeterminate risk of bias / Indeterminate level of applicability concerns, L: Low risk of bias / Low level of applicability concerns, -: Not assessed by this reader.

Table B.2 Continued: RCC Radiomics: QUADAS-2 risk of bias assessment

First author + year	Risk of bias												Applicability								
	Patient selection			Index test			Reference standard			Flow and timing			Patient selection			Index test			Reference standard		
Hoang2018	H	L	-	L	L	-	H	H	-	I	I	-	L	L	-	L	L	-	L	L	-
Hodgdon2015	H	L	-	L	L	-	L	L	-	I	I	-	L	L	-	L	L	-	L	L	-
Huhdanpaa2015	H	-	H	L	-	L	L	-	L	I	-	I	L	-	L	L	-	L	L	-	L
Khene2018	L	L	-	I	I	-	L	L	-	L	I	-	L	L	-	L	L	-	L	L	-
Kierans2014	H	L	-	L	L	-	L	L	-	L	L	-	L	L	-	L	L	-	L	L	-
Kim2008	H	-	L	H	-	H	H	-	L	I	-	I	L	-	L	L	-	L	L	-	I
Kocak2018	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
Kunapuli2018	H	H	-	L	H	-	L	L	-	I	I	-	L	L	-	L	L	-	L	L	-
Lee2017	L	L	L	I	I	I	L	L	-	I	I	I	L	L	L	L	L	L	L	L	L
Lee2018	H	L	-	I	I	-	H	H	H	I	I	-	L	L	-	L	L	-	L	L	-
Leng2017	H	L	-	L	L	-	L	L	-	I	I	-	L	L	-	L	L	-	L	L	-
Li2018	H	H	-	H	H	H	L	L	L	I	I	-	L	L	-	L	L	-	L	L	-
Li2019	L	L	L	H	H	H	L	L	L	I	I	I	L	L	L	L	L	L	L	L	L
Linguraru2011	H	-	H	I	-	I	L	-	L	I	-	I	L	-	L	L	-	L	L	-	L
Liu2017	H	L	-	I	I	-	L	L	-	I	I	-	L	L	-	L	L	-	L	L	-
Lubner2016	H	L	-	H	H	-	L	L	-	L	L	-	L	L	-	L	L	-	L	L	-
Mains2018	L	-	L	I	-	I	H	-	L	I	-	I	L	-	L	L	-	L	L	-	L

The three sub columns correspond to the assessment by each of the three readers. H: High risk of bias / High level of applicability concerns, I: Indeterminate risk of bias / Indeterminate level of applicability concerns, L: Low risk of bias / Low level of applicability concerns, -: Not assessed by this reader.

Table B.2 Continued: RCC Radiomics: QUADAS-2 risk of bias assessment

First author + year	Risk of bias												Applicability								
	Patient selection			Index test			Reference standard			Flow and timing			Patient selection			Index test			Reference standard		
Paschall2018	H	-	H	H	-	H	L	-	L	L	-	L	L	-	L	L	-	L	L	-	L
Raman2014	H	H	-	I	I	-	L	L	-	I	I	-	L	L	-	L	L	-	L	L	-
Ramesh2018	H	H	-	I	I	-	H	L	-	I	I	-	L	L	-	L	L	-	L	L	-
Reynolds2018	H	-	H	H	-	I	H	-	L	L	-	L	L	-	L	L	-	L	L	-	L
Sasaguri2015	H	L	-	H	H	-	L	L	-	I	I	-	L	L	-	L	L	-	L	L	-
Schieda2018	H	H	-	L	L	-	L	L	-	H	L	-	L	L	-	L	L	-	L	L	-
Schieda2015	H	L	-	L	L	-	L	L	-	I	I	-	L	L	-	L	L	-	L	L	-
Scrima2019	H	L	-	I	I	-	L	L	-	L	L	-	L	L	-	L	L	-	L	L	-
Simpfendorfer2009	H	-	H	L	-	L	L	-	L	I	-	I	L	-	L	L	-	L	L	-	L
Takahashi2015	H	L	-	H	H	-	L	L	-	I	I	-	L	L	-	H	H	-	L	L	-
Takahashi2016	H	L		H	H	-	L	L	-	I	I	-	L	L	-	L	L	-	L	L	-
Tanaka2011	H	-	H	L	-	L	L	-	L	I	-	I	L	-	L	L	-	L	L	-	L
Varghese2018AJR	L	L	L	L	L	L	L	L	L	I	I	I	L	L	L	L	L	L	L	L	L
Varghese2018BJR	L	L	L	I	I	I	L	L	L	I	I	I	L	L	L	L	L	L	L	L	L
Vendrami2018	H	H	H	L	L	L	L	L	L	I	I	I	L	L	L	L	L	L	L	L	L
Wang2016	L	-	L	H	-	H	L	-	L	L	-	L	L	-	L	L	-	L	L	-	L
Xi2018	H	-	L	L	-	L	L	-	L	L	-	L	L	-	L	L	-	H	L	-	L

The three sub columns correspond to the assessment by each of the three readers. H: High risk of bias / High level of applicability concerns, I: Indeterminate risk of bias / Indeterminate level of applicability concerns, L: Low risk of bias / Low level of applicability concerns, -: Not assessed by this reader.

Table B.2 Continued: RCC Radiomics: QUADAS-2 risk of bias assessment

First author + year	Risk of bias												Applicability								
	Patient selection			Index test			Reference standard			Flow and timing			Patient selection			Index test			Reference standard		
Yan2015	H	H	H	L	L	L	H	H	L	I	I	I	L	L	L	L	L	L	L	L	L
Yap2018	H	H	H	H	H	I	L	L	L	I	I	I	L	L	L	L	L	L	L	L	L
Yin2018	H	H	-	H	H	-	L	L	-	I	I	-	L	L	-	L	L	-	L	L	-
Yin2017	H	L	-	L	L	-	L	L	-	L	L	-	L	L	-	L	L	-	L	L	-
Yu2017	H	L	-	L	L	-	L	L	-	I	I	-	L	L	-	L	L	-	L	L	-
Zhang2015	H	-	H	H	-	H	L	-	L	L	-	L	L	-	L	L	-	L	L	-	L

The three sub columns correspond to the assessment by each of the three readers. H: High risk of bias / High level of applicability concerns, I: Indeterminate risk of bias / Indeterminate level of applicability concerns, L: Low risk of bias / Low level of applicability concerns, -: Not assessed by this reader.

Table B.3 RCC Radiomics: Inter-rater correlation coefficients for the RQS.

RQS Item	S* from modified Fleiss' κ statistic mean (95% confidence interval)
Image protocol	0.45 (0.20-0.67)
Multiple segmentations	0.93 (0.82-1.00)
Phantom study	1.00 (1.00-1.00)
Multiple time points	0.93 (0.82-1.00)
Feature reduction	0.93 (0.82-1.00)
Non-radiomic features	0.67 (0.49-0.85)
Biological correlates	0.93 (0.82-1.00)
Cut-off analysis	0.93 (0.82-1.00)
Discrimination statistic and resampling	0.82 (0.71-0.92)
Calibration statistic	0.96 (0.89-1.00)
Prospective study design	1.00 (1.00-1.00)
Validation	1.00 (1.00-1.00)
Gold standard comparison	0.76 (0.63-0.88)
Clinical utility	0.60 (0.38-0.82)
Cost-effectiveness analysis	1.00 (1.00-1.00)
Open science	1.00 (1.00-1.00)

RQS: Radiomics Quality Score

B.2 Radiomic features in PyRadiomics

Table B.4 RCC Radiomics: Features in the PyRadiomics package

Feature	Feature Code
Shape Elongation	Shape 1
Shape Flatness	Shape 2
Shape Least Axis Length	Shape 3
Shape Major Axis Length	Shape 4
Shape Maximum 2D Diameter Column	Shape 5
Shape Maximum 2D Diameter Row	Shape 6
Shape Maximum 2D Diameter Slice	Shape 7
Shape Maximum 3D Diameter	Shape 8
Shape Mesh Volume	Shape 9
Shape Minor Axis Length	Shape 10
Shape Sphericity	Shape 11
Shape Surface Area	Shape 12
Shape Surface Volume Ratio	Shape 13
Shape Voxel Volume	Shape 14
First-order 10 Percentile	FO 1
First-order 90 Percentile	FO 2
First-order Energy	FO 3
First-order Entropy	FO 4
First-order Interquartile Range	FO 5
First-order Kurtosis	FO 6
First-order Maximum	FO 7
First-order Mean Absolute Deviation	FO 8
First-order Mean	FO 9
First-order Median	FO 10
First-order Minimum	FO 11
First-order Range	FO 12
First-order Robust Mean Absolute Deviation	FO 13
First-order Root Mean Squared	FO 14

FO: First-order, GLCM: Grey level coocurrence matrix,GLDM: Grey-level density matrix, GLRLM: Grey-level runlength matrix, GLSZM: Grey-level size zone matrix,NGTDM: Neighbourhood grey-tone difference matrix.

Table B.4 Continued: RCC Radiomics: Features in the PyRadiomics package

Feature	Feature Code
First-order Skewness	FO 15
First-order Total Energy	FO 16
First-order Uniformity	FO 17
First-order Variance	FO 18
GLCM Autocorrelation	GLCM 1
GLCM Cluster Prominence	GLCM 2
GLCM Cluster Shade	GLCM 3
GLCM Cluster Tendency	GLCM 4
GLCM Contrast	GLCM 5
GLCM Correlation	GLCM 6
GLCM Difference Average	GLCM 7
GLCM Difference Entropy	GLCM 8
GLCM Difference Variance	GLCM 9
GLCM Id	GLCM 10
GLCM Idm	GLCM 11
GLCM Idmn	GLCM 12
GLCM Idn	GLCM 13
GLCM Imc1	GLCM 14
GLCM Imc2	GLCM 15
GLCM Inverse Variance	GLCM 16
GLCM Joint Average	GLCM 17
GLCM Joint Energy	GLCM 18
GLCM Joint Entropy	GLCM 19
GLCM MCC	GLCM 20
GLCM Maximum Probability	GLCM 21
GLCM Sum Average	GLCM 22
GLCM Sum Entropy	GLCM 23
GLCM Sum Squares	GLCM 24
GLDM Dependence Entropy	GLDM 1
GLDM Dependence Non-Uniformity	GLDM 2

FO: First-order, GLCM: Grey level coocurrence matrix,GLDM: Grey-level density matrix, GLRLM: Grey-level runlength matrix, GLSZM: Grey-level size zone matrix,NGTDM: Neighbourhood grey-tone difference matrix.

Table B.4 Continued: RCC Radiomics: Features in the PyRadiomics package

Feature	Feature Code
GLDM Dependence Non-Uniformity Normalized	GLDM 3
GLDM Dependence Variance	GLDM 4
GLDM Grey-Level Non-Uniformity	GLDM 5
GLDM Grey-Level Variance	GLDM 6
GLDM High Grey-Level Emphasis	GLDM 7
GLDM Large Dependence Emphasis	GLDM 8
GLDM Large Dependence High Grey-Level Emphasis	GLDM 9
GLDM Large Dependence Low Grey-Level Emphasis	GLDM 10
GLDM Low Grey-Level Emphasis	GLDM 11
GLDM Small Dependence Emphasis	GLDM 12
GLDM Small Dependence High-Grey Level Emphasis	GLDM 13
GLDM Small Dependence Low-Grey Level Emphasis	GLDM 14
GLRLM Grey-Level Non-Uniformity	GLRLM 1
GLRLM Grey-Level Non-Uniformity Normalized	GLRLM 2
GLRLM Grey-Level Variance	GLRLM 3
GLRLM High Grey-Level Run Emphasis	GLRLM 4
GLRLM Long Run Emphasis	GLRLM 5
GLRLM Long Run High Grey-Level Emphasis	GLRLM 6
GLRLM Long Run Low Grey-Level Emphasis	GLRLM 7
GLRLM Low Grey-Level Run Emphasis	GLRLM 8
GLRLM Run Entropy	GLRLM 9
GLRLM Run Length Non-Uniformity	GLRLM 10
GLRLM Run Length Non-Uniformity Normalized	GLRLM 11
GLRLM Run Percentage	GLRLM 12
GLRLM Run Variance	GLRLM 13
GLRLM Short Run Emphasis	GLRLM 14
GLRLM Short Run High-Grey Level Emphasis	GLRLM 15
GLRLM Short Run Low-Grey Level Emphasis	GLRLM 16
GLSZM Grey-Level Non-Uniformity	GLSZM 1
GLSZM Grey-Level Non-Uniformity Normalized	GLSZM 2

FO: First-order, GLCM: Grey level coocurrence matrix,GLDM: Grey-level density matrix, GLRLM: Grey-level runlength matrix, GLSZM: Grey-level size zone matrix,NGTDM: Neighbourhood grey-tone difference matrix.

Table B.4 Continued: RCC Radiomics: Features in the PyRadiomics package

Feature	Feature Code
GLSZM Grey-Level Variance	GLSZM 3
GLSZM High Grey-Level ZoneEmphasis	GLSZM 4
GLSZM Large Area Emphasis	GLSZM 5
GLSZM Large Area High Grey-Level Emphasis	GLSZM 6
GLSZM Large Area Low Grey-Level Emphasis	GLSZM 7
GLSZM Low Grey-Level Zone Emphasis	GLSZM 8
GLSZM Size Zone Non Uniformity	GLSZM 9
GLSZM Size Zone Non-Uniformity Normalized	GLSZM 10
GLSZM Small Area Emphasis	GLSZM 11
GLSZM Small Area High Grey-Level Emphasis	GLSZM 12
GLSZM Small Area Low Grey-Level Emphasis	GLSZM 13
GLSZM Zone Entropy	GLSZM 14
GLSZM Zone Percentage	GLSZM 15
GLSZM Zone Variance	GLSZM 16
NGTDM Busyness	NGTDM 1
NGTDM Coarseness	NGTDM 2
NGTDM Complexity	NGTDM 3
NGTDM Contrast	NGTDM 4
NGTDM Strength	NGTDM 5

FO: First-order, GLCM: Grey level coocurrence matrix,GLDM: Grey-level density matrix, GLRLM: Grey-level runlength matrix, GLSZM: Grey-level size zone matrix,NGTDM: Neighbourhood grey-tone difference matrix.

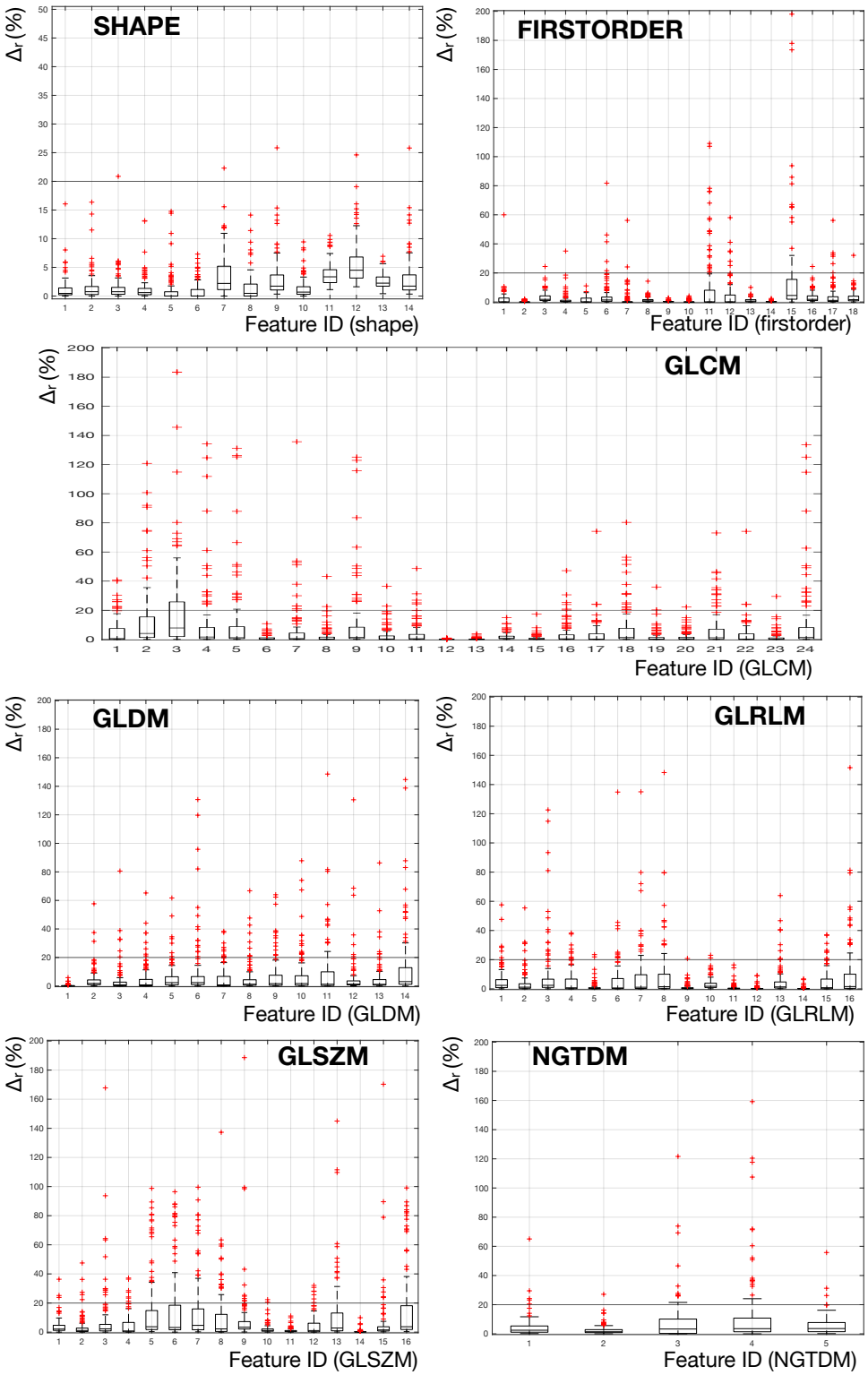


Fig. B.1 Variation in texture features in the internal dataset between the original segmentation and the region of interest following an opening operation. Boxplots reach from the 25thto the 75th percentile. Whiskers extend to the 5th and 95th percentile. Please note the different scaling of the y-axis in the top left panel. Please refer to table B.4, page 247 for the IDs of the texture features. GLCM: Grey level coocurrence matrix,GLDM: Grey-level density matrix, GLRLM: Grey-level runlength matrix, GLSZM:Grey-level size zone matrix, NGTDM: Neighbourhood grey-tone difference matrix.

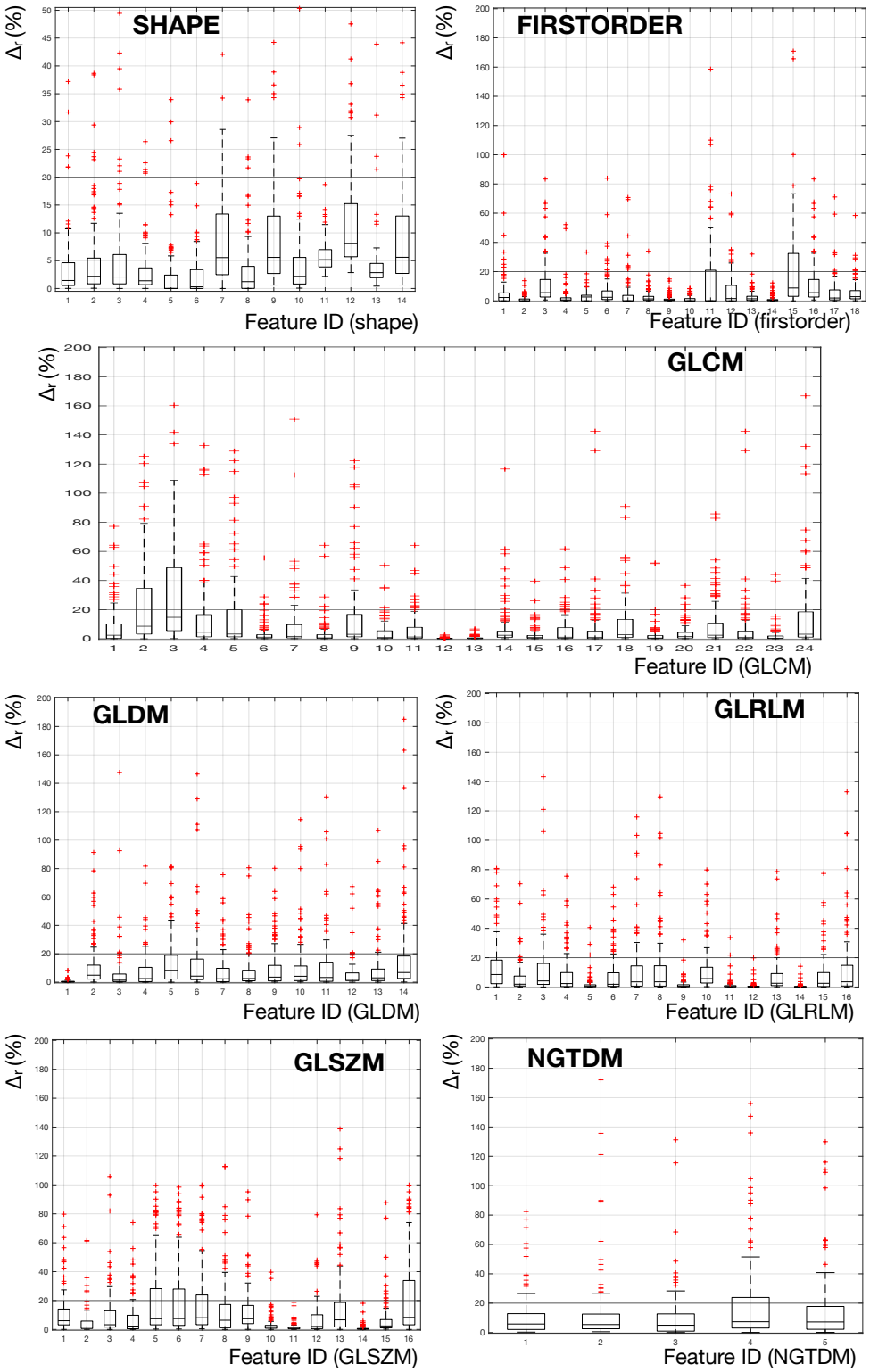


Fig. B.2 Variation in texture features in the internal dataset between the original segmentation and the region of interest following two iterations of opening operation. Boxplots reach from the 25thto the 75th percentile. Whiskers extend to the 5th and 95th percentile. Please note the different scaling of the y-axis in the top left panel. Please refer to table B.4, page 247 for the IDs of the texture features. GLCM: Grey level coocurrence matrix,GLDM: Grey-level density matrix, GLRLM: Grey-level runlength matrix, GLSZM:Grey-level size zone matrix, NGTDM: Neighbourhood grey-tone difference matrix.

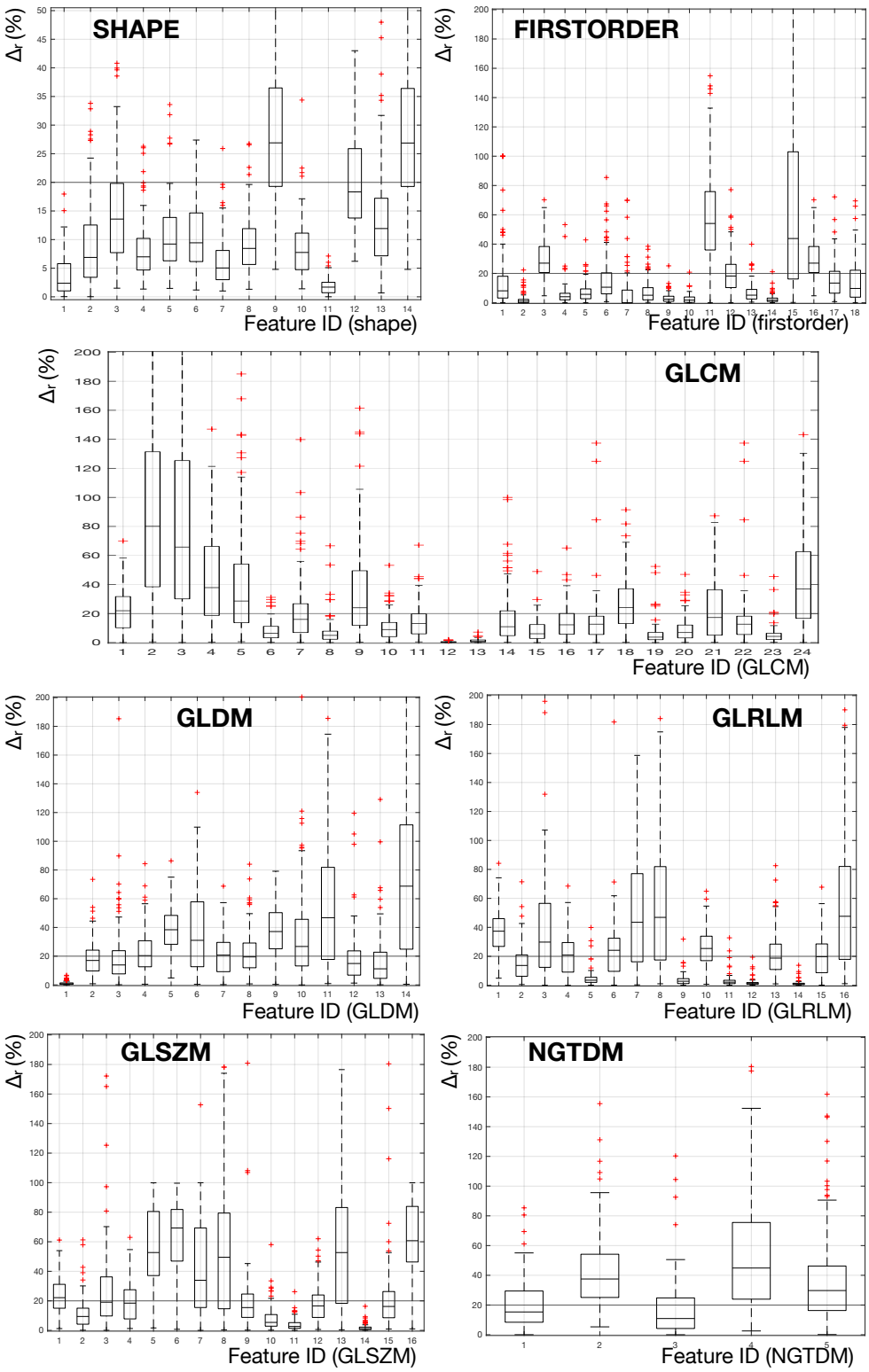


Fig. B.3 Variation in texture features in the internal dataset between the original segmentation and the region of interest following an erosion operation. Boxplots reach from the 25thto the 75th percentile. Whiskers extend to the 5th and 95th percentile. Please note the different scaling of the y-axis in the top left panel. Please refer to table B.4, page 247 for the IDs of the texture features. GLCM: Grey level coocurrence matrix,GLDM: Grey-level density matrix, GLRLM: Grey-level runlength matrix, GLSZM:Grey-level size zone matrix, NGTDM: Neighbourhood grey-tone difference matrix.

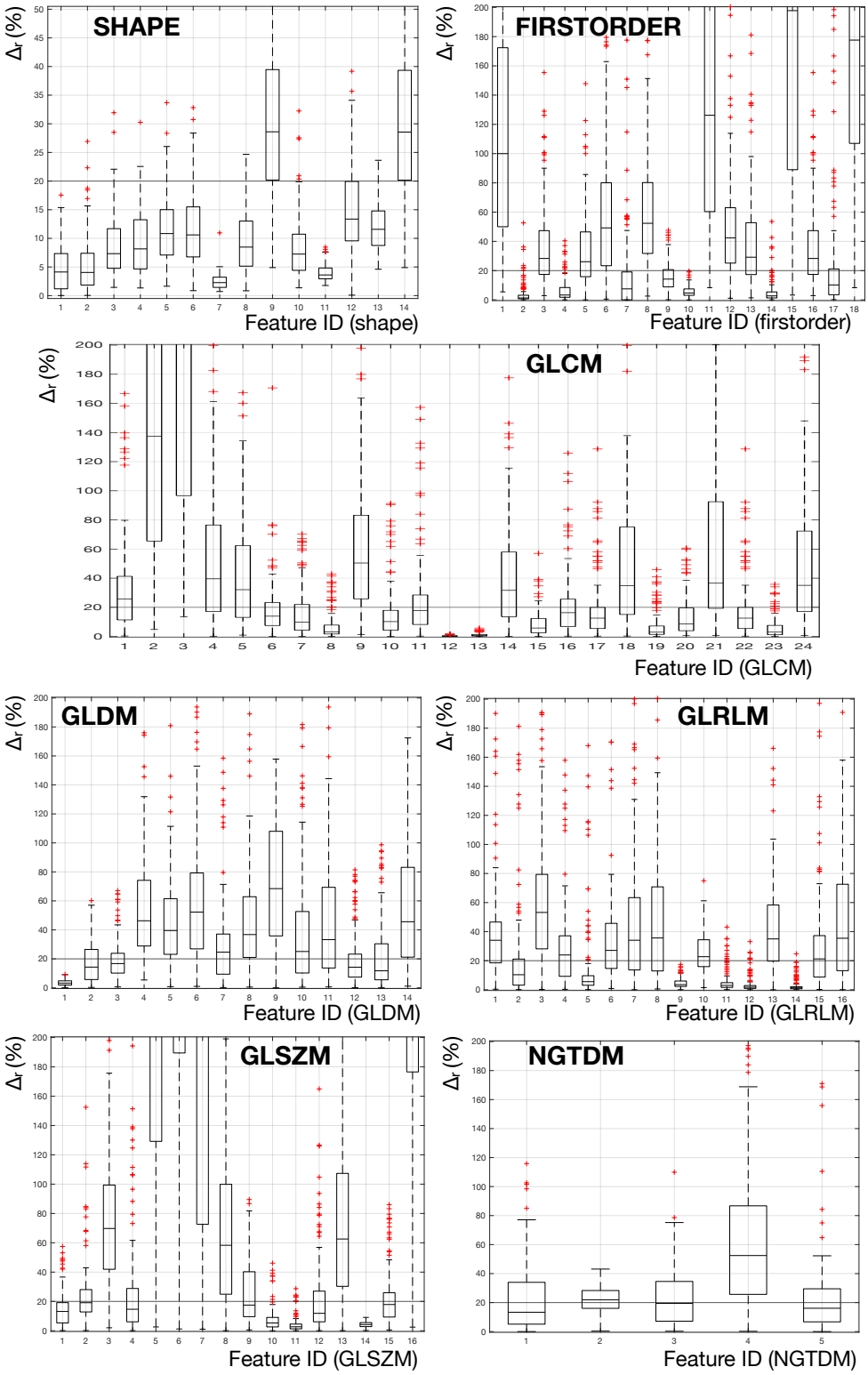


Fig. B.4 Variation in texture features in the internal dataset between the original segmentation and the region of interest following a dilation operation. Boxplots reach from the 25thto the 75th percentile. Whiskers extend to the 5th and 95th percentile. Please note the different scaling of the y-axis in the top left panel. Please refer to table B.4, page 247 for the IDs of the texture features. GLCM: Grey level coocurrence matrix,GLDM: Grey-level density matrix, GLRLM: Grey-level runlength matrix, GLSZM:Grey-level size zone matrix, NGTDM: Neighbourhood grey-tone difference matrix.

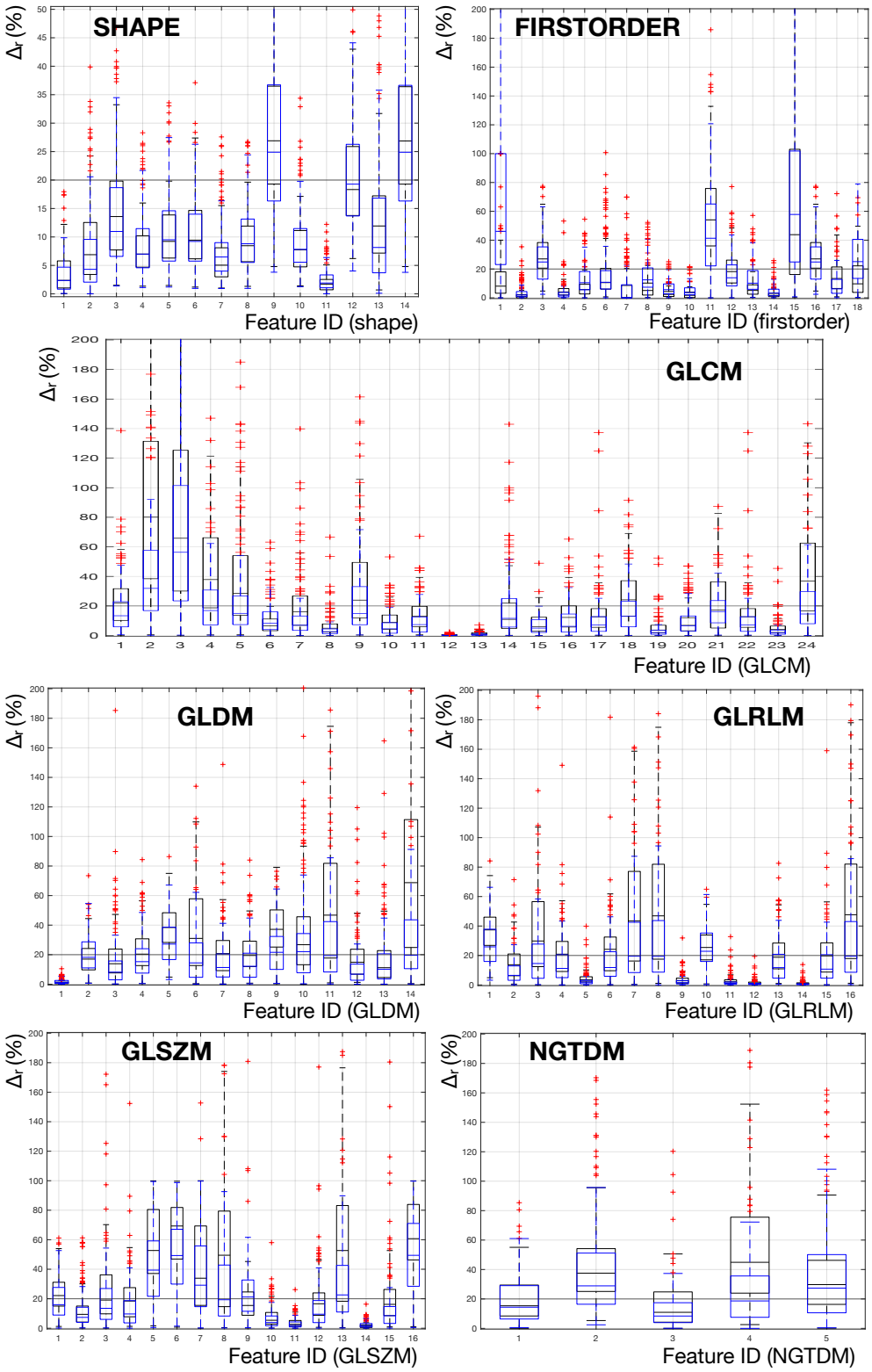


Fig. B.5 Variation in texture features the KiTS19 dataset (blue) between the original segmentation and the region of interest following an erosion operation. Boxplots reach from the 25thto the 75th percentile. Whiskers extend to the 5th and 95th percentile. Please note the different scaling of the y-axis in the top left panel. Please refer to table B.4, page 247 for the IDs of the texture features. GLCM: Grey level coocurrence matrix,GLDM: Grey-level density matrix, GLRLM: Grey-level runlength matrix, GLSZM:Grey-level size zone matrix, NGTDM: Neighbourhood grey-tone difference matrix.

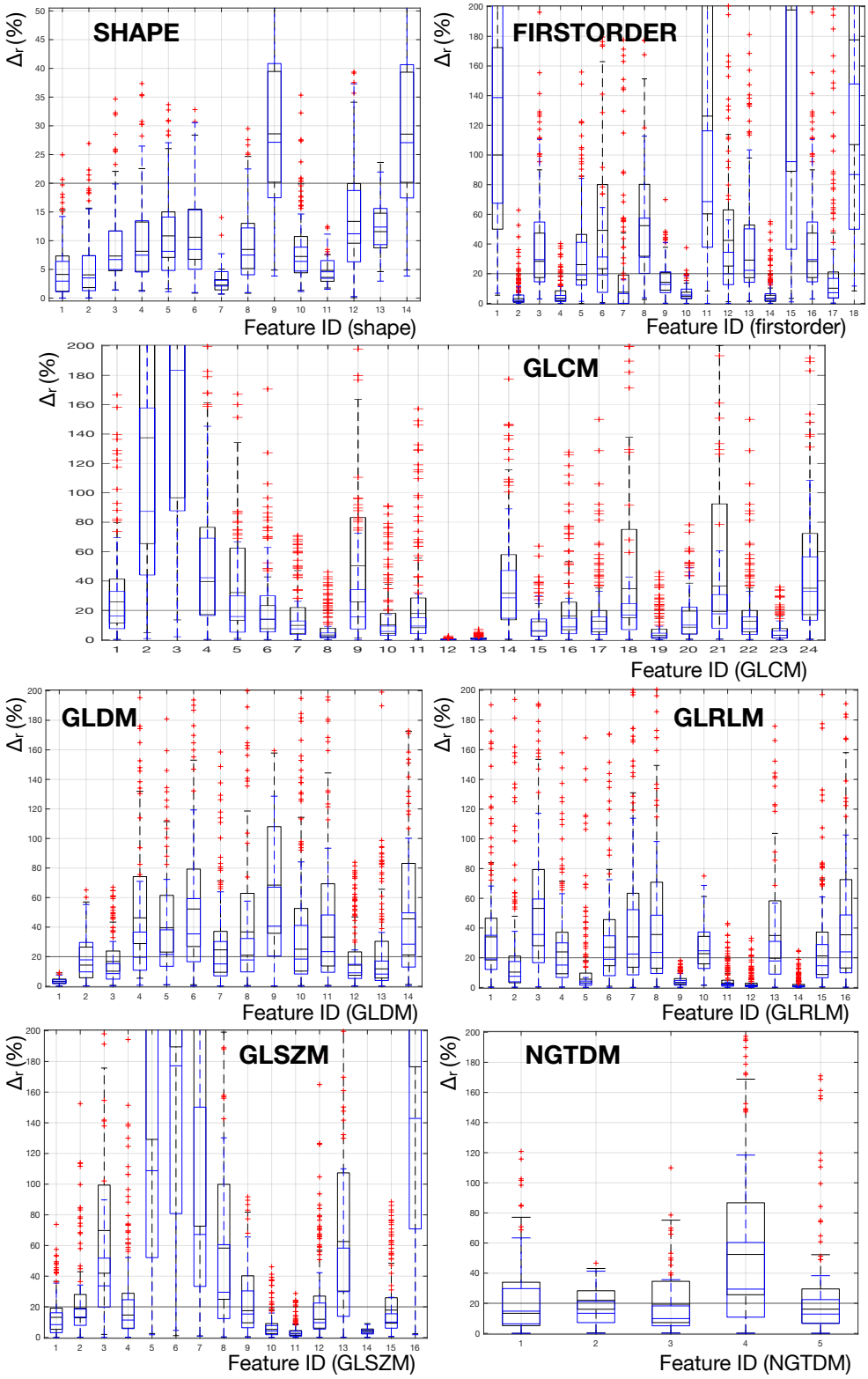


Fig. B.6 Variation in texture features in the KiTS19 dataset (blue) between the original segmentation and the region of interest following a dilation operation. Boxplots reach from the 25thto the 75th percentile. Whiskers extend to the 5th and 95th percentile. Please note the different scaling of the y-axis in the top left panel. Please refer to table B.4, page 247 for the IDs of the texture features. GLCM: Grey level coocurrence matrix,GLDM: Grey-level density matrix, GLRLM: Grey-level runlength matrix, GLSZM:Grey-level size zone matrix, NGTDM: Neighbourhood grey-tone difference matrix.

Table B.5 RCC Radiomics: Median Δ_r (%) per feature for the internal dataset after manual re-segmentation

Feature	16bins	32bins	64bins	128bins
shape Elongation	1.72	1.72	1.72	1.72
shape Flatness	1.93	1.93	1.93	1.93
shape LeastAxisLength	1.42	1.42	1.42	1.42
shape MajorAxisLength	1.16	1.16	1.16	1.16
shape Maximum2DDiameterColumn	1.54	1.54	1.54	1.54
shape Maximum2DDiameterRow	1.27	1.27	1.27	1.27
shape Maximum2DDiameterSlice	1.04	1.04	1.04	1.04
shape Maximum3DDiameter	1.19	1.19	1.19	1.19
shape MeshVolume	2.75	2.75	2.75	2.75
shape MinorAxisLength	1.11	1.11	1.11	1.11
shape Sphericity	1.15	1.15	1.15	1.15
shape SurfaceArea	1.69	1.69	1.69	1.69
shape SurfaceVolumeRatio	1.27	1.27	1.27	1.27
shape VoxelVolume	2.72	2.72	2.72	2.72
firstorder 10Percentile	3.12	3.12	3.12	3.12
firstorder 90Percentile	0.00	0.00	0.00	0.00
firstorder Energy	3.33	3.33	3.33	3.33
firstorder Entropy	2.92	2.23	1.74	1.41
firstorder InterquartileRange	2.13	2.13	2.13	2.13
firstorder Kurtosis	5.20	5.20	5.20	5.20
firstorder Maximum	0.00	0.00	0.00	0.00
firstorder MeanAbsoluteDeviation	1.83	1.83	1.83	1.83
firstorder Mean	0.80	0.80	0.80	0.80
firstorder Median	0.00	0.00	0.00	0.00
firstorder Minimum	25.58	25.58	25.58	25.58
firstorder Range	6.67	6.67	6.67	6.67
firstorder RobustMeanAbsoluteDeviation	1.43	1.43	1.43	1.43
firstorder RootMeanSquared	0.56	0.56	0.56	0.56
firstorder Skewness	17.34	17.34	17.34	17.34
firstorder TotalEnergy	3.33	3.33	3.33	3.33
firstorder Uniformity	5.90	5.62	5.62	5.28
firstorder Variance	4.39	4.39	4.39	4.39

Red (orange) indicates poorly reproducible features with median (third quantile) Δ_r above 20%.

Table B.5 Continued: RCC Radiomics: Median Δ_r (%) per feature for the internal dataset after manual re-segmentation

Feature	16bins	32bins	64bins	128bins
glcm Autocorrelation	11.65	11.98	12.09	12.15
glcm ClusterProminence	23.40	23.43	24.15	24.25
glcm ClusterShade	25.94	27.58	26.72	27.33
glcm ClusterTendency	12.50	12.46	12.53	12.53
glcm Contrast	11.28	12.05	12.34	12.40
glcm Correlation	1.85	1.78	1.72	1.71
glcm DifferenceAverage	6.39	6.39	6.39	6.41
glcm DifferenceEntropy	3.22	2.75	2.29	1.90
glcm DifferenceVariance	9.80	11.81	11.60	11.94
glcm Id	2.75	3.34	3.88	4.18
glcm Idm	3.68	4.85	5.61	5.86
glcm Idmn	0.12	0.12	0.12	0.12
glcm Idn	0.45	0.41	0.43	0.43
glcm Imc1	3.96	3.65	3.17	3.10
glcm Imc2	1.52	1.20	1.39	1.45
glcm InverseVariance	2.01	3.73	5.73	6.14
glcm JointAverage	6.33	6.58	6.68	6.74
glcm JointEnergy	12.81	11.82	11.38	9.70
glcm JointEntropy	3.23	2.45	1.88	1.39
glcm MCC	2.16	1.90	1.98	2.44
glcm MaximumProbability	13.30	12.14	5.84	4.57
glcm SumAverage	6.33	6.58	6.68	6.74
glcm SumEntropy	2.40	1.87	1.58	1.33
glcm SumSquares	11.85	12.12	12.18	11.94
gldm DependenceEntropy	0.61	0.54	0.52	0.55
gldm DependenceNonUniformity	6.09	5.74	5.73	4.53
gldm DependenceNonUniformityNormalized	5.36	5.52	4.32	3.53
gldm DependenceVariance	11.06	9.98	6.68	5.92
gldm GrayLevelNonUniformity	6.89	6.93	5.96	5.68
gldm GrayLevelVariance	11.81	12.53	13.18	12.34
gldm HighGrayLevelEmphasis	10.89	11.48	11.74	11.85
gldm LargeDependenceEmphasis	10.63	10.75	6.90	5.14

Red (orange) indicates poorly reproducible features with median (third quantile) Δ_r above 20%.

Table B.5 Continued: RCC Radiomics: Median Δ_r (%) per feature for the internal dataset after manual re-segmentation

Feature	16bins	32bins	64bins	128bins
gldm LargeDependenceHighGrayLevelEmphasis	18.99	19.46	16.49	16.15
gldm LargeDependenceLowGrayLevelEmphasis	6.90	8.16	11.24	16.11
gldm LowGrayLevelEmphasis	14.94	16.41	17.98	18.36
gldm SmallDependenceEmphasis	9.55	8.92	6.06	2.87
gldm SmallDependenceHighGrayLevelEmphasis	4.63	4.87	7.21	9.33
gldm SmallDependenceLowGrayLevelEmphasis	25.78	27.20	24.12	22.13
glrlm GrayLevelNonUniformity	5.20	5.28	5.65	5.69
glrlm GrayLevelNonUniformityNormalized	5.95	5.59	5.67	4.92
glrlm GrayLevelVariance	12.79	12.02	13.07	12.52
glrlm HighGrayLevelRunEmphasis	10.74	11.41	11.72	11.86
glrlm LongRunEmphasis	5.83	2.95	1.39	0.59
glrlm LongRunHighGrayLevelEmphasis	15.24	13.15	12.62	12.43
glrlm LongRunLowGrayLevelEmphasis	13.03	14.99	17.59	19.72
glrlm LowGrayLevelRunEmphasis	15.22	16.71	18.28	18.45
glrlm RunEntropy	0.98	1.22	1.33	1.20
glrlm RunLengthNonUniformity	5.91	3.88	3.09	2.89
glrlm RunLengthNonUniformityNormalized	3.26	1.69	0.84	0.39
glrlm RunPercentage	1.96	1.00	0.47	0.21
glrlm RunVariance	11.42	9.19	6.95	6.27
glrlm ShortRunEmphasis	1.52	0.74	0.35	0.16
glrlm ShortRunHighGrayLevelEmphasis	10.06	10.57	11.19	11.66
glrlm ShortRunLowGrayLevelEmphasis	17.32	18.07	19.40	18.67
glszm GrayLevelNonUniformity	7.45	6.24	3.61	3.29
glszm GrayLevelNonUniformityNormalized	7.52	4.92	4.98	5.06
glszm GrayLevelVariance	8.05	7.52	10.02	11.63
glszm HighGrayLevelZoneEmphasis	8.85	10.51	10.35	11.23
glszm LargeAreaEmphasis	17.58	33.86	25.44	10.99
glszm LargeAreaHighGrayLevelEmphasis	25.84	40.20	36.31	19.40
glszm LargeAreaLowGrayLevelEmphasis	12.46	34.65	15.50	14.22
glszm LowGrayLevelZoneEmphasis	19.98	17.96	19.02	19.24
glszm SizeZoneNonUniformity	11.31	11.50	10.25	6.33
glszm SizeZoneNonUniformityNormalized	3.07	1.79	2.31	2.70

Red (orange) indicates poorly reproducible features with median (third quantile) Δ_r above 20%.

Table B.5 Continued: RCC Radiomics: Median Δ_r (%) per feature for the internal dataset after manual re-segmentation

Feature	16bins	32bins	64bins	128bins
glszm SmallAreaEmphasis	1.82	1.07	1.18	1.28
glszm SmallAreaHighGrayLevelEmphasis	8.72	10.37	9.77	10.49
glszm SmallAreaLowGrayLevelEmphasis	21.83	19.77	18.42	20.18
glszm ZoneEntropy	1.51	1.06	0.42	0.38
glszm ZonePercentage	9.87	10.48	7.07	2.72
glszm ZoneVariance	17.58	34.15	27.94	15.07
ngtdm Busyness	9.03	9.86	9.02	8.79
ngtdm Coarseness	2.91	2.52	2.21	2.86
ngtdm Complexity	3.85	5.32	8.85	10.37
ngtdm Contrast	19.00	19.37	17.68	15.80
ngtdm Strength	9.06	8.32	7.93	8.44

Red (orange) indicates poorly reproducible features with median (third quantile) Δ_r above 20%.

Table B.6 RCC Radiomics: Median Δ_r (%) per feature for the internal dataset after morphological operations

Feature	Erosion	Opening (2)	Opening (1)	Intra-reader	Dilation
shape Elongation	2.37	1.44	0.46	1.72	4.13
shape Flatness	6.89	2.19	0.76	1.93	4.05
shape LeastAxisLength	13.58	2.07	0.75	1.42	7.33
shape MajorAxisLength	6.99	1.39	0.61	1.16	8.16
shape Maximum2DDiameterColumn	9.22	0.01	0.00	1.54	10.84
shape Maximum2DDiameterRow	9.44	0.30	0.00	1.27	10.60
shape Maximum2DDiameterSlice	5.04	5.55	2.20	1.04	2.27
shape Maximum3DDiameter	8.46	1.23	0.47	1.19	8.48
shape MeshVolume	26.90	5.59	1.73	2.75	28.60
shape MinorAxisLength	7.75	2.18	0.69	1.11	7.25
shape Sphericity	1.69	5.18	3.35	1.15	3.61
shape SurfaceArea	18.33	8.12	4.50	1.69	13.37
shape SurfaceVolumeRatio	11.93	2.86	2.25	1.27	11.58

Red (orange) indicates poorly reproducible features with median (third quantile) Δ_r above 20%.

Table B.6 Continued: RCC Radiomics: Median Δ_r (%) per feature for the internal dataset after morphological operations

Feature	Erosion	Opening (2)	Opening (1)	Intra-reader	Dilation
shape VoxelVolume	26.87	5.61	1.73	2.72	28.55
firstorder 10Percentile	8.11	2.27	0.00	3.12	100.00
firstorder 90Percentile	1.04	0.00	0.00	0.00	1.28
firstorder Energy	27.00	5.58	1.83	3.33	28.31
firstorder Entropy	4.07	0.63	0.39	1.74	3.31
firstorder InterquartileRange	5.71	2.63	0.00	2.13	26.09
firstorder Kurtosis	10.57	2.42	1.39	5.20	49.20
firstorder Maximum	0.00	0.00	0.00	0.00	7.52
firstorder MeanAbsoluteDeviation	5.10	1.33	0.81	1.83	52.35
firstorder Mean	2.41	0.52	0.21	0.80	14.20
firstorder Median	1.77	0.00	0.00	0.00	4.69
firstorder Minimum	54.08	0.00	0.00	25.58	126.23
firstorder Range	18.22	1.57	0.00	6.67	42.42
firstorder RobustMeanAbsoluteDeviation	5.15	1.41	0.59	1.43	29.16
firstorder RootMeanSquared	1.69	0.41	0.15	0.56	2.72
firstorder Skewness	43.81	8.80	4.51	17.34	197.78
firstorder TotalEnergy	27.00	5.58	1.83	3.33	28.31
firstorder Uniformity	13.41	1.93	1.11	5.62	10.20
firstorder Variance	9.73	2.84	1.60	4.39	177.63
glcm Autocorrelation	21.94	2.39	0.57	12.09	25.81
glcm ClusterProminence	80.06	8.56	4.10	24.15	137.49
glcm ClusterShade	65.80	14.70	7.85	26.72	324.89
glcm ClusterTendency	37.85	4.40	1.65	12.53	39.55
glcm Contrast	28.61	3.17	1.17	12.34	32.11
glcm Correlation	6.45	1.08	0.36	1.72	13.91
glcm DifferenceAverage	15.95	1.65	0.57	6.39	9.83
glcm DifferenceEntropy	4.95	0.61	0.24	2.29	3.23
glcm DifferenceVariance	23.95	3.02	1.44	11.60	50.42
glcm Id	8.99	0.81	0.34	3.88	10.20
glcm Idm	13.16	1.18	0.50	5.61	17.86

Red (orange) indicates poorly reproducible features with median (third quantile) Δ_r above 20%.

Table B.6 Continued: RCC Radiomics: Median Δ_r (%) per feature for the internal dataset after morphological operations

Feature	Erosion	Opening (2)	Opening (1)	Intra-reader	Dilation
glcm Idmn	0.28	0.04	0.01	0.12	0.24
glcm Idn	1.02	0.10	0.04	0.43	0.59
glcm Imc1	10.93	2.37	0.81	3.17	31.63
glcm Imc2	6.08	0.92	0.28	1.39	5.79
glcm InverseVariance	12.23	1.07	0.44	5.73	16.31
glcm JointAverage	12.68	1.04	0.25	6.68	12.58
glcm JointEnergy	24.14	2.70	1.31	11.38	34.78
glcm JointEntropy	3.71	0.51	0.22	1.88	2.92
glcm MCC	6.89	1.49	0.45	1.98	8.58
glcm MaximumProbability	17.23	2.39	1.15	5.84	36.55
glcm SumAverage	12.68	1.04	0.25	6.68	12.58
glcm SumEntropy	4.26	0.43	0.24	1.58	2.96
glcm SumSquares	36.93	3.08	1.52	12.18	35.10
gldm DependenceEntropy	0.79	0.21	0.13	0.52	3.27
gldm DependenceNonUniformity	17.01	4.92	1.78	5.73	14.31
gldm DependenceNonUniformityNormalized	13.88	1.40	0.83	4.32	16.74
gldm DependenceVariance	20.38	2.54	0.78	6.68	46.27
gldm GrayLevelNonUniformity	38.41	8.36	2.38	5.96	39.54
gldm GrayLevelVariance	31.03	4.28	2.39	13.18	52.26
gldm HighGrayLevelEmphasis	20.73	2.40	0.77	11.74	24.62
gldm LargeDependenceEmphasis	19.48	2.59	1.27	6.90	36.75
gldm LargeDependenceHighGrayLevelEmphasis	37.14	3.57	1.65	16.49	68.45
gldm LargeDependenceLowGrayLevelEmphasis	26.75	4.18	1.71	11.24	25.07
gldm LowGrayLevelEmphasis	46.77	3.31	1.25	17.98	33.28
gldm SmallDependenceEmphasis	14.90	2.06	1.28	6.06	14.25
gldm SmallDependenceHighGrayLevelEmphasis	11.20	3.06	1.21	7.21	11.76
gldm SmallDependenceLowGrayLevelEmphasis	68.75	6.82	3.08	24.12	45.61
glrlm GrayLevelNonUniformity	37.36	8.50	2.41	5.65	34.05
glrlm GrayLevelNonUniformityNormalized	13.71	1.89	1.15	5.67	10.35
glrlm GrayLevelVariance	29.83	4.12	2.44	13.07	53.21

Red (orange) indicates poorly reproducible features with median (third quantile) Δ_r above 20%.

Table B.6 Continued: RCC Radiomics: Median Δ_r (%) per feature for the internal dataset after morphological operations

Feature	Erosion	Opening (2)	Opening (1)	Intra-reader	Dilation
glrlm HighGrayLevelRunEmphasis	20.88	2.34	0.73	11.72	23.94
glrlm LongRunEmphasis	3.53	0.40	0.24	1.39	5.57
glrlm LongRunHighGrayLevelEmphasis	24.17	1.85	0.63	12.62	27.07
glrlm LongRunLowGrayLevelEmphasis	43.51	3.50	1.11	17.59	34.07
glrlm LowGrayLevelRunEmphasis	46.89	3.53	1.50	18.28	35.64
glrlm RunEntropy	2.83	0.47	0.29	1.33	3.52
glrlm RunLengthNonUniformity	25.43	5.72	1.78	3.09	22.75
glrlm RunLengthNonUniformityNormalized	2.17	0.28	0.17	0.84	2.98
glrlm RunPercentage	1.22	0.15	0.09	0.47	1.74
glrlm RunVariance	18.78	2.53	1.36	6.95	34.99
glrlm ShortRunEmphasis	0.89	0.12	0.07	0.35	1.27
glrlm ShortRunHighGrayLevelEmphasis	19.75	2.51	0.80	11.19	21.23
glrlm ShortRunLowGrayLevelEmphasis	47.65	3.54	1.57	19.40	35.44
glszm GrayLevelNonUniformity	22.16	6.02	2.12	3.61	13.16
glszm GrayLevelNonUniformityNormalized	9.43	1.69	1.14	4.98	19.11
glszm GrayLevelVariance	19.11	3.26	2.26	10.02	69.76
glszm HighGrayLevelZoneEmphasis	18.46	2.33	0.83	10.35	14.66
glszm LargeAreaEmphasis	52.70	7.40	3.71	25.44	380.10
glszm LargeAreaHighGrayLevelEmphasis	69.31	7.49	3.20	36.31	573.68
glszm LargeAreaLowGrayLevelEmphasis	33.92	8.10	4.66	15.50	240.50
glszm LowGrayLevelZoneEmphasis	49.50	6.36	2.22	19.02	58.30
glszm SizeZoneNonUniformity	15.45	7.41	3.41	10.25	17.50
glszm SizeZoneNonUniformityNormalized	5.41	1.62	1.09	2.31	5.30
glszm SmallAreaEmphasis	2.67	0.84	0.57	1.18	2.69
glszm SmallAreaHighGrayLevelEmphasis	16.60	2.10	1.08	9.77	11.88
glszm SmallAreaLowGrayLevelEmphasis	52.68	6.64	2.84	18.42	62.57
glszm ZoneEntropy	1.16	0.33	0.15	0.42	4.21
glszm ZonePercentage	16.20	2.44	1.52	7.07	17.88
glszm ZoneVariance	60.71	8.33	3.72	27.94	444.93
ngtdm Busyness	15.34	5.82	2.49	9.02	13.34

Red (orange) indicates poorly reproducible features with median (third quantile) Δ_r above 20%.

Table B.6 Continued: RCC Radiomics: Median Δ_r (%) per feature for the internal dataset after morphological operations

Feature	Erosion	Opening (2)	Opening (1)	Intra-reader	Dilation
ngtdm Coarseness	37.52	5.45	1.45	2.21	21.95
ngtdm Complexity	10.92	4.99	3.31	8.85	19.41
ngtdm Contrast	44.90	7.36	3.56	17.68	52.44
ngtdm Strength	29.74	7.21	3.63	7.93	16.10

Red (orange) indicates poorly reproducible features with median (third quantile) Δ_r above 20%.

Table B.7 RCC Radiomics: Median Δ_r (%) per feature for the KITS dataset after morphological operations

Feature	Erosion	Opening (2)	Opening (1)	Dilation
shape Elongation	2.35	1.67	0.64	2.96
shape Flatness	4.31	1.96	0.91	3.52
shape LeastAxisLength	10.96	1.94	0.85	6.71
shape MajorAxisLength	6.99	1.34	0.60	7.48
shape Maximum2DDiameterColumn	9.50	1.55	0.81	8.15
shape Maximum2DDiameterRow	9.22	1.01	0.27	8.47
shape Maximum2DDiameterSlice	6.48	6.09	2.78	3.11
shape Maximum3DDiameter	8.80	2.29	1.43	7.50
shape MeshVolume	24.89	5.90	2.51	27.13
shape MinorAxisLength	7.82	1.93	0.83	6.40
shape Sphericity	1.91	6.60	4.60	4.59
shape SurfaceArea	19.28	11.29	6.38	11.20
shape SurfaceVolumeRatio	8.15	4.34	3.16	12.37
shape VoxelVolume	24.89	5.89	2.52	27.07
firstorder 10Percentile	46.13	11.11	5.13	138.68
firstorder 90Percentile	1.97	0.72	0.31	2.00
firstorder Energy	24.87	6.64	2.41	29.49

Red (orange) indicates poorly reproducible features with median (third quantile) Δ_r above 20%.

Table B.7 Continued: RCC Radiomics: Median Δ_r (%) per feature for the KITS dataset after morphological operations

Feature	Erosion	Opening (2)	Opening (1)	Dilation
firstorder Entropy	1.86	0.78	0.44	2.89
firstorder InterquartileRange	10.51	2.60	1.60	19.16
firstorder Kurtosis	10.82	2.42	1.20	19.16
firstorder Maximum	0.00	0.00	0.00	0.00
firstorder MeanAbsoluteDeviation	12.48	3.13	1.62	31.05
firstorder Mean	5.67	1.09	0.58	12.80
firstorder Median	3.63	0.84	0.00	5.59
firstorder Minimum	41.33	7.64	0.00	68.58
firstorder Range	13.92	2.95	0.00	19.96
firstorder RobustMeanAbsoluteDeviation	10.18	2.76	1.33	22.25
firstorder RootMeanSquared	3.28	0.65	0.38	3.46
firstorder Skewness	57.83	10.40	5.61	95.51
firstorder TotalEnergy	24.87	6.64	2.41	29.49
firstorder Uniformity	6.21	2.46	1.39	7.17
firstorder Variance	25.11	6.46	3.50	86.84
glcm Autocorrelation	13.37	3.29	0.86	16.10
glcm ClusterProminence	31.89	11.46	5.19	87.37
glcm ClusterShade	56.31	15.04	6.71	183.39
glcm ClusterTendency	16.89	5.29	2.35	42.22
glcm Contrast	14.99	4.87	1.71	15.82
glcm Correlation	8.29	1.85	0.71	14.10
glcm DifferenceAverage	6.96	2.56	0.78	7.41
glcm DifferenceEntropy	2.61	0.89	0.34	2.24
glcm DifferenceVariance	14.93	4.50	1.93	15.65
glcm Id	4.54	1.21	0.40	5.65
glcm Idm	7.38	1.77	0.60	9.33
glcm Idmn	0.17	0.05	0.02	0.15
glcm Idn	0.53	0.16	0.06	0.53
glcm Imc1	11.62	3.33	1.22	28.54
glcm Imc2	5.01	1.18	0.47	6.24

Red (orange) indicates poorly reproducible features with median (third quantile) Δ_r above 20%.

Table B.7 Continued: RCC Radiomics: Median Δ_r (%) per feature for the KITS dataset after morphological operations

Feature	Erosion	Opening (2)	Opening (1)	Dilation
glcm InverseVariance	6.22	1.89	0.64	8.76
glcm JointAverage	7.23	1.74	0.46	7.55
glcm JointEnergy	13.17	4.15	2.08	16.91
glcm JointEntropy	1.98	0.76	0.35	1.98
glcm MCC	6.50	1.37	0.65	9.98
glcm MaximumProbability	16.20	3.98	1.77	17.66
glcm SumAverage	7.23	1.74	0.46	7.55
glcm SumEntropy	1.82	0.66	0.31	3.16
glcm SumSquares	14.57	5.20	2.23	32.86
gldm DependenceEntropy	1.30	0.31	0.17	3.02
gldm DependenceNonUniformity	18.21	5.08	2.39	17.85
gldm DependenceNonUniformityNormalized	8.20	1.45	0.83	8.54
gldm DependenceVariance	15.27	2.96	0.79	19.60
gldm GrayLevelNonUniformity	27.18	6.92	1.54	21.51
gldm GrayLevelVariance	14.45	5.93	2.90	35.44
gldm HighGrayLevelEmphasis	11.34	3.13	1.19	15.09
gldm LargeDependenceEmphasis	12.16	2.42	1.25	17.79
gldm LargeDependenceHighGrayLevelEmphasis	21.60	4.10	1.92	40.68
gldm LargeDependenceLowGrayLevelEmphasis	22.06	5.85	1.91	18.39
gldm LowGrayLevelEmphasis	19.42	9.26	4.52	23.43
gldm SmallDependenceEmphasis	6.81	2.30	1.44	9.07
gldm SmallDependenceHighGrayLevelEmphasis	10.16	2.91	1.19	8.21
gldm SmallDependenceLowGrayLevelEmphasis	19.65	13.86	6.38	28.47
glrlm GrayLevelNonUniformity	25.85	6.72	1.61	19.18
glrlm GrayLevelNonUniformityNormalized	6.50	2.51	1.43	7.50
glrlm GrayLevelVariance	14.54	5.92	2.92	35.65
glrlm HighGrayLevelRunEmphasis	11.21	3.14	1.20	15.10
glrlm LongRunEmphasis	1.88	0.44	0.24	2.85
glrlm LongRunHighGrayLevelEmphasis	11.82	2.90	1.06	19.00
glrlm LongRunLowGrayLevelEmphasis	19.50	8.91	4.12	22.53

Red (orange) indicates poorly reproducible features with median (third quantile) Δ_r above 20%.

Table B.7 Continued: RCC Radiomics: Median Δ_r (%) per feature for the KITS dataset after morphological operations

Feature	Erosion	Opening (2)	Opening (1)	Dilation
glrlm LowGrayLevelRunEmphasis	19.43	9.66	4.47	23.58
glrlm RunEntropy	1.57	0.58	0.32	3.24
glrlm RunLengthNonUniformity	22.84	5.49	2.35	24.70
glrlm RunLengthNonUniformityNormalized	1.14	0.31	0.16	1.60
glrlm RunPercentage	0.62	0.16	0.08	0.91
glrlm RunVariance	11.75	2.31	1.22	17.86
glrlm ShortRunEmphasis	0.46	0.12	0.07	0.67
glrlm ShortRunHighGrayLevelEmphasis	10.72	3.18	1.22	14.62
glrlm ShortRunLowGrayLevelEmphasis	19.53	10.09	4.52	23.97
glszm GrayLevelNonUniformity	15.97	4.67	2.06	8.41
glszm GrayLevelNonUniformityNormalized	7.38	2.40	1.37	13.32
glszm GrayLevelVariance	13.52	4.95	3.12	33.69
glszm HighGrayLevelZoneEmphasis	9.69	3.36	1.35	11.39
glszm LargeAreaEmphasis	39.40	8.18	3.91	108.79
glszm LargeAreaHighGrayLevelEmphasis	49.29	9.54	4.45	177.15
glszm LargeAreaLowGrayLevelEmphasis	29.18	9.07	3.87	67.19
glszm LowGrayLevelZoneEmphasis	19.33	13.87	5.46	29.53
glszm SizeZoneNonUniformity	21.18	7.46	3.92	15.32
glszm SizeZoneNonUniformityNormalized	3.90	1.57	1.01	4.28
glszm SmallAreaEmphasis	1.94	0.78	0.52	2.28
glszm SmallAreaHighGrayLevelEmphasis	9.51	3.75	1.22	9.68
glszm SmallAreaLowGrayLevelEmphasis	22.60	14.08	6.01	29.92
glszm ZoneEntropy	1.76	0.34	0.14	3.34
glszm ZonePercentage	8.50	2.24	1.43	10.04
glszm ZoneVariance	49.43	9.77	4.22	142.86
ngtdm Busyness	14.39	4.64	1.99	14.77
ngtdm Coarseness	28.89	4.24	1.52	13.36
ngtdm Complexity	8.39	3.97	2.30	9.82
ngtdm Contrast	18.65	8.10	3.34	29.41
ngtdm Strength	27.36	5.81	2.11	12.83

Red (orange) indicates poorly reproducible features with median (third quantile) Δ_r above 20%.

Appendix C

Response to tyrosine kinase inhibitors in RCC assessed using conventional imaging: supplementary documentation

C.1 Inclusion and exclusion criteria for the NeoSun trail

From the NeoSun trial protocol developed by Prof Tim Eisen and Dr Sarah Welsh:

Inclusion criteria:

- Presumed metastatic (Stage IV) renal cell carcinoma, as defined by the referring clinician,
- Clinically fit and scheduled for nephrectomy,
- No prior systemic therapy for renal cell carcinoma,
- Male or female, 18 years of age or older,
- Life expectancy $\geqslant$12 weeks,
- ECOG performance status 0 or 1,
- Radiologically documented, RECIST measurable, metastatic disease. Measurable disease, defined by the presence of at least one lesion outside of the kidney which can be measured in at least one dimension with the longest diameter $\geqslant$ 20mm using chest X-ray, $\geqslant$ 10mm calliper measurement by clinical exam or $\geqslant$ 10mm using spiral CT,
- Adequate end organ function:

  - Serum aspartate transaminase (AST) or serum alanine transaminase (ALT) $\leqslant$ 2.5$\times$ upper limit of normal (ULN), or AST or ALT $\leqslant 5 \times$ ULN if liver function abnormalities are due to underlying malignancy,
  - Total serum bilirubin $\leqslant 1.5 \times$ ULN,
  - Absolute neutrophil count (ANC) $\geqslant 1.5 \times 10^9$/L,
  - Platelets $\geqslant 100 \times 10^9$/L,
  - Haemoglobin $\geqslant$90 g/L (However where a transfusion is scheduled the patient may have a lower a value but this must be discussed and approved),
  - Serum creatinine $\leqslant 1.5 \times$ UL and
  - Prothrombin time (PT) $\leqslant 1.5 \times$ ULN,

- Willingness and ability to comply with scheduled visits, treatment plans, laboratory tests, and other study procedures and
- Before patient registration, written informed consent, including for preoperative biopsy, must be given according to ICH/GCP, national and local regulations

Exclusion criteria:

- Presence of $\geqslant$ grade 3 (NCI CTCAE grading version 3.0) haemorrhage within 4 weeks of registration,
- Any previous chemotherapy or investigational treatment given for renal carcinoma,
- Diagnosis of any other malignancy within the last 5 years, except for adequately treated basal cell carcinoma, squamous cell skin cancer, or in situ cervical cancer,
- Any of the following within the 12 months prior to study drug administration: myocardial infarction, severe/unstable angina, coronary/peripheral artery bypass graft, symptomatic congestive heart failure, cerebrovascular accident or transient ischemic attack, or pulmonary embolism,
- Any major surgery in the previous 4 weeks,
- Co-administration of CYP3A4 inducer or inhibitor medications that would require an increase or decrease in sunitinib dosage from the study dose of 50 mg od,
- Pre-existing thyroid abnormality with thyroid function that cannot be maintained in the normal range with medication,

- Ongoing cardiac dysrhythmias of $\geqslant$ grade 2 (NCI CTCAE grading version 3.0), atrial fibrillation of any grade, or prolongation of the QTc interval to $>$450 ms for males or $>$470 ms for females,
- Treatment with drugs of dysrhythmic potential is not allowed,
- Concurrent participation in another clinical trial using an investigational medicinal product,
- Known human immunodeficiency virus (HIV) or acquired immunodeficiency syndrome (AIDS)-related illness,
- Pregnancy or breastfeeding,
- Female patients must be surgically sterile, be postmenopausal, or must agree to use effective contraception during the period of therapy and for 6 months after. All female patients with reproductive potential must have a negative pregnancy test (serum or urine) prior to enrolment. Male patients must be surgically sterile or must agree to use effective contraception during the period of therapy and for 6 months after. Barrier contraception is recommended,
- Other severe acute or chronic medical or psychiatric condition, or laboratory abnormality that may increase the risk associated with study participation or study drug administration, or in the judgment of the investigator would make it undesirable for the patient to enter the trial or
- Uncontrolled hypertension. Patients are excluded with current blood pressure of either systolic $\geqslant$ 150 mmHg or diastolic $\geqslant$90 mmHg. Patients using anti-hypertensive medication to control blood pressure to these levels are elgible.

C.2 NeoSun ^{1}H-MRI protocol parameters

Table C.1 NeoSun: MRI Parameters

Sequence	TE / TR [ms]	Flip angle [°]	Matrix size	Slice thickness / Orientation	Comments
T_2w FRFSE	48-69 / 4000	90	320×224	4 mm coronal / 1 mm gap	ETL: 10-13, / RT, 2 Nex
T_1w FSPGR	4.8 / 139	70	256×256	4 mm coronal / 1 mm gap	BH, 0.75 Nex / ASSET 2
IVIM-EPI	73.4 / 4000	90	104×104	4 mm sagital / 1 mm gap	b: 0, 150, 500, 700 and 900 s/mm^2 6 Nex, ASSET 2
R_2^*-mapping	4.76-47.6 / spacing: 4.76 / 100	25	128×128	4 mm sagital / 1 mm gap	BH, ASSET 2
T_1-mapping	1.6 / 3.9	1, 3, 5, 10, / 15 and 20	160×160	5 mm cor-obl	BH, 1 Nex
DCE-MRI	1.6 / 3.9	18	160×160	5 mm cor-obl	multiple BH, temp resolution 4.3-6.4 s 0.5 Nex

ASSET: array coil spatial sensitivity encoding, BH: breath hold, DCE: dynamic contrast-enhanced, ETL: echo train lenght, IVIM: intravoxel incoterent motion, Nex: number of excitations, RT: respiratory triggered, TE: echo time, TR: repetition time

C.3 NAXIVA inclusion and exclusion criteria

From the NAXIVA trial protocol developed by Prof Grant Stewart and Prof Robert Jones:

Inclusion criteria:

- Age $\geqslant$18 years,
- Histologically proven clear cell RCC,
- Immediate resection of the primary tumour considered technically possible,
- Suitable for and willing to undergo nephrectomy (either cytoreductive or with curative intent),
- cT3b, cT3c, cT3a (main renal vein),
- N0, N1, or Nx,
- M0, or M1,
- ECOG performance status 0 – 1,
- Urinalysis $<$2+ protein. If dipstick is $\geqslant$2+ then a 24-hour urine collection or urinary protein creatinine ratio (PCR) should be performed and the patient may enter NAXIVA only if urinary protein is $<$2g per 24 hours or PCR $<$200mg/mmol,
- Serum creatinine $\leqslant$1.5xULN or estimated creatinine clearance $\geqslant$30mL/min as calculated using the Cockcroft-Gault (CG) equation and
- All female patients with reproductive potential must have a negative serum or urine pregnancy test within a maximum of 14 days prior to starting trial treatment.

Exclusion criteria:

- For M1 patients: poor risk on Memorial Sloan Kettering Cancer Centre (MSKCC) score and deemed suitable for cytoreductive nephrectomy at time of enrolment,
- Other invasive malignancy within the last 2 years. Patients with previous history of malignancies with a negligible risk of metastasis or death and treated with expected curative intent are eligible, for example but not exclusively:
  - Carcinoma in situ of the cervix,

  - Basal or squamous cell skin cancer or
  - Localised low to intermediate risk prostate cancer treated with curative intent and absence of prostate-specific antigen (PSA) relapse; or prostate cancer (Stage T1/T2a, Gleason $\leqslant$6 and PSA $<$10ng/mL) undergoing active surveillance and treatment,

- Women who are pregnant or are breastfeeding. Female patients must be surgically sterile, be postmenopausal, or must agree to use effective contraception during the period of therapy and up to 1 week after treatment, male patients must be surgically sterile or must agree to use effective contraception during the period of therapy and for 6 months after completion of study drug (Patients who do not meet this will not be are not eligible),

- Current signs or symptoms of severe progressive or uncontrolled hepatic, endocrine or pulmonary disease other than directly related to RCC,

- Gastrointestinal abnormalities including: a. inability to take oral medication; b. requirement for intravenous alimentation; c. prior surgical procedures affecting absorption including total gastric resection; d. treatment for active peptic ulcer disease in the past 6 months; e. active gastrointestinal bleeding, unrelated to cancer, as evidenced by hematemesis, hematochezia or melena in the past 3 months without evidence of resolution documented by endoscopy or colonoscopy; f. malabsorption syndromes,

- Current use or anticipated need for treatment with drugs that are known potent CYP3A4 inhibitors (see section 4.4, concomitant therapy),

- Current use, or anticipated need for treatment with, drugs that are known CYP3A4 inducers or substrates for CYP1A2 (see section 4.4, concomitant therapy),

- Requirement of anticoagulant therapy with oral vitamin K antagonists. Low-dose anticoagulants for maintenance of patency of central venous access device or prevention of deep venous thrombosis is allowed. Therapeutic use of low molecular weight heparin is allowed,

- Active seizure disorder, spinal cord compression, or carcinomatous meningitis,

- Any of the following within 12 months prior to study entry: myocardial infarction, uncontrolled angina, coronary/peripheral artery bypass graft, symptomatic congestive heart failure, cerebrovascular accident or transient ischemic attack,

- Uncontrolled hypertension (>160/100 mmHg despite optimised antihypertensive treatment),
- Known human immunodeficiency virus (HIV) or acquired immunodeficiency syndrome (AIDS)-related illness,
- Total serum bilirubin $\geqslant 1.5 \times$ ULN; Aspartate aminotransferase (AST) and alanine aminotransferase (ALT) $\geqslant 2.5 \times$ ULN,
- Neutrophil count $< 1.0 \times 109$/L; platelet count $< 100 \times 109$/L; Hb $\leqslant$ 90g/L,
- Known severe hepatic impairment (Child-Pugh class C) or
- Known hypersensitivity to axitinib or any of its excipients. Specifically patients with hereditary galactose intolerance, Lapp lactase deficiency or glucose-galactose malabsorption should not enter the study.

C.4 NAXIVA ^{1}H-MRI protocol parameters

NAXIVA was a multi-centre trial lead by the University of Cambridge. While the radiology manual defined guidelines for the MR acquisition protocol, differing technological capabilities at the participating sites meant that protocols varied slightly. As a reference, this section describes the protocol employed at Addenbrookes Hospital, Cambridge University Hospitals NHS Foundation Trust. All patient scans were acquired on a clinical 1.5 T MRI scanner (Discovery MR 450, GE Medical)

Table C.2 NAXIVA: MRI Parameters

Sequence	TE [ms]	TR [ms]	Flip [°]	FoV [mm]	Slice [mm]	Orientation	Type	Comments
T_1w FIESTA	1.3	3.0	45	410 × 410	8	Ax	2D	BH
T_1w FIESTA	1.5	3.4	55	340 × 340	8	Cor	2D	BH
T_2w SSFSE	60.7	1144	90	340 × 340	8	Ax	2D	BH
T_2w SSFSE	60.7	1248	90	340 × 340	8	Cor	2D	BH
T_2w FRFSE	75.4	5714	160	400 × 400	8	Ax	2D	RT, ETL: 13, ASSET 2 Nex 4
DWI	48.5	2000	90	380 × 380	8	Ax	2D	BH, b-values 0, 600 mm/s2 ASSET 2, Nex 4
T_1w Lava ±C	3.2	6.4	12	400 × 400	6	Ax	3D	BH, pre- and post-contrast
T_1w Lava ± C	3.2	6.9	12	400 × 400	4	Cor	3D	BH, pre- and post-contrast, angiographic triggering, multi-phase (8 phases)

ASSET: array coil spatial sensitivity encoding, BH: breath hold, C: contrast, ETL: echo train lenght, FoV: field of view, Nex: number of excitations, RT: respiratory triggered, TE: echo time, TR: repetition time

C.5 A-PREDICT CT protocol parameters

Table C.3 A-PREDICT: CT Parameters

	Hospital 1	Hospital 2	Hospital 3	Hospital 4
CT system	GE LightSpeed Pro 32	Siemens Somatom Definition Flash	GE Discovery CT750	GE Lightspeed VCT
Tube voltage	120	120	120	120
Reconstruction kernel	Standard	B30f	Standard	Standard
Slice thickness	5	5	5	5
Contrast agent	Omnipaque 300	Omnipaque 300	Omnipaque 300	

	Hospital 5	Hospital 6	Hospital 7
CT system	Siemens Somatom Definition AS+	Toshiba Aquilion ONE	Siemens Somatom Definition AS
Tube voltage	100	120	120
Reconstruction kernel	I30f	FC18	I26f
Slice thickness	5	3	1
Contrast agent			Omnipaque 300

Appendix D

Hyperpolarised [1-^{13C}]-pyruvate MRI is sensitive to tumour aggressiveness and adverse outcome in clear cell renal cell carcinoma: the MISSION renal study: supplementary documentation

D.1 MISSION ^{1}H-MRI protocol parameters

Table D.1 MISSION and WIRE: Acquisition parameters of the ^{1}H-MRI

Sequence	TE [ms]	TR [ms]	Flip Angle [°]	Matrix size	Field of View [mm]	Slice [mm]	Orientation	Type	Parallel imaging factor	Comments
T_1w LavaFlex	1.1 / 2.2	3.8	10	224 × 224	400 × 400	4	Coronal	3D	1.5 (ARC)	Single BH
T_2w FOCUS CUBE	100	RT	90	256 × 224	360 × 360	4	Coronal	3D	2.0 (ARC)	RT, ETL: 120
T_2* / R_2* mapping	2.3–36.2	110	30	256 × 224	400 × 400	4	Coronal	2D	2.0 (ASSET)	Multiple BH, ETL: 12 12 echos, 3.1ms spacing
IVIM-DWI	80	RT	90	96 × 96	288 × 288	4	Coronal	2D	1.5 (ASSET)	b-values: 0, 10, 20, 30, 50, 100, 300, 500, 700, 900 s/mm^2
B_1 mapping	13	31		128 × 128	400 × 400	12	Coronal	2D		Bloch-Siegert method
T_1 mapping	1.1 / 2.2	3.8	2, 3, 5, 8, 14	196 × 172	400 × 400	4	Coronal	3D	1.2 × 1.0 (ARC)	Single BH
DCE-MRI	1.1 / 2.2	3.8	18	196 × 172	400 × 400	4	Coronal	3D	2.0 × 1.5 (ARC)	0.1 ml of 1M Gadobutrol i.v. temporal resolution: 13s

ARC: autocalibrating reconstruction for cartesian imaging, ASSET: array coil spatial sensitivity encoding, BH: breath hold, ETL: echo train length, RT: respiratory triggered, TE: echo time, TR: repetition time.

D.2 Hyperpolarised ^{13}C-pyruvate MRI quality control parameters for the MISSION cohort

Table D.2 MISSION: Quality control parameters for HP-^{13}C-pyruvate preparation

Quality control parameter	Mean±S.D.
Polarization [%]	24.5±7.8
pH	7.7±0.3
EPA concentration [µmol/L]	0.8±0.4
Temperature [°C]	33.6±1.0
Pyruvate concentration [mmol/L]	261.6±7.2
Injected Volume [mL]	34.9±3.6
Time from dissolution to injection [%]	60.3±3.9

αKG: α-ketoglutarate, Cit: citrate, Gln: glutamine, Glu: glucose, Fu: fumarate, Lac: lactate, Mal: malate, Pyr: pyruvate, Suc: succinate.

D.3 Supplementary LC-MS data comparing perfused and ischaemic tissue samples

Table D.3 MISSION: Metabolite concentrations in perfused and ischaemic tissue samples

	Biopsy site	Pyr	Lac	Suc	αKG	Cit	Gln	Glu	Mal	Fu
Perfused	Normal kidney	8.0	642.1	11.3	52.8	277.3	94.7	1708	79.1	16.1
	RCC (high metabol.)	4.9	1406	24.5	35.1	292.6	92.6	1398	107.3	21.8
	RCC (low metabol.)	2.8	1126	21.7	0.5	125.1	82.9	713.5	109.8	24.7
Ischaemic	N1	6.7	2671	108.0	0.5	107.0	23.4	446.1	167.6	36.4
	N2	2.6	2413	99.9	0.8	88.5	21.3	441.8	136.0	29.6
	T1	8.5	6372	121.2	7.1	39.4	105.3	1081	179.7	43.8
	T2	22.1	6146	109.4	8.6	45.9	103.9	879.5	194.5	43.4
	T3	9.7	6439	106.7	6.8	43.7	110.0	1328	166.6	39.5
	T4	10.0	7004	122.4	8.2	37.1	102.5	1487	155.4	34.3
	T5	3.2	4644	71.3	11.4	63.8	126.4	1973	88.1	20.0
	T6	3.8	4551	69.8	1.6	62.6	112.8	1021	117.2	27.3
	T7	7.6	4185	33.3	2.0	49.3	76.5	515.1	100.3	23.8
	T8	4.6	5908	114.0	7.6	65.3	90.5	1082	112.2	26.4
	T9	6.7	6101	164.2	4.8	33.0	90.1	1123	87.9	18.5
	T10	5.0	6033	109.5	4.1	46.5	94.5	1144	148.0	34.7
	T11	6.3	6335	69.1	1.8	34.3	86.6	728.6	157.7	36.1
	T12	10.1	6224	119.7	1.9	40.5	97.1	1100	172.3	40.5
	T13	12.8	7041	43.3	1.6	37.8	91.5	870.8	198.0	47.0
	T15	10.3	3354	44.18	3.1	63.9	32.4	340.5	58.4	13.5

αKG: α-ketoglutarate, Cit: citrate, Gln: glutamine, Glu: glucose, Fu: fumarate, Lac: lactate, Mal: malate, Pyr: pyruvate, Suc: succinate.

Appendix E

WIRE Trial supplementary documentation

From the WIRE trial protocol developed by Prof Grant Stewart and Dr Sarah Welsh:

E.1 Supplementary inclusion and exclusion criteria

E.1.1 Inclusion criteria

To be included in the trial the participant must meet all of the following criteria:

- Capable of giving signed informed consent which includes compliance with the requirements and restrictions listed in the informed consent form (ICF) and in this protocol.
- Aged $\geqslant$18 years and over.
- Predicted life expectancy $\geqslant$16 weeks.
- Eastern Cooperative Oncology Group (ECOG) performance status (PS) 0 or 1.
- Have biopsy proven clear cell RCC, within 6 weeks prior to consent.
- Allow access to archival FFPE tumour tissue from biopsy and nephrectomy.
- Have a surgically resectable tumour as determined by the treating Urologist
- T1b or above, any N status, M0, OR have any T or N status, M1 (but if M1, the participant must be deemed suitable for cytoreductive nephrectomy at time of enrolment).

- No prior exposure to PARP inhibitors (including but not limited to olaparib), tyrosine kinase inhibitors (including but not limited to cediranib, sunitinib, pazopanib, axitinib, bevacizumab and cabozantinib), immunotherapy or immune checkpoint inhibitors (including but not limited to other anti-CTLA-4, anti-PD-1, or anti-PD-L1 antibodies, including durvalumab), nor prior treatment with an mammalian target of rapamycin (mTOR) inhibitor (including, but not limited to everolimus, temsirolimus, or sirolimus). Prior cytokine therapy (eg, IL-2, IFN-α) or treatment with cytotoxics is allowed.

- At least one measurable lesion according to RECIST Version 1.1 at screening that can be accurately assessed at screening by - MRI and is suitable for repeated assessment. A previously irradiated lesion cannot be considered a target lesion. Radiographic disease assessment can be performed up to 28 days prior to the first dose of trial treatment. It is acceptable for the measurable lesion to be planned for removal at surgery. CT reported RECIST assessments are acceptable at screening for participants with chest metastases.

- Have adequate organ and marrow function, as defined below (measured within 28 days of first dose of trial medication):

  - Haemoglobin $\geqslant$90 g/L
  - Platelet count $\geqslant$100 x 10^9/L
  - Neutrophil count $\geqslant$1.5 x 10^9/L
  - Creatinine clearance $\geqslant$30mL/min (calculated by Cockcroft and Gault equation: where estimated creatinine clearance = (140-age[years]) $\times$ weight (kg) (xF)a serum creatinine (mg/dL) $\times$ 72a where F=0.85 for females and 1 for males)) Participants with 2+ proteinuria on dipstick must also have UPC <0.5 on 2 consecutive samples

- Adequate hepatic function:

  - Alanine Aminotransferase (ALT) (SGPT) $\leqslant$2.5x the institutional upper limit of normal (ULN) unless liver metastases are present, in which case it must be $\leqslant$5x the institutional ULN, AND
  - AST $\leqslant$2.5x the institutional ULN unless liver metastases are present, in which case it must be $\leqslant$5x the institutional ULN, AND

  - Total bilirubin $\leqslant 1.5\times$ the institutional ULN unless in the presence of known or suspected Gilbert's syndrome- the inclusion of potential participants with known/suspected Gilbert's syndrome must be discussed with the Trial Oncologist prior to their inclusion in the trial.

- Evidence of post-menopausal status or negative serum pregnancy test for female pre-menopausal participants. Women will be considered post-menopausal if they have been amenorrhoeic for 12 months without an alternative medical cause. The following age-specific requirements apply:

  - Women <50 years of age would be considered post-menopausal if they have been amenorrheic for 12 months or more following cessation of exogenous hormonal treatments and if they have luteinizing hormone and follicle-stimulating hormone levels in the post-menopausal range for the institution or underwent surgical sterilization (bilateral oophorectomy or hysterectomy)
  - Women $\geqslant$50 years of age would be considered post-menopausal if they have been amenorrhoeic for 12 months or more following cessation of all exogenous hormonal treatments, had radiation-induced menopause with last menses >1 year ago, had chemotherapy-induced menopause with last menses >1 year ago, or underwent surgical sterilization (bilateral oophorectomy, bilateral salpingectomy or hysterectomy).
  - For women of childbearing potential a negative serum pregnancy test must be performed within 28 days of trial treatment and confirmed prior to treatment on day 1.

- Male participants must be willing to use a condom during treatment and for 3 months after the last dose of trial treatment when having sexual intercourse with a pregnant woman or with a woman of childbearing potential. Female partners of male participants should also be willing to use a highly effective form of contraception if they are of childbearing potential.

- Participant is willing and able to comply with the protocol for the duration of the trial.

- Adequately controlled thyroid function, with no symptoms of thyroid dysfunction

Specific inclusion criteria for olaparib containing arms:

- Haemoglobin must be $\geqslant$100 g/L

- If abnormalities in the Full Blood Count (and it is clinically indicated): Peripheral blood smear with no features of myelodysplastic syndrome or acute myeloid leukaemia
- Serum creatinine must be $\leqslant$1.5x the institutional ULN concurrent with creatinine clearance $\geqslant$51mL/min (calculated by Cockcroft and Gault equation as above) or based on a 24 hour urine creatinine clearance test.

Specific inclusion criteria for olaparib plus cediranib arm only:

- Urine protein:creatinine ratio (UPC) $\leqslant$1 OR $\leqslant$2+ proteinuria on two consecutive dipsticks taken no less than 1 week apart. Patients with 2+ proteinuria on dipstick must also have UPC <0.5 on 2 consecutive samples.

E.1.2 Exclusion criteria

The presence of any of the following core exclusion criteria will preclude participant inclusion:

- cT1a N0 M0-staged Renal Cell Cancer
- Participants with brain metastases. A scan to confirm the absence of brain metastases is not required.
- Participants with spinal cord compression unless considered to have received definitive treatment for this and evidence of clinically stable disease for 28 days prior to start of first dose of treatment.
- History of leptomeningeal carcinomatosis.
- Body weight $\leqslant$30kg
- Contraindication to cediranib, olaparib, durvalumab or chimeric or humanized antibodies or fusion proteins.
- Specifically participants with hereditary galactose intolerance, Lapp lactase deficiency or glucose-galactose malabsorption should not enter the trial.
- History of hypersensitivity to active or inactive excipients of cediranib, olaparib or durvalumab.

- Other invasive malignancy within the last 2 years. Participants with previous history of malignancies with a negligible risk of metastasis or death and treated with expected curative intent are eligible at discretion of clinical team, for example:
  - Carcinoma in situ of the cervix.
  - Basal or squamous cell skin cancer.
  - Localized low to intermediate risk prostate cancer treated with curative intent and absence of prostate-specific antigen (PSA) relapse; or prostate cancer (Stage T1/T2a, Gleason $\geqslant$6 and PSA <10 ng/mL) undergoing active surveillance and treatment naïve.
- Major surgery within 4 weeks prior to first dose of trial drug (excluding placement of vascular access). If participants have undergone major surgery more than 4 weeks prior to the scheduled first dose of trial drug, they must have fully recovered from the procedure.
- Minor surgery (not including the diagnostic biopsy) within 2 weeks prior to first dose of trial treatment
- Any concurrent chemotherapy, IP, biologic, or hormonal therapy for cancer treatment. Concurrent use of hormonal therapy for non-cancer-related conditions (e.g., hormone replacement therapy) is acceptable.
- Concurrent enrolment in another clinical trial unless it is an observational (non-interventional NOT involving CTIMPs) or translational clinical trial, or during the follow-up period of an interventional clinical trial.
- Receipt of the last dose of anticancer therapy or radiotherapy, chemotherapy, immunotherapy, endocrine therapy, targeted therapy, biologic therapy, tumour embolisation, monoclonal antibodies) $\leqslant$28 days prior to the first dose of trial drug.
- Gastrointestinal abnormalities including:
  - Refractory nausea and vomiting,
  - inability to take oral medication;
  - requirement for intravenous alimentation;
  - prior surgical procedures affecting absorption including total gastric resection;

  - treatment for active peptic ulcer disease in the past 6 months prior to the first dose of trial treatment;
  - active gastrointestinal bleeding, unrelated to cancer, as evidenced by haematemesis, haematochezia or melena in the past 120 days priort to the first dose of trial treatment without evidence of resolution documented by endoscopy or colonoscopy;
  - malabsorption syndromes.

- Any of the following within 12 months prior to consent:
  - Myocardial infarction,
  - uncontrolled angina,
  - coronary/peripheral artery bypass graft,
  - symptomatic congestive heart failure,
  - cerebrovascular accident or transient ischaemic attack,
  - peripheral arterial embolus.

- Current or prior use of immunosuppressive agents within 28 days prior to the first day of trial treatment , including anti-TNF and anti-IL17 agents, with the exceptions of intranasal or inhaled corticosteroids, or systemic corticosteroids at physiological doses which are not to exceed 10mg/day prednisolone (or an equivalent corticosteroid). The following exceptions are allowed:
  - Intranasal, inhaled, topical or local steroid injections (e.g. intra articular injection).
  - Systemic corticosteroids at physiological doses not to exceed 10mg/day prednisolone (or equivalent).
  - Steroids for premedication of hypersensitivity reactions (e.g. as CT premedication).

- Immunocompromised participants (e.g., participants who are known to be serologically positive for human immunodeficiency virus (HIV), or have a history of active primary immunodeficiency).

- Active infection including tuberculosis (clinical history, physical examination and radiographic findings, and Tuberculosis (TB) testing in line with local practice), hepatitis B (known positive HBV surface antigen (HBsAg) result), hepatitis C, or human immunodeficiency virus (positive HIV1/2 antibodies). Participants with a past or resolved HBV infection (defined as: presence of hepatitis B core antibody -anti-HBc- and absence of hepatitis B surface antigen –HbsAg-) are eligible.
- As judged by the Investigator, any participant considered a high medical risk due to a serious uncontrolled medical or psychiatric disorder, non-malignant systemic disease or on-going or active infection.
- Persistent toxicities ($\geqslant$ Common Terminology Criteria for Adverse Event (CTCAE) grade 2) caused by previous cancer therapy, excluding alopecia and vitiligo.
  - Participants with Grade $\geqslant$2 neuropathy will be evaluated on a case-by-case basis after consultation with the Chief Investigator.
  - For durvalumab-containing arms only: Participant with irreversible toxicity not reasonably expected to be exacerbated by treatment with durvalumab may be included only after consultation with the investigator.
- Women who are pregnant, or are lactating or breastfeeding. Women of childbearing potential and male participants who are unwilling to use adequate contraception from consent and for 3 months after the last dose of trial drug.
- Participants with contraindication to MRI including; contraindicated metallic implants, contraindicated coronary stents and pacemakers. Inability to lie flat or still in an MRI scanner for whatever reason (e.g., claustrophobia).
- Judgement by the Investigator that the participant should not participate in the trial.
- Involvement in the planning and/or conduct of the trial
- Previous allogeneic bone marrow transplant or double umbilical cord blood transplantation.

Specific exclusion criteria for cediranib and/or olaparib containing arms:

- Current use or anticipated need for treatment with drugs that are known potent CYP3A4 inhibitors, or inducers or substrates for CYP1A2.

- Concomitant medications known to prolong the QT interval (see Appendices 4, 5 & 6 , Concomitant therapy) or with factors that increase the risk of QTc prolongation or risk of arrhythmic events (such as heart failure, hypokalaemia, congenital long QT syndrome, family history of long QT syndrome or unexplained sudden death under 40 years-of-age), history of Torsades de pointes.

- Resting ECG indicating uncontrolled, potentially reversible cardiac conditions as judged by the investigator (eg., unstable ischaemia, uncontrolled symptomatic arrhythmia, congestive heart failure, QTcF prolongation >500ms, electrolyte disturbances, resting QTc of $\geqslant$470ms (Fridericia; as per local reading), etc.) on two or more time points within a 24 hour period or family history of long QT syndrome.

- Requirement of anticoagulant therapy with oral vitamin K antagonists. Therapeutic use of low molecular weight heparin is allowed.

- Poorly controlled hypertension (persistently elevated > 150/100mmHg (or > 140/90 for olaparib plus cediranib arm only), either systolic or diastolic or both, despite anti-hypertensive medication).

- Clinically significant signs and/or symptoms of bowel obstruction within 3 months prior to starting treatment.

- History of intra-abdominal abscess within 3 months prior to starting the first dose of trial treatment.

- History of GI perforation. Participants with a history of abdominal fistula will be considered eligible if the fistula was surgically repaired, there has been no evidence of fistula for at least 6 months prior to starting the first dose of trial treatment, and participant is deemed to be at low risk of recurrent fistula.

Specific exclusion criteria for cediranib monotherapy or combination therapy arms only:

- Left ventricular ejection fraction (LVEF) < lower limit of normal (LLN) per institutional guidelines, or <55%, if threshold for normal not otherwise specified by institutional guidelines, for participants with the below risk factors (an Echocardiogram should be performed at baseline and if clinically indicated):

  - Prior treatment with anthracyclines

  - Prior treatment with trastuzumab
  - Prior central thoracic RT, including exposure of heart to therapeutic doses of ionising RTM
  - Prior history of other significant impaired cardiac function

Specific exclusion criteria for durvalumab containing arms only:

- Receipt of live, attenuated vaccine within the last 30 days. Note: enrolled participants should not receive live, attenuated vaccine while receiving durvalumab nor within 30 days of last dose of durvalumab.
- Active or prior documented autoimmune or inflammatory disorders (except vitiligo), for example:
  - Intestinal: Inflammatory Bowel Disease (Colitis (including ulcerative colitis), Crohn's Disease), Diverticulitis (with the exception of Diverticulosis), Coeliac Disease (except participants with coeliac disease controlled by diet alone), irritable Bowel Disease.
  - Vascular: any type of vasculitic disorder, e.g. Wegener syndrome, granulomatosis with polyangiitis.
  - Endocrine: any endocrine alteration related to an autoimmune process e.g. Hashimoto syndrome, Grave's disease. NOTE: participants with hypothyroidism (eg, following Hashimoto syndrome) stable on hormone replacement treatment may be included.
  - Respiratory: Active Pneumonitis (of any origin: inflammatory or infectious), Sarcoidosis syndrome.
  - Dermatological: Psoriasis, Lupus/SLE (unless the skin condition has never required systemic therapy).
  - Other: Rheumatoid Arthritis, Hypophysitis, Uveitis.
  - History of organ transplant that requires use of immunosuppressive medications or any medical condition in which immunosuppressive agents were administered, including but, not limited to: Systemic corticosteroids, methotrexate, azathioprine.
  - Tumour necrosis factor alpha (TNF-4α) blockers

  - Participants with autoimmune conditions without active disease in the past 5 years may be included but only after discussion with the Trial Medical Oncologist.

- Participants with persistent toxicities (⩾Common Terminology Criteria for Adverse Event (CTCAE v 5.0) grade 2) caused by previous cancer therapy, excluding alopecia and vitiligo, which are not reasonably expected to be exacerbated by treatment with durvalumab may be included only after consultation with the Chief Investigator.

E.2 Permitted and prohibited concomitant treatment

From the WIRE trial protocol developed by Prof Grant Stewart and Dr Sarah Welsh: Any concurrent chemotherapy, IMP, biologic, or hormonal therapy for cancer treatment is not allowed. Concurrent use of hormonal therapy for noncancer-related conditions (e.g. hormone replacement therapy) is acceptable. Receipt of the last dose of previous anticancer therapy (radiotherapy, chemotherapy, immunotherapy, endocrine therapy, targeted therapy, biologic therapy, tumour embolisation, monoclonal antibodies) must be at least 28 days prior to the first dose of IMP. Major surgery must be completed 4 weeks prior to the first dose of IMP (excluding placement of vascular access) and if participants have undergone major surgery more than 4 weeks prior to the scheduled first dose of IMP, they must have fully recovered. Minor surgery (not including the diagnostic biopsy) is allowed within 2 weeks prior to first dose of IMP. In arms containing cediranib or olaprarib, drugs not allowed include known potent CYP3A4 inhibitors, or inducers or substrates for CYP1A2, concomitant medications known to prolong the QT interval, and anticoagulant therapy with oral vitamin K antagonists (therapeutic use of low molecular weight heparin is allowed). For durvalumab containing arms receipt of live, attenuated vaccine within the 30 days before or after durvalumab is not allowed, nor use of immunosuppressive agents within 28 days prior to the first dose of trial treatment, including anti-TNF and anti-IL17 agents, with the exceptions of intranasal, inhaled, topical or local steroid injections (e.g. intra articular injection), or systemic corticosteroids at physiological doses which are not to exceed 10mg/day prednisolone (or an equivalent corticosteroid or use of steroids for premedication of hypersensitivity reactions (e.g. as CT premedication).