INSTITUTE OF CLINICAL TRIALS AND METHODOLOGY

Improving Outcomes for Patients with Higher Risk Locally Advanced Renal Cancer

Dr Bhavna Oza

[redacted]

A thesis submitted for the degree of Doctor of Medicine by Research

University College London

Declaration

I, Dr Bhavna Oza declare that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

……

Dr Bhavna Oza

Acknowledgements

I would like to thank my supervisors Dr Duncan Gilbert and Professor Ruth Langley for their supervision and guidance over the four years. To Dr Angela Meade, I would like to express my gratitude, for her unwavering support and mentorship. I am very grateful to Professor James Larkin for his clinical support and teaching.

I would also like to thank the RAMPART and SORCE Trial Management Groups, particularly Professor Tim Eisen, Professor Grant Stewart, Professor Axel Bex and Professor David Harrison for their advice and support and expertise. I am grateful for the opportunity to work alongside Professor Patrick Royston and to learn from his expert oversight over the statistical analysis of the Leibovich validation study. I am particularly grateful to Elena Frangou who provided statistical support throughout my time at the MRC and to Matthew Burnell for his support with the oligometastatic study.

I would like to acknowledge the RAMPART trial team particularly Ben Smith, Hanna Bryant and Claire Shakeshaft for their patience in helping me to understand the detailed working of the trials.

Finally, I would like to thank my colleagues at the MRC CTU, Nalinie Johratnam-Hogan, Hannah Rush, Archie Macnair, Siobhan Cleary and Sophie Merrick for their feedback, inspiration and friendship.

I would especially like to thank my family, Nila Oza, Jatin Oza, Punam Oza and Vivek Oza for their unwavering support and encouragement, always.

1.8 My involvement at the MRC CTU at UCL

By October 2018 when I joined the MRC CTU at UCL, eleven years after opening to recruitment, overall survival data from the SORCE trial had matured and the TMG were preparing for the data lock prior to analysis and publication of primary results. Simultaneously, the RAMPART trial had recently opened to participant accrual in the UK. I managed the clinical aspects of the trial on a day-to-day basis with support from Professor James Larkin, the Chief Investigator. I was the first point of contact for eligibility queries and clinical review of adverse events between the trial opening and February 2022.

I led on the authorship of the RAMPART protocol paper, which highlighted key mid-trial changes including those made during the COVID-19 pandemic to optimise safety for participants. In addition, I co-authored the RAMPART design paper 'A Model for a Regulatory-Ready Academic-led Phase 3 Trial in the Adjuvant Renal Cell Carcinoma Setting'. Both papers have been published in Contemporary Clinical Trials [1] [2]

During my fellowship, I led on key RAMPART TMG discussions and was involved in the development of key protocol decisions, including the justification behind the use of the 2003 Leibovich Score for risk stratification of participants in RAMPART. I developed the statistical analysis plan for the external validation of the 2003 score using data from SORCE (chapter two). The manuscript and subsequent invited comment were published in the Journal of Clinical Oncology in January 2022 and May 2020 respectively [3, 4]. I was awarded first authorship of both.

In addition, I led on a number of significant eligibility discussions and amendments that extended RAMPART inclusion to participants with adrenal involvement at time of primary nephrectomy and to those who develop resectable metastatic disease later on. These were directly influenced by findings from the sub-studies that have formed the basis of this thesis (see chapters three and four).

I presented the SORCE trial primary results to the MRC CTU at UCL Clinical trial group in an oral presentation, which formed the basis of the late breaking abstract at the European Society for Medical Oncology Annual Conference in Spain in September 2019. I co-authored the primary analysis publication (3rd author) which was accepted by the Journal of Clinical Oncology in October 2020 [5].

Gene list

AGR3	Anterior gradient 3, protein disulphide isomerase family member
ALK	Anaplastic Lymphoma Kinase
AMACR	Alpha-Methylacyl-CoA Racemase
AXL	AXL Receptor Tyrosine Kinase
BAP1	BRCA1 Associated Protein 1
BIRC5	Baculoviral IAP repeat containing 5
BRCA	BReast CAncer gene 1
CAIX	Carbonic Anhydrase IX
CDKN2A/B	Cyclin-Dependent Kinase Inhibitor 2A/B
CEACAM6	CEA cell adhesion molecule 6
CIMP	CpG island methylator phenotype
CK7	Cytokeratin-7
CKS2	CDC28 Protein Kinase Regulatory Subunit 2
C-MET	C- MNNG HOS Transforming
DKN2A	Cyclin-Dependent Kinase Inhibitor 2A
EGFR	Epidermal Growth Factor Receptor
FH	Fumarate Hydratase
FLCN	Folliculin
HGF	Hepatocyte Growth Factor
KIT	KIT Proto-Oncogene, Receptor Tyrosine Kinase
MET	MNNG HOS Transforming
MiT	Microphtalmia-associated Transcriptional factor
MTOR	Mammalian Target of Rapamycin
NF2	Neurofibromatosis Type 2
NRF2- ARE	Nuclear Factor Erythroid 2–related Factor 2-Antioxidant Responsive Element
PBRM1	Polybromo 1
PIK3Ca	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha
PTEN	Phosphatase and tensin homolog
PTTG1	PTTG1 Regulator Of Sister Chromatid Separation, Securin
RASGEF1A	RasGEF domain family, member 1A
RET	Rearranged During Transfection
SDHB	Succinate Dehydrogenase Complex Iron Sulfur Subunit B
SDHC	Succinate Dehydrogenase Complex Iron Sulfur Subunit C
SDHD	Succinate Dehydrogenase Complex Iron Sulfur Subunit D
SETD2	SET Domain Containing 2, Histone Lysine Methyltransferase
SETD2	SET Domain Containing 2
SF	Steroidogenic Factor
TCEB1	Transcription elongation factor B polypeptide 1
TERT	Telomerase Reverse Transcriptase
TFE2	Transcription Factor E2
TFE3	Transcription Factor Binding To IGHM Enhancer 3

TFEB	Transcription Factor Binding To IGHM Enhancer B
TP53	Tumour Protein 53
TSC1	Tuberous sclerosis Complex 1
TSC2	Tuberous sclerosis Complex 2
VEGF	Vascular Endothelial Growth Factor
VEGFA	Vascular Endothelial Growth Factor A
VHL	Von Hippel-Lindau

List of abbreviations

ADT	Androgen deprivation therapy
AJCC	American Joint Committee on Cancer
APCCC	Advanced Prostate Cancer Consensus Conference
APC	antigen presenting cell
ASSURE	Adjuvant Sorafenib or Sunitinib for Unfavorable Renal Carcinoma
ATP	Adenosine triphosphate
bd	twice a day
BMI	body mass index
ccRCC	clear cell renal cell carcinoma
cdRCC	collecting duct renal cell carcinoma
CEA	carcinoembryonic antigen
chRCC	chromophobe renal cell carcinoma
CI	confidence interval
c-index	concordance index
CKD	chronic kidney disease
CSR	complete surgical resection
CRUK	Cancer Research UK
CSS	Cancer-specific survival
CT	Computerised tomography
CT CAP	Computerised tomography, chest/abdomen/pelvis
CTLA-4	cytotoxic T-lymphocyte associated antigen 4
CTU	Clinical Trials Unit
DFS	Disease-free survival
EAU	European Association of Urology
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Association
ESMO	European Society of Medical Oncology
EVEREST	Everolimus for Renal Cancer Ensuing Surgical Therapy
FDA	Food and Drug Association
FDGF	fibroblast growth factor
FDGFR	fibroblast growth factor receptor
FGFR	fibroblast growth factor receptor
FUSCC	Fudan University Shanghai Cancer Center
FWER	family-wise error rate
GLOBOCAN	Global Cancer Incidence, Mortality and Prevalence
GRANT score	GRade, Age, Nodes and Tumour score
HDI	human development index
HIC	high income countries
HIF	hypoxia inducible factor
HIF1-α	hypoxia inducible factor-alpha
HIF1-β	hypoxia inducible factor-beta
HR	hazard ratio

ICI	immune checkpoint inhibitor
IFN-α	interferon-alpha
IL2	interleukin-2
IMDC	International Metastatic RCC Database Consortium
IPD	individual participant data
IQR	interquartile range
ISUP	International Society of Urological Pathology
ITT	intention to treat
KIT	tyrosine-protein kinase
KM	Kaplan-Meier curve
LU	Leuven-Undine
LMIC	Low- and middle-income countries
M0	TMN- No metastasis
M1	TNM- Metastatic disease
M1 NED	Metastatic with no evidence of disease after resection of metastases
MAMS	Multi-arm, multi-stage
mdRCC	Medullary renal cell carcinoma
MFS	Metastatsis-free survival
MHC	Major histocompatibility complex
MRC	Medical Research Council
MRC CTU at UCL	Medical Research Council, Clinical Trial Unit at University College London
MSKCC	Memorial Sloan Kettering Cancer Center
N stage	TNM- Nodal status
NICE	National Institute for Health and Care Excellence
non-ccRCC	Non-clear cell renal cell carcinoma
NR	Not reached
od	Once a day
OS	Overall survival
PD-1	Programmed cell death protein 1
PDGF	Platelet derived growth factor
PDGFR	Platelet derived growth factor receptor
PDL-1	Programmed cell death protein ligand 1
PFS	Progression-free survival
PI	Prognostic index
pRCC	Papillary renal cell carcinoma
PS	Performance status
RAMPART	Renal Adjuvant MultiPle Arm Randomised Trial
RCC	Renal cell carcinoma
RFS	Relapse-free survival
ROC	Receiver operating curve
RR	Relative risk
SEERS-17	Surveillance, Epidemiology and End Results-17 programme
SORCE	Sorafenib in Treating Patients at Risk of Relapse after Undergoing to Remove Kidney Cancer

sRCC	sarcomatoid renal cell carcinoma
SSIGN	tumour Stage, Size, Grade, Necrosis
STRAC	Sunitinib Versus Placebo for The Treatment of Patients At High Risk Of Recurrent Renal Cell Cancer
T stage	TNM- Tumour size
TKI	tyrosine kinase inhibitors
TMG	trial management group
TNM	(T)umour size, (N)ode, (M)etastases
TTR	Time to recurrence
TRIPOD	transparent reporting of a multivariable prediction model for individual prognosis or diagnosis
UCL	University College London
UCLA	University of California Los Angeles
UISS	University of California Los Angeles Integrated Staging System
UK	United Kingdom
uRCC	unclassified renal cell carcinoma
US	United States of America
VEGF	vascular endothelial growth factor
VEGFR	vascular endothelial growth factor receptor
VENUSS score	VEnous tumour thrombus, NUclear grade, Size, T and N Stage
VHL	Von-Hippel-Lindau
WHO	World Health Organisation
α	alpha
β	beta
SE	Standard error

List of Tables

List of Figures

Table of Contents

Abstract

Background: Renal Cell Carcinomas (RCCs) are a heterogeneous group of malignancies. Although patients presenting with locally advanced tumours may be cured by surgery alone, many subsequently relapse and succumb to their disease. Studies of adjuvant tyrosine-kinase-inhibitors (TKIs) did not meet their primary goals but provide a rich data source to refine prognostication in contemporary patients and inform the design of clinical trials testing new therapeutic approaches.

Methods: This study included patients treated surgically with curative intention from two international clinical trials testing the addition of TKIs in intermediate and high-risk patients; SORCE (n=1711) and ASSURE (n=1943). Three questions were addressed; A validation of the 2003 Leibovich prognostic score (a widely used scoring system) in contemporary patients with clear-cell and non-clear-cell RCCs, also comparing this with the Tumour/Nodes/Metastases (TNM) classification. Discrimination and calibration were assessed by comparing data from SORCE to original data used to derive the Leibovich score using Harrell's concordance-indexes, Kaplan-Meier curves and hazard ratios (HRs). Secondly data from SORCE and ASSURE were combined, generating a large dataset to examine clinical characteristics of higher risk non-clear-cell RCC variants (papillary RCC (pRCC), chromophobe RCC (chRCC) and sarcomatoid RCC (sRCC)). The impact of histology on disease-free-survival and overall survival were presented using Kaplan-Meier curves and adjusted Cox regression models. Finally, a retrospective cohort study examining data from SORCE compared outcomes of those relapsing first at single anatomical sites to those relapsing at multiple sites using Kaplan-Meier methodology. The prognostic impact of organ site and time-to-relapse (TTR), performance status and treatments upon relapse were evaluated using Cox regression models.

Results: The 2003 Leibovich score demonstrated discriminative accuracy in patients with clear-cell (c-index 0.63, 95% CI 0.61 to 0.65) and non–clear-cell RCCs (c-index 0.64, 95% CI 0.59 to 0.69). Discrimination by the 2003 Leibovich score exceeded that of 2002 TNM (c-indexes of 0.67 (SE 0.01) vs 0.56 (SE 0.01)). Distinct patterns of relapse for patients with chRCC, pRCC and sRCCs were shown. Notably, the median TTR for patients with pRCC was five months less than patients with ccRCC; (1.34 years (IQR 0.76, 2.59) vs 1.78 years (IQR 0.96, 3.38, p=0.012)). Those with pRCC

relapsing in the abdomen had almost double the risk of death (HR 1.7 (95% CI 1.15-2.5 p <0.001), compared to those with ccRCC. Patients with ccRCC relapsing at a single anatomical site exhibited better RCC-specific survival than those relapsing first in multiple sites, (HR 0.56 95% CI, 0.43-0.72, p<0.001). Prognostic significance of TTR was demonstrated with a median survival-after-recurrence of 3.1 years, 5.6 years and 'not reached' in patients relapsing at <12 months, 12-36 months and >36 months, respectively (p < 0.003).

Conclusion: The 2003 Leibovich score discriminates between intermediate and high-risk patients in multi-subtype RCC populations. Outcomes for patients with non-clear-cell RCCs are heterogeneous; those with pRCC with intra-abdominal first relapses had particularly poor survival. Prognostic groups were defined for patients relapsing after nephrectomy based on number of anatomical sites involved and TTR. These results guide prognostication, future translational work and clinical trial designs for patients presenting with locally advanced RCC.

Impact statement

SORCE was the fifth and final adjuvant trial evaluating tyrosine kinase inhibitors after nephrectomy in the context of locally advanced RCC. The granularity and duration of data collected during SORCE trial follow-up presented an opportunity to enhance understanding of the clinical behaviours of patients with higher risk RCCs.

External validations of the 2003 Leibovich Score have previously been conducted in small, retrospective datasets. I conducted the first evaluation of the score in prospectively collected individual participant data in an international cohort of patients with ccRCC and non-ccRCC. Through collaboration with Professor Bradley Leibovich and colleagues at the Mayo clinic, USA and working alongside Professor Patrick Royston (senior statistician, MRC CTU at UCL), I developed methodology for a validation in which the score's original performance was compared to that in contemporary data. Findings were presented to the MRC Cancer Clinical Group in September 2021 and subsequently published in the Journal of Clinical Oncology in 2022 [3].

This study showed that although not perfect, discrimination between intermediate and high risk RCC by the 2003 Leibovich score, was maintained. Furthermore, the 2003 Leibovich score allows superior discrimination when compared to 2002 TNM staging thus supporting its preferential use for risk stratifying participants in future adjuvant RCC trials. Outcome prediction for contemporary patients using clinico-pathological risk scores requires refinement. An important next step is improved understanding of the molecular characteristics that drive worse outcomes in RCC.

Chapter two detailed the largest contemporary analyses of clinical outcomes following nephrectomy in patients with higher risk papillary, chromophobe and sarcomatoid RCCs, clinically important and classically understudied patient groups. The use of phase three clinical trial data from two international trials facilitated the precise delineation of their clinical profiles. Results provide robust support for histology specific surveillance for patients with higher risk RCCs. Secondly, by linking overall prognosis to site of initial relapse, findings are guiding current translational studies on SORCE nephrectomy samples aimed at searching for potentially targetable oncogenic drivers of poor prognosis disease.

Adjuvant checkpoint inhibitor trials in RCC have expanded their eligibility to include participants with fully resected oligometastatic relapse. Their inclusion criteria vary widely between trials in terms of number of metastases, site and timing of oligometastatic relapse. Chapter three delineated a clinically relevant oligometastatic phenotype in RCC providing evidence to support the specific inclusion of these patients into adjuvant clinical trials alongside participants with locally advanced non-metastatic disease. The criteria presented is currently being considered for adoption as a mid-trial protocol addition in the RAMPART trial and is a blueprint for future adjuvant trial inclusion strategy.

Chapter 1: Introduction

1.1 Burden of renal cancer

From a global perspective, the incidence of renal cancer, of which the commonest type is Renal Cell Carcinoma (RCC), is rising, both in high-income countries (HICs) and low to middle income countries (LMICs). Overall, RCC accounts for 2-3% of world cancer diagnoses [6]. In 2020, the annual global incidence had risen from 338,000 in 2012 to 431,288, deaths rising from 143,405 to 179,368 [6]. Of those, 271,249 cases were diagnosed in men and 160,039 in women, reflecting a relative risk (RR) of 1.7 for men compared to women [7]. RCC is a global disease. In 2020, the highest incidence was reached in Eastern Asia with 108,503 new cases and 54,658 deaths, followed by North America and Central and Eastern Europe (incidence 76,975 and 49,772 respectively) [7].

1.2 Updated classification of RCC

RCC comprises a heterogeneous group of parenchymal tumours that arise from the kidney. The recent 2022 World Health Organization (WHO) classification [8] characterises RCC subtypes based on predominant cytoplasmic or morphological features. Clear-cell RCC (ccRCC) is by far the commonest and most studied morphological subtype followed by papillary RCC and chromophobe RCC. Tumour types with distinct molecular components include eosinophilic solid and cystic RCC. Tumours are classified according to anatomical location (e.g. collecting duct carcinomas), and also into those with a specific renal disease background (e.g. acquired cystic disease-associated RCC).

The updated classification has introduced molecularly-driven renal tumour types. For example Microphthalmia-associated Transcriptional factor (MiT) family translocations, Factor H deficiency and succinate dehydrogenase deficiency RCCs that may show very heterogeneous morphology. Newly classified entities include TFEB amplified, TCEB1 mutated, RCC with ALK rearrangement and renal cancers with SMARCB1 mutations. The widespread integration of classic histologic diagnosis with advanced molecular techniques provides a promising avenue for future drug- personalised therapeutic strategies.

1.3 Molecular pathways involved in RCC

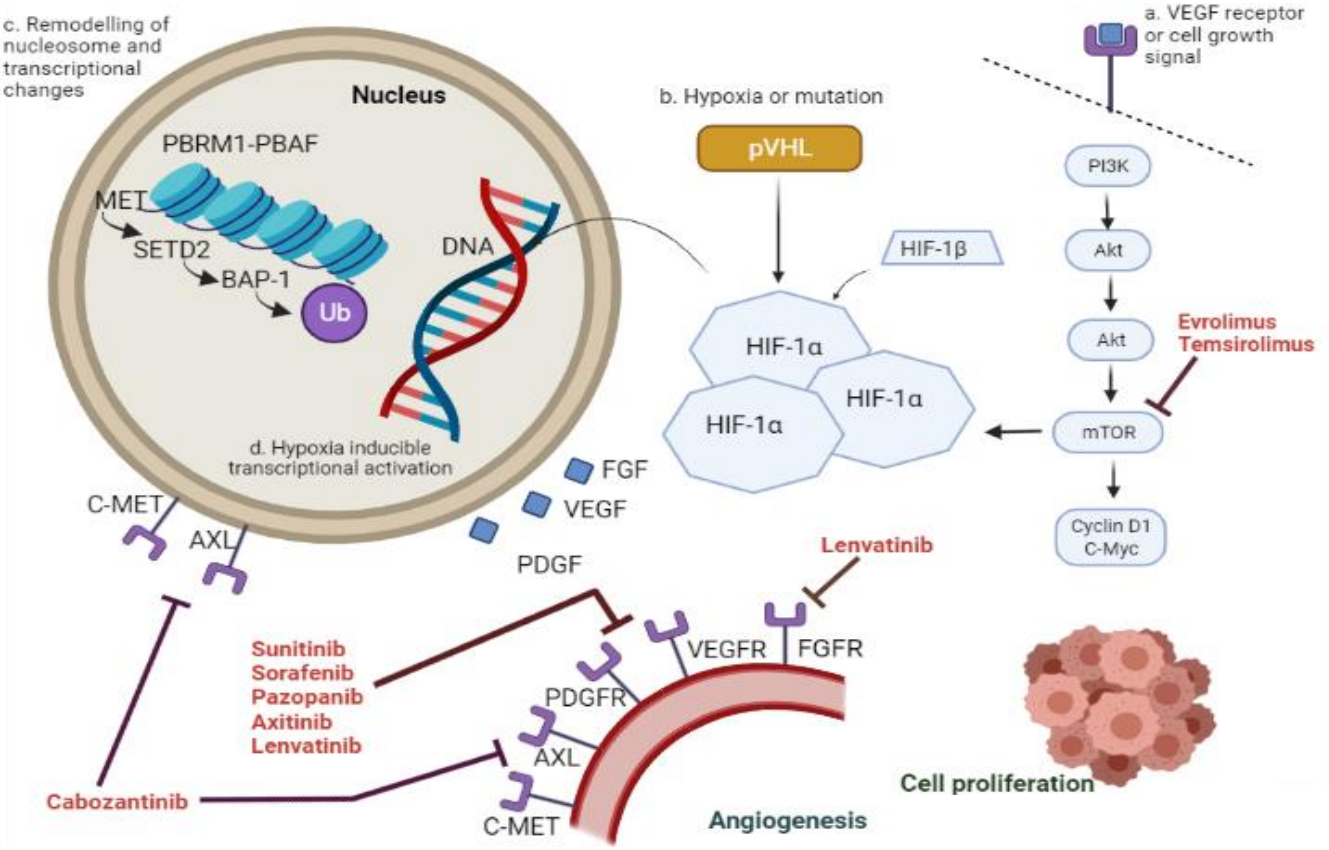
Investigation of familial forms of RCC (approximately 5% of all cases) has revealed associated germline mutations in at least eleven genes (namely BAP1, FLCN, FH, MET, PTEN, SDHB, SDHC, SDHD, TSC1, TSC2, and VHL*)* some of which have also been implicated in the development of sporadic RCC [9]. A prominent example is the VHL tumour suppressor gene located on chromosome 3p, the well-studied mutation underlying von Hippel-Lindau disease which is also a hallmark of sporadic ccRCC tumours. A summary of the molecular pathways involved in RCC tumorigenesis is shown in **Figure 1**. Loss of 3p underlies the inactivation of VHL tumour-suppressor protein via mechanisms including promoter methylation, mutations and deletions. In ccRCC, chromothripsis of chromosome 3p, characterized by a single 'catastrophic' event leading to simultaneous 3p loss and 5q gain is the most common structural abnormality [10]. The loss of functioning VHL protein results in the upregulation and constitutive activation of a multi-protein complex that acts as a signalling mark (otherwise known as ubiquitination) that triggers activation of hypoxia inducible factor *(*HIF*)* transcription factor. HIF is a heterodimeric protein consisting of an unstable alpha (α) subunit and a stable beta (β) subunit. Under low oxygen conditions HIF-1α accumulates and binds to HIF-1β. This in turn actives the transcription and membrane transport of hypoxia inducible factor *(*HIF*)*. Up to 100 HIF-responsive genes have been described [11], many of which are involved in adapting to acute or chronic hypoxia including vascular endothelial growth factor (VEGF) and other molecules implicated in angiogenesis, (a crucial step in tumour progression, cell proliferation and survival). In cells lacking functional VHL protein, HIF-1α remains constitutively activated leading to dysregulated vascular growth, which has been linked to the formation of sporadic and familial forms of RCC [12]. Therapies that target pro-angiogenic factors including tyrosine kinase inhibitors (TKIs) like sunitinib and pazopanib, are treatment options for patients with metastatic ccRCC, see **Figure 1.**

The MTOR pathway, which plays a crucial role in HIF activation, intersects the HIF pathway upstream of the VHL protein. MTOR pathway mutations (PIK3Ca, MTOR and PTEN) in ccRCC generally result in missense and functionally activating mutations leading to upregulation of MTOR [13]. This explains the rationale for MTOR pathway inhibitors, including everolimus and temsirolimus, which have shown efficacy in treatment of advanced ccRCC [14], see **Figure 1.**

Alongside VHL inactivation, secondary inactivating mutations in genes close to VHL with important tumour suppressor roles include SETD2, BAP1 and PBRM1 have been identified, see **Figure 1**. Mutations of BAP1 a gene involved in cellular apoptosis, occur in 10% of ccRCCs and encode for the histone deubiquitinase BRCA1 associated protein-1. Mutations in SETD2, a histone methyltransferase, involved in chromatin remodelling, also occur in 10% of ccRCC [15]. BAP1- or SETD2-mutated ccRCCs have been linked to poor prognosis in ccRCC tumours, while mutations in PBRM1, a gene associated with cellular senescence and genome stability, has been associated with favourable overall survival (OS) for example in RCC pancreatic metastases [15].

Despite an increasing understanding of the molecular diversity underpinning RCC, a therapeutic role for manipulating many of these genes and gene targets has not yet been elucidated. Furthermore, the current global disparity of access to molecular resources and genetic testing means that basic morphological analysis remains the gold standard initial diagnostic tool.

Figure 1 Tumorigenesis signalling in renal cell carcinoma and drug targets



a. VEGF receptor activation at the tumour cell surface leads to downstream activation of MTOR pathway signalling which leads to HIF activation, intersecting the HIF pathway upstream of VHL gene. *MTOR* pathway mutations (PIK3Ca, MTOR and PTEN) in ccRCC result in missense and functionally activating mutations leading to upregulation of MTOR.

b. Under low oxygen conditions or in cells lacking functional VHL protein, HIF1-α sub-unit accumulates and binds to HIF1-β sub-unit forming HIF.

c. Recurrent mutations of histone modifying and chromatin remodelling tumour suppressor genes including PBRM1, BAP1 and SETD2, alter ubiquitination of target histones leading to genomic instability and uncontrolled tumour growth.

d. Up to 100 HIF-responsive genes including vascular endothelial growth factor (VEGF), platelet derived growth factor (PDGF) and fibroblast growth factor (FGF), AXL and C-MET receptor tyrosine kinase inhibitors are implicated in angiogenesis, tumour cell proliferation and cell survival through stimulating corresponding blood vessel surface receptors. Therapies that target these pro-angiogenic factor receptors including tyrosine kinase inhibitors (TKIs) and MTOR inhibitors have been utilised in the metastatic setting in ccRCC.

Adapted from Fig.1 'Tumorigenesis-related signalling with drug development implications in renal cell carcinoma from: Towards individualized therapy for metastatic renal cell carcinoma' [16]. Created with BioRender.com

1.4 Locally advanced RCC

1.4.1 Surgical strategies for locally advanced RCC

At presentation, the incidence of locally advanced RCC (including those localised to the kidney) accounts for 65% of cases. Extension into the renal vein or inferior vena cava is reported in 4-10% of cases [17]. The standard curative treatment for patients with primary locally advanced RCCs is either partial (nephron sparing) or radical nephrectomy (removal of the kidney). Laparoscopic and robotic assisted techniques offer the advantage of lower analgesic use, shorter hospital stay, reduced blood loss and quicker recovery times. Partial nephrectomy, which can be performed, with an open, laparoscopic, or robot-assisted approach [18], better preserves kidney function and is the treatment of choice for clinically stage T1a-b[1] renal masses, for younger patients, those with pre-existing chronic kidney disease (CKD) or those at risk of tumour growth in the contralateral kidney [19]. In contrast to bladder and head and neck cancer guidance, no data clearly demonstrates a role for empirical lymph-node dissection for patients with RCC and it is therefore not routinely performed.

An option for medically inoperable patients, for those with poor performance status or with low risk of significant tumour spread is active monitoring (surveillance of tumour growth with periodic radiographic studies). Other options include percutaneous radiofrequency ablation[2] and laparoscopically assisted or percutaneous cryoablation which may also be appropriate for those with a solitary kidney and a high risk of complete loss of renal function following partial nephrectomy [20]. Alternatively, patients assessed as unsuitable for surgery, who present with massive haematuria or flank pain may be offered palliative procedures such as embolization[3],microwave ablation and stereotactic radiosurgery[4] [20].

1.4.2 Heterogeneity of outcomes after nephrectomy

Despite advances in radical surgical techniques that improve upon tumour removal, patients with intermediate or high-risk locally advanced RCC, (see below for more

[1] Component of TNM; tumour size (T)-nodal status (N)-presence of metastasis (M) staging system developed by the American Joint Committee on Cancer (AJCC)

[2] Ablation is the surgical removal of a tissue or body part, which can be achieved using a variety of techniques.

[3] Particles, such as tiny gelatine sponges or beads used to block flow of blood to a tumour or abnormal area of tissue

[4] A non-surgical high intensity radiation therapy used to treat functional abnormalities and small tumours

details on clinico-pathological prognostic scores), are at significant risk of relapse after tumour resection. 20-30% patients with intermediate risk and 40-60% patients with high risk RCC develop metastatic disease following nephrectomy [21, 22]. Furthermore, upon relapse, RCC presents a range of clinical trajectories from those who develop single sites of recurrence years after surgery with favorable prognosis to those who progress rapidly to oligometastatic or incurable metastatic disease.

The heterogeneity of outcomes in RCC has driven a continued focus on developing prognostic tools to assist both patients and their physicians in understanding the likely prognosis and individualising their post-surgical surveillance. In the absence of validated molecularly driven predictive assays, clinicopathological scoring systems remain central for prognostication of patients with locally advanced RCC after nephrectomy. They are also widely employed to guide the selection and stratification of patients onto adjuvant clinical trials.

1.4.3 Risk prediction for patients with locally advanced RCC

Components of TNM have been combined with clinical and/or histopathological variables to develop RCC specific predictive models that have improved prognostic accuracy. Current widely cited models include the Mayo clinic's 2003 Leibovich score, the University of California Los Angeles (UCLA) Integrated Staging System (UISS) and the Kattan postoperative nomogram [23-26]. They all rely on clinical and/or histopathological variables but vary regarding the number and type of covariates, tool characteristics (nomogram or prognostic categories), and end points (OS, cancer-specific-survival (CSS), and relapse-free survival (RFS)). See **Table 8, Chapter 2** for more detail on the variables included and endpoints used in each model. A common statistical approach for deriving the scores has been the use of univariable and multivariable Cox proportional hazards models in retrospective datasets of patients with locally advanced RCC. Using a pre-specified level of statistical significance (p-value), these models evaluate the clinical, surgical, and pathologic features significantly associated with chosen outcome measures, in order to determine features that may comprise a final risk prediction score.

1.5 Tyrosine kinase inhibitors in metastatic RCC

Since 2006, small molecule oral tyrosine kinase inhibitors (TKIs) have been the standard first line treatment for patients with metastatic RCC. The landmark trial data

for single agent TKIs including their receptor targets, is outlined in **Table 1**. TKIs provide therapeutic effect by targeting pro-angiogenic HIF-responsive receptors, for example vascular endothelial growth factor receptor (VEGFR), platelet derived growth factor receptor (PDGFR) and fibroblast growth factor receptor (FDGFR). All members of this family are surface tyrosine kinase receptors with a protein structure consisting of five extracellular, immunoglobulin-like domains, a hydrophobic transmembrane domain, and an intracellular tyrosine kinase domain. They are upregulated following VHL gene loss and HIF-α accumulation, see **Figure 1**. TKIs primarily function by blocking Adenosine triphosphate (ATP) binding at the intracellular catalytic binding site thereby inhibiting downstream tumour endothelial growth and survival signalling [27]. TKIs differ from each other in the spectrum of tyrosine kinases they inhibit, their pharmacokinetics, their anti-tumour effect and their spectrum of side-effects [27].

Single agent sunitinib; a small molecule inhibitor of tyrosine kinases VEGFR, PDGFR, KIT and RAF, pazopanib; inhibitor of VEGFR, PDGFR, FGFR, KIT, RET and tivozanib, inhibitor of VEGFRs are globally available TKIs for the first line treatment of metastatic RCC regardless of International Metastatic RCC Database Consortium (IMDC) risk groups [28]. Cabozantinib, inhibitor of kinases c-MET, VEGFR2, MET, AXL and RET, is an alternative for patients with IMDC intermediate- and poor-risk disease. There are now four TKI and immune check-point inhibitor (ICI) combinations [29-32] available including axitinib (TKI) with pembrolizumab and nivolumab with cabozantinib (TKI) which are favoured in ESMO guidance as acceptable first line strategies for patients of all IMDC risk groups [14]. Recent data from the CLEAR trial (NCT02811861) showed a significant OS advantage of the combination of Lenvatinib (TKI) and pembrolizumab (Programmed death 1, PD-1 inhibitor) compared to sunitinib, (HR 0.66, 95% CI 0.49-0.88, p=0.005, median OS not reached (NR))[32]. The combination has now been NICE approved for patients with intermediate and poor IMDC risk RCC. Axitinib with avelumab is another combination available via the UK CDF for patients in all IMDC risk categories following recent data showing superiority over sunitinib (median PFS 13.3 (95% CI, 11.1-15.3) v 8.0m (95% CI, 6.7-9.8), HR 0.69 (95% CI, 0.574-0.825) p<0.00011) [30]. See **Table 2** for summary of trial results evaluating ICIs and ICI/TKI combinations in metastatic RCC.

1.6 Tyrosine kinase inhibitors in locally advanced RCC

Since the late 1990s, several adjuvant strategies after nephrectomy have been examined including cytokines, radiotherapy, hormones and TKIs, with limited success. TKIs, have been extensively investigated due to their ease of administration, response rates and ability to prolong survival in the metastatic setting. Five large adjuvant phase three TKI trials have reported results. ASSURE (Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinoma NCT00326898), PROTECT (Randomized Phase III Trial of Adjuvant Pazopanib Versus Placebo After Nephrectomy in Patients With Localized or Locally Advanced Renal Cell Carcinoma NCT01235962), S-TRAC (Adjuvant Sunitinib in High-Risk Renal-Cell Carcinoma after Nephrectomy, NCT00375674), ATLAS (Adjuvant Axitinib Therapy of Renal Cell Cancer in High Risk Patients, NCT01599754) and SORCE (Adjuvant Sorafenib for Renal Cell Carcinoma at Intermediate or High Risk of Relapse NCT00492258) [5, 33-36]. All the adjuvant TKI trials examined the intention to treat (ITT)[5] population using DFS as the primary outcome measure and OS as the secondary outcome measure and up to three years of treatment after nephrectomy. **Table 3** compares the baseline characteristics and **Table 4** compares the survival and toxicity outcomes of participants within each trial.

[5] ITT population includes all patients who were randomly assigned within the trial and ignores noncompliance, protocol deviations, participant withdrawal, and anything that happens after patients are randomly assigned. This preserves the prognostic balance generated by the original random treatment allocations.

Table 1: **Landmark phase III trial data for TKIs in advanced RCC, their receptor targets and current use**

TKI	Receptor target	Trial	Comparator	Median PFS (months)	HR for PD (95% CI, p-value)	Median OS (months)	Hazard Ratio for Death (95% CI, p-value)	Current use
Sorafenib	VEGF, PDGFR, KIT, RAF	TARGET [37]	Placebo	5.5 vs 2.8	0.44 (0.35–0.55) <0.01	17.8 vs 15.2	0.88 (0.74–1.04), P=0.15	Superseded by other TKIs
Sunitinib	VEGF, PDGF	Motzer et al [38]	IFNα	11 vs 5	0.42 (0.32–0.54) <0.001	26.4 vs 21.8	0.82 (0.67–1.00), P=0.05	A 1st and >1st line standard of care
Pazopanib	VEGFR1-3, PDGFR, FGFR 1, 3 4, KIT, RET	COMPARZ [39]	Sunitinib	8.4 vs 9.5	1.05 (0.90–1.22), NR	28.3 vs 29.1	0.91 (0.76–1.08), P=0.28	A 1st and >1st line standard of care
Axitinib	VEGFRs 1, 2 3	AXIS [40]	sorafenib	6.7 vs 4.7	0.66 (0.55–0.81) <0.001	20.1 vs 19.2	0.97 (0.80–1.17), P=0.37	Single agent and in combination with ICIs
Cabozantinib	MET, AXL, VEGFR, RET	CABOSUN* [41]	sunitinib	8.2 vs 5.6	0.66 (0.46–0.95) 0.012**	26.6 vs 21.2	0.80 (0.53–1.21), p=NR	A 1st and >1st line standard of care
Tivozanib	VEGFRs 1, 2 3	TIVO-1 [42]	sunitinib	11.9 vs 9.1	0.80 (0.64–0.99), 0.042	28.8 vs 29.3	1.25 (0.95–1.62), P=0.105	A 1st and >1st line standard of care

CI; confidence interval, NR; not reported, TKI: tyrosine kinase inhibitor, HR; hazard ratio, OS; overall survival, VEGFR; vascular endothelial growth factor receptor, PDGFR; platelet derived growth factor receptor, KIT; KIT Proto-Oncogene, Receptor Tyrosine Kinase, RAF; rapidly accelerated fibrosarcoma, RET; rearranged during transfection, MET; mesenchymal-epithelial transition factor, AXL; AXL Receptor Tyrosine Kinase *phase II, **one-sided p-value

Table 2 Outcomes of trials of immune checkpoint inhibitors in advanced RCC

Trial Name	Patient group	Intervention	Comparator	Summary of Efficacy data
Checkmate 025 [43]	Clear cell aRCC (≤ 2nd line)	Nivolumab (anti-PD-1) N=410	Everolimus N=411	Median OS 25 m v 19.6m, HR 0.73 (95% CI, 0.7-0.93) p= 0.0018
Checkmate 214 [44]	Treatment naïve clear cell aRCC	Nivolumab + Ipilimumab (anti-CTLA-4) N=422	Sunitinib N= 422	IMDC intermediate/high; OS HR 0.63, (95% CI, 0.44-0.89) p<0.0001 PFS HR 0.82 (95% CI, 0.64-1.05) IMDC low; OS HR 1.13 (95% CI, 0.64-1.99) p=0.6710
IMmotion151 [45]	Treatment naïve clear cell or sarcomatoid aRCC	Atezolizumab (anti-PD-L1) + Bevacizumab (anti-VEGFA) N=454	Sunitinib N=461	mPFS (PD-L1+) 11.2m v 7.7m HR 0.74 (95% CI, 0.57-0.96) p=0.02 OS (ITT) HR 0.93 (95% CI 0.76-1.14) ns OS (PD-L1+) HR=0.74 (95% CI, 0.57-0.96) p=0.02
JAVELIN Renal [30]	Treatment naïve clear cell aRCC	Avelumab (anti-PD-L1) + Axitinib N=442	Sunitinib N=444	mPFS 13.3 (95% CI, 11.1-15.3) v 8.0m (95% CI, 6.7-9.8) HR 0.69 (95% CI, 0.574-0.825) p<0.00011
Keynote 426 [46]	Treatment naïve clear cell aRCC	Pembrolizumab (anti-PD-1) + Axitinib N=432	Sunitinib N=429	mPFS 15.4m v 11.1m HR 0.71 (95% CI, 0.60-0.84) p=0.001 mOS NR vs 35.7m (95% CI, 33.3- NR) HR 0.68 (95% CI, 0.55-0.85) p=0.0003
Checkmate-9ER [31]	Treatment naïve clear cell aRCC	Nivolumab (anti-PD-1) + Cabozantinib N=323	Sunitinib N=328	mPFS 16.6m v 8.3 m HR 0.51 (95% CI, 0.41- 0.64) p=0.0001) mOS NR either arm OS HR 0.60 (98.89% CI 0.4-0.89) p=0.0010

CLEAR [32]	Treatment naïve clear cell aRCC	Pembrolizumab + Lenvatinib (TKI) (P/L) N=355 Everolimus (MTORi) + Lenvatinib (E/L) N=357	Sunitinib N=357	mPFS P/L 23.9m v 9.2m HR 0.39 (95% CI, 0.49-0.88) p=0.005 OS P/L HR 0.66 (95% CI, 0.49- 0.88) p = 0.005 mPFS E/L 14.7m vs. 9.2 m HR 0.65 (95% CI, 0.53- 0.80) p<0.001 OS E/L HR 1.15 (95% CI, 0.88 -1.50) p = 0.30

aRCC; advanced RCC, NS; not significant, NR; not reached, OS; overall survival, mPFS; median progression-free survival; CI; confidence interval, HR; hazard ratio, m; month, IMDC; International Metastatic RCC Database Consortium

Table 3 Baseline Characteristics of participants in adjuvant TKI trials

	ASSURE n=1943	S-TRAC n=615	PROTECT n=1538	ATLAS n=724	SORCE n=1711
Treatment	Sunitinib or Sorafenib vs placebo 1:1:1	Sunitinib vs placebo 1:1	Pazopanib vs placebo 1:1	Axitinib vs Placebo 1:1	Sorafenib 1Y or Sorafenib 3Y vs Placebo 1:1:1
Histology	Clear cell (79%) Papillary (6%) Chromophobe (6%) Mixed (31%) Unclassified (4%) Sarcomatoid (8%)	100%	100% >50% clear cell	100% >50% clear cell	Clear cell (84%) Papillary (8%) Chromophobe (7%) Collecting duct (<1%) Other (1%)
Region	United states and Canada	21 countries	26 countries	China, France, India, Japan, Korea, Spain, Taiwan	UK, Aus, Belgium, Denmark, France, The Netherlands, Spain
Performance status	ECOG 0-1	ECOG 0-2	Karnofsky performance score of ≥ 80	ECOG 0-1	ECOG 0-1
Tumour stage	T1bN0M0 (G3-4), T2-4N1-3M0 (Gany)	T2-4N0M0, TxN+M0	T2N0M0 (G3-4), T3-T4N0M0 or TxN1M0 (Gany)	T2-4 N0M0 or any TxN+ M0	T1aN0M0 (G4), T1bN0M0 (G3-4), T2-3N0M0, T1b-4N1M0
Risk stratification score	UISS intermediate high to very high risk	UISS high risk to very high risk	SSIGN Intermediate- high risk	High risk defined by TMN 2010 and Fuhrman grade	Leibovich intermediate to high
Dose	Sunitinib 50mg od (4 weeks on 2 weeks off (54w) Sorafenib 400 → 200mg bd (54w)	Sunitinib 50mg od 4 weeks on 2 weeks off (54w)	Pazopanib 600mg od reduced from 800mg od	Axitinib 5mg bd	Sorafenib 400mg bd
% Patients receiving starting dose	70%	100%	26%	100%	13%
Minimum dose permitted	25mg od	37.5mg od	400mg od	NR	400mg od

ECOG; Eastern Co-operative Oncology Group performance status, UISS; UCLA Integrated Staging System, T; tumour size, N; nodal status, M; presence of metastasis, NR; not reported. ASSURE (Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinoma NCT00326898), S-TRAC (Adjuvant Sunitinib in High-Risk Renal-Cell Carcinoma after Nephrectomy, NCT00375674), PROTECT (Randomized Phase III Trial of Adjuvant Pazopanib Versus Placebo After Nephrectomy in Patients With Localized or Locally Advanced Renal Cell Carcinoma, NCT01235962), ATLAS (Adjuvant Axitinib Therapy of Renal Cell Cancer in High Risk Patients, NCT01599754), SORCE (Adjuvant Sorafenib for Renal Cell Carcinoma at Intermediate or High Risk of Relapse NCT00492258)

Table 4 Toxicity and survival outcomes for adjuvant TKI trials

	ASSURE	S-TRAC	PROTECT	ATLAS	SORCE
Median follow up	5.8 years (IQR 4.9, 6.9) overall	5.4 years (95% CI, 5.2-5.6) overall	30.4 months (pazopanib 600mg) 47.9 months (pazopanib 800mg)		6.5 years (IQR 4.9, 8.0 years)
Outcomes in ITT population					
DFS	**Investigator** No improvement Sunitinib v placebo: HR 1.02 (97.5% CI 0.85-1.23 p=0.8038) Sorafenib vs placebo: HR 0.97 (97.5% CI 0.80-1.17 p=0.7184)	**Investigator** No improvement Sunitinib v placebo: HR 0.81 (95% CI, 0.64 to 1.02; p=0.08) **Independent** Improvement Sunitinib v placebo: HR 0.76 (95% CI 0.59-0.98 p=0.03)	**Investigator** No improvement ITT600mg: HR 0.94 (95% CI, 0.77-1.14) p = 0.51 Improvement ITT800mg; HR 0.69; 95% CI, 0.51- 0.94 p = 0.02	**Investigator** No improvement HR 0.776; 95% CI, 0.599–1.005 p = 0.0536 **Independent** No improvement HR 0.87 (95% CI 0.660–1.147; p=0.3211)	**Investigator** No improvement Sorafenib 1Y: HR 0.94 (p=0.509) Sorafenib 3Y: HR 1.01 (p= 0.946)
OS	No improvement Sunitinib v placebo: HR 1.17 (97.5% CI 0.90-1.52 p=0.1762) Sorafenib vs placebo: HR 0.98 (97.5% CI 0.75-1.28 p=0.8577)	No improvement Sunitinib v placebo HR 1.01 (95% CI, 0.72- 1.44; p=0.94)	No improvement ITT600mg HR 1.0 (CI 95% 0.80 -1.26, p>0.9) (updated 76month follow-up	No improvement HR 1.026 (95% CI 0.6–1.756) P= 0.92	No improvement 3Y sorafenib vs placebo HR 1.06 (95% CI, 0.82-1.38; p =0.638 1Y sorafenib vs placebo HR 0.92; (95% CI, 0.71-1.20; p = 0.541).
Toxicity					
Any grade (% patients)	not reported	99.7% sunitinib 88.5% placebo (98.4% and 75.8% treatment related)	98.6% pazopanib, (600mg group) 90% placebo	99% axitinib 56% placebo	100% 1Y sorafenib 99.2% 3Y sorafenib 97.4% placebo

Grade ≥3 (at least 1)	Starting dose; 63% suninitib 72% sorafenib 25% placebo Amended dose; > 55% sunitinib and sorafennib	63% sunitinib 22% placebo Dose not reported	66% pazopanib, 21% placebo (600mg group)	49% axitinib 12% placebo	58.6% sorafenib 1Y, 63.9% sorafenib 3Y 29.2% placebo
Treatment discontinuation	Starting dose; 44% sunitinib 45% sorafenib Amended dose; 34% sunitinib 30% sorafenib 10% placebo	28.1% sunitinib 5.6% placebo	35% 600mg group 39% 800mg group	Not reported	Not reported

IQR; inter quartile range, CI; confidence interval, HR; hazard ratio, ITT; Intention to treat, Y; year, ASSURE (Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinoma NCT00326898), S-TRAC (Adjuvant Sunitinib in High-Risk Renal-Cell Carcinoma after Nephrectomy, NCT00375674), PROTECT (Randomized Phase III Trial of Adjuvant Pazopanib Versus Placebo After Nephrectomy in Patients With Localized or Locally Advanced Renal Cell Carcinoma, NCT01235962), ATLAS (Adjuvant Axitinib Therapy of Renal Cell Cancer in High Risk Patients, NCT01599754), SORCE (Adjuvant Sorafenib for Renal Cell Carcinoma at Intermediate or High Risk of Relapse NCT00492258)

1.6.1 ASSURE trial (**Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinom**a **NCT00326898)**

ASSURE was a double blind, placebo-controlled, randomly assigned, phase three trial that published results in 2016**.** 1943 intermediate and very high-risk patients with ccRCC and non-ccRCC histological subtypes from 226 study centres in North America were assigned to one of two intervention: sunitinib (n=647) or sorafenib (n=649) for a year, or placebo (n=647) in a ratio of 1:1:1. Participants received fifty-four weeks of either sunitinib; 50mg once a day (od) for the first four weeks of a six week cycle and a total of nine cycles or sorafenib 400 mg twice a day (bd) for twenty-eight day cycles for up to thirteen cycles or placebo. Significant treatment related toxicities for patients on all arms led to a mid-trial protocol amendment allowing the reduction of starting doses to 37.5mg for sunitinib or 400mg od for sorafenib for the first one or two cycles of therapy. Further dose reductions and dose escalations were permitted.

Results from ASSURE found no significant differences in DFS or OS for either of the interventions relative to the placebo, (**Table 4)**. For sunitinib versus placebo the HR for OS was 1.17 (97.5% CI 0.90-1.52 p=0.1762) and for sorafenib versus placebo the HR for OS was 0.98 (97.5% CI 0.75-1.28 p=0.8577). Treatment discontinuation due to toxicity was reported for both sorafenib (45%) and sunitinib (44%) which was significantly reduced after the mid-trial starting dose reductions (p=0·0142 for sunitinib and p=0·0001 for sorafenib). However, even after dose reductions, the proportion of grade 3 or worse adverse events still exceeded 55% in both the sunitinib and sorafenib groups. 30% of sorafenib treated, and 34% of sunitinib treated participants discontinued treatment. See **Table 4** for summary of toxicity outcomes.

1.6.2 S-TRAC trial (**Adjuvant Sunitinib in High-Risk Renal-Cell Carcinoma after Nephrectomy, NCT00375674)**

S-TRAC was a randomly assigned double-blind phase three trial that also published results in 2016. It was a smaller trial including 615 patients, assigned in a 1:1 ratio to either a year of sunitinib (n=309) or to placebo (n=306). All participants were high risk by the modified UCLA UISS [25] prognostic model and a clear cell histological component was mandated. Sunitinib was administered at 50mg od for the first four weeks of a six-week cycle to a total of nine cycles. Dose interruptions or dose reductions to a lower limit of 37.5 mg per day were allowed, depending on the type and severity of toxicity.

In S-TRAC unlike ASSURE, there was modest a DFS benefit of 1 year of sunitinib when DFS events were subject to blinded independent review (HR 0.76; 95% CI, 0.59–0.98; p=0.03) [35]. However, no improvement in final OS was shown for sunitinib (HR 1.01 (95% CI, 0.72- 1.44; p=0.94). In practice, in both STRAC and ASSURE the associated toxicity of sunitinib, (**Table 4)**, prevented optimal dosing for many participants and necessitated significant dose reductions and treatment discontinuation.

1.6.3 PROTECT trial (Randomized Phase III Trial of Adjuvant Pazopanib Versus Placebo After Nephrectomy in Patients with Localized or Locally Advanced Renal Cell Carcinoma, NCT01235962)

The PROTECT trial, completed in 2017, was a phase three randomly assigned, placebo-controlled study that evaluated the efficacy of adjuvant pazopanib for one year in patients with intermediate-to-high risk locally advanced RCC. Patients were stratified by partial versus radical nephrectomy and their 2010 TNM staging and Fuhrman nuclear grades. The starting dose was reduced from 800mg to 600mg od after a mid-trial protocol change, following higher than expected treatment discontinuation rates, on blinded safety review. The primary analysis was revised to focus on the ITT group receiving pazopanib 600mg (ITT-600mg). Sample sizes were re-estimated to allow sufficient power to detect a 30% reduction in disease recurrence in the (ITT-600mg) group. The primary outcome analysis showed no significant DFS benefit in the ITT-600mg group compared with placebo (HR 0.86; 95% CI, 0.70 to 1.06; p =0.165), and mirrored an insignificant OS difference between the two groups [47] **(Table 4)**.

1.6.4 ATLAS trial (Adjuvant Axitinib Therapy of Renal Cell Cancer in High-Risk Patients, NCT01599754)

The ATLAS trial was a phase three, double-blinded trial comparing up to three years (minimum 1 year) of axitinib versus placebo in patients with predominantly clear-cell RCC (ccRCC). Randomly assigned patients were stratified by country and by risk group, using a combination of TNM with Fuhrman nuclear grading to receive (1:1) oral axitinib 5mg bd or placebo for up to three years. The trial was stopped due to futility at a pre-planned interim analysis at 203 DFS events. There was no significant difference in DFS (HR = 0.870; 95% CI, 0.66–1.15; p = 0.3211) and 23% of patients in the axitinib arm discontinued treatment due to adverse events **(Table 4)** [33].

1.6.5 SORCE trial (**Adjuvant Sorafenib for Renal Cell Carcinoma at Intermediate or High Risk of Relapse NCT00492258)**

SORCE [5] was an international randomised, phase three trial led by the Medical Research Council (MRC), Clinical Trial Unit at University College London (CTU at UCL). It evaluated the efficacy and toxicity of up to three years of adjuvant sorafinib compared to placebo, in patients with resected RCC with an intermediate or high risk for disease recurrence, as assessed by the 2003 Leibovich score [23] (more details of score components below). SORCE recruited 1711 patients from July 2007 to April 2013 (see **Figure 2** for SORCE trial consort diagram), from 147 centres in 7 countries, the majority from the UK (1331), then from Australia (168), France (122), Belgium (36), The Netherlands (19), Spain (15) and Denmark (20). Like ASSURE and PROTECT, patients with non- ccRCC were permitted. Participants were randomly assigned to one of three arms in a 2:3:3 ratio to receive either three years of placebo (Arm A), one year of sorafenib followed by two years of placebo (Arm B) or three years of sorafenib (Arm C). Stratification into arms was based on 2003 Leibovich risk group **(Table 10)** (intermediate vs high risk) and country. The starting dose was reduced in November 2008 from 400mg bd to 400mg od to address a higher than expected discontinuation rate due to toxicity **(Table 4)**.

SORCE along with ATLAS were the only two trials to investigate up to three years of treatment with a TKI. In light of S-TRAC showing efficacy of treatment in the highest risk categories, a subgroup analysis was pre-specified in the trial protocol; DFS in participants with a high-risk Leibovich score (score, 6-11) and in participants with clear cell histology.

No difference in DFS between the three-year sorafenib versus placebo group (HR, 1.01; 95% CI, 0.82 to 1.23; p=0.946) or OS was shown (HR, 1.06; 95% CI, 0.82 to 1.38; p= .638). Comparing one year of sorafenib versus placebo, no difference in DFS (HR, 0.94; 95% CI, 0.77 to 1.14; p = 0.509) and no difference in OS was observed (HR, 0.92; 95% CI, 0.71 to 1.20; *p* = .541).

Mirroring the other adjuvant TKI trials, despite permitting dose modifications, many participants failed to complete assigned protocol treatment, even in the placebo arm (Arm A 43%, Arm B 66%, Arm C 62% discontinued treatment). Hand-foot skin reaction and hypertension caused the majority of grade three events in both treatment cohorts,

despite the implementation of a reduced starting dose. Approximately half of participants withdrew from treatment by the end of the first year in both treatment arms and despite offering treatment adaptations, excessive toxicity was the reason for stopping treatment in 10-45% of cases.

1.6.6 Everest trial (Everolimus in Treating Patients with Kidney Cancer Who Have Undergone Surgery NCT01120249)

MTOR inhibitors have also shown evidence of response and the ability to prolong survival in the metastatic setting but have largely been superseded by immune checkpoint inhibitors (ICIs) and TKIs in the first- and second-line setting. Only one trial to date explores adjuvant therapy with an MTOR inhibitor, EVEREST-SWOG 0931 (NCT01120249). Participants were randomly assigned to everolimus or placebo for a period of 1 year. The study has yet to report results.

1.7 Lessons learnt form SORCE and the other adjuvant TKI trials

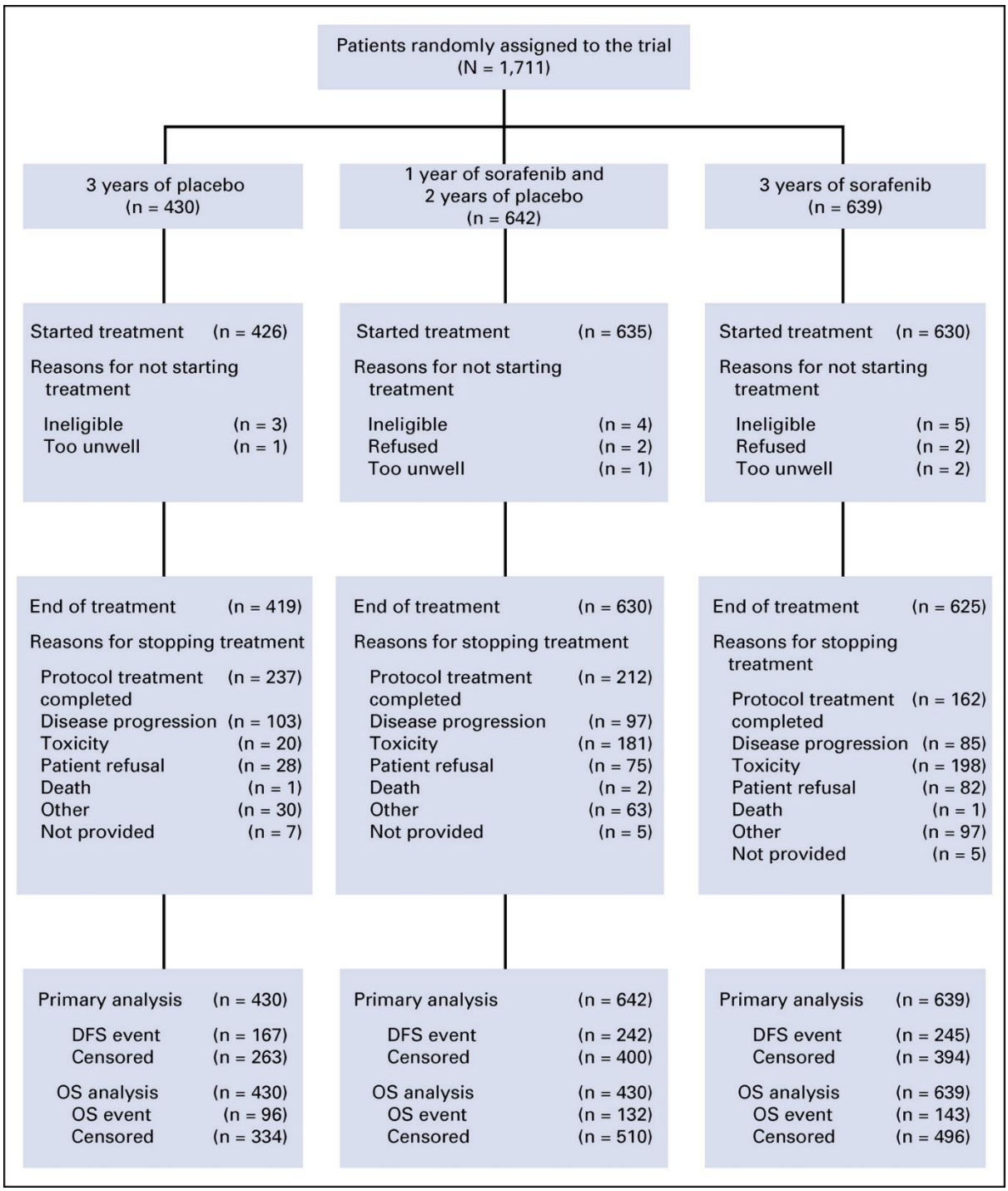
Despite the well documented dependency of VEGF/VEGFRs-driven angiogenesis in advanced ccRCC, it does not appear to play the same role in the locally advanced setting.

This incongruence is seen with other tumour types. Adding bevacizumab (VEGF monoclonal antibody) to chemotherapy in metastatic colorectal, breast, and non-small-cell lung cancers results in improved outcomes for patients compared to chemotherapy alone [48-50]. These improvements are not seen when bevacizumab is used in the adjuvant setting [51]. By targeting angiogenesis, anti-VEGF treatments are postulated to be directly cytostatic by interfering with tumour blood and oxygen supply. Another mechanism may be in improving chemotherapy delivery by stabilising structurally aberrant vessels [52]. The exact mechanisms of action remain uncertain but are unlikely to be relevant in the setting of fully excised or micro-metastatic diseases.

ASSURE and S-TRAC presented contrasting DFS results for sunitinib treated participants compared to placebo. Possible explanations for this may be differences in trial designs, participant populations and variability in doses participants were exposed to **(Table 3).** Ultimately, all TKIs tested in the adjuvant setting in RCC were associated with a higher than anticipated number of common toxicities, which were difficult for participants to tolerate. Toxicities that were reported in all trials included hypertension, hand-foot syndrome, fatigue and nausea. Across adjuvant TKI trials,

approximately 60% of patients on treatment arms experienced Grade ≥3 with a significant number of patients unwilling or unable to complete treatment due to toxicity even with dose reductions permitted**.** Sunitinib received FDA but not European Medicines Association (EMA) approval, after balancing cost, toxicity and efficacy of sunitinib in this setting. Lack of corroboration from data from the subsequent TKI trials means that the S-TRAC results have not been globally practice changing.

Figure 2 SORCE Consort diagram



Taken from the SORCE primary publication [5] in the Journal of Clinical Oncology showing patient flow in Arm A (left), B (middle) and C (right). DFS; disease-free survival. OS; overall survival

1.8 Rationale for a trial of immune checkpoint inhibitors (ICIs) in adjuvant RCC

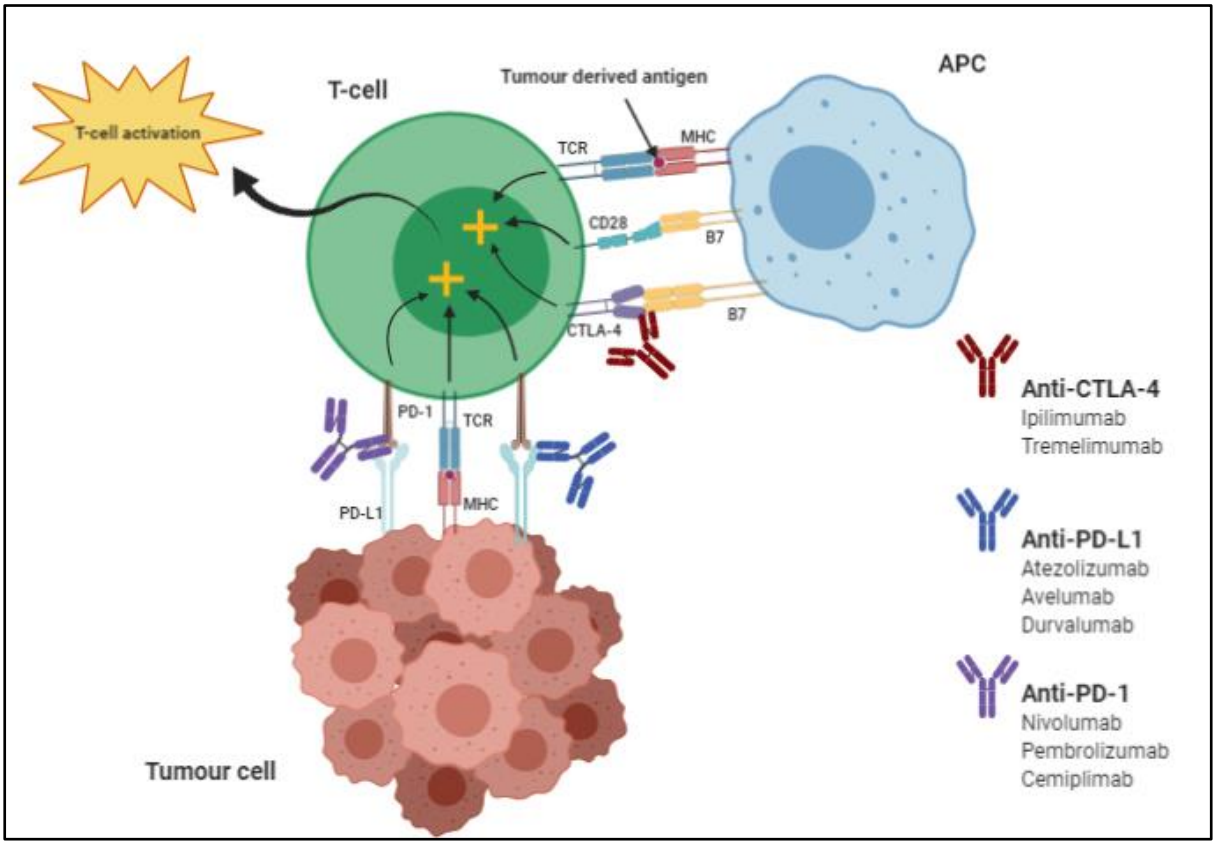
In recent years, immunotherapy with ICIs has revolutionised the management of patients with advanced RCC and other tumour types in both the adjuvant and advanced setting including lung cancer and melanoma. ICIs exert their influence by releasing the immunological breaks on T-cell activation and proliferation, and in doing so they reverse tumour mediated immune tolerance, (**Figure 3)**.

Specifically, antibodies against cytotoxic T-lymphocyte associated antigen 4 (CTLA-4) and programmed cell death protein 1 / programmed cell death protein ligand 1 (PD-1/PDL-1) have been approved and are standard of care treatment for patients with metastatic RCC and other emerging ICIs and ICI combinations are under review.

1.9 ICIs in advanced RCC

The combination of ipilimumab (anti-CTLA-4) with nivolumab (anti-PD-1) is now a first line treatment for patients with intermediate and poor risk RCC by IMDC [28] risk scoring. This followed findings from Checkmate-214 (NCT02231749) [53], which showed an 18 month OS of 75% (95% CI, 70-78) for the combination versus 60% (95% CI, 55-65) with sunitinib (hazard ratio for death, 0.63; 99.8% CI, 0.44 to 0.89; $p<0.001$). Single agent nivolumab, is also routinely available in the second line setting for patients who have progressed on TKI therapy, based on a 5.4 month OS benefit of nivolumab over everolimus [43]. As previously mentioned, there are four ICI and TKI combination strategies which have shown efficacy compared to sunitinib and some which are routinely available in the first-line treatment of metastatic RCC [30-32, 46]. The first phase three study, COSMIC-313 (NCT03937219) to evaluate upfront triplet therapy with Ipilimumab plus nivolumab and cabozantinib in patients with previously untreated advanced RCC was presented at the 2022 European Society of Medical Oncology (ESMO) congress [54]. It demonstrated significant PFS benefit in patients with IMDC intermediate and poor risk compared to those receiving ipilimumab and nivolumab (median PFS-not reached (triplet therapy) versus 11.3 months (doublet therapy)). Without data supporting OS benefit, findings have not yet changed clinical practice. See **Table 2** for summary of trial results evaluating ICIs in metastatic RCC.

Figure 3 Immune checkpoint inhibitor activity



Created using BioRender.com.

PD-1 and CTLA-4 are immune checkpoints expressed on activated T-cells. The expression of programmed cell death-ligand-1 on tumour cells and subsequent binding to PD-1 on the T-cell results in suppression of T-cell mediated anti-tumour activity. Monoclonal antibodies against PD-1 and PD-L1 inhibit PD-L1 on tumour cells. This releases the immunological breaks on T-cell activation and proliferation, in doing so promoting anti-tumour T cell activity.

Tumour derived antigens are presented by the antigen presenting cell (APC) via the major histocompatibility complex (MHC) and are recognized by the corresponding T cell receptor (TCR). A co-receptor interaction involving the CD28 protein and the corresponding B7 molecule causes T cell activation and subsequent anti-tumour response.

1.10 ICIs in locally advanced RCC

Propelled by their success in treating patients with metastatic RCC, ICIs are under investigation as adjuvant treatments after nephrectomy (See **Table 5** for current adjuvant RCC trials evaluating ICIs).

The phase three, international Keynote-564 trial (NCT03142334) was the first to report results in 2021 [55]. 994 participants with histologically confirmed RCC with intermediate (pT2, grade 4, N0 M0; or pT3, any grade, N0 M0) or high risk (pT4, any grade, N0 M0; or any T any grade, N+, M0), using TNM and Fuhrman grade criteria were included. They were randomly assigned to either pembrolizumab (anti-PD-1) or placebo after curative nephrectomy. Eligibility included those with sarcomatoid component and patients with resected oligometastatic disease within twelve months of nephrectomy. At a median follow-up of twenty-four months, adjuvant pembrolizumab was associated with significantly higher disease-free events (77% versus 68%) and a HR for DFS of 0.68 95% (CI, 0.53-0.87 p=0.0010) compared to placebo. Overall survival (OS) showed a non-statistically significant trend towards a benefit for pembrolizumab (HR 0.54, 95% CI 0.30-0.96, P=0.0164). Grade 3-5 all-cause adverse events occurred in 32% versus 18% of patients for pembrolizumab and placebo, respectively. Although limited by incomplete long-term overall survival data, the notable signal in DFS and tolerable safety profile means that these results are likely to be practice-changing, globally. Based on these results, pembrolizumab was in 2022 approved by the European Medicines Agency (EMA) and US Food and Drug Administration (FDA) for the treatment of RCC in the post-nephrectomy adjuvant setting.

Contrary to the KEYNOTE-564 results, data from two other large, randomised trials did not support the use of either atezolizumab monotherapy (IMmotion010-NCT03024996) or a combination of nivolumab and ipilimiumab (Checkmate-914-NCT03138512) in the post-nephrectomy setting. In IMotion010, median investigator-assessed DFS was 57·2 months (95% CI 44·6 to not evaluable) with atezolizumab and 49·5 months (47·4 to not evaluable) with placebo (HR 0·93, 95% CI 0·75-1·15, p=0·50) [56]. In Checkmate- 914, with a median follow-up of 37·0 months (IQR 31·3, 43·7), median DFS was not reached in the nivolumab plus ipilimumab group and was 50·7 months (95% CI 48·1 to not estimable) in the placebo group (HR 0·92, 95% CI 0·71–1·19; p=0·53) [57]. The 2022 EAU guidelines, consequently, report a weak

recommendation for the use of adjuvant pembrolizumab for patients with high-risk ccRCC until final OS results are available [20].

Ongoing efficacy data is required to address several important questions including the impact of adjuvant ICIs in those with non-ccRCCs notably the sarcomatoid RCC group and those with fully resected metastatic disease, termed M1NED. In addition, different ICIs and their combinations may have variable efficacies as already shown in localised and advanced RCC. Therefore additional data from the ongoing adjuvant ICI trials will be important.

RAMPART is a parallel international phase three randomised multi-arm, multi-stage (MAMS) platform trial led by the MRC CTU at UCL, evaluating the addition of ICIs after curative nephrectomy for patients with locally advanced RCC. The experience in delivering SORCE and lessons learnt from the TKI trial results have played a central role in shaping the development of RAMPART.

Durvalumab is a monoclonal antibody against PD-L1 which was chosen for investigation in RAMPART. RAMPART, (trial schema: **Figure 4)** investigates durvalumab as monotherapy (Arm B) and in combination with tremelimumab (arm C) after complete surgical excision of locally advanced RCC alongside patients on active surveillance, (Arm A). Durvalumab was chosen as it is has proven efficacy in the adjuvant setting in patients with non-small-cell lung cancer after completing definitive chemo-radiotherapy. The PACIFIC trial (NCT03519971) [58] showed a PFS advantage of 11.2 months for durvalumab treated patients compared with those taking a placebo. Durvalumab in combination with tremelimumab has shown benefit to OS and PFS in the third-line treatment of patients with metastatic non-small cell lung cancer. It is also being evaluated in the advanced setting in other tumour type settings (NCT03298451), (NCT03994393) and (NCT02516241).

1.11 RAMPART Trial Design

Similar to SORCE, a multi-arm design was used in RAMPART. This allows the simultaneous investigation of durvalumab alone and in combination with tremelimumab within the same trial. The MAMS approach enables more than one question to be addressed, in a way that controls the overall type 1 error rate or the family-wise error rate, (the chance of incorrectly rejecting the null hypothesis across all arms) [59]. Another advantage of the multi-arm design is that the family-wise error

rate can be controlled if a further research arm were to be added at a later time point. For example, combinations of TKI and ICIs that show promise in the advanced setting in the next few years, may potentially be assessed in an additional arm without increasing the chance of generating a false result when assessing the original arms [2, 60].

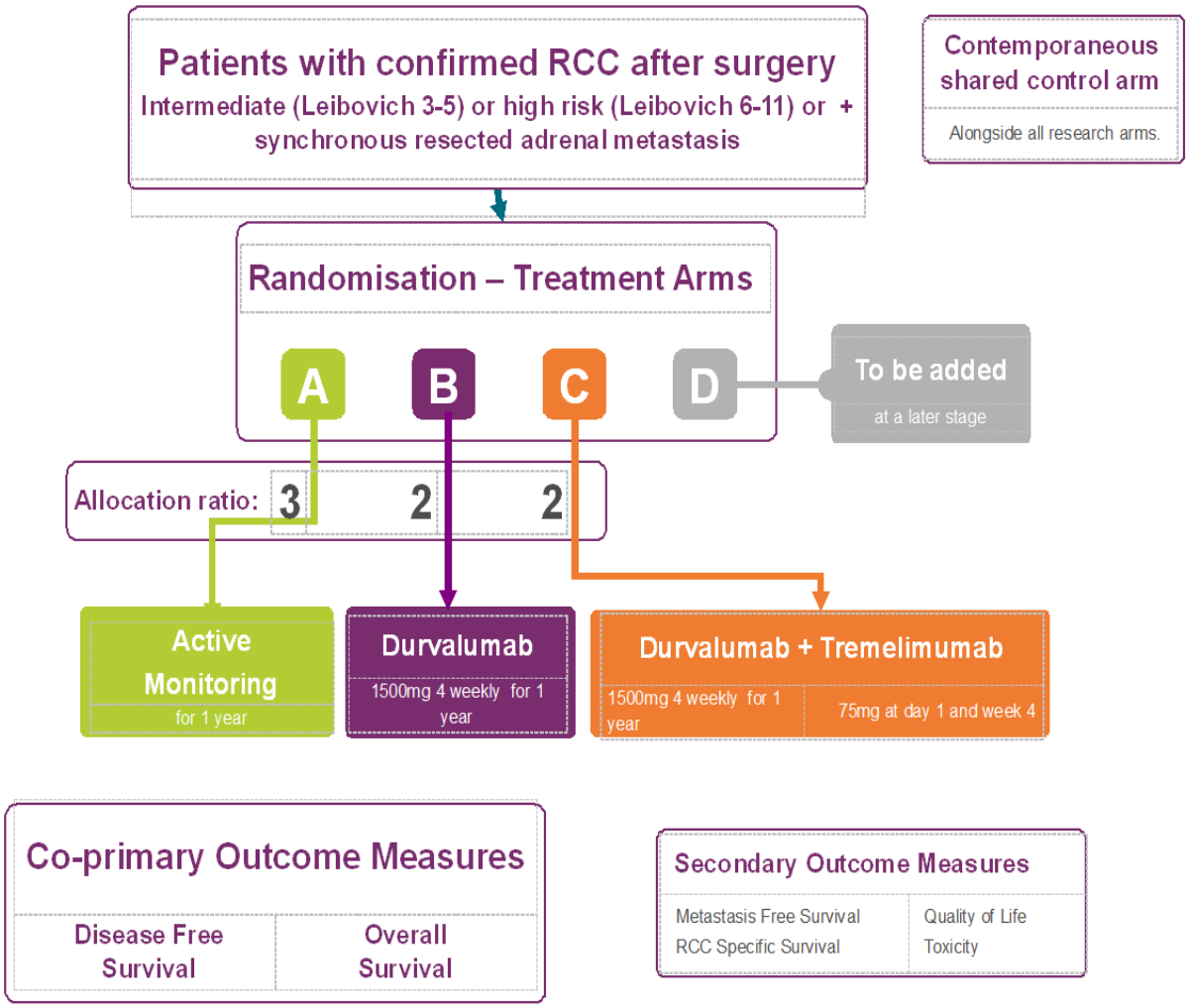
Another benefit of RAMPART's multi-stage design is the inclusion of a number of pre-planned, time-to-event driven interim analyses, that assess for both overwhelming benefit and lack of benefit of treatment. If treatments evaluated offer no effect at the interim analysis stage, recruitment to the corresponding treatment arms may be ceased and accrual switched to other more promising research arms or the control arm [60]. This means that fewer patients will be exposed to futile treatments and their toxicities. Conversely, including a planned overwhelming benefit analysis allows the early reporting of results if sufficient benefit is observed.

Table 5 Adjuvant RCC trials evaluating ICIs after curative nephrectomy

	KEYNOTE 564	RAMPART	IMMOTION010	CHECKMATE 914	PROSPER
No. participants	n=994	n=1750	n=778	n=800	n=766
Treatment arm(s)	Pembrolizumab (PD-1)	Durvalumab PD-L1 +/- tremelimumab (CTLA-4)	Atezolizumab (PD-L1)	Nivolumab (PD-1) +ipilimumab (CTLA-4)	Nivolumab(X1) → nephrectomy → nivolumab (X9)
Comparator arm	Placebo	Active monitoring	Placebo	Placebo	Nephrectomy alone
Duration	12m/ 17 cycles	12m/17 cycles	12m/16 cycles	24m	10 cycles total (1 + 9 months)
Histology	ccRCC +/- sarcomatoid component	ccRCC and non-ccRCC	component of ccRCC or sarcomatoid	ccRCC +/- sarcomatoid component	Any histology
Risk group	T2a, G3-4, N0M0; T2b-T4,G any, N0M0 Tany, Gany, N1M0	Intermediate (LS 3-5) or high (LS 6-11) or M1NED	T2NxM0 or TanyN+ or M1NED	T2a, G3-4,NXM0 T2b-T4,G any,N0M0 Tany, Gany, N1M0	T2NXM0 Tany, N1M0 M1 NED
Risk score	7th TNM 2010, Fuhrman grade	2003 Leibovich score	7th TNM 2010, Fuhrman grade	7th TNM 2010, Fuhrman grade	7th TNM 2010

N; nephrectomy, ccRCC; clear cell renal cell carcinoma, T; tumour size, N; nodal status, M; presence of metastasis, G; grade, M1NED; fully resected metastatic disease, KEYNOTE 564; Adjuvant Pembrolizumab after Nephrectomy in Renal-Cell Carcinoma NCT03142334, RAMPART; Renal Adjuvant Multiple Arm Randomised Trial NCT03288532, IMMOTION 010; Atezolizumab as Adjuvant Therapy in Participants With Renal Cell Carcinoma (RCC) at High Risk of Developing Metastasis Following Nephrectomy NCT03024996, CHECKMATE 914; nivolumab or nivolumab plus ipilimumab in patients with localized renal cell carcinoma at high-risk of relapse after radical or partial nephrectomy, PROSPER; Nivolumab in Treating Patients With Localized Kidney Cancer Undergoing Nephrectomy NCT03055013

Figure 4 RAMPART trial schema



Taken from RAMPART trial eposter presentation at NCRI Conference 2020

1.12 Identified Research Priorities

1.12.1 Choice of study population; stratifying patients according to risk of relapse after nephrectomy

In the absence of validated molecularly driven predictive assays, clinico-pathological scoring systems, developed over two decades ago, are used for selecting participants with RCC to adjuvant clinical trials. The ideal prognostic score for this purpose should allow the selection of patients with the highest risk of relapse whilst avoiding recruitment of and therefore potential treatment related toxicity in patients who will gain little benefit.

During the design of RAMPART, the 2003 Leibovich risk score [23] (**Table 10)** was chosen by the SORCE and later the RAMPART TMG for selection and random allocation of patients to trial arms. Firstly, its simplicity was favoured. As a five component score it allows straightforward classification of patients into three prognostic groups; low, intermediate and high risk of which the intermediate and high-risk groups were deemed suitable for clinical trial inclusion. Secondly, the 2003 Leibovich Score had demonstrated the ability to discriminate between risk groups in various datasets over the past two decades with consistent accuracy [61-64] and is recommended in international guidance for follow-up after curative nephrectomy [65]. The use of updated scores that focus purely on risk stratifying patients with papillary and chromophobe subtypes were considered when designing RAMPART. However, none had been externally validated at the time, so their accuracy in datasets other than those used to develop the scores was unknown. Therefore, as for SORCE, the 2003 Leibovich score was chosen for participant selection and risk stratification of patients considered for RAMPART with all included histological subtypes.

In order to justify the continued use of the 2003 Leibovich score, data was prospectively collected to perform an updated external validation. This was specified, in the SORCE trial protocol. The additional advantage of using data from SORCE participants was the potential to explore the accuracy of the score in patients with non-clear cell RCCs, which has not previously been done.

1.12.2 Inclusion of clear cell and non-clear cell histology's

Although there has been a move towards further molecular and histology-based classification of RCC tumours, this has not yet translated to diversification of treatment

paradigms in RCC. There exists a wide variability in clinical phenotypes exhibited by the RCC subtypes, including timescales of relapse after nephrectomy, sites of metastases and overall prognosis including response to systemic therapy. Despite this, studies characterising the histological subtypes individually are generally based on small retrospectively collected cohorts. In most of the adjuvant phase three TKI trials and current ICI trials participants were required to have a component of ccRCC histology. Designing trials for participants with ccRCC, the most prominent histology, led to more homogeneous populations but rendered the effect of treatment on those with non-clear cell histology's unclear. Consequently, current clinical guidance for the surveillance and treatment of patients with non-ccRCCs is based largely on data from patients with ccRCC.

There is a growing need to generate subtype specific clinical guidance particularly for the commoner subtypes like pRCC and chRCC. The highest level of evidence comes from prospectively collected multi-institutional datasets that provide randomised data with long-term follow-up. RAMPART's eligibility criteria includes any participant with fully resected intermediate or high risk RCC with any histology except for those with pure oncocytoma, (very favourable prognosis) and those with collecting duct, medullary and transitional cell cancers, (rare poor prognosis RCCs). Including a large multi-subtype ITT population allows inferences to be drawn on the clear cell and non-clear cell RCCs with high statistical power. However, including diverse characteristics and clinical behaviours in this way risks diluting the overall treatment effect of a particular subtype. Conversely, conducting separate analyses on each individual non-ccRCC subtype, risks losing statistical power for each analysis. How best to assess the effect of adjuvant strategies in rarer but clinically relevant sub-populations remains uncertain.

1.12.3 Extension of eligibility criteria to include patients with fully resected metastases

Systemic treatment currently represents the standard of care for patients with metastatic RCC. A percentage of these have single or oligometastatic disease. Oligometastatic RCC is a variably defined cancer state characterised by the development of a limited number of metastatic deposits that do not initially progress to a widespread distribution of cancer. The optimal treatment strategy for patients with oligometastatic RCC, particularly those that may be amenable to curative

metastasectomy or other radical treatments is under debate. Currently, systemic therapy with PD-1-based combination therapy is an accepted standard of care for patients who relapse within 1 year of nephrectomy [20]. For some patients, radically treating local or isolated synchronous or oligometastatic disease has delayed the initiation of systemic treatments until further progressive disease. Whether this optimises their overall survival outcomes, is uncertain and more data is required to underpin management guidance in this setting.

The role of early upfront systemic treatment after curative metastasectomy is of current scientific interest with many of the adjuvant ICI trials now including a cohort of patients with resected metastatic disease (M1NED). In 2021, a mid-trial protocol amendment was considered by the RAMPART TMG and finalised in September 2021, allowing patients with ipsilateral adrenal metastases resected at the time of nephrectomy to be included. The RAMPART TMG were also keen to widen eligibility further to include those with single fully resected metastatic deposits. The choice of patients for inclusion onto adjuvant trials, for example, the precise timings and extent of resected oligometastatic disease that may benefit from early upfront ICIs, lacks high-level evidence base. This uncertainty is highlighted by a variability in the precise definition of the resected metastatic cohort included across adjuvant RCC trials.

1.13 Summary of research objectives

I formulated several research questions using data from the SORCE and ASSURE trials. The overarching aim was to future-proof critical elements of adjuvant RCC trial design. These datasets provided a large cohort of data with long-term follow-up from patients with non-metastatic RCC who had undergone nephrectomy including participants with clear cell and non-clear cell histology's.

The first aim was to evaluate whether contemporary data from patients enrolled in SORCE could be used to conduct an up-to-date validation of the 2003 Leibovich score. I had the unique opportunity to compare contemporary outcomes of patients after nephrectomy to data from patients used to generate the score. I was also able to perform the first external validation of the 2003 Leibovich score in a multi-subtype population, thereby assessing its appropriateness for continued use in contemporary adjuvant RCC trials.

Secondly, combining data from SORCE and ASSURE, two previously reported international phase three studies provided a large dataset with long-term follow-up. The first aim was to characterise and compare the clinical trajectories of patients with non-ccRCCs, in detail. The second aim was to delineate potential drivers of poor prognostic disease. The third aim was to inform subtype specific surveillance recommendation after nephrectomy.

Finally, through analysing data from patients enrolled on SORCE, I was able to compare outcomes of patients who developed widely metastatic disease to those who initially developed limited relapse. The aim was to explore whether the timing and extent of first recurrences may be used as a prognostic tool for patients with higher risk RCC and to precisely define a group of patients that may be appropriate for evaluation in adjuvant trials.

Chapter 2: Accurate risk stratification and prognostication for patients with locally advanced RCC: A Validation of the 2003 Leibovich Score

2.1 Introduction

2.1.1 Prognostication and risk of recurrence after nephrectomy for renal cell carcinoma (RCC)

Locally advanced RCC is a heterogeneous disease. Clinical phenotypes range from those that remain in long-term remission post nephrectomy, those that develop resectable oligometastatic deposits over time and patients that progress rapidly with incurable metastatic disease. Despite the heterogeneity of outcomes, the current widely accepted treatment of locally advanced RCC post-surgery until recently has been the same for everyone - active surveillance post nephrectomy. The absence of novel biomarkers to guide prognostication after curative surgery in RCC, means that clinico-pathological scoring systems- some of which were generated over two decades ago, are still used to stratify patients into their risk of relapse or death post nephrectomy.

The tumour-node-metastasis (TNM) staging system developed by the American Joint Committee on Cancer (AJCC), has classically played the central role in staging and risk prediction in RCC. It anatomically classifies malignancies based on the extent of the primary tumour (T number), regional lymph node involvement (N number), and distant metastases (M number). Since its first publication in 1977, AJCC TNM has evolved to improve its predictive accuracy. As of January 1st, 2018, its 8th edition (**Table 6**) is recommended. All carcinomas of the renal cortex are covered by this staging system.

Staging the tumour is crucial for facilitating the exchange of standardised information amongst clinicians and researchers and allows the comparison of cases across regions, time periods and treatment modalities [66]. Staging is also used to evaluate changes in cancer incidence, disease extent at initial presentation, and the impact of various policy and treatment interventions. One of the primary dogmas of using AJCC TNM classification for staging is that a higher overall stage predicts a worse outcome. A report by Shao and colleagues compared the predictive accuracy of the 8th TNM in 2,120 RCC patients treated at Fudan University Shanghai Cancer Center (FUSCC) between 2000 and 2015 [67]. Results from this study revealed that the five-year overall

survival (OS) for T1-3N1M0 (AJCC stage III) was similar to T4N0M0 (AJCC stage IV), and lower than T3N0M0 (AJCC stage III) (38.1% *vs.* 36.2% *vs.* 72.7%) in the FUSCC cohort. Using the FUSCC cohort as a training set, validation of these findings was performed using the Surveillance, Epidemiology, and End Results (SEER) RCC cohort, which included data from 74,506 patients diagnosed between 2004 and 2014, yielding similar results. Therefore, although TNM is the widely accepted staging system, it lacks consistent predictive accuracy for all patients with RCC.

In order to improve the prediction of relapse risk for patients with RCC, components of TNM have been built upon to develop RCC specific prognostic models to better counsel patients after surgery and they have been used to individualise post-surgical surveillance. Additionally, accurate categorisation of risk groups is required to design and adequately power clinical trials of novel therapies. An additional important role for RCC risk prediction models is their use in clinical trials to guide the selection and random allocation of participants into trial arms. In the adjuvant setting, a number of prognostic models have been used in trials that assess potentially beneficial therapies after nephrectomy for patients with locally advanced RCC (**Table 8**).

Table 6 AJCC TNM 8th edition and corresponding staging

Primary tumour (T)	
T0	No evidence of tumour
T1a	≤4cm limited to kidney
T1b	>4cm to 7cm limited to kidney
T2a	T2a; Tumour > 7cm to 10cm limited to kidney
T2b	Tumour >10cm limited to kidney
T3a	Tumour extends to renal vein/ peri-nephric tissues not beyond Gerota's fascia
T3b	Tumour extends into vena cava below diaphragm
T3c	Tumour extends into vena cava above diaphragm or invades vena cava wall
T4	Tumour extends beyond Gerota's fascia (+/- ipsilateral adrenal)
Regional Lymph-nodes (N)	
N0	No nodes
N1	Involved nodes
Distant metastases (M)	
M0	No distant metastases
M1	Distant metastases
Stage	
I	T1N0M0
II	T2N0M0
III	T1-2, N1, M0
	T3, any N, M0
VI	T4, any N, M0
	Any T, any N, M1

2.1.2 Validating a prognostic model

Evidence of maintained performance of a model comes from validation studies [68]. Internal validation essentially means reusing parts or all of the 'derivation' dataset on which a model was developed, in order to assess the performance of the model [69]. External validation evaluates the performance of a model in a sample independent of that used to develop the model [69]. It allows for transferability of the model to patients beyond those included in the original dataset to be assessed. Any score should publish data on its internal and external validation before it can be used [70].

Successful validation of a model has come to mean achieving satisfactory *discrimination* and *calibration* in the chosen sample. Discrimination is the extent to which a model can assign a risk category that matches the patient prognosis e.g. patients predicted to be at higher risk should exhibit higher event rates than those deemed at lower risk [69]. Discriminative accuracy can be measured using data from an independent sample of patients who are assigned a risk category using the score in question and by observing the separation of Kaplan-Meier survival curves between the risk group cohorts. It can also be measured quantitatively using the Harrell's concordance index (c-index) (**Figure 5**) which quantifies the proportion of all patient pairs in which the predictions and observed outcomes are concordant [71]. C-indexes range from 0.5 (chance) to 1 (perfect). Calibration measures the agreement between the observed and predicted event rates for groups of patients [69, 72]. A risk score that is well-calibrated to contemporary outcomes, assigns the correct event probability at all levels of predicted risk.

2.1.3 Prognostic models and nomograms used for prognostication and follow-up post nephrectomy

The pursuit of improved prognostication in RCC has led to the development of at least ten models **(Table 8)** aimed at predicting tumour recurrence after nephrectomy or death. However the Kattan, Leibovich, UISS (University of California-Los Angeles Integrated Staging System), Karakiewicz and SSIGN (tumour Stage, Size, Grade, and Necrosis), risk prediction models are the most widely used post-operative scoring systems and nomograms and will be focused on here [23, 25, 26, 73-76].

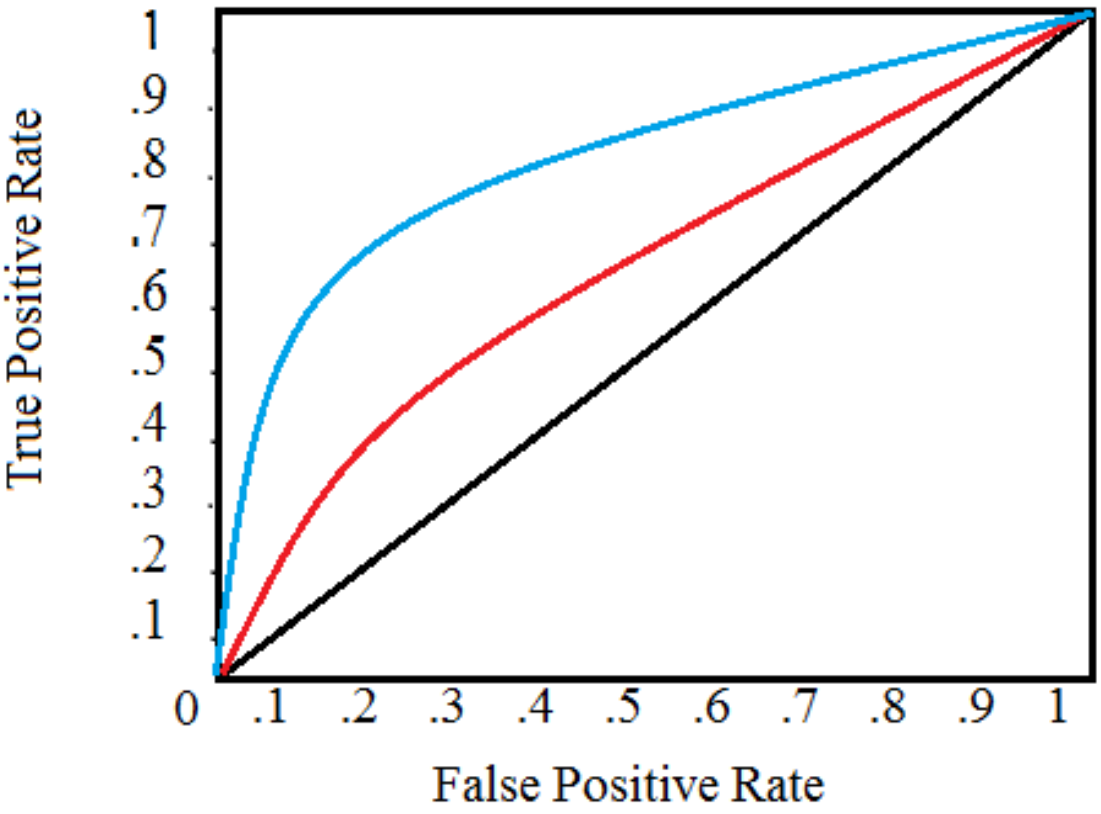
The European Society of Medical Oncology (ESMO) 2016 and the European Association of Urology (EAU) 2022 guidelines highlight three particular models-the 2003 Leibovich score, SSIGN and the UISS scores for assessing the risk of recurrence

after curative surgical treatment for ccRCCs [14, 20, 65]. Despite having undergone more than thirty published external validations between them (**Table 8)** the use of any one of these scores to guide prognostication and follow-up is not mandated. One potential reason for this is that although the models contain more variables than TNM, their discriminative abilities (by directly comparing their c-indexes) have recently been shown to only marginally outperform 2002 TNM staging **(Figure 6)**. On comparative analysis of eight predictive models, SSIGN was shown to perform the best- (c-index 0.688) compared to 2002 TMN (C-index, 0.60). Two of the models proposed (UISS and Yaycioglu) demonstrated decreased predictive accuracy (see **Figure 6**). This study also demonstrated statistically significant declines in predictive ability with respect to events beyond the second year post diagnosis, through charting c-index estimates for all the eight predictive models over time. This was shown for all models, most marked with SSIGN and included TNM [64].

Importantly, the performance of these models has typically be tested in retrospectively assembled datasets and rarely in large, prospectively collected data from multicentre cohorts. This would provide the highest level of evidence (Levels one and two) for their use (**Table 7**), according to the hierarchy of evidence applicable for adopting a prediction model [77, 78].

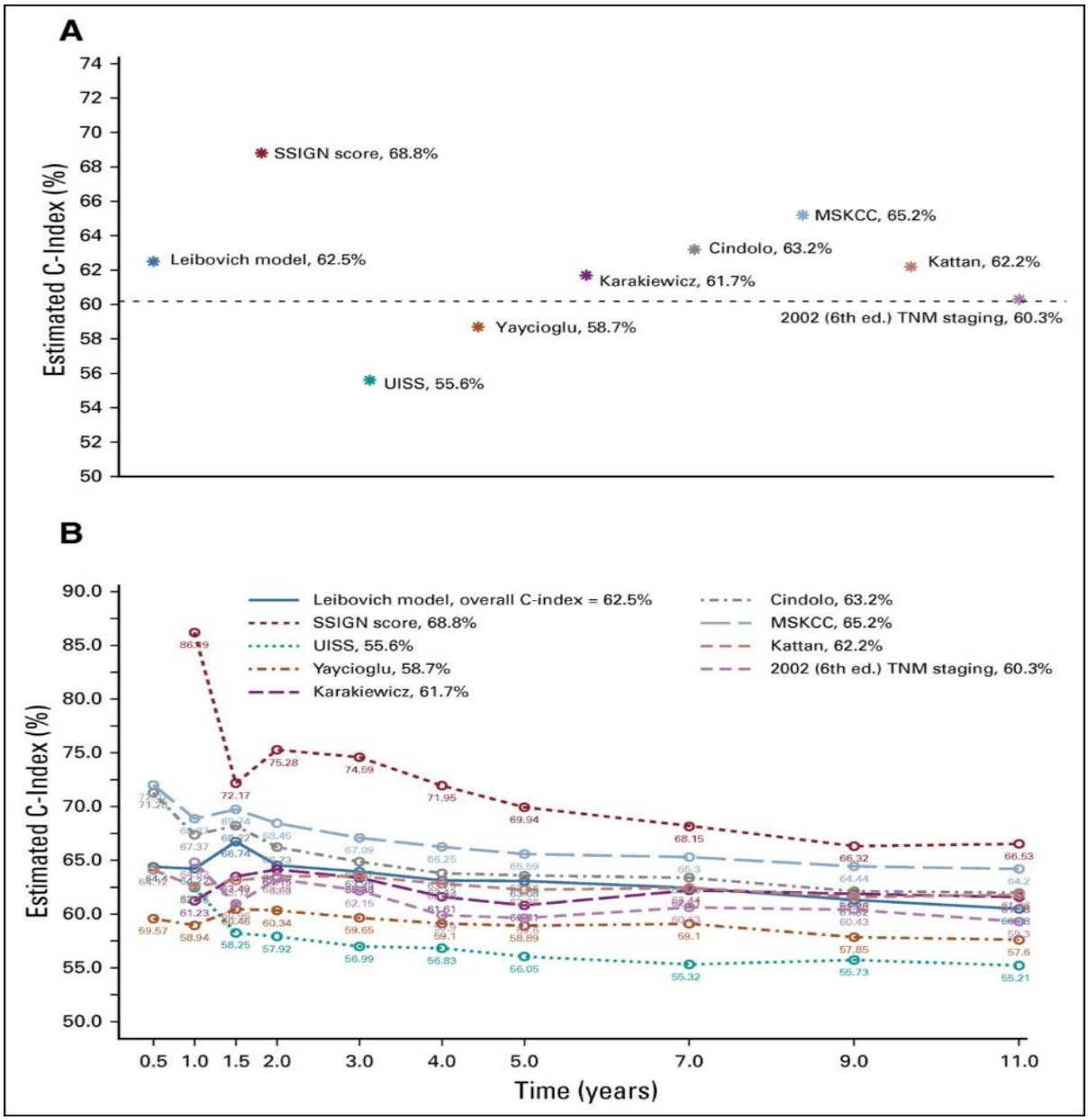
Figure 5 ROC curve showing accuracy of two tests.

The closer the graph is to the top and left-hand borders (blue test), the more accurate the test. The closer the graph to the diagonal (black test), the less accurate the test.



Credit to https://www.statisticshowto.com/receiver-operating-characteristic-roc-curve/

Figure 6 C-index estimates for eight prognostic models using data from the assure trial.



A. Overall C-index estimates; dashed line delineates the TNM C-index threshold. B. The C-index estimates over time. MSKCC, Memorial Sloan Kettering Cancer Center; SSIGN, State, Size, Grade, Necrosis Score; UISS, University of California at Los Angeles Integrated Staging System.

Taken from 'Predicting Renal Cancer Recurrence: Defining Limitations of Existing prognostic Models With Prospective Trial-Based Validation' Correa *et al* 2019 [64]

Table 7 Levels of Evidence for Prognostic Studies*

Level	Type of Evidence
I	High quality prospective cohort study with adequate power or systematic review of these studies
II	Lesser quality prospective cohort, retrospective cohort study, untreated controls from an RCT or systematic review of these studies
III	Case-control study or systematic review of these studies
VI	Case series
VI	Expert opinion; case report or clinical example

Adapted from the American Society of Plastic Surgeons [77]

Table 8 Validated c-indexes and prognostic factors of common RCC prognostic models

Model	Outcome	Histology	No. factors	TNM	Factors	Reported C-index	No. external validations	External c-indexes
UISS (2001) [24]	OS	Clear cell RCC/ papillary RCC/ chromophobe RCC/ sarcomatoid RCC	3	1997	For localised; T-stage Fuhrman grade ECOG PS	0.73	5	0.55-0.86
SSIGN (2002) [74]	CSS	Clear cell RCC	6	1997	T stage Fuhrman grade Size <5 or ≥5cm Nodal status Tumour Necrosis M0/M1	0.84	6	0.69-0.88
Leibovich (2003) [23]	MFS	Clear cell RCC	5	2002	T stage Modified grading* Size <10 or ≥10cm Nodal status Tumour Necrosis	0.82	4	0.625-0.82
Kattan (2001) [26]	RFS	Clear cell RCC/Papillary RCC/chromophobe RCC	6	1997	T stage M0/M1 Symptoms at presentation Age Gender	0.74	8	0.27-0.84
Yaycioglu [75]	RFS	ccRCC/pRCC/chRCC	2		Symptoms at presentation	0.65	4	0.58-0.7

					Tumour size			
Karakiewicz (2009) [79]	CSS	ccRCC/pRCC/chRCC	6	2002	T stage Fuhrman grade Tumour size Nodal status M0/M1 Symptoms at presentation	0.86	3	0.617-0.88
Cindolo [76]	RFS	ccRCC/pRCC/chRCC	2		Tumour size Symptoms at presentation	0.67	4	0.63-0.75

UISS; University of California, Los Angeles, integrated staging system, SSIGN; Mayo Clinic Stage, Size, Grade, and Necrosis Score, OS; Overall survival, CSS; Cancer-specific survival, MFS Metastasis-free survival, RFS; Relapse-free survival, T; Tumour size, ECOG PS; Eastern Cooperative Oncology Group Performance status. M0; no metastases, M1; presence of metastases

*See appendix Table A

Adapted from 'Predicting Renal Cancer Recurrence: Defining Limitations of Existing prognostic Models With Prospective Trial-Based Validation' Correa *et al* 2019 [64]

2.1.4 Models used in Adjuvant Trials in RCC

The inclusion of patients to previously reported adjuvant phase three trials in RCC (**Table 9**) and to those that are currently accruing patients (**Chapter 1, Table 5**), rely on the prognostic abilities of a number of post-operative risk score models; UISS [25], SSIGN [74] and the 2003 Leibovich score [23]. Pre-operative models, for example, Yaycioglu [75] and Cindolo [76] models, (**Table 8**), have consistently shown lower discriminating accuracy compared to post-operative counterparts and are not recommended in clinical guidelines or in adjuvant RCC trials. RCC nomograms, for example the Kattan and Karakiewicz models (**Table 8**) outperform algorithm-based models [80] in terms of survival prediction. However, the nomogram format is less practical for generating risk groups to stratify patients into trial arms and is not favoured.

The UISS model was used in adjuvant Tyrosine Kinase Inhibitor (TKI) and mechanistic target of rapamycin (mTOR) trials; ASSURE (ECOG-ACRIN E2805; Adjuvant Sorafenib or Sunitinib for Unfavorable Renal Carcinoma), STRAC (Sunitinib Versus Placebo For The Treatment Of Patients At High Risk Of Recurrent Renal Cell Cancer) and EVEREST (Everolimus for Renal Cancer Ensuing Surgical Therapy). Published in 1997, UISS was generated from 814 patients having undergone a radical or partial nephrectomy and delineated five survival groups which were later modified to stratify patients into clinically useful categories; low, intermediate, and high risk [25]. UISS predicts OS in patients with locally advanced and metastatic RCC and included clear cell and non-clear cell histological subtypes. A disadvantage is its inclusion of Eastern Cooperative Oncology Group (ECOG) Performance status (PS) a factor which may be subject to inter-user variability.

The SSIGN model, in contrast comprises histopathological components (**Table 8**), each with strict international guidance for their standardised reporting [81]. SSIGN was published in 2002 based on a cohort of 1,801 patients with ccRCC and predicts CSS [74] in patients with localised and advanced disease. The ease of standardisation of SSIGN components together with its consistently high discriminative ability (c-index 0.76-0.88) supports its continued recommendation in RCC surveillance guidelines.

2.1.5 2003 Leibovich Score

The SSIGN model was adapted by Leibovich *et al* in 2003, benefitting from the same histological features, but modified for additional practical applicability in the locally advanced setting, (SSIGN stratifies CSS based on scores of 0-17 whereas Leibovich stratifies metastasis-free survival (MFS) based on scores of 0-11, tiered into 3 risk strata). For these reasons, the 2003 Leibovich score, (**Table 10**), has superseded SSIGN for characterising and stratify patients in adjuvant RCC clinical trials and is also used by clinicians to guide the follow-up of patients with ccRCC after nephrectomy. It was used in the recently published SORCE trial (Sorafenib in Treating Patients at Risk of Relapse after Undergoing Surgery to Remove Kidney Cancer) trial [5] and is currently used in RAMPART, (Renal Adjuvant Multiple Arm Randomised Trial: NCT03288532), to guide patient recruitment and randomisation.

Leibovich and colleagues developed the 2003 score using retrospectively collected clinico-pathological details from 1671 ccRCC patients, from a single US institution, who underwent radical nephrectomy between 1970 and 2000. Univariate and multivariate cox proportional hazards regression models were used to determine the association between the time to distant metastases or death and a number of clinical and pathological features [23]. Five features were found to be significantly associated with time to distant metastases ($P < 0.001$) and comprised the final multivariable model: T-stage (TNM 2002), N stage, maximum diameter of tumour (cm), Fuhrman nuclear grade and presence of tumour necrosis. It was possible to identify a high-risk cohort (374 patients or 22% of the sample) an intermediate risk cohort (608 patients or 36% of the sample) and a low-risk cohort (689 patients or 41.2% of the sample). This was based on comparing MFS probabilities at individual risk scores (**Table 10**) and amalgamating them into the three clinically useful groups; scores 0-2, (low risk), scores 3-5 (intermediate risk) and scores 6 or higher, (high risk). Five-year metastasis-free probabilities were reported as 97.1%, 73.8% and 31.2% respectively [23] **(Figure 7**).

Ongoing clinical use of the 2003 Leibovich score is supported by maintained discrimination over time, mostly from retrospective, single institution studies. C-indexes [71] for the 2003 score range from 0.62 to 0.82 [61-63, 82], and on direct comparison it has been shown to outperform a number of other prognostic scores, notably those that include baseline patients characteristics [64]. Furthermore, the 2003 Leibovich is a pragmatic score. Clinical markers such as patient's performance status are not included, reducing bias in its calculation. In addition, the 2003 Leibovich score uses MFS as the outcome measure which is considered a practical outcome measure in the adjuvant setting as compared with overall survival (UISS) or CSS (SSIGN). The score also provides reliable information over a long median follow-up of 5.4 years, (UISS 2.5 years, Kattan 3.3 years, Yaycioglu 4 years) providing accuracy of long-term prognostication.

2.1.6 RCC Subtype Specific Prognostic Scores

Several newer subtype specific prognostic scores have been developed [72, 83] but not all have been externally validated in large representative cohorts. In 2018, Leibovich and colleagues published five scoring systems, modelling progression-free-survival (PFS) and CSS for ccRCC and pRCC and PFS for chRCCs [72]. An internally validated c-indexes of 0.83 was achieved. The 2018 score was externally validated in a single centre cohort from Singapore which included 829 patients with ccRCC (c-index of 0.81 for PFS, 0.83 for CSS) and 113 patients with pRCC (c-index of 0.72 for PFS and 0.74 for CSS) [84].

Several pRCC prognostic models have been published over the past years. A nomogram predicting CSS was developed and validated in 2010 [85], but included patients with and without distant metastases, which limits its specificity for the adjuvant setting. Buti *et al*. developed the GRade, Age, Nodes and Tumour (GRANT) score and used an independent multi-subtype cohort from the ASSURE trial (NCT00326898) in a validation exercise yielding a c-index for DFS and OS of 0.589 (95% CI, 0.571–0.6074) and 0.613 (95% CI, 0.589–0.636) respectively [86]. In 2019 the VENUSS score was developed by Klatte and colleagues for patients with non-metastatic pRCC based on routine pathological variables; presence of VEnous tumour thrombus, NUclear grade, Size, T and N Stage to predict low, intermediate and high risk of disease recurrence [83]. Internal validation showed that both discrimination and calibration from VENUSS appeared to be superior to the UISS, TNM and 2018

Leibovich models. In 2021 the VENUSS risk groupings were externally validated using data from 1085 consecutive patients with pRCC from seven European institutions and noted a discrimination of 0.786 (95% CI, 0.748-0.824) [87]. Despite the increasing number of subtype specific scores being developed and recently being externally validated, trial protocols continue to favour the simplicity and standard application of single scores even though most of them recruit patients with a variety of RCC subtypes.

2.1.7 RAMPART (Renal Adjuvant Multiple-Arm Randomised Trial)

The MRC CTU led RAMPART trial [1] (NCT03288532) is one of the ongoing phase three adjuvant renal trials that examines the efficacy of adjuvant immune checkpoint inhibitors (ICIs) after nephrectomy in the setting of locally advanced RCC (see **Figure 4** Chapter 1 for trial schema). It includes patients with intermediate and high-risk disease as defined by the Leibovich 2003 score. For RAMPART specifically, given the associated costs, resource implications and potential toxicity associated with testing ICIs, the optimal selection of participants based on risk of relapse is critical. In addition, information from the outcomes and Leibovich scores from SORCE participants, was used in the power calculations that underpin the patient numbers required in each of the RAMPART trial arms. For this, the Leibovich score needs to show a reasonable degree of calibration to contemporary outcomes particularly within the intermediate risk group, (where events generally occur later) in order to ensure that enough events occur for meaningful statistical comparisons to be made by the time the trial has fully accrued. Finally, the 2003 Leibovich score selects intermediate and high-risk patients with non-clear cell histology's into RAMPART, with no evidence of its discriminative accuracy in these cohorts.

Therefore, even though the 2003 Leibovich score has been validated before, its ongoing use in contemporary clinical trials requires its concurrent validation in up-to-date clinical datasets. validation exercise using the SORCE trial cohort (a similar population to that of RAMPART), ensures that the score remains fit for this specific purpose and in doing so confirms the integrity of adjuvant trial designs such as RAMPART, going forward.

Table 9 Comparison of completed adjuvant TKI trials

	ASSURE	S-TRAC	PROTECT	ATLAS	SORCE
No. participants	**n=1943**	**n= 615**	**n= 1538**	**n=724**	**n=1711**
Treatment arm(s)	Sunitinib or Sorafenib	Sunitinib	Pazopanib	Axitinib	Sorafenib or Sorafenib
Comparator arm	Placebo	Placebo	Placebo	Placebo	Placebo
Duration	12m	12m	12m	36m	12m or 36m
Histology	ccRCC and non-ccRCC	ccRCC (≥50%)	ccRCC (≥50%)	ccRCC (≥50%)	ccRCC and non-ccRCC
Risk group	Intermediate high to very high		Intermediate to high risk		Intermediate to high risk
Risk score	Modified UISS intermediate high-very high	Modified UISS high risk	SSIGN Intermediate-high risk	TNM and Fuhrman grade	2003 Leibovich Score
TNM edition	TNM 2002	TNM 2002	TNM 2010	TNM 2010	TNM 2002

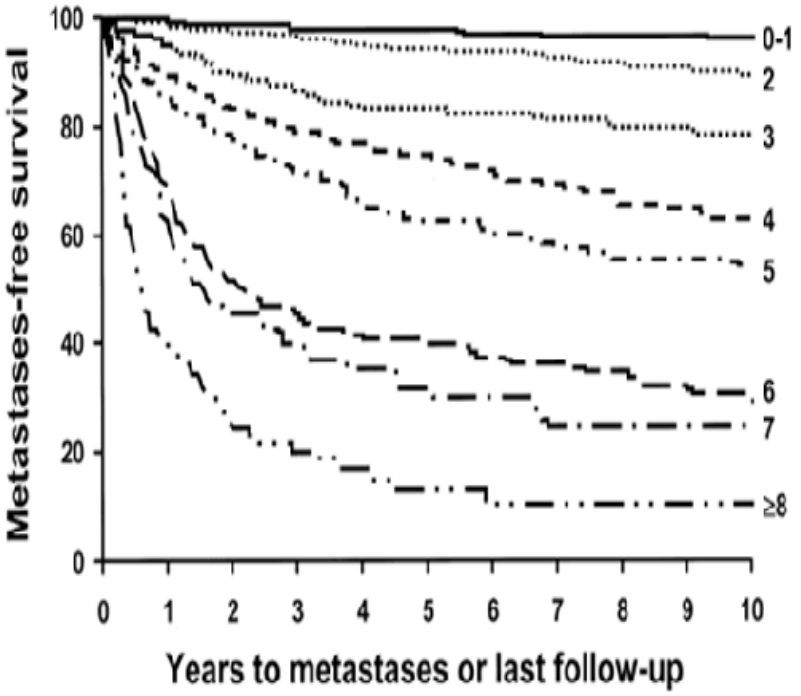
TNM; tumour size (T), nodal status (N), Metastases (M); ccRCC; clear cell renal cell carcinoma, non-ccRCC; clear cell renal cell carcinoma M1NED; fully resected metastatic disease, no evidence of disease, LS; Leibovich score, m; month, ASSURE; Adjuvant Sorafenib or Sunitinib for Unfavorable Renal Cell Carcinoma (NCT00326898), S-TRAC; Sunitinib vs. Placebo for the Treatment of Patients at high risk for Recurrent Renal Cell Cancer (NCT00375674), PROTECT; Pazopanib as an Adjuvant Treatment for Locally Advanced Renal Cell Carcinoma (NCT01235962), ATLAS; Adjuvant Axitinib Treatment of Renal Cancer (NCT01599754), SORCE; Sorafenib for Patients with Resected Primary Renal Cell Carcinoma (NCT00492258), EVEREST; Everolimus for Renal Cancer Ensuing Surgical Therapy (NCT01120249), UISS; University of California, Los Angeles, integrated staging system. SSIGN; Mayo Clinic Stage, Size, Grade, and Necrosis Score.

Table 10 Leibovich Score 2003

Feature		Score
Pathological T category of primary tumour	pT1a	0
	pT1b	2
	pT2	3
	pT3a-4	4
Regional lymph node status	pNx or pN0	0
	pN1-pN2	2
Tumour size	<10cms	0
	10cms or more	1
Nuclear grade	1 or 2	0
	3	1
	4	3
Histological tumour necrosis	No	0
	Yes	1

Scores	Group	Mean time to disease progression (years)
0-2	Low-risk	7.4
03-May	Intermediate risk	4
6 or more	High-risk	1.7

Figure 7 How individual 2003 Leibovich scores relate to 5-year metastasis survival



Leibovich, B.C., et al., Prediction of progression after radical nephrectomy for patients with clear cell renal cell carcinoma: a stratification tool for prospective clinical trials. Cancer, 2003. 97(7): p. 1663-71 [23].

2.1.8 Rationale for study

The performance of a score, developed two decades using clinico-pathological characteristics of patients who underwent nephrectomy between 1970-2000, may change over time. To better understand the implications of using the 2003 Leibovich score for contemporary adjuvant clinical trial design, I conducted a validation exercise in data from participants from SORCE [88]. Although the 2003 Leibovich score has previously been validated, its performance using randomised individual participant data (IPD), prospectively collected for this purpose has not previously been achieved. SORCE [88] was one of the largest multi-institutional phase three trials in RCC to report on the effect of TKIs after nephrectomy in patients with locally advanced RCC. The 2003 Leibovich score was used to stratify participants into intermediate and high-risk groups. Given that active surveillance remains a standard of care for patients with resected locally advanced RCC, a validation exercise in data from SORCE; where treatment arms showed no difference in survival to that of placebo, provides a, clinically relevant, dataset, of individual participant data with detailed follow-up, for this purpose.

I was granted access to the data originally used to derive the 2003 Leibovich score by Professor Bradley Leibovich and colleagues at the US Mayo Clinic. Unique to this study, this allowed two 'matched' datasets (in-terms of baseline clinical characteristics and follow-up time) using data from intermediate and high risk ccRCC patients to be generated, one from SORCE data and the other from the original Leibovich dataset. This patient group are of particular interest for consideration of adjuvant treatments after nephrectomy (low risk patients are generally not included because they are usually cured by surgery or ablation). As such, focusing on this cohort specifically, is directly applicable to adjuvant RCC trials.

The use of IPD rather than summary level data allowed a much more detailed analysis to be conducted than previously achieved. For example, I was able to truncate follow-up time in both derivation and validation datasets at ten years, which enhanced the comparability between the derivation and validation datasets. Using the matched dataset, new estimates of discrimination and calibration were derived.

Some of the immune-oncology focussed adjuvant RCC trials including IMMOTION010 (NCT03024996) and KEYNOTE-564 [55], used the TNM staging system and Fuhrman

nuclear grading rather than a validated RCC specific prognostic model for participant eligibility and stratification to trial arms. Perhaps simplicity rather than accuracy dictated this choice (TNM comprises two relevant variables; Tumour size and presence of nodal disease). As aforementioned, the analysis by Correa *et al.* published in 2019 compared the discriminative accuracy of 2002 TNM staging, to a number of RCC prognostic scores by conducting a retrospective analysis of data collected within the ASSURE trial [64]. Their validation noted a c-index for the 2003 Leibovich score of 0.625 which was superior to the 2002 TNM Staging (c-index 0.60) (**Figure 6**). Variables comprising the 2002 TNM staging components for each SORCE participant were collected. To support the use the 2003 Leibovich Score for risk stratification, I was able to directly compare the discriminative accuracy of 2002 TNM to that of the 2003 Leibovich score analysing the IPD from SORCE and from the original Leibovich dataset.

Finally, a secondary aim was to explore the ability of the 2003 Leibovich score to discriminate between risk categories within important histological sub-populations: any non-ccRCC, pRCC only, and chRCC only carcinomas.

2.1.9 Research question and aims

In this chapter, I address the following research question and objectives.

Research question

Can contemporary data from the SORCE trial be used to validate the 2003 Leibovich Score in a multi-subtype population and thereby justify its continued use in contemporary adjuvant RCC trials.

Research aims

Primary aim

An external validation the Leibovich 2003 score using contemporary data from participants with intermediate and high risk RCC recruited to the SORCE trial. Comparing its performance to a matched dataset derived from the original 2003 Leibovich dataset.

Secondary aim

1. An exploration of the 2003 Leibovich Score's discriminative accuracy compared to that of the 2002 TNM staging system using IPD from the SORCE trial and from the 2003 Leibovich dataset.

2. An exploration of the 2003 Leibovich score's ability to discriminate between high and intermediate risk categories within important SORCE sub-populations: patients with any non-ccRCC, pRCC, and chRCC.

2.2 Methods

2.2.1 Design

This study is an external validation of the 2003 Leibovich score using data from the SORCE trial [5]. The IPD used to derive the Leibovich score was provided by Professor Bradley Leibovich and colleagues, Mayo Clinic, US.

2.2.2 Obtaining the Original Leibovich data

For this analysis, I completed a data release request in January 2019 for the raw data used to derive the 2003 Leibovich score, including baseline characteristics, histopathological variables and survival outcomes (times to event and censoring indicator). This ensured that the Leibovich scores could be re-calculated using their data for this analysis.

2.2.3 Participants – Leibovich Score Calculation and Randomisation

SORCE participants were recruited from July 2007 to April 2013 from 147 centres in 7 countries; the UK, Australia, France, Belgium, The Netherlands, Spain and Denmark and followed up until July 2019. Participants with intermediate (3-5) and high (≥6) Leibovich scores were eligible for SORCE and low-risk participants were excluded. Participants with all histologies apart from oncocytoma were eligible. As part of the screening procedures for each potential patient entering the SORCE trial, the Leibovich score was calculated via the predefined features derived from histopathological examination of post-nephrectomy tumour specimens**.** Values for components of the 2003 Leibovich score were collected for each patient and recorded on the SORCE Randomisation Case Report Form.

The Leibovich dataset included US-based patients with ccRCC whose primary surgery occurred between 1970 and 2000. 1671 patients had complete data for the pathologic features of interest in this study. In SORCE and the original Leibovich publication, 2002 TNM staging was used to calculate the Leibovich score. Both used a nuclear grading system that simplified nuclear grading by selecting the worst WHO/ International Society of Urologic Pathologists (ISUP) features at each grade [89] (Appendix Table A). All tumours, regardless of histological subtype, were graded according to this updated nuclear grading system and assigned a Leibovich score.

2.2.4 Primary Analysis Cohorts

The two cohorts were matched as closely as possible to optimise comparability. A ‘derivation cohort’ was derived from the 2003 Leibovich dataset which included patients with ccRCC only and excluded the low-risk group (scores 0-2). The ‘validation cohort’ was derived from intermediate and high-risk participants with ccRCC recruited to SORCE. All patients in the 2003 Leibovich cohort underwent radical nephrectomy. Participants who underwent partial or radical nephrectomy from SORCE were chosen, as this more closely reflects contemporary surgical practice.

2.2.5 Secondary Analysis Cohorts

Non-clear cell RCC subtypes are recruited to adjuvant clinical trials using the 2003 Leibovich score. Its performance firstly in SORCE patients with any non-ccRCC sub-types and secondly in those with pRCC only and thirdly in those with chRCC only was evaluated and compared to the score’s performance in the primary analysis derivation cohort.

2.2.6 Outcome Measures

The primary outcome for the analysis was time to MFS, defined as the interval between nephrectomy and the date of distant metastases. The time origin was chosen as the 'date of nephrectomy' to match the time origin used in the original Leibovich analysis [23], where deaths preceding metastasis were censored (C. Lohse, personal communication, 4.11.2020) and so for this analysis, MFS was defined identically. MFS in both cohorts was censored at ten years to reflect available follow-up data in SORCE.

Secondary analyses using the same MFS definition were carried out in the three SORCE sub-populations defined above.

2.2.7 Statistical Methods

Missing data

The validation exercise was performed in adherence to transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) guidelines [90].

A nuclear grade assignment was missing for one participant, which was imputed singly by substituting the most common nuclear grade value (3), to ensure completeness of the validation dataset. In SORCE, there were 56/1711 (3%) missing dates of surgery. This was handled by taking a random selection of 56 values from the distribution of observed intervals between surgery and randomisation.

Statistical approach

For this external validation, I developed the clinical questions and wrote the statistical analysis plan which was presented and accepted by the SORCE trial management group (TMG). The statistical analysis was carried out by Eleni Frangou supported by Professor Patrick Royston, whose expertise in prognostic score validation was invaluable for the design and implementation of the statistical analysis. All analyses were carried out using STATA version 16.1 (StataCorp LLC, College Station, TX).

The performance of the 2003 Leibovich score was evaluated using surrogate measures of discrimination and calibration. Discrimination was assessed graphically by observing the degree of separation between the Kaplan-Meier survival curves for intermediate and high-risk groups comparing the derivation and validation cohorts. Discrimination was quantified in each cohort according to Harrell's c-index.

Discrimination was also quantified by the hazard ratio (HR) between the two risk groups in each cohort, with intermediate-risk as the baseline category.

Calibration measures the extent to which predicted and observed outcomes align. Kaplan-Meier survival curves were graphically compared. Good calibration was inferred if the curves for risk groups in the derivation and validation cohorts were aligned. Calibration was also quantified by the HR of the indicator variable for the two cohorts (0 = derivation cohort, 1= validation cohort) separately for each of the two risk groups (0 = intermediate risk, 1 = high risk). A HR near 1 suggests accurate calibration.

Going further I was interested in assessing the ability of the Leibovich score to discriminate patients with individual risk scores. To do this, the analysis was conducted using ungrouped data, (Leibovich scores 3 to 11), within the two datasets. Firstly, Kaplan-Meier plots of MFS split by Leibovich score in both the derivation and validation cohorts, were constructed to further delineate the effect of consecutive increases in Leibovich score on survival. Then HRs and c-index values for individual scores in the validation dataset were compared to the corresponding HR in the derivation dataset. Leibovich score 3 was the reference category in this analysis. Cox models were fitted with each individual score as the explanatory variable and the results were graphed.

Components of the 2002 TNM staging were prospectively collected in both the Leibovich and the SORCE dataset. Given that several recent trials have reverted to using the TNM system to risk stratify participants into trial arms, I compared the discriminatory accuracy of the 2002 TNM staging system with that of the 2003 Leiboivch Score using the two matched (derivation and validation) datasets. To do this the ungrouped score data was used as a single entity to compare the discrimination (c-index) of the 2003 Leibovich score with that of the 2002 TNM system.

The secondary, exploratory analyses were conducted with the three SORCE sub-populations using the same procedures for evaluating discrimination and calibration as with the primary analysis.

All measures were reported with 95% confidence intervals (CIs) and p-values were two sided.

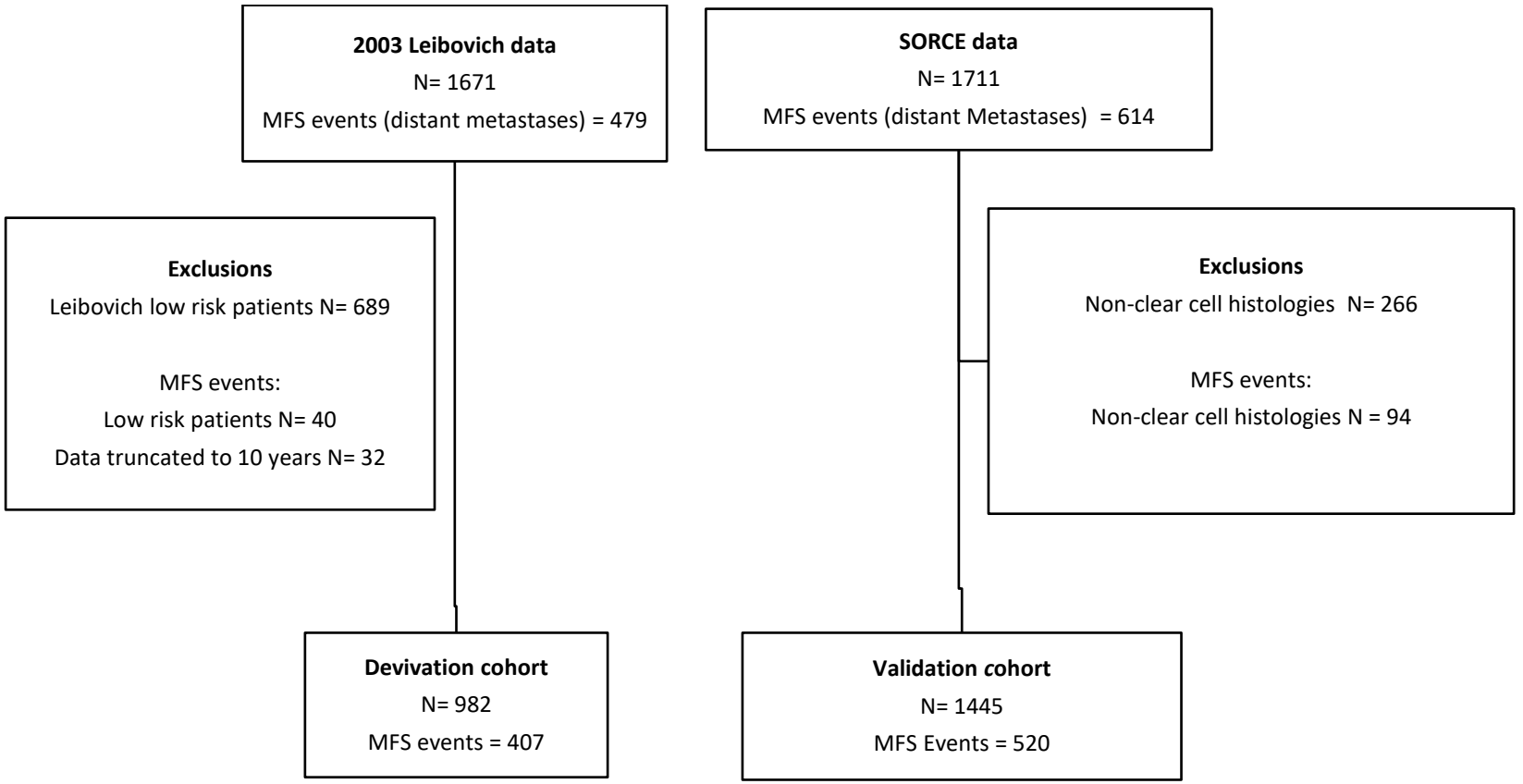
2.3 Results

An external validation the Leibovich 2003 score using contemporary data from the SORCE trial was carried out by comparing its performance to a matched dataset derived from the original 2003 Leibovich dataset.

2.3.1 Overview

The cohort selection is shown in **Figure 8**. The 2003 Leibovich data included 479 MFS events in 1671 US-based patients who had radical nephrectomies between 1970 and 2000. Of this, 689 patients were excluded from the primary analysis because they were of low Leibovich risk. The SORCE data recorded 614 MFS events in 1711 patients enrolled between 2007 and 2013. Of this 266 patients were excluded from the primary analysis because they were of non-clear cell histology. The derivation cohort included 407 MFS events in 982 patients with median follow-up 7.3 years (interquartile range IQR 3-10) while the validation cohort included 520 MFS events in 1445 patients with median follow-up 7.2 years (IQR 6.1, 8.4) (**Table 12**).

Figure 8 The Primary Analysis Cohorts



MFS Metastasis-Free Survival: time from nephrectomy to the date of distant metastases, deaths preceding metastasis were censored.

Table 11 and **12** describe the clinical, demographic and histological characteristics of patients in the original Leibovich data, the Leibovich derivation cohort, the original SORCE cohort and the SORCE validation cohort respectively. The original Leibovich data consisted solely of ccRCC histologies. Non-clear histological subtypes in the SORCE data included; chRCC (6%), pRCC (7%), collecting duct (<1%) and others (2%); sarcomatoid, translocation, mixed and unclassified RCCs. The non-clear cell histological subtypes were evaluated separately in the secondary analysis.

Comparing the patient characteristics in the derivation to the validation datasets, the numbers of male (71% vs 66%) to female (29% vs 34%) patients and those with a history of smoking (50% vs 55%) did not exhibit much variation (**Table 11**).

The validation cohort included more high-risk patients (46% vs 38%). It included 652 (47%) patients who had a laparoscopic nephrectomy and 43 (3%) patients who had a partial nephrectomy, whereas all patients in the Leibovich cohort underwent radical open nephrectomy (**Table 11**). The median time to MFS in the validation cohort was not reached within 10 years of follow-up compared to a median time to MFS of 9.2 years in the derivation cohort.

Table 11 Patient Characteristics in Leibovich and SORCE data.

Continuous data are presented as Mean/ (Standard deviation (SD)) while categorical data as N/ (%)

Variable at Baseline	Original Leibovich Data (N= 1671)	Leibovich Derivation Cohort (N= 982)	Original SORCE Data (N= 1711)	SORCE Validation Cohort (N=1445)
Demographic Characteristics				
Age at Randomisation (years)	-		58·22 (10·6)	58.5 (10.3)
Age at nephrectomy (years)	63.2 (11.3)	63.3 (11.0)	58.1 (10.7)	58.3 (10.3)
Gender				
Female	610 (37%)	337 (34%)	495 (29%)	426(29%)
Male	1061 (63%)	645 (66%)	1216 (71%)	1019 (71%)
History of Smoking				
No	744 (45%)	438 (45%)	675 (39%)[6]	577 (40%)
Yes	927 (55%)	544 (55%)	866 (51%)	726 (50%)
Missing	-	-	170 (10%)	142 (10%)
Type of Nephrectomy				
Open	-	-		
No			716 (42%)	606 (42%)
Yes			908 (53%)	761 (53%)
Missing			87 (5%)	78 (5%)
Laparoscopic	-	-		
No			852 (50%)	717 (50%)
Yes			774 (45%)	652 (45%)
Missing			85 (5%)	76 (5%)
Radical nephrectomy	1671 (100%)	982 (100%)	1591 (93%)	1349 (94%)
Partial nephrectomy	-	-	59 (3%)	43 (3%)
Missing	-	-	61 (4%)	53 (3%)
Events				
MFS Events				
Distant metastases	479 (29%)	436 (44%)	614 (36%)	520 (36%)
Death due to RCC	-	-	21 (1%)	18 (1%)
Ungrouped Leibovich Scores				
Leibovich Score				
0-1	368 (22%)	-	-	-
2	321 (19%)	-	-	-

[6] History of smoking in SORCE is taken from the following question on the case report form: Has patient ever smoked regularly (most days for at least 6 months)?

3	162 (19%)	162 (17%)	151 (9%)	130 (9%)
4	246 (15%)	246 (25%)	347 (20%)	303 (21%)
5	200 (12%)	200 (20%)	412 (24%)	343 (24%)
6	182 (11%)	182 (18%)	369 (22%)	320 (22%)
7	93 (6%)	93 (9%)	201 (12%)	163 (11%)
8	48 (3%)	48 (5%)	131 (8%)	105 (7%)
9	35 (2%)	35 (4%)	82 (5%)	69 (5%)
10	8 (0.5%)	8 (1%)	8 (<1%)	4 (<1%)
11	8 (0.5%)	8 (1%)	10 (1%)	8 (<1%)

MFS; metastasis-free survival

Table 12 Histopathological Characteristics in Leibovich and SORCE data

Categorical data are presented as numbers / (%)

Variable at Baseline	2003 Leibovich data (N= 1671)	Derivation cohort (N= 982)	SORCE data (N= 1711)	Validation cohort (N=1445)
Histological Characteristics				
Histology				
Clear Cell	1671 (100%)	982 (100%)	1445 (84%)	1445 (100%)
Papillary	-	-	128 (7%)	-
Chromophobe	-	-	96 (6%)	-
Collecting Duct	-	-	4 (<1%)	-
Other	-	-	38 (2%)	-
Other histologies				
Mixed	-	-	8 (21%)	-
Sarcomatoid	-	-	23 (61%)	-
Unclassified	-	-	5 (13%)	-
Translocation	-	-	2 (5%)	-
Pathological T category				
pT1a	384 (23%)	6 (<1%)	7 (<1%)	5 (1%)
pT1b	440 (26%)	129 (13%)	197 (12%)	170 (12%)
pT2	335 (20%)	335 (34%)	400 (23%)	298 (20%)
pT3a-4	512 (31%)	512 (52%)	1107 (65%)	972 (67%)
Regional lymph node status				
pNx/ pN0	1605 (96%)	916 (93%)	1637 (96%)	1405 (97%)
pN1/ pN2	66 (4%)	66 (7%)	74 (4%)	40 (3%)
Tumour size				
<10 cm	1312 (79%)	623 (63%)	1152 (67%)	996 (69%)
≥10 cm	359 (21%)	359 (37%)	559 (33%)	452 (31%)
***Nuclear grade**				
1	182 (11%)	47 (5%)	89 (5%)	68 (5%)
2	786 (47%)	284 (29%)	440 (26%)	374 (26%)
3	600 (36%)	548 (56%)	859 (50%)	735 (51%)
4	103 (6%)	103 (10%)	322 (19%)	268 (18%)
****Histological tumour necrosis**				
No	1232 (74%)	561 (57%)	774 (45%)	671 (46%)
Yes	439 (26%)	421 (43%)	937 (55%)	774 (54%)
Leibovich Score Groups in the Leibovich and SORCE data				
Leibovich score group				
Low risk	689 (41%)	-	-	-

Intermediate risk	608 (36%)	608 (62%)	910 (53%)	776 (54%)
High risk	374 (22%)	374 (38%)	801 (47%)	669 (46%)
Median follow-up				
Years (IQR)	7.5 (3.2,10)	7.3 (3.0,10)	7.3 (6.1, 8.4)	7.2 (6.1, 8.4)

* The Leibovich and SORCE grading system used simplifies and selects the worst ISUP features at each grade. For details of grading components, see supplementary material, Table D.
** For the definition of histological tumour necrosis outlined in SORCE trial protocol, see supplementary material; Figure 1

2.3.2 Primary Analysis

Discrimination

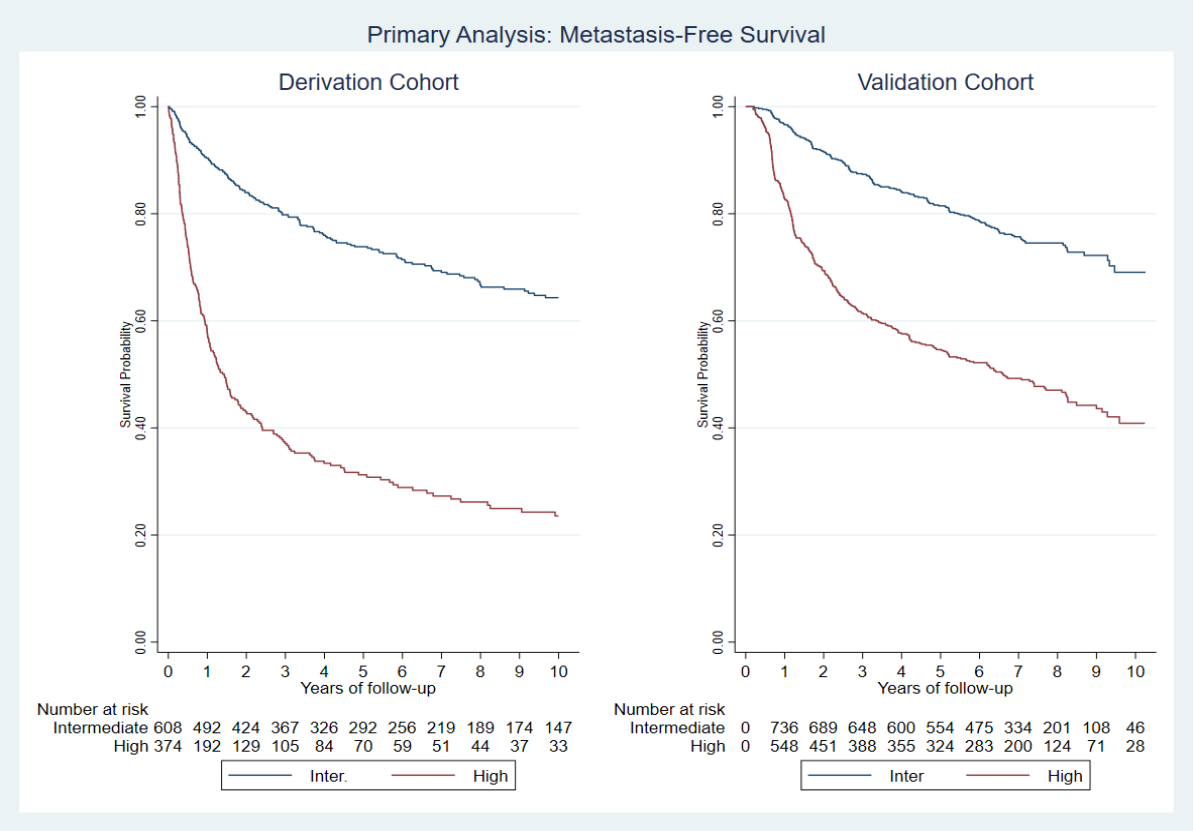
Discrimination between intermediate and high-risk groups in both the validation and derivation cohorts is shown graphically by Kaplan-Meier curves of MFS for each cohort (**Figure 9**). By considering the extent of the separation of the curves, the ability of the Leibovich score to discriminate between risk groups is shown to be substantial but not entirely maintained in the validation cohort. The c-index quantified the discrimination between risk groups in each cohort. When interpreting the c-index it is useful to consider that a value of 0.5 represents a performance no better than chance and 1 a perfectly discriminating score. The c-index in the derivation cohort was 0.67 (95% CI, 0.65 to 0.69) compared to 0.63 (95% CI, 0.61 to 0.65) in the validation cohort ($p = 0.01$, Chi squared test). The c-indexes derived indicate that the 2003 Leibovich score achieves good discrimination in both datasets. Calculation of HRs with intermediate risk as the baseline category is another way of quantifying and comparing discrimination between high and intermediate-risk groups. In the derivation cohort a HR of 3.88 (95% CI, 3.18 to 4.74), compared with 2.74 (95% CI, 2.29 to 3.28) in the validation cohort. Overall, the results show that discrimination is maintained in the validation cohort, albeit reduced compared with that achieved in the derivation cohort.

Calibration

The degree of calibration between the derivation and validation cohorts can be evaluated through observing the alignment between the survival curves for intermediate and high-risk groups across the two cohorts in **Figure 9.**

Overall, the MFS rate was 26% lower in the validation than the derivation cohort shown by a HR of 0.74, (95% CI 0.65 to 0.85). For the intermediate risk group, the reduction in MFS rate was 24% (HR = 0.76, 95% CI 0.61 to 0.94), compared with 46% (HR = 0.54, 95% CI 0.45 to 0.64) in the high-risk group. The results confirm that calibration between the two datasets is poor which is more marked between the high-risk groups. Five-year MFS rates are favourable in the validation cohort compared with the derivation cohort, for both risk groups. The five-year MFS was 30% (95% CI, 25-35) in the Leibovich high risk cohort versus 52% (48-56) in the SORCE cohort, and in the intermediate groups; 72% (67-75) in the Leibovich cohort and 78% (75-81) in the SORCE cohort.

Figure 9 Kaplan-Meier curves for MFS by Leibovich risk group in the derivation and validation cohorts



NB: In the Validation cohort Kaplan-Meier plot, the number of patients entering at time 0 is given as 0 in the at-risk tables. It is a consequence of the late entry character of the follow-up data. Patients were not deemed at risk until they were randomised into SORCE, which occurs after t = 0.

2.3.3 Analysis of ungrouped Leibovich score

In order to assess discrimination in more detail, the following analysis was conducted on the ungrouped derivation and validation datasets. **Figure 10** shows the HRs, comparing individual scores in the derivation and validation datasets with the reference category (Leibovich score 3). **Table 13** presents a table of the HR values for each score in both cohorts. They show that the HRs increase markedly as the score increases in both the derivation and validation cohorts. Of note, the estimated HRs are imprecise in categories 10 and 11 due to few patients scoring 10 or 11 therefore scores 9-11 have been combined in **Figure 10**. The distribution of the c-index values for each Leibovich score in the two cohorts is shown in **Figure 11**. Similarly to the pattern for HR values, c-indexes are shown to increase with consecutive increases in Leibovich score reflecting better discrimination with higher scores. Comparing the HRs at each score in the validation cohort compared to the corresponding deviation cohort score supports the earlier conclusion that discrimination is maintained albeit attenuated in the contemporary dataset. Recalling that the HR comparing intermediate and high-risk groups was 3.88 in the derivation cohort and 2.69 in the validation cohort, it is possible to conclude that the collapse of scores 3-5 and 6-11 into two larger prognostic groups (intermediate and high-risk groups) results in the loss of additional discrimination achieved particularly by higher individual Leibovich scores. This finding applies to both the derivation and validation datasets. It is a compromise necessary to achieve a more clinically applicable risk stratification tool.

Figure 10 HRs estimated from a Cox model for ungrouped Leibovich scores in the derivation dataset and in the validation dataset

Values are presented with 95% confidence intervals. The lowest score (3) in the validation dataset is the reference category.

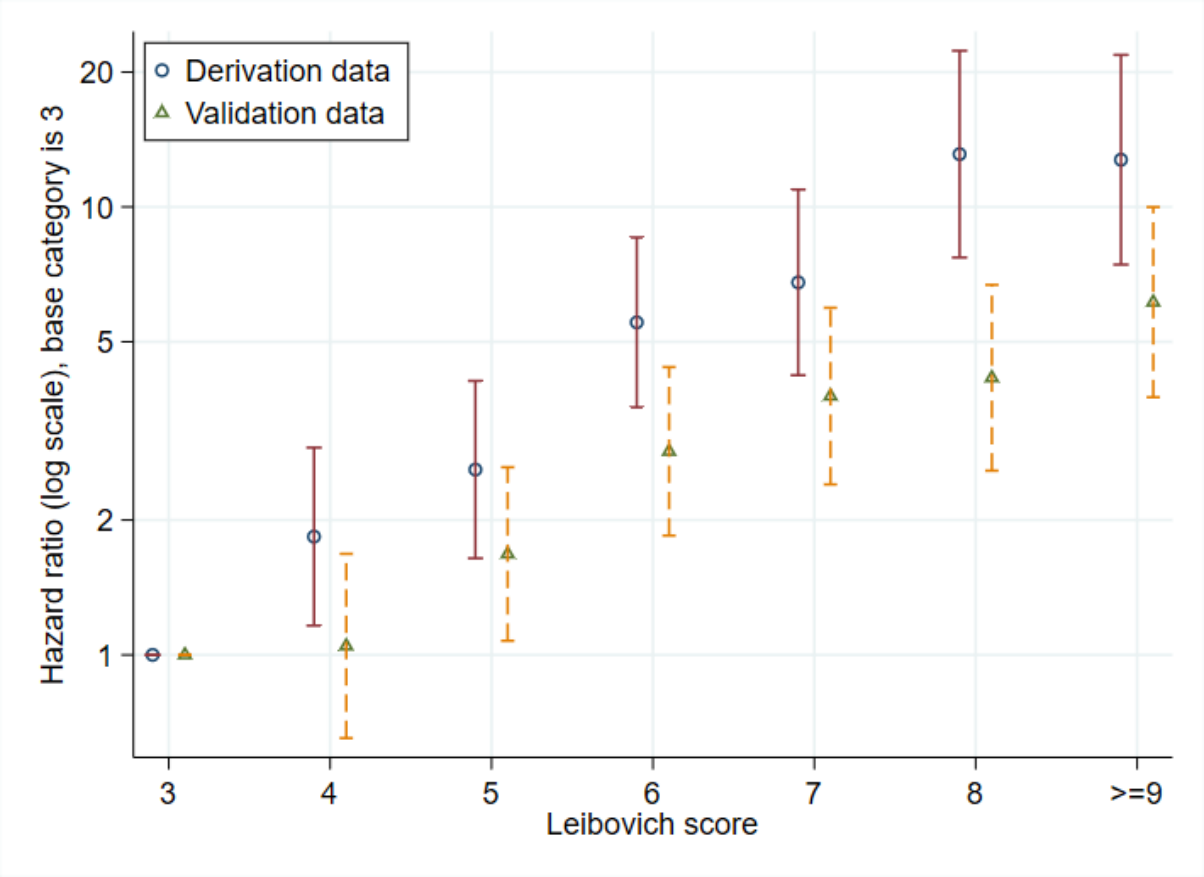


Table 13 HRs estimated from a Cox model in the derivation dataset and another in the validation dataset

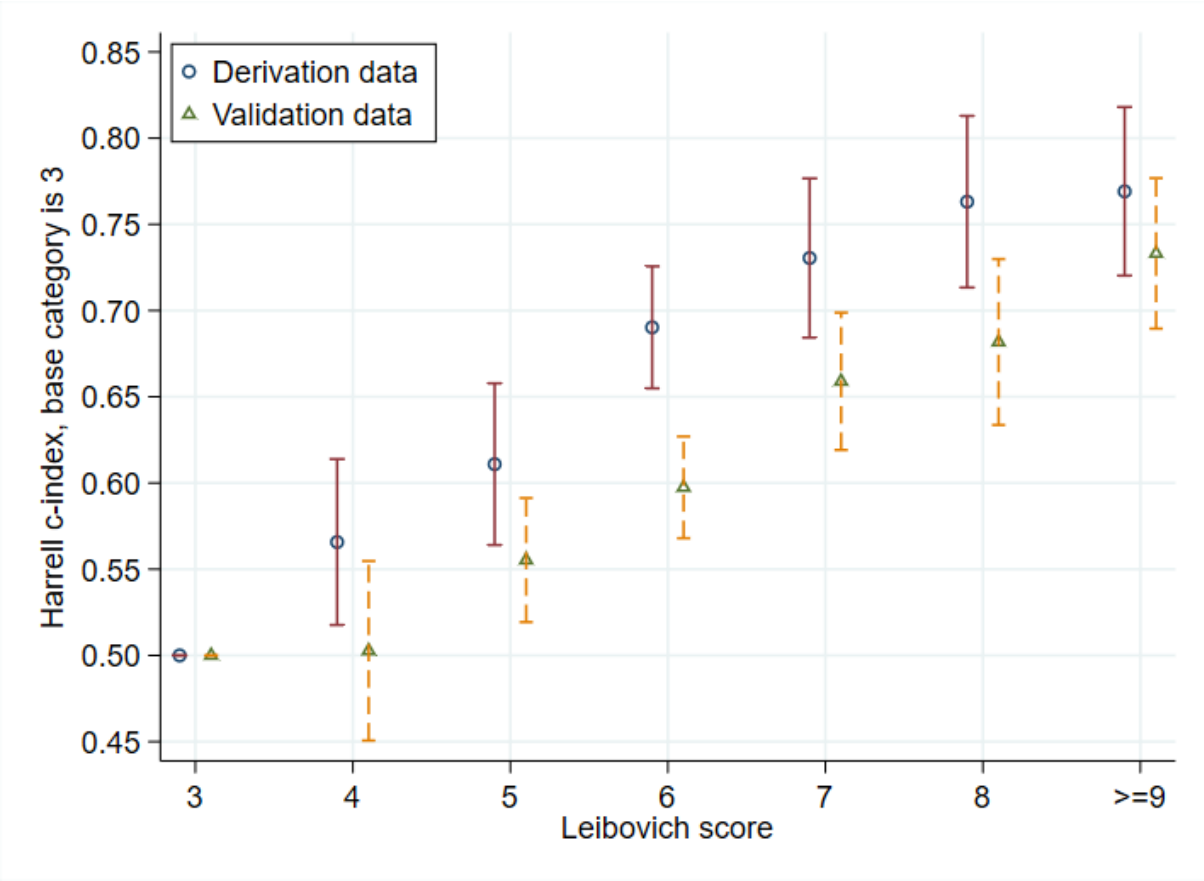
Values are presented for individual ungrouped Leibovich scores. The lowest score (3) in the validation dataset is the reference category.

Score	No. derivation patient/events	Derivation HR	No. validation patients/events	Validation HR	95% CI
0	318/10	0.12	-	-	-
1	50/2	0.29	-	-	-
2	321/31	0.41	-	-	-
3	162/32	(1.00)	151/25	(1.00)	-
4	246/76	1.84	348/68	1.17	0.33-1.21
5	200/79	2.58	411/117	1.85	0.38-1.34
6	182/112	5.62	370/164	3.22	0.31-1.05
7	93/59	6.93	201/103	4.14	0.31-1.14
8	48/41	13.11	132/75	4.81	0.19-0.72
9	35/27	16.62	81/48	6.20	0.18-0.78
10	8/3	5.27	8/7	15.57	0.68-12.76
11	8/7	29.15	9/7	11.86	0.12-1.34

HR; hazard ratio, CI; confidence interval

Figure 11 Distribution of the c-index for each Leibovich score in the two cohorts

The lowest score (3) in the validation dataset is the reference category



2.3.4 Secondary Analyses

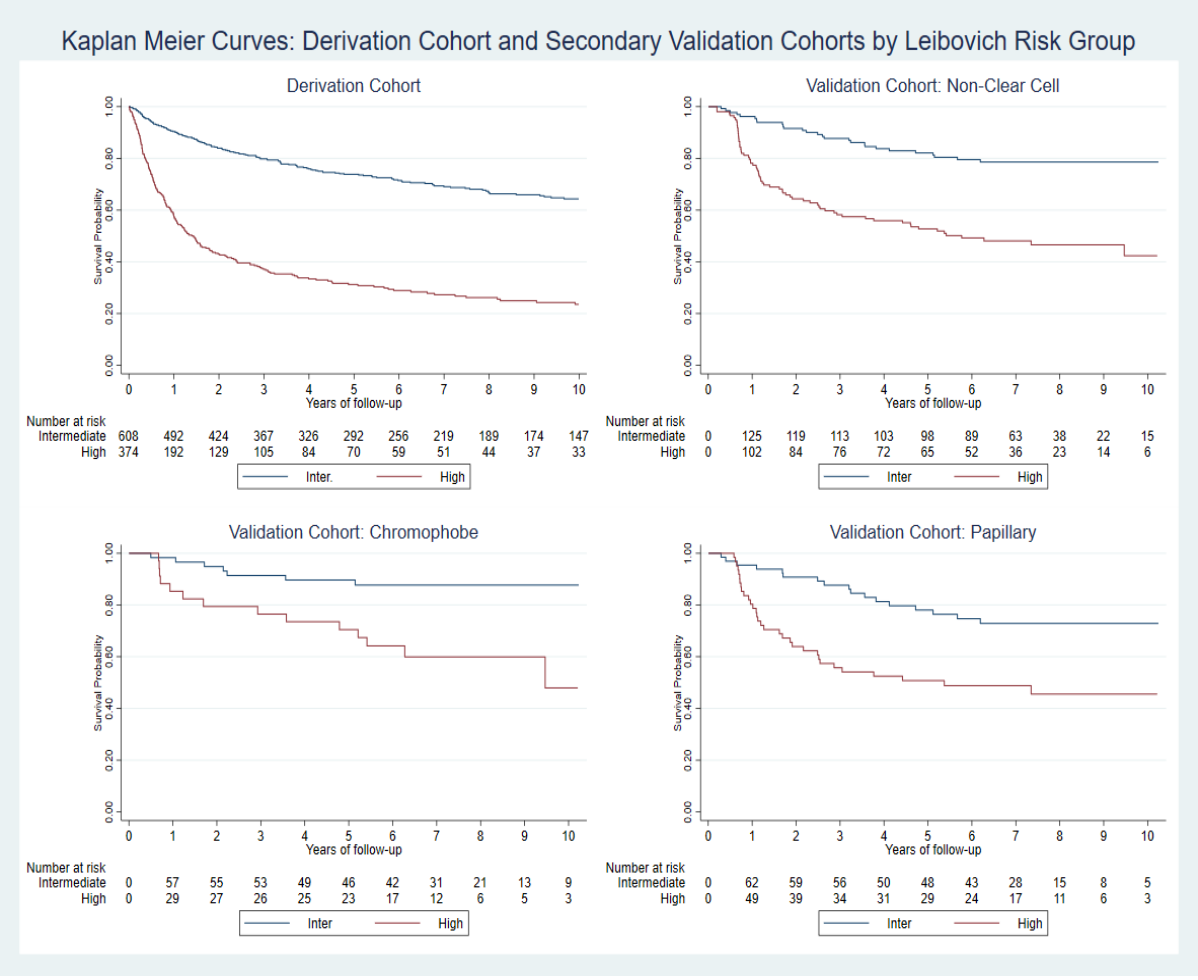
Comparison of discrimination by 2002 TNM and the 2003 Leibovich score

To compare the discrimination of the ungrouped Leibovich score with that of 2002 TNM staging, c-indexes were calculated using the primary analysis datasets**.** The 2003 Leibovich score was shown to improve upon discrimination by 2002 TNM staging in the derivation cohort (Leibovich c-index of 0.72 (standard error SE 0.01) vs 2002 TNM c-index of 0.56 (SE 0.01). This improvement is retained in the validation cohort (c-indexes of 0.67 (SE 0.01) vs 0.56 (SE 0.01)).

Discrimination between risk groups in SORCE sub-populations

To evaluate the ability of the 2003 Leibovich score to discriminate between intermediate and high-risk groups for patients with non-ccRCCs, three SORCE subpopulations; any non-ccRCC (N=266, MFS events 94), pRCC (N=128, MFS events 49) and chRCC (N=96, MFS events= 21). Discrimination between intermediate and high-risk groups within each SORCE sub-cohort was compared to that of the derivation cohort.

Figure 12 Kaplan-Meier Curves for MFS in the derivation cohort compared to the SORCE Non-Clear Cell, Chromophobe and Papillary cohorts stratified by Leibovich risk group



NB: In the Validation cohort Kaplan-Meier plot, the number of patients entering at time 0 is given as 0 in the at-risk tables. It is a consequence of the late entry character of the follow-up data. Patients were not deemed at risk until they were randomised into SORCE, which occurs after t = 0.

Discrimination between intermediate and high-risk groups in each of the SORCE sub-groups is shown graphically by Kaplan-Meier curves of MFS for each cohort. **Figure 12** shows clear separation between the risk group Kaplan-Meier curves for the three SORCE sub-cohorts beyond six months, which indicates that the 2003 Leibovich score retains long-term discriminative capability specifically in these populations. Compared to the derivation cohort c-index of 0.67, values of 0.64 (95% CI, 0.59 to 0.69) for the SORCE non-ccRCC cohort, 0.63 (95% CI, 0.56 to 0.69) for SORCE pRCC, and 0.65 (95% CI, 0.55 to 0.76) for the SORCE chRCC group were found.

Evaluating the HR between risk groups a value of 3.88 (95% CI, 3.18 to 4.74) is observed in the derivation cohort, compared with 3.06 (CI, 1.95 to 4.79) for SORCE patients with non-ccRCCs, 2.51 (95% CI, 1.38 to 4.56) for pRCC, and 3.56 (95% CI, 1.44 to 8.83) for the chRCC cohort.

2.3.5 Calibration between outcomes in the derivation cohort and SORCE sub-cohorts

The degree of calibration between outcomes in the derivation cohort and SORCE sub-populations can be observed graphically in **figure 12** by comparing the Kaplan-Meier curves across the cohorts. Attenuated calibration is shown for all SORCE subpopulations compared to the derivation cohort by observing the misalignment of the survival curves corresponding to each risk group. HRs for MFS after fitting a Cox regression model to each risk group separately (**Table 14**) show that MFS outcomes in the intermediate (HR 0.36 CI, 0.23-0.69) and high-risk (HR 0.40 CI, 0.23-0.69) SORCE chromophobe patients are particularly poorly calibrated to MFS in the derivation cohort. Whereas MFS outcomes of the of the intermediate-risk pRCC cohort (HR 0.84 (95% CI, 0.51-1.38) are more aligned to those of the derivation cohort. Five-year relapse probabilities (**Table 15)** show that MFS is higher in all SORCE subgroups compared to the corresponding deviation cohort. The difference is most marked between the high-risk groups.

Table 14 HRs for MFS for SORCE sub-populations compared to the derivation cohort in intermediate risk and high-risk groups

Presented with 95% CI.s

	SORCE Non-Clear Cell	**95% CI.**	**SORCE Papillary**	**95% CI**	**SORCE Chromophobe**	**95% CI**
Number of patients	266		128		96	
MFS Events	94		49		21	
FUP time (Median years, (IQR))	7.3 (6.1, 8.3)		7.3 (6.1, 8.4)		7.3 (6.3, 8.3)	
Intermediate risk HRs	**0.65**	0.43-0.97	**0.84**	0.51-1.38	**0.36**	0.17-0.77
High risk HRs	**0.57**	0.44-0.76	**0.59**	0.40-0.85	**0.40**	0.23-0.69

CI; confidence interval, MFS; metastasis-free survival. FUP; follow up, IQR; interquartile range. HR; hazard ratio

Table 15 Five Year Survival probabilities for MFS in the derivation cohort, the validation cohort and SORCE sub-cohorts

Presented with 95% CI.s

	SORCE Validation cohort (clear-cell)	SORCE Non-Clear Cell	SORCE Papillary	SORCE Chromophobe
Derivation, Intermediate risk	72% (67 - 75)			
Intermediate Risk	78% (75-81)	79% (71 - 85)	75% (62 - 84)	87% (75 - 94)
Derivation, High Risk	30% (25 - 35)			
High Risk	52% (48-56)	50% (41 - 58)	49% (36 - 60)	64.9% (45 - 78)

2.4 Discussion

The main clinical uses of RCC specific scoring systems in the post-resection setting are two-fold. Firstly, to indicate to patients and clinicians what the risk of future relapse is and secondly to select patients suitable for inclusion to adjuvant trials or for standard adjuvant therapy if this becomes available. Clinical trialists need to be confident that the prognostic score chosen, can reliably select patients with a high enough risk of relapse to warrant exposure to potentially toxic adjuvant treatments. The 2003 Leibovich score was chose by the SORCE and RAMPART TMGs to select and randomise patients with intermediate and high-risk into trial arms.

This analysis shows that the 2003 Leibovich score retains the ability to discriminate between intermediate and high risk in the SORCE validation cohort (c-index 0.63, 95% CI 0.61 to 0.65), albeit marginally reduced when compared to its performance in the derivation dataset (c-index 0.67, 95% CI 0.65 to 0.69). The AJCC have published a sixteen-item tool with which to consider the approval of a risk model including nine specific considerations for its optimal validation. The design and conduct of this validation adhere to all nine of these, and as such, represents the highest quality of validation exercise according to their guidelines [70]. Although developed almost two decades ago, it is possible to support the use of the 2003 Leibovich score, for discriminating risk between patients with intermediate and high-risk RCC, a group of patients of specific interest for recruitment to adjuvant trial recruitment.

The second aim was to compare discrimination achieved by the 2003 Leibovich score to that of 2002 TNM staging. This analysis shows that discrimination by Leibovich exceeds that of 2002 TNM in both the derivation cohort (c-indexes of 0.72 (SE 0.01) vs 0.56 (SE 0.01)) and in the validation cohort (c-indexes of 0.67 (SE 0.01) vs 0.56 (SE 0.01)). The improvement is noteworthy when considering that a c-index of 0.5 represents selection no better than that offered by a coin-flip. Correa *et al.* interrogated the ASSURE trial dataset to evaluate whether the 2003 Leibovich score improves upon discrimination by 2002 TNM. They showed a Leibovich score c-index of 0.625 compared to a c-index of 0.60 for 2002 TNM. Therefore, results from both studies support the discriminative superiority of the 2003 Leibovich score. Correa *et al.* go on to compare the discriminative capability of 2002 TNM (c-index 0.602) to the 1997 TNM (c-index 0.603), and 2010 TNM (c-index 0.605) systems [64, 82]. They show that successive adaptations of TNM do not vastly improve upon discrimination achieved.

Overall, given the inferior discrimination exhibited by TNM 2002 in this analysis, it is possible to challenge the recent use of TNM staging for patient randomisation by some of the recent immune-oncology focussed adjuvant RCC trials.

The third aim was to evaluate discrimination and calibration of the 2003 Leibovich score in three important SORCE sub-populations, participants with any non-ccRCC, pRCC and chRCC RCCs. The 2003 Leibovich score was shown to discriminate comparably between intermediate and high-risk in the non-ccRCC SORCE group (c-index 0.64, 95% CI 0.59 to 0.69). Acknowledging the inherent variability in clinical trajectories within the non-ccRCC subgroup, the two largest non-ccRCC subtypes were explored individually. The modelling exercise resulted in c-indexes of 0.63 (95% CI 0.56 to 0.69) for participants with pRCC and 0.65 (95% CI 0.55 to 0.76) for those with chRCC. Although the latter analyses are limited by smaller patient numbers, and correspondingly larger statistical imprecision, little reduction in discrimination compared with the derivation cohort is shown. Therefore, this analysis provides support for the use of the 2003 Leibovich score to stratify patients with non-clear cell and clear-cell subtypes in adjuvant RCC trials.

In 2018, Leibovich and colleagues published separate scoring systems for patients with ccRCC, pRCC and chRCC carcinomas [72]. However, like other subtype specific scores, its use has yet to be recommended in clinical guidelines and there are no adjuvant trials that use this updated score to risk stratify participant. A trade-off for histological specificity is added complexity in terms of the number of scoring systems for different sub-types and models that comprise many more components for ccRCC (nine components for PFS and twelve for CSS) in the 2018 score compared to the 2003 score. In addition, internally validated c-indexes for PFS and CSS for the 2018 ccRCC score are 0.83 and 0.86 respectively, which represents only marginal improvement compared to that published by Leibovich and colleagues in 2003 (c-index of 0.82 for MFS) [23]. The challenge is in judging the practical clinical importance of a small difference in c-index and in assessing whether the gain in discriminative accuracy provided by the 2018 score justifies the added complexity in its calculation. Arguably, a single, easy to derive score which continues to discriminate between risk groups in contemporary data in a multi-subtype population, is from a clinical trial standpoint, more practical and readily standardisable across trial sites. Overall, this analysis provides support for the use of the 2003 Leibovich score, for risk stratification

in adjuvant RCC trials over and above recently published, yet to be widely externally validated, subtype-specific scores [72, 83].

A robust calibration analysis was performed matching risk groups from the original Leibovich study, using IPD and unifying MFS definitions across cohorts. It clearly demonstrated longer MFS in contemporary intermediate and high-risk clear cell RCC patients (5-year MFS; 78% (CI, 75% -81%) and 52% (CI, 48% -56%) respectively), compared with the historical data; (5-year MFS; 72% (CI, 67%-75%) and 30% (CI, 25%-35%). Of note, there are some differences in patient and tumour characteristics between the two cohorts. The median age of validation patients was five years younger and included higher rates of T3a-4 tumours (67% versus 52%), higher rates of histological tumour necrosis (18% vs 10%) and higher numbers of nuclear grade 4 (54% vs 43%).

Improved outcomes in patients with seemingly worse prognostic features, may reflect an evolution in renal tumour biology over time that has shifted the significance of histopathological features, classically linked to prognosis. This may in part be driven by changing rates of modifiable risk factors such as obesity, smoking and hypertension or as yet unidentified novel carcinogens. It may also be confounded by 'stage migration' a phenomenon that may be linked artificially inflated cancer survival rates associated with advances in diagnostic techniques, improved processing of surgical specimens or with evolving methods of cancer staging. In the case of RCC, the introduction of minimally invasive techniques such as laparoscopic nephrectomy that improve on kidney and vascular mobilisation, mean that tumours with poor prognostic features can now be removed with radical surgery, and therefore this may also contribute to improved outcomes [91]. Interestingly, improved survival over time is seen in other tumour types. For example in locally advanced oesophageal cancer, the OE05 trial [92] showed an 8% improvement in median overall survival when compared to a similar cohort of patients receiving exactly the same chemotherapy ten years previously in OE02 [93].

Overall, data from SORCE has provide detailed, IPD with long-term follow-up, and therefore provides a high quality substrate for this prognostic score validation. It is possible to recommend the use of phase three trial data for future validation exercises particularly when an assessment of discrimination is required. However, a key point to consider is that patients enrolled into clinical trials often have better outcomes than

those from real-world data, at least in part owing to strict trial eligibility criteria that selects narrow patient groups including those with optimal baseline performance status. Therefore, caution must be applied when assessing the extent of calibration deficits between trial data and non-trial cohorts. Importantly, though, the improvement in contemporary outcomes for patients with locally advanced RCC that are shown in this analysis, are corroborated in findings from several other non-trial contemporary dataset [61, 64, 94].

This analysis has some limitations. Firstly, SORCE pathology samples used for this analysis were not centrally reviewed, although clear guidance was provided in the SORCE trial protocol for their review by local sites. Secondly, patients with low Leibovich risk (score 0-2) were not included in this analysis, because they are usually cured by surgery or ablation and not usually considered for recruitment to adjuvant trials. It is acknowledged that excluding the low-risk group is likely to have resulted in loss of some discrimination compared to that possible by the complete Leibovich data, (c-index for the derivation cohorts is 0.67 compared to its original publication of 0.82). However, the purpose of this study was to evaluate the Leibovich score in a clinically relevant population, therefore datasets including intermediate and high-risk participants only, were compared. Thirdly, this validation could be criticised for using data from the whole SORCE cohort rather than restricting to the placebo group. This was duly considered during the statistical design stage. As SORCE showed a definite lack of benefit of sorafenib after nephrectomy, I decided in favour of retaining as much data as possible by including patients from all experimental arms, as this would likely have no detrimental impact.

2.5 Conclusion

At present the 2003 Leibovich Score remains the optimum staging system for use in enrolling patients to the current generation of adjuvant trials in RCC. However it has been demonstrated that MFS rates among patients have improved over time, which means that clinico-pathological prognostic scores developed years ago, need to be regularly reviewed and caution should be applied when using them for prognosticating and for counselling patients after nephrectomy. With the accumulation of data from recent adjuvant RCC trials, there is an opportunity to improve upon the 2003 Leibovich score to better reflect the current landscape of RCC (Chapter 5). Until the development

of molecularly-based prognostic tools that markedly improve predictive accuracy, the 2003 score should remain the clinical standard.

Findings from this chapter were published in the Journal of Clinical Oncology on February 25th 2022 [3]. A reply to comments by Capitano *et al.* was published in the Journal of Clinical Oncology in October 2022 [4].

Chapter 3: Clinical outcomes in non-clear cell RCC, a pooled analysis from SORCE and ASSURE trials

3.1 Introduction

3.1.1 Overview

Over the past two decades, the treatment of metastatic renal cell carcinoma (RCC) has evolved from a non-specific immunostimulatory approach (cytokines, interferon alpha or interleukin) to targeted agents against vascular endothelial growth factor (VEGF) and now immune checkpoint inhibitors (ICIs)). Most landmark clinical trials included cohorts of patients predominantly with clear cell RCC (ccRCC). In the real world however, RCC comprises a heterogeneous group of conditions exhibiting a range of clinical behaviours and prognoses.

The 2022 World Health Organization (WHO) classification of renal cell malignant tumours [95] recognises increasingly more histological and molecular classifications. The most common and most studied variant is ccRCC which accounts for 70-75% of RCCs. Non-clear cell renal cell carcinomas (non-ccRCC) account for 15-25% of primary renal malignancies of which papillary (pRCC) and chromophobe (chRCC) subtypes comprise 80%. Clear cell RCC, pRCC and chRCC comprise the 'major' subtypes. RCCs with sarcomatoid features (sRCC) account for ~5% of RCCs. Although not technically a histological subtype, histology comprising sRCC is a clinically important distinction carrying a poor prognosis overall [96]. Other RCCs are classified by anatomical location such as medullary RCC and collecting duct RCC. Some classifications of RCC associate with renal disease, for example acquired cystic disease-associated RCC. Other RCCs are classified according to pathognomonic molecular alterations (e.g., microphthalmia transcription factor, family translocation RCC and succinate dehydrogenase-deficient RCC, and familial predisposition (e.g. hereditary leiomyomatosis and RCC-associated RCC) [95].

The move towards subdividing RCC is in line with the current momentum behind precision medicine. Precision, or "personalized" cancer medicine, as defined by the National Cancer Institute, is the use of "specific information about a person's tumour to help diagnose, plan and find out how well treatment is working, or make a prognosis". In RCC however, predominance of small retrospective non-randomised data for the non-ccRCC histological and molecular subtypes provides an imprecise

and often conflicting understanding of their clinical trajectories. In the adjuvant setting, subtype specific surgical strategies, their post nephrectomy follow-up, prognostication and treatment lacks consensus or a high-level evidence base.

SORCE (NCT00492258) [5] and ASSURE (NCT00326898) [97] were two large international adjuvant tyrosine kinase inhibitor (TKI) trials that included patients with non-clear cell histology's. I examined individual patient data from both trials, to provide a comprehensive comparison of the disease trajectories of patients with fully resected non-ccRCCs to those with ccRCC. I focused on the two commonest subtypes; pRCC and chRCC, and on those with an sRCC component,(as they are associated with poor outcomes). The aims were three-fold. Firstly to depict the precise clinical trajectories of the non-clear cell subtypes from the point of radical resection to first relapse and death, to delineate potential drivers of poor prognostic disease and finally to inform on subtype specific follow-up guidance after nephrectomy.

Table 16: Molecular, immune-histological and genetic characteristics of clear-cell, papillary, chromophobe and sarcomatoid RCCs

	Incidence	Origin	Macroscopic	Cytogenetic Alterations	Molecular Alterations	Microscopic	Immunohistochemistry	Associations
ccRCC[98]	75-85%	Proximal nephron, tubular epithelium	Solid, yellowish colour. Cystic change, haemorrhage, necrosis	95% Sporadic: Loss of function VHL,chr 3p, inappropriate stabilization of HIFs	Genetic mutations in PI3K/AKT, PBRM1, pathway, mutations of SETD2, BAP1, MTOR, MET, AXL, p53 [99] 5% Autosomal dominant	Clear cells +/- eosinophilic granular cytoplasm, particularly if high grade	Pan-cytokeratin and vimentin, CAIX +ve CD10 +ve, CK7 -ve, AMACR -ve	Von Hippel-Lindau (25– 45%), tuberous sclerosis (2%)
pRCC [98]	10-15%	Distal nephron, tubular epithelium	Grey or brown colour Grossly solid, +/- cystic change or encapsulation, Necrosis and haemorrhage.	Sporadic; Gain of Chr 7 and/or Chr 17, loss of Chr Y, Del Chr 9p	Sporadic; MET, SETD2, EGFR, CDKN2A, NF2, TERT, BAP1, NRF2/ARE mutations [100] Autosomal dominant; hereditary pRCC syndromes, germline MET, TCA, FH, NRF2/ARE pathway mutations [100]	Papillary structure, foamy macrophages; Basophilic; single layer of basophilic cells surrounding basal membrane. scanty cytoplasm. Eosinophilic; abundant granular eosinophilic	CD10+, CK7+, AMACR+, CAIX -ve	Acquired cystic kidney disease or hereditary pRCCs
chRCC [98]	5%	Distal nephron, intercalated cells of the distal tubules	Tan in colour, grey after fixation	Loss of Chr 1, 2, 6, 10, 13, and 17,	Somatic mutation in mitochondrial DNA, TP53 and PTEN, mTOR mutations, high TERT expression [101]	Large cells, prominent cell membranes, pale cytoplasm, crinkled 'raisinoid' nuclei, perinuclear halos	KIT and CK7+ve, CAIX, CD10 -ve S100A1, CD82 +ve	Birt–Hogg–Dubé syndrome; FLCN gene mutations on chromosome 17, Cowdens syndrome (PTEN) [101]
sRCC	3% (pure)	Variable	Dense grey, firm fleshy	gains of chr 1, 2, 6, 10, and 17, gains of chr 1q and 8q, losses of 9q, 15q, 18p/q, and 22q [96],	EMT mutations; TNF, TGF-B, Wnt, MAPK, P13K/AKT [102]	spindle cells, fibrous, leiomyomatou, rhabdoid, osteoid, or chondroid transformation	Keratin and PAX +ve	

Key: ccRCC, clear cell renal cell carcinoma, pRCC, papillary renal cell carcinoma, chRCC, chromophobe RCC, sRCC, sarcomatoid RCC, Chr, chromosome; CIMP, CpG island methylator phenotype; HIF, hypoxia-inducible factor; VHL, von Hippel Lindau; EMT; epithelial-mesenchymal transition, all genes listed on page 6.

3.1.2 Clear Cell RCC

Clear cell RCC is the commonest RCC subtype (75-85%) (**Table 16**). 95% of cases occur sporadically, and the remaining 5% are associated with hereditary syndromes (von Hippel-Lindau disease, tuberous sclerosis) [98]. Macroscopically, ccRCCs originate from the epithelium of the proximal convoluted tubules (renal cortex) and commonly extend toward the peri-renal fat, renal sinus fat and into the renal vein and inferior vena cava [103] (**Table 16**). Microscopically, ccRCCs are defined by cytoplasmic clear cells and staining features positive for cytokeratin, vimentin and carbonic anhydrase. Somatic mutations of the von Hippel Lindau (VHL) gene located on chromosome 3p are found in up to 90% of sporadic ccRCC tumours, which explains the relevance of anti-VEGF targeted treatments for these patients. Molecular studies reveal a variety of germline mutations in Von Hippel Lindau (VHL) genes, including substitutions, deletions, methylations and insertions are associated with hereditary ccRCCs [104].

Survival outcomes for patients with locally advanced and metastatic ccRCC have been well described, with a median five year survival of 75-85% reported [105, 106]. In the metastatic setting the median five-year survival drops to approximately 10% [106], although with newer therapies, the overall survival for selected patients has improved [107]. Few studies directly compare survival outcomes between patients with ccRCC and non-ccRCC. One retrospective analysis of Surveillance, Epidemiology and End Results Program (SEER) data included 15,015 patients with metastatic RCC from the United States, diagnosed between 2000 and 2013 [108]. Participants were stratified into those treated with or without TKIs. Patients with ccRCC had a median survival of 10.0 months in the TKI cohort and 8.0 months in the pre-TKI cohort (Hazard ratio (HR) for death 0.86; 95% CI, 0.84–0.91, P<0.0001). For those with non-ccRCC, the median survival was 7.0 months in both cohorts, (HR 0.98; 95% CI, 0.88–1.09, P=0.714). Although limited by the heterogeneity of histology within the non-clear cell group, this study exemplifies the better overall survival (OS) of patients with ccRCCs compared to patients with non-ccRCC and the differential effect of TKIs between the groups.

3.1.3 Papillary RCC

Papillary RCC (pRCC) is the second commonest RCC histological subtype, representing 10-15% of cases [8] (**Table 16**). Similarly to ccRCC it presents typically in the 6th to 8th decade of life and has a similar male predominance with male-to-female ratios reported between 1.5-2.1:1 [109, 110]. Papillary RCC is characterised histologically by papillary cellular patterns and foamy macrophages with areas of internal haemorrhage and cystic alterations, particularly in larger lesions [98]. Papillary RCC generally occurs sporadically but can be associated with familial conditions, for example hereditary pRCC. Papillary RCC has classically been subdivided into two forms; type 1 and type 2 based on histological appearances, genetics and biological behaviour. Current imaging techniques cannot preoperatively differentiate type 1 and 2 pRCC and therefore histological differentiation is necessary [103]. Type 1 pRCCs are characterised by a single layer of basophilic cells surrounding the basal membrane. Type 2 pRCCs are identified by papillae covered by cells with abundant granular eosinophilic cytoplasm and with prominent nucleoli associated with areas of necrosis [98]. Various other non-ccRCC subtypes mimic type 2 pRCC by displaying prominent papillary architecture, resulting in contention from pathologists when considering type 2 pRCC as a distinct entity of its own [111].

In contrast to ccRCC, the role of the loss of chromosome 3p leading to Von Hippel Lindau (VHL)-mediated oncogenesis in pRCC is less well characterised. Case-series indicate a lower expression of VEGF and VEGF micro-RNA in pRCC tumours compared to ccRCCs [112], which goes some way to explaining the mixed effectiveness of VEGF targeted treatments in patients with pRCC (see below for treatment options for non-clear cell RCCs). Type 1 pRCCs have classically been associated with activating germline mutations of the Mesenchymal Epithelial Transition (MET) gene on chromosome 7q31 in more than 80% of cases [113]. MET is a tyrosine kinase receptor of hepatocyte growth factor/scatter factor (HGF/SF) involved in tissue repair and regeneration. Genetic alterations in MET have been associated with tumour invasion, anti-apoptototis, angiogenesis, and metastasis and have led to inhibitors of MET being evaluated in patients with pRCC. Other alterations associated with type 1 pRCC include TERT, CDKN2A/B and EGFR genes. Type 2 pRCC appears to be more genetically

heterogeneous, with MET copy number gain found in up to 50% of cases, and recurrent alterations of SETD2, EGFR, CDKN2A, NF2, TERT, FH, TFE2 fusions and CIMP also being described [100]. Given the significant clinical and genetic heterogeneity within the pRCC group, the type 1 and 2 distinction although still referenced, was removed from the 2022 WHO classification [8]. An updated classification delineating the heterogeneity within the pRCC subgroup is required.

The prognosis of early stage locally advanced pRCC is considered to be favourable compared to ccRCC, with five-year cancer-specific survival (CSS) and OS ranging from 79-83% after nephrectomy [114, 115], largely from cohort studies. Limited data suggests that patients with locally advanced pRCCs exhibiting high risk features may have equal or comparatively worse survival to those with similar risk ccRCCs. In 2018, Leibovich and colleagues examined 607 patients with pRCC and 2726 with ccRCC. They found five-year progression-free survival (PFS) in those deemed to have high-risk locally advanced pRCCs to be 60% (95% CI, 49-69), which was comparable to that PFS of patients with high risk (Leibovich score 8-9) ccRCCs. The five-year OS in those with high risk pRCCs was comparatively worse at 74% (95% CI, 59-84), similar to those with *very* high risk-(score 12) ccRCCs (70% 5-year OS 95% CI, 67-73) [72]. In the metastatic setting, the prognosis for patients with metastatic pRCC is notably worse than those with ccRCC with 7% five-year CSS being reported [116].

Despite the potential for survival outcomes to be worse for patients with high risk pRCCs, the surgical management and surveillance of patients with pRCC currently follows ccRCC guidance. In addition, the European Association of Urology (EAU) recommends the use of TNM stage and Fuhrman grade to stratify patients with pRCC into post nephrectomy surveillance groups [20], although currently there remains no evidence for this approach to risk stratification.

3.1.4 Chromophobe RCC

Chromophobe RCC (chRCC) represents approximately 5% of all malignant renal epithelial tumours (**Table 16**). It tends to carry the best prognosis amongst the major subtypes of RCCs [8, 117]. Histologically, these tumours present with large pale cells with

reticulated cytoplasm and perinuclear halos [98]. In contrast to those with ccRCC the immune profile is positive for CK7, and AMACR and negative for CAIX.

The genetic components of chRCC appear strikingly distinct from that of pRCC and ccRCC. One study comparing the VHL mutation rates in various RCC subtypes found alterations in VHL in only 10% of chRCCs [118]. Chromophobe RCC has instead been associated with recurrent structural breakpoints within TERT gene promoter region which correlates with highly elevated TERT expression [119]. TERT over activation is hypothesised to result in cellular immortalization and malignant transformation by stabilizing telomere length and inhibiting cellular growth arrest [119]. Other cancer sites found to be associated with increased TERT gene copies include breast, thyroid and lung cancers [120]. Chromophobe RCCs have also demonstrated frequently mutated genes involved in the mTOR pathway (MTOR, TSC1, and TSC2) TP53 and PTEN [121].

Aggregated evidence suggests that chRCCs behave more indolently than ccRCC [89, 117, 122] with 5-year and 10-year survival rates reported between 78-100% and 80-95%, respectively. In 2018, Leibovich characterised risk of progression for patients with chRCC according to low, intermediate and high groups based on presence or absence of three histopathological features: fat invasion, nodal involvement and sarcomatoid differentiation. 5 and 10-year progression-free survival (PFS) were reported as 94% and 91% (low- risk), 71% and 59% (intermediate-risk) and 13% and 4% (high-risk) [72]. As with pRCCs, updated European Association of Urology (EAU) guidance recommends the use of TMN and Fuhrman grade to stratify patients with chRCC into post nephrectomy risk groups [20], despite no evidence for this approach.

3.1.5 Sarcomatoid RCC (sRCC)

Studies investigating sRCCs are few and generally small owing to the rarity of the disease. Sarcomatoid dedifferentiation usually represents high-grade histological transformation and can be present within any subtype of RCC. It can involve between 1 and 100% of the total tumour, with a mean and median extent of 40%–50% [123]. It appears in 4% of all RCCs increasing to 20% in metastatic RCCs and generally confers a poor prognosis [123]. Sarcomatoid RCC's are characterised immunohistologically by the presence of spindle cells, positive for keratin and PAX8 [124] (**Table 16**). They are

distinct from primary renal sarcomas which are rare in adults (<1%) and do not contain any classic areas of RCC. In published series, sarcomatoid tumours are usually large, (mean tumour size of 9–10 cm), 90% are symptomatic at presentation [125] and 45%–84% are metastatic at presentation [126].

The definitive molecular drivers of sarcomatoid transformation in RCC are poorly understood. Several genes involved in epithelial–mesenchymal transition show increased expression in the sarcomatoid component of ccRCCs, which supports a single 'cell of origin' theory where progenitor cells undergo transformation from epithelial to mesenchymal phenotype [127]. Additionally, some sarcomatoid associated genes, for example, AURORA KINASE-1, involved in regulating chromosomal alignment in mitosis and also linked to overactivation of the mTOR pathway, have been found [128]. Sarcomatoid RCCs also retain expression of genes associated with the underlying epithelial histology. For example, ccRCCs with sarcomatoid transformation have shown maintained expression of hypoxia-inducible factor (HIF) pathway components for example HIF-1α, VEGF and CAIX [129].

The majority of case series confirm the presence of sRCC to be associated with poor prognosis, rapid progression and death from RCC [126]. A large study including 230 patients with sRCC and 2056 non-sRCC found that patients with sRCC had significantly worse International Metastatic RCC Database Consortium (IMDC) prognostic scores compared with non-sRCC ($p<0.0001$) [130]. Time from original diagnosis to relapse (excluding synchronous metastatic disease) was shorter in the sRCC group (18.8 vs. 42.9 months; $p<0.0001$). Sarcomatoid histology was associated with a significantly worse progression-free-survival (PFS) and OS after adjusting for individual IMDC risk factors on multivariable analysis (HR, 1.5; $p< 0.0001$ for both) [130]. Despite sRCCs being associated with consistently poor prognoses, there remains a sparsity of research into relapse patterns, no sarcomatoid specific surveillance guidance after nephrectomy and known actionable driver mutations or specific treatment options in the adjuvant or metastatic setting.

3.1.6 Treatment strategies for Metastatic non ccRCC

3.1.6i Targeted therapies

The large phase three trials heralding the era of VEGF-TKI targeted therapy in metastatic RCC typically only included patients with ccRCC or tumours with at least a component of ccRCC (**Table 17**). Although this practice is justified by the prominence of the VHL pathways in ccRCC, it has led to a lack of supporting data for the use of targeted therapies specifically for patients with non-ccRCCs.

A meta-analysis of TKI data published in 2015 including 7544 patients, found 384 patients (5.1%) were from studies conducted exclusively for patients with non-ccRCC. Out of those including patients with mixed histology's, only 860 patients (12%) had non-ccRCC and 6300 (88%) had ccRCC. The study showed that overall, TKIs approved for metastatic ccRCC are demonstrably less effective in the setting of non-ccRCC, It highlighted comparably lower response rates (ORR 0.52; 95% CI, 0.40-0.68 $p<0.001$), worse median PFS (7.4 versus 10.5 months) and OS (13.4 versus 15.7 months) for patients with non-ccRCC [131]. A phase two study of first-line everolimus (an mTOR pathway inhibitor) followed by second-line sunitinib (TKI) compared with the reverse, analysed a small subset of patients with metastatic non-ccRCC [132]. The median PFS was shorter for those with non-ccRCC (7.2 vs 10.8 months respectively). Only three first-line randomised studies have evaluated sunitinib compared to everolimus in patients with metastatic non-ccRCC; ASPEN (NCT01108445), ESPN (NCT01185366) and RECORD-3 (NCT00903175) [133-135]. A meta-analysis of these studies favoured sunitinib over everolimus for first-line therapy. The response rate for sunitinib was less than 20% with a PFS between 6 to 12 months [131].

There have been several small inconclusive studies testing MET-targeted therapies in pRCC based on MET being a key driver of oncogenesis in this setting [136-138]. The most recent of these was the SAVOIR trial (NCT03091192) [138], a randomised study evaluating savolitinib (MET inhibitor) compared with sunitinib in MET-driven pRCCs. SAVOIR was halted prematurely due to presumed futility after results from a concurrent epidemiological study showed that MET- driven status was not a predictive factor for poor treatment outcome. Interestingly, analysis of the few randomised patients (33 to salvotanib, 27 to sunitinib) showed that salvoitinib demonstrated encouraging efficacy

with numerically greater PFS, OS, and objective response rate (ORR) compared to sunitinib alone with a favourable side-effect profile.

Combined MET-VEGF-AXL inhibitors such as cabozantinib and formetinib have also been evaluated in patients with pRCC [136, 139]. AXL, a TKI receptor that belongs to the TAM family, is overexpressed in both ccRCC and pRCC. Although patient numbers and follow-up are limited in these studies, together they suggest that pRCC like ccRCC is likely to be driven by a combination of VEGF, MET and AXL-signalling with a skew towards MET and AXL in pRCC and VEGF signally in ccRCC.

In the case of chRCC, data regarding effective therapies are extremely limited, but small studies show that TKIs are minimally efficacious in this subtype. An example being a retrospective case series by Choueiri *et al*. that included 12 patients with chRCC who received sorafenib (*n* = 5) or sunitinib (*n* = 7) [138]. It reported that three patients (25%) achieved a partial response. Patients receiving sorafenib tended towards a longer median PFS (27.5 months for sorafenib vs. 8.9 months for sunitinib).

Despite the patchy evidence supporting the use of TKIs for patients with pRCC and chRCC, current recommendations follow the wider guidance issued for ccRCC, favouring sunitinib and pazopanib for their frontline treatment. In 2019, the UK Cancer Drugs Fund (CDF) approved single agent cabozantinib as another first line options for patients with intermediate or poor risk advanced papillary and chromophobe subtypes even though all patients enrolled in the referencing trial were required to have a clear cell component [140]. Salvotinib is an additional TKI option available in MET driven pRCC tumours.

For sarcomatoid differentiated RCCs, sunitinib and pazopanib are also recommended, although patients rarely experience clinical benefit. Several small studies support the use of TKIs in sRCC despite low response rates and response durations reported as between nine to twelve months [126, 141].

3.1.6ii Immune checkpoint Inhibitors

All the large phase three trials evaluating immune checkpoint inhibitors (ICIs) in the advanced RCC setting included participants with at least a component of ccRCC or ccRCC exclusively except for Immotion151 (first line atezolizumab and bevacizumab)

which included patients with sRCC (see **Chapter 1, Table 2** for summary of landmark phase three ICI trials in advanced ccRCC). Non-clear cell RCCs also express the immune checkpoint PD-L1 on tumour infiltrating mononuclear cells and on tumour cells [142]. However, evidence for the activity of PD-1 blockade for patients with advanced non-ccRCC, is inconclusive (see **Table 18** for summary of trials) partly owing to the predominance of non-randomised studies, small patient numbers and heterogeneous multi-subtype populations comprising a combination of treatment naïve and pre-treated groups.

Despite sparse evidence supporting the use of ICIs in the treatment of metastatic non-ccRCCs, the UK CDF has approved the combination of Ipilimumab and Nivolumab, as a first line option for patients with intermediate or poor risk advanced pRCC. This followed results from Checkmate-214 trial (NCT02231749) where patients with non-ccRCC were enrolled but all were required to have a clear cell component [143]. The combination of axitinib and avelumab, is currently available via the CDF for those with pRCC and chRCCs.

Accumulating data suggests that advanced sRCC tumours are sensitive to ICIs. A meta-analysis of patients with sRCC recruited to ICI phase three trials, (including 226 treated with ICI combinations and 241 receiving sunitinib) showed that ICI-based combinations were associated with more than 40% reduction in progression (HR = 0.56; $p < 0.0001$) and mortality (HR = 0.56; $p = 0.001$) in this group [144]. Moreover, those receiving ICI combinations had double the response rate compared to those receiving sunitinib (relative risk [RR] = 2.15; $p < 0.00001$). Finally, ICIs significantly increased the complete response rate (RR = 8.15, $p = 0.0002$) in the sRCC group with an incidence of 11% [144]. Based on this data, the use of nivolumab and ipilimumab for sRCC is considered a valid treatment option for patients with metastatic sRCC although current National Comprehensive Cancer Network (NCCN) do not provide specific guidance for treating RCCs with sarcomatoid component. Given the cost and the potential for severe and/or long-term irreversible toxicities of ICIs, there is a strong rationale, for the development of subtype specific guidance on their use, rather than simply adopting the same standards of care for ccRCC in those with non-ccRCCs, without robust evidence base.

Table 17 Phase III trials of VEGF-TKI inhibitors in the metastatic setting in RCC showing inclusion of patients with ccRCC histology

Treatment	Comparator	Treatment line	Patient number	Patients	Median PFS, months	Median OS, months
Sunitinib	IFN-α	1st	750	ccRCC component	11.0 vs. 5.1	26.4 vs. 21.8
Sorafenib	Placebo	2nd	903	ccRCC only	5.5 vs. 2.8	17.8 vs. 15.2
Pazopanib	placebo	1st or 2nd	435	ccRCC component, (91%) ccRCC only	9.2 vs. 4.2	22.9 vs. 20.5
Pazopanib	Sunitinib	1st	1100	ccRCC only	non-inferior	non-inferior
Tivozanib	sorafenib	1st and 2nd	517	ccRCC component	11.9 vs 9.1	29.3 vs 28.8
Tivozanib	sorafenib	3rd or 4th	350	ccRCC component, (94%) ccRCC only	5.6 vs 3.9	16.4 vs 19.7
Axitinib	Sorafenib	2nd	723	ccRCC only	8.3m vs 5.7	20.1 vs 19.2m
Cabozantinib	Sunitinib	1st	157	ccRCC component	8.3 vs 5.3	26.6 vs 21.2
Cabozantinib	Everolimus (mTORi)	2nd or more	658	ccRCC component	7.4 vs 3.9	21.4 vs 16.5
Axitinib/Pembrolizumab (anti-PD-1)	Sunitinib	1st	861	ccRCC only	15.1 vs 11.1	NR
Axitinib/Avelumab (anti-PD-L1)	Sunitinib	1st	886	ccRCC component	13.8 vs 8.4m	NS
Cabozantinib/nivolumab (anti-PD-1)	Sunitinib	1st	651	ccRCC only	16.6m 8.3m	NR
Lenvatinib/Pembrolizumab (anti-PD-1) or Lenvatinib +Evrolimus (mTORi)	Sunitinib	1st	1069	ccRCC component	L/P 22.1m vs 9.5m L/E 14.6m vs 9.5m	L/P NR L/E NR

VEGF-TKI; Vascular endothelial growth factor inhibitor, tyrosine kinase inhibitor, PD-1; programmed cell death protein 1, PD-L1; programmed cell death protein ligand 1, mTORi; Mechanistic target of rapamycin inhibitor, ccRCC; clear cell RCC, m; month, PFS; progression-free survival, OS; overall survival, NS; not specified, NR; not reached

Table 18 summary of trials investigating immune checkpoint inhibitors in patients with non-clear cell RCCs

Trial acronym or author	Patient group	Randomised comparison (n)	Treatment	Treatment naïve (%)	Efficacy data
CHECKMATE-374 [145] Phase III/IV	pRCC n=24 chRCC n=8 uRCC n=8 other n=5	nil	Nivolumab	69	ORR 13.6% (95% CI, 5.2-27.4) Median OS 16.3m (95% CI, 9.2-NE)
KEYNOTE-427 [146] Phase II	pRCC n=118 chRCC n=21 uRCC=26	nil	Pembrolizumab	100	ORR All; 26.7% (95% CI, 17.9 to 34.3). pRCC; 28.8% chRCC; 9.5% uRCC 30.8%
Gupta *et al.* [147]	pRCC n=6 chRCC n=5 uRCC n=3 adenoc n=2 tRCC n=1 mRCC n=1	nil	Nivolumab + ipilimumab	100	ORR 33.3% PR n=8 SD n=3 PD n=9
CALYPSO [148] Phase I/II	MET driven pRCC n=40	nil	Durvalumab (PD-L1) + salvotinib (MET)	71	ORR 27% mOS 12.3m RR 27% mPFS 3.3m (95% CI, 1.5-NR)
COSMIC-021 [149] Phase Ib/II	pRCC n=15 chRCC n=7 other n=8	nil	Cabozantinib + Atezolizumab (PD-L1)	NS	ORR All; 33% pRCC; 40% chRCC; 14% other; 60%

pRCC; chRCC, uRCC, tRCC, mRCC; adenoc; adenocarcinoma PD-L1; programmed cell death protein ligand 1 MET; mesenchymal-epithelial-transition, ORR; objective response rate, mOS; median overall survival, mPFS; median progression-free-survival

3.1.7 Locally advanced non-ccRCC

In the setting of locally advanced RCC, the standard of care after curative nephrectomy for patients with non-ccRCCs (apart from sRCCs) is currently active monitoring regardless of risk group classification. This is based on findings from several large randomised controlled phase three trials that evaluated the use of adjuvant TKIs post nephrectomy, (Chapter 1, **Table 4**).

However, only, ASSURE [97], SORCE [5] and ATLAS (NCT01599754) [36] trials included histologically diverse patients. S-TRAC (NCT00375674) [35] and PROTECT (NCT01235962) [34] required all patients to have a clear cell histological component. In ASSURE, a post-hoc exploratory subgroup analysis of 267 participants with non-clear cell histology's found a HR for DFS of 1.07 (0.69-1.67) comparing sunitinib with placebo and for sorafenib with placebo of 0.94 (0.59-1.48). The precise treatment effect of TKIs in patients with non-ccRCCs remains unclear.

The post-nephrectomy standard of care for UK patients with high risk ccRCC and those with sarcomatoid component now includes adjuvant pembrolizumab following results from Keynote-564, the first group to publish results from an adjuvant ICI trial. Keynote-564 recruited 994 participants with histologically confirmed ccRCC including those with co-existing sarcomatoid histology, and randomly assigned them to either pembrolizumab (anti-PD-1) or placebo. At a median follow-up of twenty-four months, adjuvant pembrolizumab, significantly improved DFS compared with placebo, with a 32% reduction in the risk of recurrence in the pembrolizumab group, (HR 0.68 95% CI, 0.53-0.87 p=0.0010). Mandating a ccRCC component for all participants is akin to the situation in the metastatic setting- the effect of ICIs in patients with non-ccRCC subtypes remains uncertain. Several other phase III adjuvant RCC trials will report in the next few years. RAMPART and PROSPER (NCT03055013) plan to include patients with non-ccRCCs in the intention-to-treat (ITT) analysis. Neither have specified subtype specific analyses in their protocols. Therefore, the optimum adjuvant strategy for patients with non-sarcomatoid non-clear cell RCCs will remain uncertain.

3.1.8 Follow-up of higher risk RCCs after nephrectomy

The optimal follow-up after nephrectomy for patients with non-ccRCCs lacks international standardisation and exhibits variability between European and American institutions in terms of the recommended imaging frequency, modality and duration [20] [150]. Uniquely, EAU provides follow-up guidance specifically for non-ccRCCs but not individually for the histological subtypes. Therefore, the current guidance may lead to over-imaging of patients with indolent RCCs and fail to detect very early relapses in patients with faster growing tumours. Optimal surveillance requires specificity and in the current absence of prognostic molecular biomarkers, histological subtype is a pragmatic delineation in RCC. In order to achieve this, an accurate and comprehensive understanding of relapse timings and sites for each subtype is required.

Narayan and colleagues reported a retrospective analysis of the relapse patterns of 403 patients with mixed non-ccRCCs [151]. It found that patients with non-ccRCCs were at additional risk of abdominal relapse compared to those with ccRCC, (HR 1.22, 95% CI, 0.96–1.53, p = 0.099) and a lower risk of recurrences to the chest HR 0.57, 95% CI, 0.43–0.76, p < 0.001). The study noted that 10% of non-ccRCC relapses occurred beyond 5 years suggesting the potential need for prolonged imaging surveillance in this group. Although limited by examining the non-ccRCCs as one group, the study highlighted the need for separate surveillance guidance for non-ccRCCs.

3.1.9 Relapse patterns for patients with non-ccRCCs

There is limited published data (**Table 20**) regarding the relapse sites and ensuing clinical behaviour of the non-ccRCCs. This is largely due to a sparsity of large, well-designed cohort studies with follow-up depicting relapses over a long enough duration. The Abu-Ghanem *et al.* RECUR[7] study [152] is the largest analysis comparing outcomes of patients with various RCC histology's using a retrospectively collected registry of 3331 patients from clinical trial data and hospital databases. Patients were treated across 15 centres from 10 countries, from 2006 to 2011, with radical or partial nephrectomy. A total of 2565 patients (77.0%) had ccRCC, 535 (16.1%) had pRCC, and 231 (6.9%) had

[7] RECUR; European association of urology RCC guidelines panel Collaborative multicentre consortium for the studies of follow-Up and recurrence patterns in Radically treated renal cell carcinoma

chRCC. Authors reported relapse sites and patterns for the individual histology's and divided patients into low, intermediate and high-risk groups based on clinical risk scores validated for each subtype. The majority of patients were of low relapse risk. Predictably, they found that most of the recurrences occurred in the intermediate to high-risk groups. Overall, patients with ccRCC had significantly poorer five-year RFS than patients with pRCC and chRCC (78% vs 86% vs 91%, p = 0.001). **Table 19** summaries the key subtype specific results from the trial. Although the study was limited by poor standardisation of the data-points collected across centres for example different starting time-points, surveillance imaging and frequencies, it was the first large multi-subtype comparative cohort study to evaluate outcomes, including relapse sites, and to explore patients by clinico-pathological risk group.

Another study by Dudani *et al.* went further to explore the relationship between site of relapse and OS (**Table 20**). They reported a multicentre retrospective cohort of patients with metastatic RCC taken from the International Database Consortium (IDC) database and compared the four commonest sites of relapse (bone, liver, lymph-node and lung) in patients with ccRCC, pRCC and chRCC. They showed that for each subtype survival exhibited variability associated with site of initial metastatic involvement. For patients with pRCC, metastasis to lymph-nodes and bone and for patients with chRCC metastases to lung conferred the worst prognosis (**Table 20**). The study was limited to patients who received systemic treatment for metastatic disease, so patients who were treated curatively with initially locally advanced disease and those that were managed in other ways, for example, active surveillance or metastasectomy, were not captured. A similar study in a large cohort of contemporary patients followed up from nephrectomy, through relapse and undergoing a variety of treatments would be informative.

Table 19 Data from RECUR study [152]- The Impact of Histological Subtype on the Incidence, Timing, and Patterns of Recurrence in Patients with Renal Cell Carcinoma After Surgery

	ccRCC	pRCC	chRCC
No patients	2565	535	231
No patients recurred/%	NS	59 (11)	19 (8.2)
Median TTR	21.2m (IQR: 7.9, 41.1)	19m (IQR: 8.5, 41)	37.4m (IQR: 11.1, 64.6)
No patients recurring >5 years from surgery (%)	65 patients (13.5)	9 patients (1.7)	6 patients (18.2)
Relapse sites			
Low risk group	Local (27%) Contralateral kidney (23%)	Contralateral kidney (31.3%) Lung (31.3%)	
Intermediate and high-risk groups	Lung (58%) Liver (17%)	Lung (40%), Retroperitoneal lymph nodes (35.5%)	NR
TTR for abdominal sites only compared to thorax recurrence only	31.1m (+/-2.43) and 26m (+/-1.8) p = 0.008	NR	NR

RECUR; European association of urology RCC guidelines panel Collaborative multicentre consortium for the studies of follow-Up and recurrence patterns in Radically treated renal cell carcinoma, RCC; renal cell carcinoma, ccRCC; clear cell RCC, pRCC; papillary RCC, chRCC; chromophobe RCC, TTR; time to relapse, NR; not reported, m; months, IQR; interquartile range

Table 20 Studies reporting on sites of first recurrence in patients with ccRCCs and non-ccRCCs

	ccRCC			pRCC			chRCC				sRCC
	Abughanem *et al.* 2020 [152]	Dudani *et al.* 2021 [153]		Abughanem *et al*. 2020	Dudani *et al.* 2021		Abughanem *et al.* 2020	Dudani *et al.* 2021		Neves *et al.* 2021 [154]	Thomas *et al.* 2016 [155]
Number	2565	9252		357	667		231	186		295	273
Site	**% sites**	**% sites**	**Median OS (95% CI)**	**% sites**	**% sites**	**Median OS (95% CI)**	**% sites**	**% sites**	**Median OS (95% CI)**	**% patients**	**% sites**
Lung	50.5	70	25.1 (24.1-26.0)	35.8	39	15.6 (12.5-19.0)	11.8	36	14.1 (8.2-23.8)	0.1	45
Pleura	3.9	4	15.6 (13.7-18.8)	3.8	3		5.9	0.7		0.003	
Mediastinum	11.4			17			0				
Retroperitoneum	15.7	2	23.3 (11.8-36.3	29.6	5		5.9	4			
Liver	12.6	18	17.6 (16.0-19.2)	9.3	22	11.8 (9.6-13.9)	22.2	34	26.0 (12.9-36.8)	0.1	13
Pancreas	3.9	5	50.1 (41.1-55.5)	1.9	1		0	2			
Adrenal	9.6	10	27.3 (24.5-31.9)	5.7	7		5.9	6		0.003	
Contra Kid	10.5			13			11.8				
Bone	16	32	19.4 (18.1-20.5)	20.8	29	11.0 (9.8-14.1)	29.4	33	26.7 (18.4-35.6)	0.2	13
Brain	3.9	8	16.5 (13.2-18.7)	3.8	3		5.9	2		0.003	
Thyroid		0.7	44.0 (17.5-59.6)		0.2			0			
Lymph-node		45	21.4 (20.2-22.5)		69	14.3 (12.8-17.2)		51	28.1 (21.2-36.6)		
Bowel		0.7	29.0 (23.3-41.9)		0.2			1			
Spleen		0.9	19.8 (12.1-34.9)		0.6			0.8			
Local*										0.006	25

RCC; renal cell carcinoma, ccRCC; clear cell RCC, pRCC; papillary RCC, chRCC; chromophobe RCC, OS; overall survival

3.1.10 Rationale for analysing data from the RCC Histological Subtypes from Adjuvant Clinical trials

Despite the vast heterogeneity of prognostic features, survival outcomes, and recurrence patterns exhibited by the RCC subtypes, studies conducted on them individually are generally small-scale, retrospective in nature with variable follow-up. As a result, inconsistencies across the literature are seen and clinical practice for treatment and follow-up for the non-clear cell subtypes still largely follows guidance for ccRCC.

In this study, I used contemporary individual patient data with long-term follow-up from two recently reported adjuvant trials; SORCE and ASSURE. It provided the largest cohort study from phase three trial data with which to individually characterise the relapse patterns of patients with ccRCC, pRCC, chRCC and sRCC and correlated them with survival. In doing so, this study directly improves upon prognostication and supports follow-up guidance for patients with non-clear cell RCCs. This in-turn guides clinical trial design and translational work aimed at further understanding the key drivers of poor prognosis non-ccRCCs.

3.1.11 SORCE Trial

The SORCE trial [5] is described in detail in Chapter 1. Patients with ccRCC (n=1445) and non-ccRCC (pRCC (n=128), chRCC (n=96), collecting duct (n=4) and other (n=38)) were eligible apart from those with pure oncocytoma. A pre-specified analysis of the intention-to-treat (ITT) patients and separately those with ccRCC showed no difference in DFS for either three years or one year of adjuvant sorafenib compared with placebo [5]. Prospective baseline and follow-up individual participant data (IPD) was captured for patients over a median follow-up duration of 6.5 years (interquartile ratio IQR 4.9, 8.0).

3.1.12 ASSURE trial

The ASSURE trial [97], described in Chapter 1, was a double-blind, placebo-controlled, phase three trial led by the ECOG-ACRIN E2805 group at the Dana Faber Cancer Institute, US. 1943 American and Canadian patients with ccRCC and non-ccRCC from 226 centres, were included in the ITT analysis. 1541 patients with ccRCC, 150 with pRCC, 111 with chRCC, 83 classified as ‘mixed’ RCC, 56 with unclassified RCC and 170 with sarcomatoid features were evaluated. The primary analysis showed no significant

differences in disease-free survival or OS for the ITT group and separately those with ccRCC treated with either sunitinib or sorafenib compared to placebo. Prospective baseline and follow-up IPD were captured for patients over a median duration of 5·8 years (IQR 4·9, 6·9).

3.1.13 Pooling data from two trials

Pooling individual participant data from multiple phase three studies increases the sample size and statistical power to perform subgroup analyses [156] and provides a large international dataset to assess and compare the long-term clinical trajectories of patients with low prevalence higher-risk RCC subtypes.

3.1.14 Research question and aims

In this chapter, I addressed the following research question and objectives.

Research question

Can pooled data from SORCE and ASSURE, two prospective randomised controlled trials (RCTs) be used to explore the characteristics, clinical behaviours and relapse patterns of patients with non-ccRCCs.

Research aims

Primary aim

A comprehensive evaluation of the clinical characteristics of patients with locally advanced RCC separated by histological subtype, including a comparison of:

a. baseline clinical characteristics
b. relapse patterns (including location, number of sites and timing of first metastases)
c. DFS, OS and survival associated with sites of first relapse.

Secondary aims

To evaluate whether relapse pattern data can be used to enhance histology-specific prognostication upon relapse. To evaluate whether relapse pattern data can be used to inform histology-specific follow-up surveillance guidance.

3.2 Methods

3.2.1 Study population; SORCE and ASSURE trial participants

In both the SORCE and ASSURE trials, participants were eligible within twelve weeks of nephrectomy. Patients who relapsed within ninety days after primary surgery were not eligible as they were considered probably metastatic at the time of primary surgery. Baseline patient and tumour characteristics, site and timing of recurrence events, subsequent treatments and survival outcomes were captured prospectively via trial specific Case Report Forms (CRFs). Both trial designs and results have been summarised in Chapter 1.

I developed and presented this study to the SORCE TMG and in January 2020, I was granted access to the SORCE trial dataset. I submitted a data release request for the data from the ASSURE trial after presenting the study aims and statistical analysis plan to Professor Naomi Hass (ASSURE trial Chief Investigator) and the ECOG-ACRIN E2805 group at the Dana Faber Cancer Institute on 2nd January 2019. I received the complete data set on 4th May 2021.

3.2.2 Histological sub-groups analysed

For this analysis, I summarised the data by histological subtypes. The categories included were the four largest histological groups, those with ccRCC, pRCC, chRCC and sRCC. Participants with mixed histology's (that did not include a sarcomatoid component) and those with any other histology's were excluded. In keeping with the overall SORCE and ASSURE trial findings, there were no significant differences in DFS or OS in those who receive treatment compared to placebo for patients with ccRCC, pRCC, chRCC or sRCC for either trial. (Appendix Figure B). Therefore, for this study, data from the whole cohort as opposed to just the placebo arms of both trials were included to maximise the sample size.

3.2.3 Classification of 'sarcomatoid' histology

In the SORCE and ASSURE datasets, delineation of 'Sarcomatoid' RCC differed. In SORCE 'Sarcomatoid RCC' was listed as a stand-alone group on the randomisation CRF. Any patients with mixed histological features were listed in an 'other' category which I manually text sorted to extract those with mixed sarcomatoid features. Any SORCE

participant listed as 'sarcomatoid' or with 'mixed sarcomatoid features' was included in the "sRCC" category. In ASSURE information on presence of sarcomatoid features was collected for each participant as a Y/N tick box alongside details of their underlying histology. Any participant with 'Y' tick was included in the sRCC category.

3.2.4 Statistical Methods

I wrote the statistical analysis plan for this analysis and worked alongside Elena Frangou, a statistician at the MRC CTU at UCL who wrote the statistical code and carried out the statistical analysis. Professor Mahesh Parmar (senior statistician and director of the MRC CTU at UCL) checked the statistical methodology. All analyses were performed using STATA 16.1 (StataCorp LP, College Station, TX, USA). P-values were given to two significant figures. All statistical tests were 2-sided.

3.2.5 Determining heterogeneity between SORCE and ASSURE prior to pooling the datasets

Dataset heterogeneity was considered by comparing the eligibility criteria, randomisation criteria, and follow-up and outcome measures used in both trials. The eligibility criteria, (**Appendix Table E**), were broadly similar apart from the risk scores used to select and stratify participants into trial arms (see below).

3.2.6 Participant risk stratification for selection and randomisation to trial arms

In SORCE, all tumours, regardless of histological subtype were assigned a 2003 Leibovich score (**Table 10**). Patients whose tumours had an intermediate (3-5) or high (≥6) score were eligible for SORCE [5, 23]. Participants were stratified into treatment groups by country and Leibovich risk group.

In ASSURE patients were eligible with at least T1b N0 M0 and Fuhrman grade 3 or 4 disease. Both trials used the 2002 AJCC TNM staging system. In ASSURE participants were stratified to 'intermediate high risk' or 'very high risk' based on grouping from the modified University of California Los Angeles Integrated Staging System (UCLA UISS) prognostic score [25]. Participants were randomly assigned stratified by histology (clear *vs* non-clear cell), surgery (laparoscopic *vs* open), Eastern Cooperative Oncology Group (ECOG), performance status (0 *vs* 1), and risk category (intermediate high-

risk *vs* high or very high risk). 2003 Leibovich Score components were collected for patients recruited to ASSURE.

3.2.7 Primary Outcome Measures

In SORCE, the primary outcome measure was DFS, defined as the interval from random assignment to first evidence of local or distant metastases, or death from RCC. In ASSURE the primary outcome measure was DFS, defined as the time from random assignment to recurrence, development of second primary cancer, or death from any cause. OS was a secondary outcome measure defined as time from randomisation to death from any cause, in both trials.

3.2.8 Participants Follow-up

SORCE participants were assessed every three months with alternating computerised tomography (CT) and chest x-ray (CXR) until the end of year three on study, six monthly CXRs until year five, then annual CXRs until year ten.

Participants from ASSURE had imaging every 4·5 months during treatment, then every six months for two years, and then once a year for ten years. Follow-up imaging was non-contrast chest CT and gadolinium enhanced magnetic resonance imaging (MRI) of the abdomen and pelvis or contrast enhanced CT of the chest, abdomen and pelvis.

In order to optimally inform on histology-specific surveillance guidance, the original statistical analysis plan for this study included a comparison of the timing of re-staging imaging on relapse detection and intended to compare survival between similar (Leibovich) risk patients in ASSURE and SORCE. However, I was unable to obtain the ASSURE data for this analysis.

3.2.9 Determining how to evaluate the two datasets

I considered the optimal way to examine the two datasets in the context of some heterogeneity in eligibility, randomisation, follow-up and DFS definition between the trials. Three approaches; a pooled analysis, an analysis comparing the trials side-by-side and a meta-analysis were considered. This analysis sought to describe and compare the characteristics and behaviours of patients within the four histological subtypes using data from two phase three trials. A meta-analysis approach was rejected as I did not seek to

evaluate the effect size of TKI treatments for patients within the subgroups. A pooled analysis approach was favoured over a side-by-side approach, to increase the sample size of cohorts. The combined SORCE and ASSURE dataset was defined as the 'pooled' dataset.

3.2.10 Assigning 2003 Leibovich scores

2003 Leibovich score components were extracted from the ASSURE dataset, using STATA programming, allowing Leibovich scores to be assigned to the ASSURE patients to match the SORCE cohort. Only participants from ASSURE with a complete set of Leibovich components were used to generate Leibovich scores, which led to 20% missing data points. One participant from SORCE had a missing nuclear grade which was substituted with a 3 (the commonest nuclear grade).

3.2.11 Definition of sites of recurrence

'Local' recurrence was defined as recurrence in the ipsilateral renal bed, remnant kidney or local lymph-nodes. An abdominal cavity deposit was defined as peritoneal or abdominal recurrences that was not within local lymph-nodes and not within a retroperitoneal or abdominal organ. A non-local lymph-node was defined as a distant node, which could include one originating from the abdomen, retroperitoneum, or thoracic cavity.

For the analysis of 'chest vs abdominal' sites, sites of relapse were classified into two localisation groups: chest (lung, pleura, and mediastinal lymph nodes), abdomen (liver, pancreas, adrenal gland, contralateral kidney, and local recurrence). I chose to focus on chest and abdominal sites to compare findings to the recent literature [151, 152]. Metastases to bone, brain, and other non-chest miscellaneous sites were therefore excluded. The groups for the 'chest vs abdomen' analysis were as follows: chest only, abdominal only, or both (two or more sites were involved, including both chest and abdominal locations).

3.2.12 Definition of oncological outcomes for pooled analysis

For comparability, the SORCE definition of DFS was used in this analysis. Relapse-free survival (RFS) was the interval from nephrectomy to first evidence of local recurrence or distant metastases or death from RCC. RCC-specific survival was the time from

randomisation to death from RCC. OS was the time from randomisation to death from any cause, including RCC. Time-to-relapse (TTR) in months was calculated according to 'date of recurrence event minus date of primary nephrectomy'.

3.2.13 Comparing and pooling the datasets

Descriptive statistics were used to compare the baseline clinical and histopathological characteristics of the SORCE and ASSURE datasets and then the pooled dataset stratifying by histological subtype. Categorical variables were depicted with percentages and continuous variables as median and interquartile range (IQR) for each histological subtype.

3.2.14 Histology-specific survival outcomes

DFS and OS were presented for each histological cohort. For analysis of RCC-specific survival (**Appendix Figure C**), deaths due to other causes were regarded as censored observations. Time-to-event data was presented graphically using the Kaplan-Meier method for each histology. In all time-to-event analyses, patients that had not experienced the event in question (e.g. death) were censored on the date last seen. The log-rank test was used to compare the survival distributions between two groups (ccRCC was the reference group).

3.2.15 Histology and time to relapse

The percentage of relapses at any anatomical site occuring per year was presented in bargraphs stratified by histology. RFS was used to calculate TTR to enable evaluation of relapses occuring after nephrectomy. TTRs for each subtype were compaerd using the log-rank test. The absolute numbers of relapses at particular anatomical sites comparing those occuring prior to and after the median TTR was represented in vertical barplots.

3.2.16 Histology-specific sites of recurrence

First organ sites of recurrence were presented in pie charts reflecting the percentage of total relapses in each of the nine organ sites identified in ASSURE. The nodal anatomy was further delineated in data from SORCE, (**Appendix Tables J & K**).

3.2.17 Sites of relapse and survival outcomes

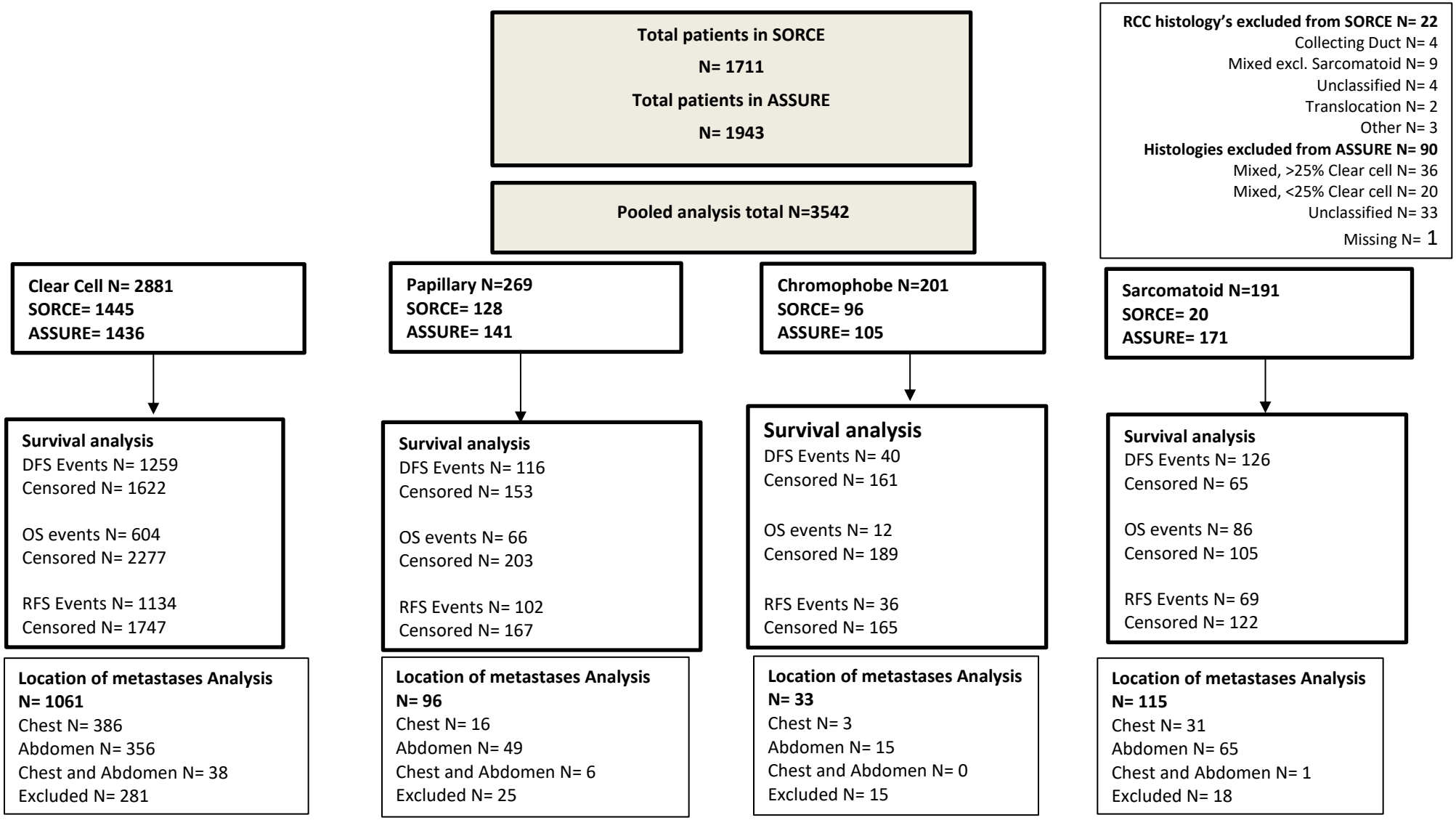
Survival outcomes of the four cohorts comparing the five most common sites of recurrence were evaluated using the Kaplan-Meier method.

3.2.18 Evaluation of the pattern of relapses according to location

Adjusted Cox models for DFS and OS comparing location of relapse (chest vs abdomen vs chest only and abdomen only) were presented. Each model contained histology and was adjusted for study (SORCE or ASSURE) and TKI treatment (sorafenib or sunitinib). In the same model, the interaction between histology and study was presented to obtain the HRs for each trial, (ccRCC was the reference category). Results were presented as HRs, 95% C.Is. and p-values.

3.3 Result

Figure 13 Consort diagram showing participants from SORCE and ASSURE in the pooled analysis



3.3.1 Comparison of the SORCE and ASSURE cohorts

I set out to evaluate whether pooled data from SORCE and ASSURE, can inform the characteristics, clinical behaviours and relapse patterns of patients with intermediate and high-risk non-ccRCC. The cohort selection is shown in **Figure 13**. Of 1711 patients randomly assigned to the SORCE trial between 2007 and 2013, 1689 (99%) were eligible for inclusion in this analysis. 22/1711 patients were excluded with other histology's and those that were mixed without a sarcomatoid component.

Of the 1943 non-metastatic patients randomly assigned to ASSURE trial between 2006 and 2010, 1853 (95%) were eligible for inclusion in this analysis. 90/1943 patients from the ASSURE dataset were excluded; 1 with missing histology and 89 with either mixed histology's without a sarcomatoid component or unclassified RCCs were excluded. **Table 21** shows that clinic-pathological and surgical features for SORCE and ASSURE patients were broadly well matched. Patient characteristics stratified by trial and by histology are displayed in **Appendix: Tables F and G**.

3.3.2 The pooled Dataset

3542 patients from SORCE and ASSURE were included in the pooled analysis cohort. The median post-operative follow-up was 9.5 years (IQR 2.2, NR). 62% (2197/3542) of patients had no relapse event at the last follow-up date. Of the total patients, 38% (1329/3542) were of intermediate Leibovich risk and 52% (1850/3542) were of high risk. In terms of histology, 2881/3542 (81%) of patients had ccRCC, 269/3542 (8%) pRCC, 201/3542 (6%) chRCC and 191/3542 (2%) sRCC. The three non-ccRCC subtypes had distinct clinical and histological characteristics (**Table 22**). For the ccRCC, and pRCC groups, the proportion of patients with Leibovich high risks were comparable at approximately 50% (1468/2881 for ccRCC, 147/269 for pRCC), compared to 77% of patients with sRCCs (148/191) and 43% (87/201) of patients with chRCC.

Table 21 Baseline characteristics of participants

Variable at Baseline	SORCE Total (N=1689)	ASSURE Total (N=1853)
Mean age at Randomisation (SD)	58.2 (10.6)	56 (10.6)
Sex		
Male	1198 (71%)	1253 (68%)
Female	491 (29%)	600 (32%)
Performance Status*		
0	1346 (79%)	1479 (80%)
1	330 (19%)	331 (18%)
2	1 (0%)	2 (0%)
3	12 (1%)	3 (0%)
Missing		38 (2%)
Pathological T cat. of Primary Tumour (2002 TNM)		
pT1a	7 (0%)	8 (0%)
pT1b	194 (11%)	139 (8%)
pT2	396 (23%)	501 (27%)
pT3a-4	1092 (64%)	1166 (63%)
Missing		39 (2%)
Regional Lymph Node Status		
pNx/ pN0	1618 (96%)	1703 (92%)
pN1/ pN2	71 (4%)	150 (8%)
Tumour Size		
<10	1133 (67%)	1049 (57%)
>10	556 (33%)	468 (25%)
Missing		336 (18%)
Nuclear Grade**		
1 or 2	525 (31%)	630 (34%)
3	851 (50%)	855 (46%)
4	312 (18%)	354 (19%)
Missing	1 (0%)	14 (1%)
Histological Tumour Necrosis		
No	767 (45%)	939 (51%)
Yes	922 (55%)	634 (34%)
Missing		280 (15%)
Leibovich Risk Group***		
Intermediate	904 (54%)	425 (23%)
High	785 (46%)	1065 (57%)
Missing		363 (20%)
Histology		
Clear Cell	1445 (85%)	1436 (77%)
Papillary	128 (8%)	141 (8%)
Chromophobe	96 (6%)	105 (6%)
Sarcomatoid	20 (1%)	171 (9%)
Type of Nephrectomy		
Radical	1569 (93%)	1753 (95%)
Partial	59 (3%)	100 (5%)
Missing	61 (4%)	
Type of Operation		

Open	898 (53%)	1056 (57%)
Laparoscopic	704 (42%)	797 (43%)
Missing	87 (5%)	

*The SORCE trial used ECOG performance status. ASSURE used WHO performance status. Appendix Table C
** The SORCE grading system used simplifies and selects the worst ISUP features at each grade. ASSURE used Fuhrman grading system. For details of grading components, see Appendix Table A.
*** 2003 Leibovich Scores were derived from baseline histological data provided in the ASSURE dataset.

Table 22 Baseline characteristics of participants in the pooled sample stratified by RCC histological subtype

Variable at Baseline	Clear Cell (N=2881) 81%	Papillary (N=269) 8%	Chromophobe (N=201) 6%	Sarcomatoid (N=191) 5%	Total (N=3542)
Age at Randomisation, Mean (SD)	57.4 (20 - 85.6)	57.6 (18.7 - 80)	52.2 (24 - 84)	56.5 (27 - 83)	57.1 (18.7 - 85.6)
Sex					
Male	1991 (69%)	210 (78%)	117 (58%)	133 (70%)	2451 (69%)
Female	890 (31%)	59 (22%)	84 (42%)	58 (30%)	1091 (31%)
Performance Status					
0	2305 (80%)	213 (79%)	167 (83%)	140 (73%)	2825 (80%)
1	529 (18%)	54 (20%)	31 (15%)	47 (25%)	661 (19%)
2	2 (<1%)			1 (1%)	2 (<1%)
3	2 (<1%)	1 (<1%)			3 (<1%)
Missing	43 (2%)	1 (<1%)	3 (2%)	3 (1%)	50 (1%)
Pathological T cat. of Primary Tumour					
pT1a	10 (0%)	2 (1%)	1 (0%)	2 (1%)	15 (<1%)
pT1b	285 (10%)	28 (10%)	10 (5%)	10 (5%)	333 (9%)
pT2	661 (23%)	92 (34%)	100 (50%)	44 (23%)	897 (25%)
pT3a-4	1898 (66%)	143 (53%)	86 (43%)	131 (69%)	2258 (64%)
Missing	27 (1%)	4 (2%)	4 (2%)	4 (2%)	39 (1%)
Regional Lymph Node Status					
pNx/ pN0	2767 (96%)	215 (80%)	189 (94%)	150 (79%)	3321 (94%)
pN1/ pN2	114 (4%)	54 (20%)	12 (6%)	41 (21%)	221 (6%)
Tumour Size					
<10 cm	1844 (64%)	158 (59%)	95 (47%)	85 (44%)	2182 (62%)
>10 cm	776 (27%)	88 (33%)	86 (43%)	74 (39%)	1024 (29%)
Missing	261 (9%)	23 (8%)	20 (10%)	32 (17%)	336 (9%)
Nuclear Grade					
1 or 2	967 (34%)	99 (37%)	77 (38%)	12 (6%)	1155 (33%)
3	1432 (50%)	143 (53%)	96 (49%)	35 (18%)	1706 (48%)
4	480 (16%)	23 (9%)	21 (10%)	142 (75%)	666 (19%)
Missing	2 (<1%)	4 (1%)	7 (3%)	2 (1%)	15 (<1%)
Histological Tumour Necrosis					
No	1459 (51%)	92 (34%)	112 (56%)	43 (23%)	1706 (48%)
Yes	1209 (42%)	158 (59%)	72 (36%)	117 (61%)	1556 (44%)
Missing	213 (7%)	19 (7%)	17 (8%)	31 (16%)	280 (8%)
Leibovich Risk Group					
Intermediate	1141 (40%)	93 (35%)	88 (44%)	7 (4%)	1329 (38%)

High	1468 (51%)	147 (55%)	87 (43%)	148 (77%)	1850 (52%)
Missing	272 (9%)	29 (11%)	26 (13%)	36 (19%)	363 (10%)
Type of Nephrectomy					
Radical	2712 (94%)	234 (87%)	194 (97%)	182 (95%)	3322 (94%)
Partial	116 (4%)	29 (11%)	5 (2%)	9 (5%)	159 (4%)
Missing	53 (2%)	6 (2%)	2 (1%)		61 (2%)
Type of Operation					
Open	1569 (54%)	154 (57%)	113 (56%)	118 (62%)	1954 (55%)
Laparoscopic	1234 (43%)	109 (41%)	85 (42%)	73 (38%)	1501 (42%)
Missing	78 (3%)	6 (2%)	3 (1%)		87 (2%)
Median follow-up time (years) (IQR)	6.1 (4.9,7.7)	6.1 (4.9,7.3)	6.0 (5.0,7.2)	5.8 (4.5,6.7)	

3.3.3 Histology-specific survival outcomes

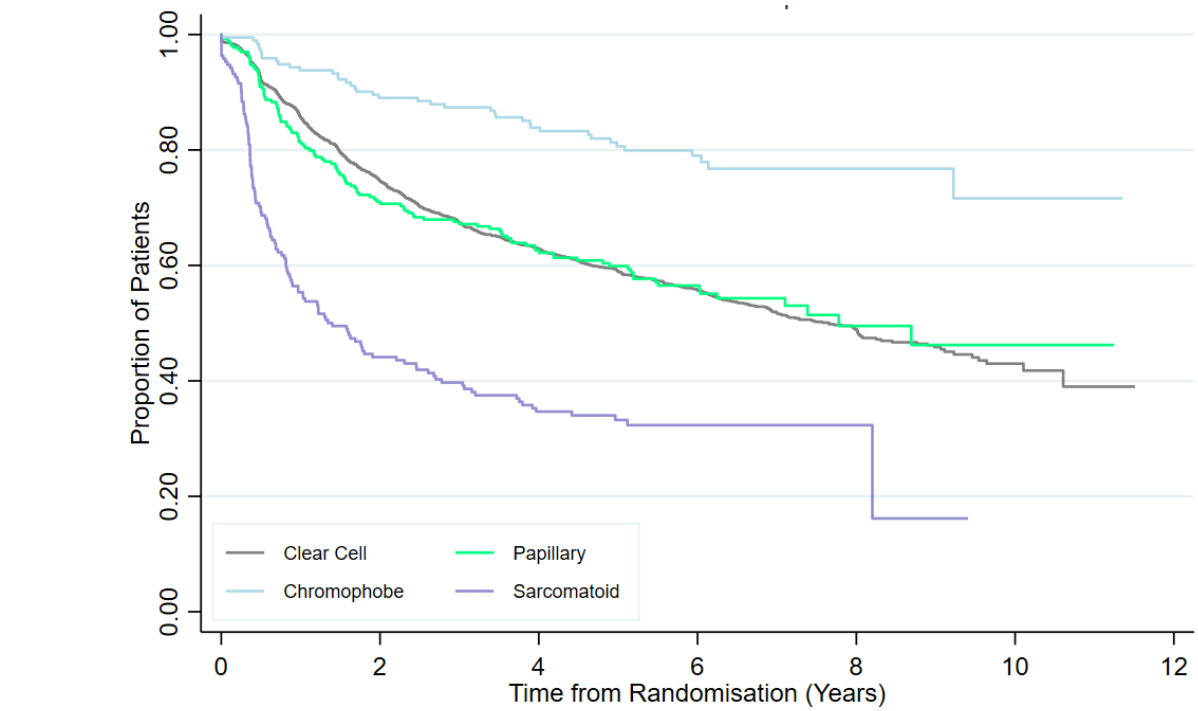
All log-rank tests used ccRCC as the reference group.

Figure 14 shows the DFS Kaplan-Meier curves for the four histological subtypes. The curves for the pRCC and ccRCC cohorts align (log-rank p=0.983). The median DFS for patients with pRCC was 7.78 years (IQR 1.56, NR), similar to 7.82 years (IQR 1.84, NR) for patients with ccRCC.

Predictably, patients with chRCC exhibited favourable DFS (log rank p<0.001) with a five-year DFS of 78.6% (95% CI, 71.7% - 84.1%). The sRCC cohort showed the steepest DFS decline over time (log-rank p-value<0.001), with a five-year DFS of 29.6% (95% CI, 23-36.5%).

Figure 15 shows the survival of patients by histological subtype. The OS Kaplan-Meier curve for the pRCC and ccRCC cohorts align (log rank p=0.110). OS was the worst for patients with sRCC (log-rank p<0.001 sRCC) and the most favourable for patients with chRCC (log-rank p <0.001).

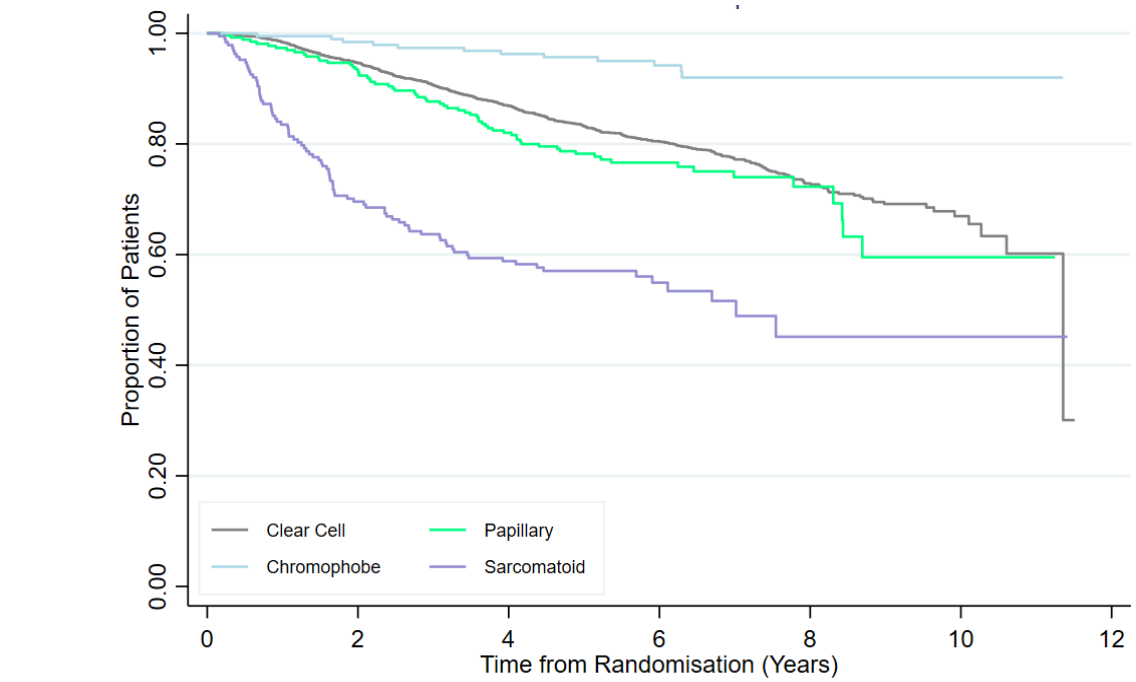
Figure 14 DFS for patients with papillary RCC, chromophobe RCC and sarcomatoid RCC compared to those with clear-cell RCC



Number at risk

Clear Cell	2881	2001	1563	897	288	57	0
Papillary	269	183	146	80	21	6	0
Chromophobe	201	167	140	80	26	9	0
Sarcomatoid	191	82	61	25	4	0	0

Figure 15 Overall survival for patients with papillary RCC, chromophobe RCC and sarcomatoid RCC compared to those with clear cell RCC



Number at risk

Clear Cell	2881	2592	2249	1343	430	68	0
Papillary	269	242	199	114	33	6	0
Chromophobe	201	188	172	106	31	10	0
Sarcomatoid	191	130	107	42	8	2	0

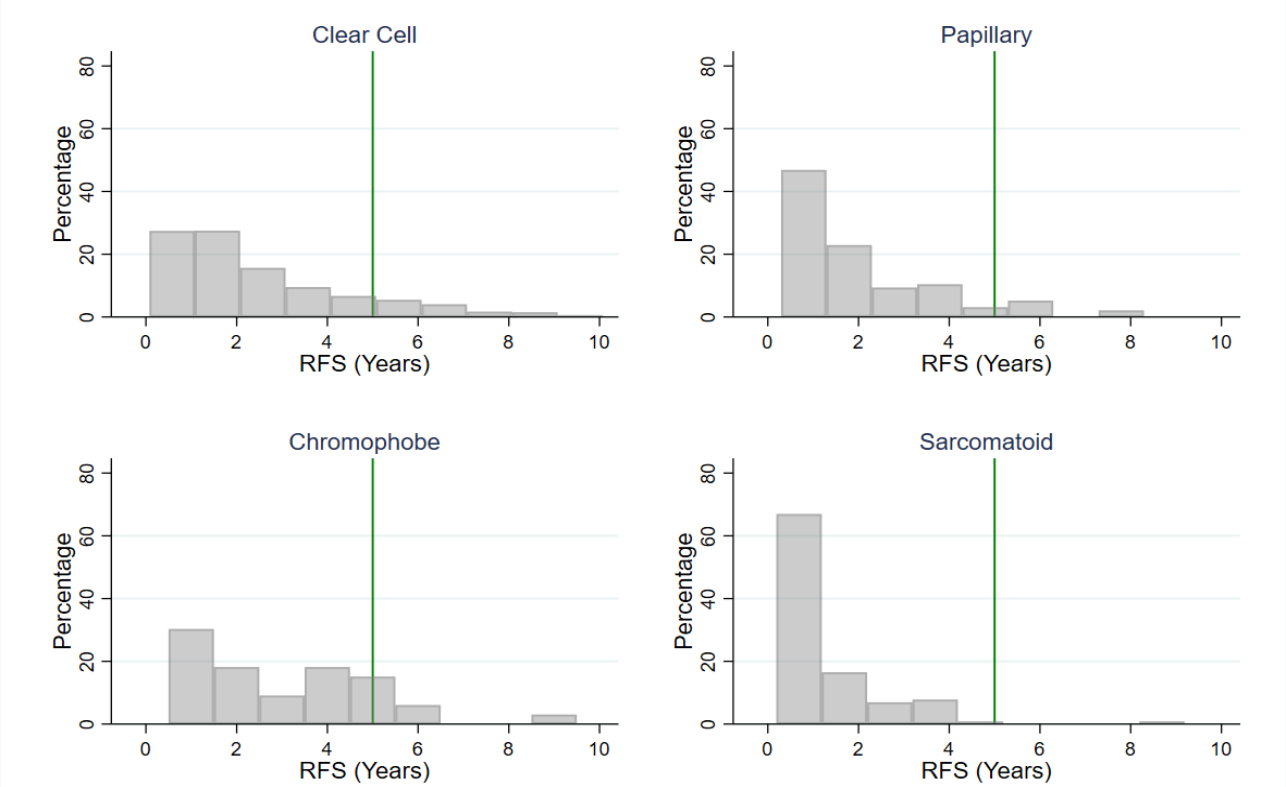
3.3.4 Histology and time to relapse

The rates and timing of relapses were compared between the four histologies. 43.7% of patients with ccRCC recurred within a median time of 1.78 years (IQR 0.96, 3.38 years). The median TTR for patients with pRCC cohort was five months less than those with ccRCC; (1.34 years (IQR 0.76, 2.59, p=0.012)). The relapse rate for the chRCC cohort was 17.9% and they exhibited the longest TTR at 2.72 years (IQR 1.07, 4.11, p=0.192)). The TTR for patients with sRCC was the shortest at 0.79 years (IQR 0.50, 1.55, p<0.001)).

The percentage of relapses occurring each year after nephrectomy for each histology is depicted in **Figure 16**. Compared to patients with ccRCC, (13.3% (142/1061)), fewer patients with pRCC (7.3% (7/96)) recurred after five years. Most patients with pRCC recurred within two years; 66.6% (64/96) for pRCC vs 53.6% (569/1061) for ccRCC. 100% of patients with sRCC relapsed within 5 years with most relapses occurring within the first year. 42% (14/33) of patients with chRCC relapsed within two years and 18% (6/33) relapsed after 5 years.

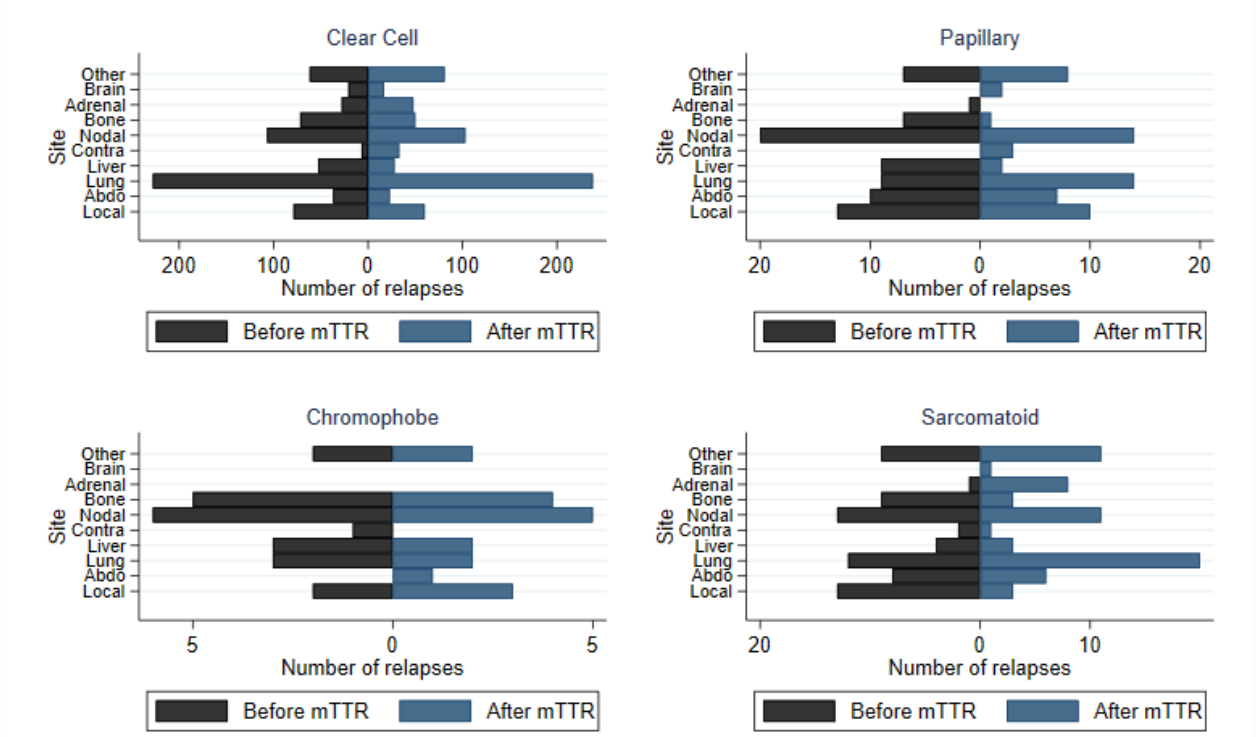
In order to inform on relapse patterns and optimal follow-up imaging for each histology, the anatomical sites of relapse prior to and after the median TTR was evaluated in **Figure 17**. Of note, frequencies of relapses to the lung were well balanced before and after the median TTR for all subtypes suggesting the need for ongoing imaging of the thorax beyond the median TTR timepoint. No relapses to the brain occur before the median TTR for patients with pRCC and sRCC therefore the need for brain imaging before this time point for these patients needs to be justified symptomatically.

Figure 16 Percentage of patients relapsing each year after curative nephrectomy, stratified by histological subtype. Green line indicates recurences before and after five years



RFS; relapse-free survival

Figure 17 Recurrences occuring prior to (dark blue) and after (light blue) median TTR relapses, by site and stratified by histological subtype



Note that median TTR is point 0 on the x-axis for each subtype although the specific value of mTTR varies according to histological subtype.
mTTR; median TTR

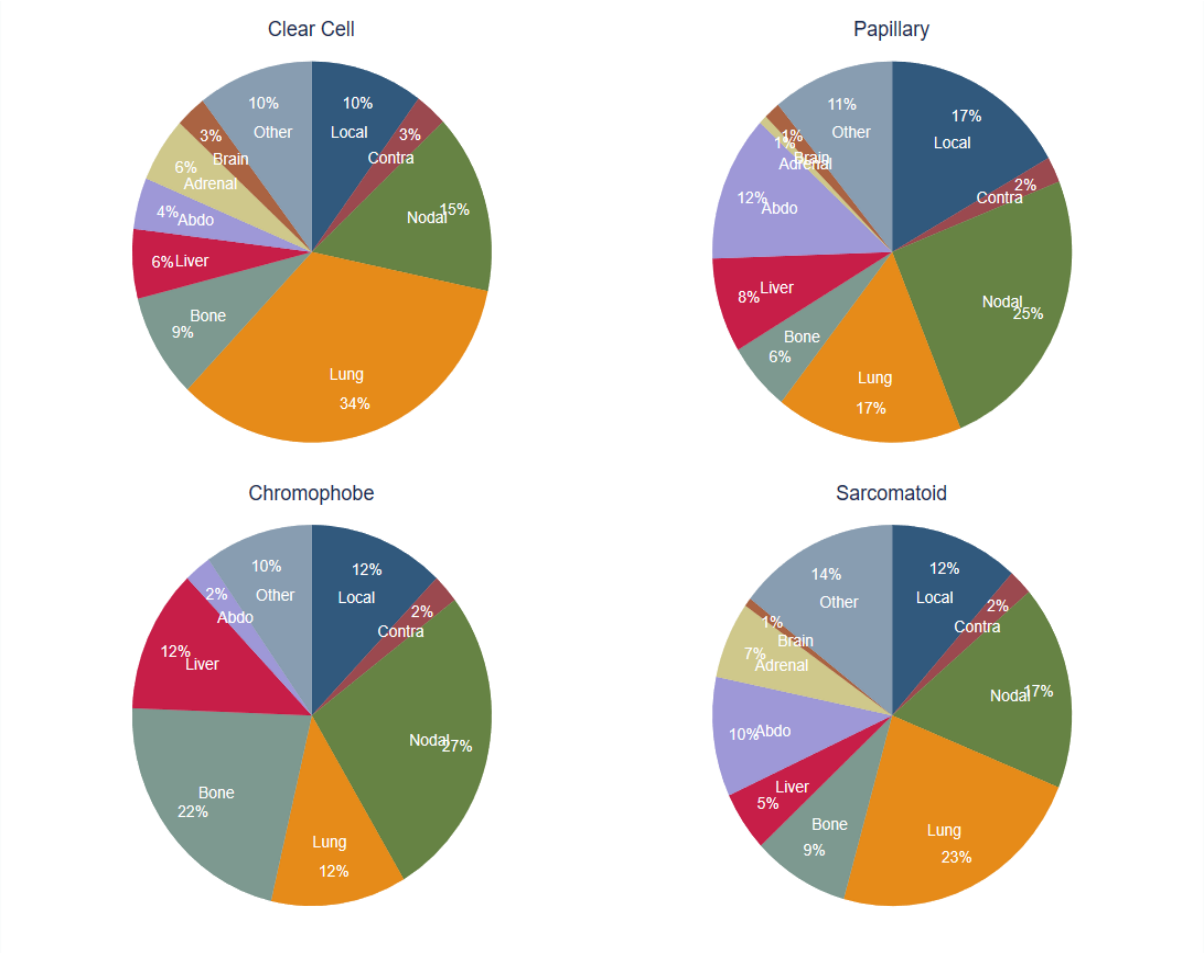
3.3.5 First relapses sites and overall survival

A breakdown of recurrence by site, stratified by histological subtype are shown in **Figure 18**. Although limited by small numbers, OS Kaplan-Meier curves for patients recurring at specific relapse sites comparing patients with ccRCC, pRCC, chRCC and sRCC are shown in **Figures 19 and 20**.

The lung, nodes, bone and liver were the commonest single sites of relapse amongst the four histological subtypes, which is consistent with the literature. For patients with pRCC (25% 34/137) and chRCC (27% 11/41) relapses involving distant nodal sites were commonest compared to other sites and proportionally more prevalent compared to the ccRCC cohort (15% 210/1375). In patients with pRCC, distant nodal relapses were associated with poorer OS compared to distant nodal relapses in the ccRCC cohort. Of note, relapses at all abdominal and lung sites excluding local sites were associated with poorer OS for patients with pRCC compared to those with ccRCC.

In the chRCC group, relapses to bone (22% 9/41) and liver (12% 5/41) were proportionally high whereas relapses to lung (12% 5/41) were proportionally low compared to those with ccRCC, pRCC and sRCC. Relapses at most organ site in the chRCC cohort exhibited favourable OS compared to relapses at the same sites for the other histology's. Of notable exception were relapses to the lung, which exhibited uncharacteristically poor OS**.**

Figure 18 Sites* of first recurrences by histological subtype



	SORCE and ASSURE N (%)										
Histology	**Local**	**Lung**	**Nodal**	**Liver**	**Bone**	**Adrenal**	**Intra-abdominal deposit**	**Contra**	**Brain**	**Other****	**Total**
Clear Cell	139 (10%)	466 (34%)	210 (15%)	81 (6%)	122 (9%)	76 (6%)	60 (4%)	40 (3%)	38 (3%)	143 (10%)	1375
Papillary	23 (17%)	23 (17%)	34 (25%)	11 (8%)	8 (6%)	1 (1%)	17 (12%)	3 (2%)	2 (1%)	15 (11%)	137
Chromophobe	5 (12%)	5 (12%)	11 (27%)	5 (12%)	9 (22%)	0 (0%)	1 (2%)	1 (2%)	0 (0%)	4 (10%)	41
Sarcomatoid	16 (12%)	32 (23%)	24 (17%)	7 (5%)	12 (9%)	9 (7%)	14 (10%)	3 (2%)	1 (1%)	20 (14%)	138

*Relapse site at first report of recurrence. Some patients recurred at more than one site

Figure 19 Overall survival Kaplan-Meier curves comparing histological subtypes, stratifying by site of first relapse

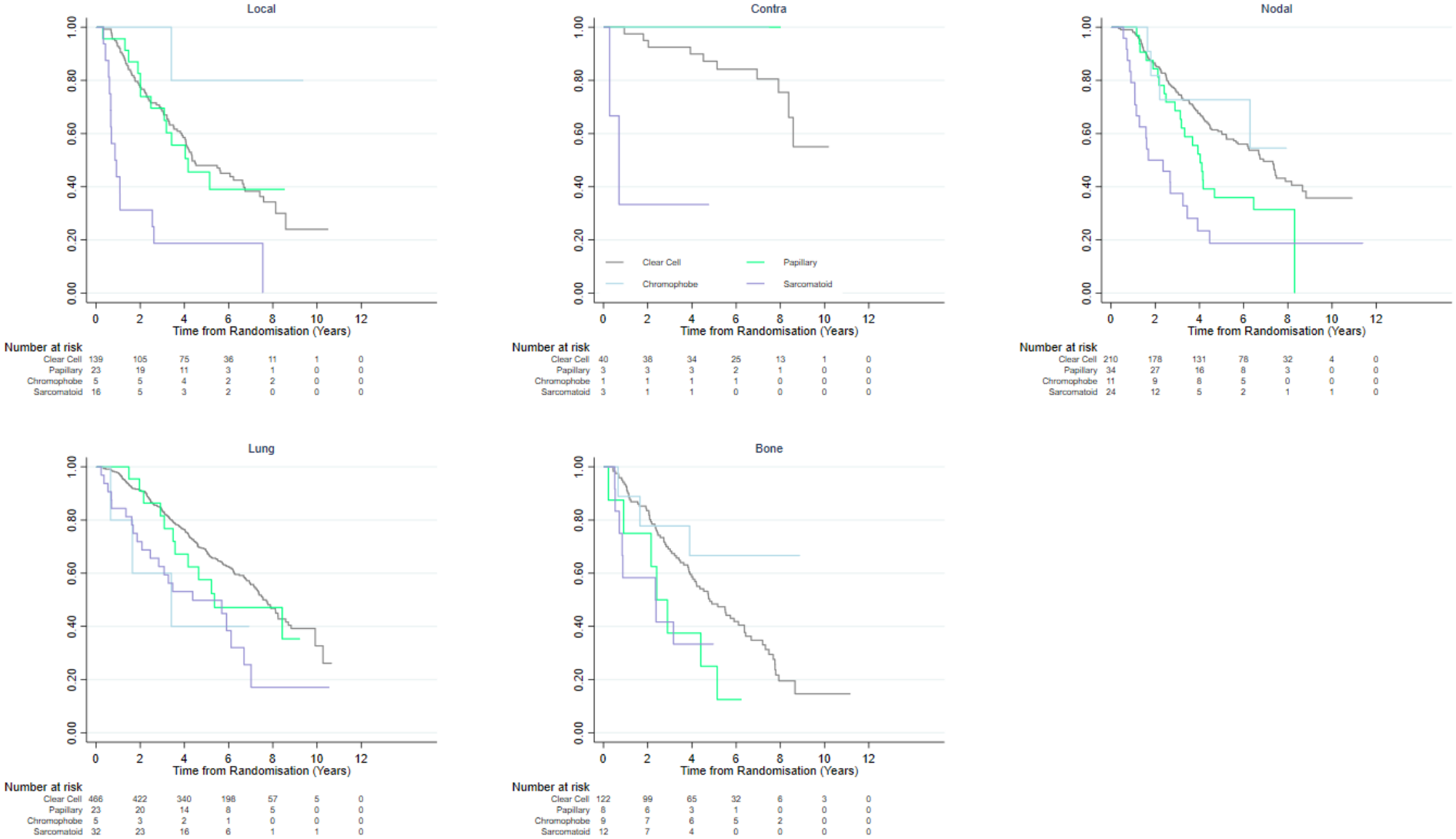
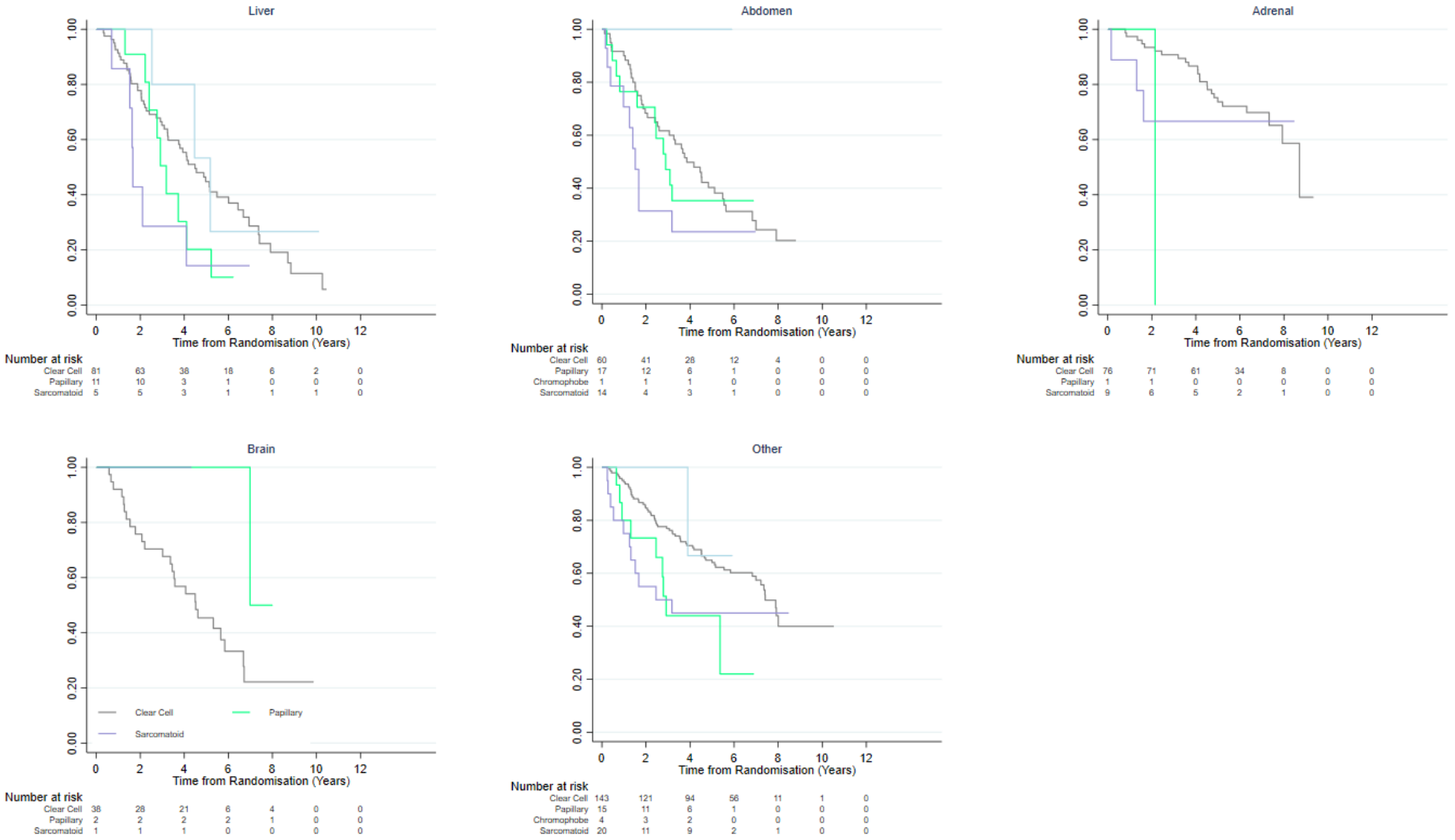


Figure 20 Overall survival Kaplan-Meier curves comparing histological subtypes, stratifying by site of first relapse



3.3.6 Evaluation of the pattern of relapses according to location

Disease free survival

In order to enhance prognostication in the post progression setting for patients with pRCC, chRCC and sRCCs, cause-specific multivariate cox models (**Appendix Table H**) were used to evaluate the effect of histology on DFS with respect to chest or abdominal relapses.

Sarcomatoid histology was associated with double the risk of relapsing to the abdomen compared to those with ccRCC; HR 2.13 (95% CI 1.59-2.86 p<0.001) and over double the risk; HR 2.65 (95% CI 1.81-3.88 p<0.001) of relapses to the chest compared to ccRCC, when adjusting for T stage, nodal involvement, performance status and study. For patients with pRCC the risk of relapsing to the abdomen (HR of 1.33 (95% CI, 0.984-1.91 p=0.063)) and chest (HR 1.1 95% CI, 0.706-1.71 p=0.105) were statistically similar to that of ccRCC. For chRCCs a HR of 0.663 (95% CI, 0.383-1.15 p=0.142) suggested a trend towards a reduced risk of abdominal relapse compared to ccRCC.

DFS was compared between those relapsing in the chest and abdomen for each subtype. **Figure 21** shows that for patients with ccRCC, DFS did not vary statistically dependent on location of first relapse (chest vs abdomen HR 0.92 CI, 0.81-1.07 p=0.289). This pattern was mirrored when comparing chest vs abdominal relapses in the sRCC cohort (HR 0.95, CI, 0.62-1.47, p=0.827). In contrast, for patients with pRCC, relapses involving the abdomen tended towards a worse DFS compared to those who relapsed in the chest, shown by a HR of 0.66 (CI, 0.37-1.17 p=0.155). In the case of chRCC, although chest relapses were rare, leading to imprecise survival estimates, they were associated with worse DFS compared relapses to the abdomen (HR 3.17 CI, 0.83-12.15 p= 0.093).

Overall survival

The effect of histology on OS, with ccRCC as the reference category, was examined in patients who relapsed first in the abdomen, the chest, the abdomen and chest. Multivariable cox models were again adjusted for T-stage, age, performance status and study, (**Appendix Table I**). Those with sRCC relapsing first in the abdomen were associated with almost four times the risk of death compared to those with ccRCC; HR 3.61 (95% CI, 2.53-5.15) p<0.001) and those relapsing in the chest were associated with over twice the risk of death (HR 2.61 (95% CI, 1.33-3.42 p<0.001).

Notably, patients with pRCC who relapsed to the abdomen were associated with almost double the risk of death compared to those with ccRCC (HR 1.7 (95% CI, 1.15-2.5 p<0.001) whereas chest relapses were not associated with a statistically significant increased risk of death; HR 1.21 (95% CI, 0.656-2.24 p 0.639). For patients with chRCC those relapsing in the abdomen were associated with a reduced risk of death; HR of 0.388 (95% CI 0.123-1.22 p=0.105) although statistical significance was not reached.

For each histology, OS was compared in patients who relapsed first in the chest versus the abdomen. **Figure 22** shows that for patients with ccRCC, relapses first to the lungs were associated with better OS (median OS 7.78 (IQR 4.43, NR)) compared to those who relapsed in the abdomen (median OS 7.42 (IQR 3.27, NR), HR 0.76 (05% CI, 0.61-0.95, p=0.016). This pattern was mirrored in patients with pRCC who relapsed in the chest compared to those relapsing in the abdomen (HR 0.54 (95% CI, 0.23-1.26 p=0.152)) and in patients with sRCC, (HR 0.59 95% CI, 0.35-1.00, p=0.051).

Figure 21 Comparison of disease- free-survival in patients who first relapsed to a. the chest b. the abdomen c. the chest and abdomen, stratified by histological subtype

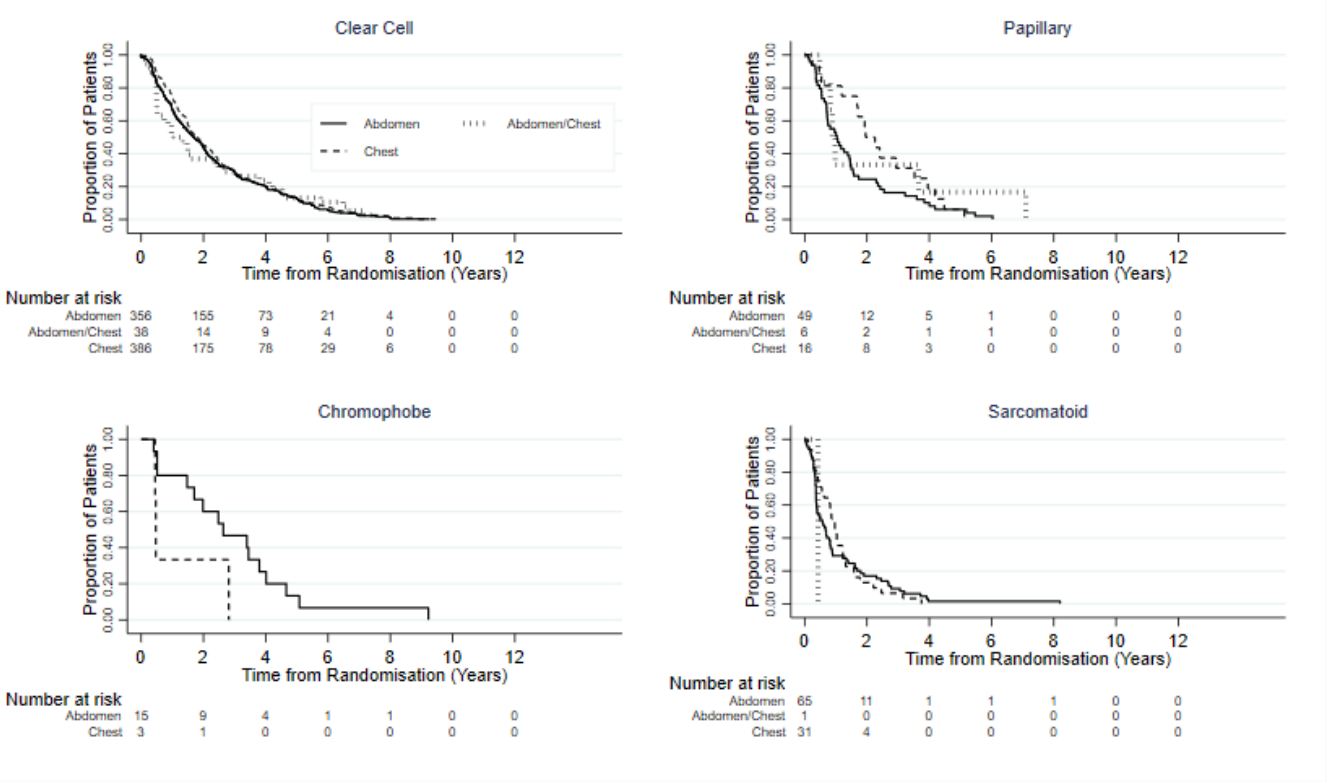
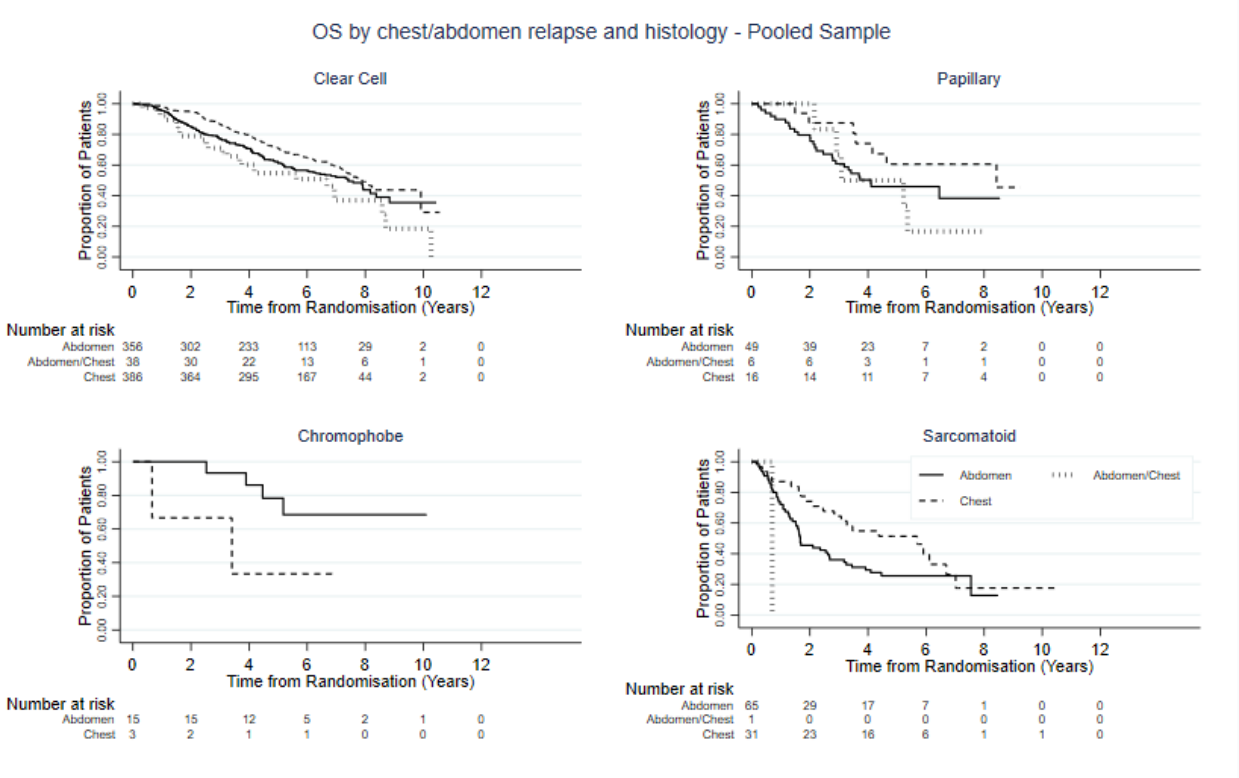


Figure 22 Comparison of overall survival in patients who first relapsed to a. the chest b. the abdomen c the chest and abdomen, stratified by histological subtype



3.4 Discussion

Relapse patterns and survival outcomes in patients with non-ccRCCs are sparsely and inconsistently reported in the current literature. This study represents the largest contemporary comparison of intermediate and high-risk patients with ccRCC, pRCC, chRCC and sRCCs, the four commonest non-ccRCC subgroups recruited to clinical trials. The use of phase three clinical trial data, provided detailed and standardised protocol driven follow-up, enabling the clinical profiles for each histological subtype to be clearly defined.

The pRCC cohort comprised a distinctly smaller proportion of the total cohort (8%) than that noted in the literature (10-15%) (**Table 16**). Patients with pRCC were the same age at diagnosis (57 years) as those with ccRCC. They exhibited a higher male to female ratio (3.5:1) than previously reported [110]. More patients with pRCC were of lower stage, (34% T2, 54% T3a-4) compared to ccRCC (66% T3a-4). This may explain why 11% of pRCCs were resected with partial nephrectomy compared to 4% of ccRCCs. Interestingly, the pRCC cohort exhibited similar histological characteristics to those with sRCC, including high rates of poor prognostic features such as baseline nodal involvement and coagulase tumour necrosis. Correspondingly, the pRCC cohort exhibited relapse rates, median TTR, disease free and overall survival outcomes distinctly worse than previously reported for the subtype [152]. Patients with pRCC who relapsed in the abdomen were associated with almost double the risk of death compared to patients with ccRCC (HR 1.7 (95% CI 1.15-2.5 $p<0.001$). Sites with worse prognosis included the liver, abdominal deposits and lymph-nodes. Although these are exploratory findings, they show that the intermediate and high risk pRCCs, although fewer in number, have notably poor prognosis. In addition, the location of their initial relapse may provide additional clinically useful prognostic information for counselling patients.

In contrast, patients with intermediate and high risk chRCCs exhibited indolent histological features, being the smallest in size, (55% of patients were T2 or less, versus 33% of ccRCCs, 45% of pRCCs, 29% of sRCCs) and less than 10% had involved nodes. They had the lowest relapse rate (17% chRCC compared to 37% ccRCC, 36% pRCC, 60% sRCC) occurring over the longest period and the best overall survival, which is consistent with the literature. Arguably, an important consideration given their favourable prognosis is whether patients with chRCC should be excluded

from future adjuvant RCC clinical trial cohorts, where the risk of severe toxicities particularly in the context of combination ICI-ICI and ICI-TKI agents, are likely to outweigh the added survival benefit. Inclusion of patients with chRCC should perhaps be limited to those with poor prognosis phenotypes within the intermediate and high-risk group. This study showed that, chest relapses, although rare, were associated with a trend towards worse DFS and OS (**Figure 21 and 22**) and were shown to mirror the survival of patients with sRCC chest relapses. Rates of 8-9% of sarcomatoid differentiation in chRCCs have previously been reported [122]. In this present study, 8/201 (4%) of patients with chRCC had sarcomatoid features. Within this group, the relapse rate was 38% (3/8), (compared to 17% of chRCCs overall), of which 33%, (1/3) relapsed first to the lung. Therefore, the association between sarcomatoid differentiation and poorer prognosis lung metastases in chRCC warrants further investigation.

This study confirms that tumours exhibiting sarcomatoid differentiation are associated with accelerated progression of disease and death (**Figure 14 and 15**) compared to patients with pure ccRCC, pRCC and chRCC. Presence of sarcomatoid RCC was associated with over two times the chance of relapsing to the chest and to the abdomen and over three times the risk of death compared to those with ccRCC, regardless of relapse site. The recent EMA and FDA approval of adjuvant pembrolizumab for patients with sRCC, may improve upon their survival outcomes. A post-hoc survival analysis of patients with sRCC enrolled on the Keynote-564 trial would be informative.

There remains a poor understanding of the molecular drivers linked to poor prognostic disease in patients with high-risk RCC and also those underpinning sarcomatoid transformation in RCC. TRANSORCE was a sample collection that took place alongside SORCE, in which nephrectomy samples from 1500 enrolled participants were collected. The next step will be to conduct longitudinal gene transcription and digital histopathological analyses on the baseline tumour and metastases from patients who relapsed at sites conferring poor prognosis e.g. nodal, liver and bone in pRCCs, and from those with sRCC. It will be possible to compare them to samples from patients who remained relapse free. This may in-turn reveal driver mutations associated with tumour invasion and metastasis and potentially targetable immuno-histochemical prognostic markers.

In terms of relapse patterns, each of the non-ccRCC subtypes exhibited characteristic behaviours. The results are broadly consistent with the literature reported for patients with pRCC [152, 153] and chRCC [152-154] and sRCC [126, 155]. Of interest, relapses to lymph-nodes were proportionally high across all non-ccRCC subgroups compared to other sites. This finding is particularly relevant for the pRCC and sRCC cohort where relapses to lymph-nodes were common and conferred a poorer prognosis compared to nodal involvement in patients with ccRCC (**Figure 19**). In the pRCC group, metastases found in distant nodes were proportionally higher in patients with nodal involvement at baseline; 28% (15/54) of patients with N1+ disease versus 10% (22/215) of patients with clear nodes at nephrectomy. Therefore, in the setting of pRCC, preferential metastatic seeding via regional lymph-nodes to distant nodal groups may be hypothesised. This highlights the need to further investigate the anatomy and pattern of spread in patients with pRCC with early nodal involvement. Studies evaluating the use of diffusion-weighted imaging to detect functionally active nodal involvement in head and neck cancers may be of value [157]. Going further, studies may inform a prognostic role for sentinel lymph node biopsy for selected patients with high risk pRCC, akin to the practice in breast cancer, and may also support a therapeutic role for upfront local lymph-node dissection [151].

There is currently no consensus on the optimal surveillance schedule for patients with non-ccRCCs and no studies to date that compare the efficacy of various surveillance intervals and imaging modalities. Current EAU recommendations [20] stratify prognostic risk for patients with non-clear cell histology's based on TNM and Fuhrman grade without clear evidence base, (**Table 23**)**.** Findings from this study support the need for histology-specific post nephrectomy surveillance. Following discussions with members of the SORCE TMG, preliminary suggestions have been made for patients with pRCC, chRCC and sRCC at intermediate and high Leibovich risk of relapse, (**Table 24**). For patients with chRCC, it was proposed that imaging can be streamlined to reflect their low risk of relapse and favourable prognosis overall. For patients with intermediate and high-risk pRCCs, enhanced surveillance in the first year should focus on detection of abdominal relapses, which present early and confer a poor prognosis. For patients with pRCCs and sRCCs, relapse rates drop after the fourth year, therefore the frequency of proposed surveillance should reduce accordingly. Of note, it remains uncertain whether altering the timing of relapse detection impacts on the overall course

of disease. I was originally intending to conduct an analysis comparing imaging schedules of patients with similar Lebovich risk from ASSURE and SORCE, correlating timing of relapse detection with overall survival. However, I was unable to acquire the relevant data from ASSURE. A head-to-head evaluation of relapse detection and survival in a cohort of patients randomly assigned to the current generic EAU surveillance schedule compared to a cohort following my enhanced histology-specific recommendations would be informative future work.

There are several limitations of this study. Firstly, as patients with low risk of relapse or poorer PS were excluded from trial entry, the applicability of results to all patients is uncertain. Arguably focusing on the higher risk cohorts, enhancing the evidence base for their post-nephrectomy surveillance and prognostication is of highest clinical impact. Importantly, as patients with poor performance and those relapsing within ninety days of nephrectomy status were excluded from trial recruitment, sites that confer poor prognoses (e.g. brain) may have been underrepresented. Secondly, in some cases, small event numbers lead to imprecise statistical estimates particularly for the chRCC group, despite pooling data from two large trials.

Finally, there are inherent limitations to pooling dataset from clinical trials. Firstly, variability in the way data was captured in the two trials led to inconsistencies in the reporting of factors such as sarcomatoid histology and relapse sites between the trials. This may explain why there were significantly more patients with sRCC in the ASSURE cohort. There were also differences in the scoring systems used in each trial. For example, SORCE used a modified nuclear grading system that simplified and selected the worst WHO/International Society of Urologic Pathologists (ISUP) features at each grade, whereas ASSURE used Fuhrman grading. ASSURE did not collect Leibovich score components for all participants, so I was unable to calculate Leibovich scores for the entire cohort. However, overall, the use of protocol driven follow-up data from two clinical trials vastly reduced the data point variability and missing data and this is a model for future subtypes-specific analyses.

3.5 Conclusion

This study shows that phase three trial data can be used to delineate the clinical behaviours of patients with higher risk non-ccRCCs. It also shows how it can be used to guide subtype-specific surveillance recommendations.

In RCC, histological subtype remains an important predictor of relapse behaviour, and of survival. The intermediate/high risk pRCC group, although proportionally small, represents a poorer prognostic group than previously reported, particularly those relapsing in abdominal sites. Chromophobe histology confers a favourable prognosis, therefore their inclusion in adjuvant clinical trials should be carefully considered. Poorer outcomes for patents with higher risk pRCC and sRCC compared to those with ccRCC suggest that current standards of care for surgical resection, systemic therapy and surveillance largely based on ccRCC data are insufficient for these patients.

Finally, within each subtype, a heterogeneity of survival outcomes exists and shows variability with sites of initial relapse. Parallel translational studies are required to evaluate whether unique genetic and molecular signatures that correspond to relapses at favourable and poor prognosis sites can be identified. These may eventually generate biomarkers with prognostic utility or useful molecular targets that pave the way for molecularly stratified adjuvant RCC trial designs.

Table 23 EAU guidance for surveillance following treatment for RCC [20]

	3 m	6 m	12 m	18 m	24 m	30 m	36 m	>3y	>5y
Low risk For ccRCC: Leibovich Score 0-2 For non-ccRCC: pT1a-T1b pNx-0 M0 and histological grade 1 or 2.	-	CT	-	CT	-	CT	-	CT every 2 years	-
Intermediate risk For ccRCC: Leibovich Score 3-5 For non-ccRCC: pT1b pNx-0 and/or histological grade 3 or 4.	-	CT	CT	-	CT	-	CT	CT every year	CT every 2 years
High risk For ccRCC: Leibovich Score ≥ 6 For non-ccRCC: pT2-pT4 with any histological grade or pT any, pN1 cM0 with any histological grade	CT	CT	CT	CT	CT	-	CT	CT every year	CT every 2 years

Table 24 Histology-specific surveillance recommendations for patients with intermediate and high risk RCC based on data from this study and SORCE TMG discussion

	3 m	6 m	12m	18 m	24 m	30 m	36 m	42 m	48m	>4 to 6y
pRCC* **Leibovich scores >3** mTTR 1.34 years (IQR 0.76, 2.59)	CT AP	CT CAP	CT AP	CT CAP	CT CAP	CT CAP	CT CAP	-	CT CAP	CT CAP every 2 years
	3 m	6 m	12 m	18 m	24 m	30 m	36 m	42 m	48m	>4 to 6y
sRCC** **Leibovich scores >3** mTTR 0.79 years (IQR 0.50, 1.55)	CT CAP	CT CAP	CT CAP	CT CAP	CT CAP	CT CAP	CT CAP	CT CAP	CT CAP	CT CAP every 2 years
	6m	12m	24m	36m	48m	60m	>5-10y			
chRCC*** **Leibovich scores >3** mTTR 2.72 years (1.07- 4.11)	-	CT CAP	CT CAP	CT CAP	CT CAP	CT CAP	CT every 2 years			

*pRCC; majority of relapses occur within two years, and the rest occur steadily until year six. Abdominal relapses are associated with worse outcomes compared to chest relapses, supporting intensified abdominal surveillance with an additional CT abdomen/pelvis 3 months after nephrectomy followed by a CT CAP at 6 months, then 6 monthly CTCAPs thereafter for 3 years, yearly CTCAPs to five years and two yearly CTCAPs to six years. Most relapses are captured by following this schedule including the 8% of pRCC relapses occurring beyond 5 years.

**sRCC; majority of metastases occur within the first year necessitating three monthly imaging in the first six months, reducing to six monthly to year four. Relapses exhibit a steep drop off beyond four years, justifying reduced surveillance beyond this point.

***chRCC; relapse rates are low and confer good prognosis. yearly CTs to year five and two yearly CTs to year ten are sufficient for this group and captures relapses that occur beyond the five-year point.

For all non-ccRCCs, relapses to abdominal and chest sites occur both before and after the median time to relapse, justifying use of a CT CAP throughout the surveillance period, unless specified.

Chapter 4: Patterns of relapse and clinical outcomes in patients with RCC treated within SORCE

4.1 Introduction

4.1.1 Overview

20-40% of patients with renal cell carcinoma (RCC) will develop recurrent disease within five years of radical nephrectomy [158]. Many of these develop widespread metastases (multiple sites in several locations) for which palliative systemic treatment currently represents the standard of care [159]. RCC metastases commonly involve the lungs, distant lymph-nodes, bones, and liver and (more rarely) the other viscera and endocrine organs (**Chapter 3, Figure 18**).

Some patients develop single or few metastases, limited in location, so-called oligo metastases for which optimal treatment strategies are evolving. The concept of oligometastatic disease was initially developed in 1995 by Hellman and Weichselbaum [160]. They defined an intermediate disease state, clinically and biologically distinct from those with widespread progression, in which the patient develops less than five metastatic deposits and exhibits limited metastatic capacity that may be amenable to complete removal [160]. Niibe and Hayakawa subsequently categorised the oligometastatic state not by number of deposits but according to whether the primary tumour is controlled and also according to the timing of the recurrence [161]. 'Synchronous metastases' are those present at the time of surgery and 'metachronous metastases' are those detected at least three months after the removal of the primary tumour [161]. The time interval differentiating between synchronous and metachronous relapse is debated. Broadly, oligo-recurrence represents a condition defined as a limited number of recurrent metastatic sites, in which the primary tumour is controlled and recurrent metastatic sites are potential candidates for curative treatment approaches [162].

The precise definition with respect to number of metastases, timing and extent of organ involvement that encompasses the oligometastatic state in RCC and in other tumour types lacks clarity [163, 164]. Clinical and biological drivers of the oligometastatic phenotype are uncertain and the optimal treatment for patients with RCC who develop oligo-recurrence lacks high-level evidence in RCC [165]. For some patients with oligo-recurrence whose metastatic disease can be fully resected (or at

least locally treated) many centres opt for focal modalities; surgery, ablation or radiosurgery, to completely remove the secondary recurrence before initiating systemic treatments or they delay systemic treatments until further progression of disease.

The term ‘M1NED’ refers to the category of patients with no evidence of disease (NED) after complete resection of oligo-recurrence (M1) [55]. For these patients the benefit (or otherwise) of ‘adjuvant’ therapy with immune-checkpoint inhibitors (ICIs) (i.e. after complete local therapy to all metastatic sites but before further systemic progression) is currently being evaluated (**Table 44**). Recent outcome data for patients with M1NED are conflicting, in part due to inconsistent definitions for the M1NED category in all of the referencing phase three clinical trials [55] [56].

Data from SORCE provided an opportunity to precisely characterise patients with RCC who developed oligo-recurrence after radical nephrectomy and compare their outcomes to those who developed widespread metastatic disease. In this chapter I defined an oligometastatic cohort as those with first metastases in a single anatomical organ. I evaluated their survival outcomes compared to patients with metastatic disease in multiple anatomical organs at first relapse. I examined clinical features at the point of progression, aiming to unpick the differential outcomes noted between the patient groups. I compared survival outcomes of patients who received local treatments to those who received palliative or systemic anti-cancer treatments upon first relapse, in order to examine the added utility of including treatment intent in prognostic guidance for patients who relapse. Finally, I defined an oligometastatic RCC cohort suitable for inclusion in adjuvant clinical trials.

4.1.2 The oligometastatic phenotype in other cancers

Improvements in accessing minimally invasive surgical procedures (e.g Video Assisted Thorascopic Surgery (VATS) and liver metastatectomies) and radiotherapy techniques (e.g. Stereotactic Body Radiotherapy – SBRT) that can treat targeted areas to high doses whilst sparing adjacent normal tissues have shaped the definitions and clinical approach to patients with oligometastatic disease. For certain tumour types like colorectal and lung cancers, weak to moderate evidence underpins established oligometastatic treatment protocols. Whereas in other cancer types for

example prostate and renal cancer, the evidence base is sparser and treatment paradigms continue to evolve.

Colorectal cancer (CRC)

In the setting of CRC, oligometastatic disease is defined in practice as up to five predominantly visceral and occasionally nodal metastases in up to two anatomical sites which can be completely resected/ablated. The precise size and volume of individual metastases is not included [166]. In the absence of randomised trials comparing surgical with nonsurgical disease management, surgery remains the established standard approach rather than active monitoring for further progression for patients with resectable pulmonary or hepatic metastases (sites conferring better prognosis than other locations in CRC) [167].

Despite this, the prognostic impact of developing oligometastatic disease in CRC remain poorly defined. Broadly, features that favour longer survival in metastatic CRC have been identified as the presence of a single metastasis, a long interoperative interval [168] and a low/normal (<2.9 ng/mL) blood carcinoembryonic antigen (CEA) level [169]. None of the informing multivariate analyses evaluated modality of treatment upon progression or the extent of oligometastatic disease.

A 2013 meta-analysis of the twenty-five largest (all single arm) studies comprising patients undergoing lung metastasectomy at an average interval of two years after primary resection from 2000 to 2011 reported a five-year overall survival (OS) of 41% [170]. Four parameters negatively associated with survival; short (timepoint not specified) disease-free-interval (DFI) between primary resection and development of lung metastases, multiple lung metastases, positive hilar and/or mediastinal lymph nodes and elevated pre-thoracotomy CEA.

The PulmiCC trial; pulmonary metastasectomy in colorectal cancer (NCT01106261) was the first RCT to systematically investigate lung metastatectomy versus active monitoring in patients with potentially resectable CRC lung metastases [171]. Unfortunately, the trial was stopped early after 65 patients were enrolled due to slow recruitment. There were no other interventions in the first six months, no crossovers from control to treatment and no treatment related deaths or adverse events. It found a non-significant improvement in median survival after metastatectomy (3.91 years (95% CI, 2.99- not reached (NR)) compared with 3.38 years (95% CI, 3.11-NR) in

matched controls, (hazard ratio (HR) 0.82 (95% CI, 0.43-1.56)). Given the small participant numbers and large overlap in confidence intervals, the benefit of metastatectomy for patients with CRC and limited pulmonary metastases remains uncertain.

In the setting of CRC with resectable (leaving no tumour at the margin) liver metastases, there have been no trials directly comparing outcomes for those undergoing liver resections versus matched controls. In patients with resectable liver metastases, surgery or ablation is frequently employed alongside radical management of the primary tumour and is quoted to be associated with 40% five-year survival rates [172]. This compares very favourably to those with widely metastatic CRC where the 5 year survival is reported as less than 20% [173]. On this basis, National Institute for Health and Care Excellence (NICE) recommend a radical approach for resectable CRC and liver metastases. In addition, these patients are included in some adjuvant CRC trials. The AddAspirin trial (NCT02804815) [174], a large, placebo-controlled basket RCT examining the effect of aspirin after completion of standard adjuvant treatments is an example.

Prostate Cancer

The standard of care for metastatic prostate cancer is androgen deprivation therapy (ADT) alone with or without the addition of abiraterone or docetaxel chemotherapy. Radical treatment of oligometastatic prostate cancer is the subject of ongoing debate. Typically, it has been defined based on the number of metastases and involved sites, but several other variables including presence of synchronous (<3 months of primary diagnosis) versus metachronous metastases (3-6 months after completing curative treatment) and whether the patient is hormone-sensitive (HSPC) or hormone-resistant (HRPS) have also been included. For patients with metachronous oligometastatic HSPC, prospective clinical trial data are available, albeit from relatively small non-randomised phase two datasets. Metastasis directed therapy with SABR in STOMP (NCT01558427) and ORIOLE (NCT02680587) reported a survival advantage compared to surveillance alone [175] [176]. Both trials defined oligometastatic disease as between 1-3 metastases, ORIOLE stipulated a size limit of 5cm in largest axis and STOMP excluded extracranial metastases. The trials used different imaging for screening; STOMP used choline positron emission tomography (PET) while ORIOLE used prostate-specific membrane antigen (PSMA) PET. SABR-COMET another 99

patient phase two trial, enrolled a multi-tumour site cohort (prostate n=16/99), stipulating a controlled primary site and up to five metastases amenable to SABR [177]. It randomised participants to either palliative standard of care (SOC) or to SOC with SABR. The five-year OS was 17.7% (95% CI, 6-34%) versus 42.3% respectively (95% CI, 28% to 56%; log-rank p= 0.006). Based on these data, the 2022 Advanced Prostate Cancer Consensus Conference guidance supports the use of ADT plus local treatment of oligometastatic lesions in metastatic HSPC [178]. The definition of oligometastases provided is imprecise; an example of three bone lesions is given. In the setting of synchronous oligometatstaic HSPC, there is currently no consensus on the management. Radiotherapy to the primary and all known sites of metastatic disease (≤5 metastases), is being prospectively evaluated against systemic treatments alone, in several phase three trials including a new comparison within the STAMPEDE trial (NCT00268476) [177].

4.1.3 Biological basis of the oligometastatic phenotype in RCC

The discovery of genetic features linked to distinct progression phenotypes in RCC is a step towards understanding the diversity of biological behaviour exhibited by RCCs after curative nephrectomy.

A small number of studies support a molecular rationale for an oligometastatic state in RCC. Wuttig *et al.* conducted transcriptome wide expression profiles of samples from laser resected RCC pulmonary metastasis. They identified 135 genes that were differently expressed in two groups characterised by ‘few’ (<8) or ‘many’ (>16) pulmonary metastases [179]. Among the upregulated genes were those known to be involved in metastatic spread like RASGEF1A, AGR3 or CEACAM6 and also positive cell cycle regulators BIRC5, PTTG1 and CKS2. The TracerX Renal Consortium analysis (NCT03226886, NCT03004755) was the largest systematic study of 575 primary and 335 metastatic ccRCC biopsies across 100 patients over multiple time-points [180]. Tumours associated with ‘rapid progression’ were defined as those with multiple sites of progression within six months of primary surgery. They were compared to those with ‘attenuated progression’, defined as presenting with a single site of progression in less than 6 months; or multi-site progression more than 6 months after surgery, often limited to a single organ that were able to be controlled with additional surgery or radiotherapy. The two distinct patterns of metastatic behaviour

were linked to unique genetic and molecular characteristics. The rapid progressors were enriched for certain genetic components including multiple clonal drivers, VHL wild-type, loss of 9p, BAP1 alterations and lower intertumoral heterogeneity (ITH), compared to those with attenuated progression. Whereas the attenuated progressors were characterised by higher ITH and were enriched for PBRM1→SETD2 and PBRM1→PI3K genetic markers.

4.1.4 Current guidelines in RCC

With improvements in surgical, radiotherapy and ablative techniques, there are now more options for the radical treatment of renal metastases, prior to or alongside systemic treatments. Although RCC is traditionally considered a radioresistant malignancy, novel techniques such as SBRT that deliver highly collimated radiation at higher-doses per fraction in fewer fractions, to a precisely defined target area have produced curative results [181] [182]. Additionally, ablative techniques may be an option for local control in certain situations [183]. For single or oligometastatic relapses, both European Society of Medical Oncology (ESMO) and European Association of Urology (EAU) guidance are vague and promote the consideration of radical strategies by local multi-profession teams on an individual patient basis, with treatment being guided by local consensus and also the availability of expertise [14, 159]. ESMO guidance highlights that there 'might be a role for delaying systemic treatment and associated toxicity' but it offers no guidance on the identification of cases that should be prioritised for local treatment of metastases.

Robust data supporting a potential curative benefit and prolonged OS following radical tumour removal for patients with resectable oligometastatic RCC is dominated by retrospective, non-randomised studies and systematic reviews that focus on the use of surgical modalities [184] [185]. ESMO guidance cites one systematic review of sixteen studies including 2350 patients that pointed to an OS and cancer-specific survival (CSS) benefit for patients after complete surgical resection (CSR) compared with incomplete or no metastasectomy [186]. Conclusions were limited by low quality studies and high confounding between results. A study by Pogrebniak *et al*. [187] evaluated twenty-three metastatic patients with RCC who underwent pulmonary metastatectomy, fifteen of whom had previously been treated with interleukin-based immunotherapy. Patients who underwent CSR of metastatic disease had better survival (median not reached) than those with incomplete resection (median

16 months p=0.02). Favourable subgroups included those patients with a solitary site of metastases and a DFI of greater than one year. In the context of pulmonary disease, the results were used to support a potential benefit of pursuing complete resection.

Another retrospective review of 138 mainly metastatic ccRCC patients who underwent CSR at several organ sites evaluated prognostic features within this group. It found that tumour size at metastatectomy (HR 1.18 per 1 cm, 95% CI, 1.07-1.29, p = 0.001) and the presence of sarcomatoid histology (HR 3.70, 95% CI 1.09, 12.62, p= 0.037) were significantly associated with worse CSS [188]. One earlier 278 patient retrospective analysis of surgically resected oligometastatic RCCs treated from 1980 to 1993 found that favourable features for survival were a DFI of less than twelve months compared to greater than twelve months (55% v 9% five-year OS rate; p<0.0001) [189]. The presence of solitary versus multiple sites (54% v 29% five-year OS; p<0.001) and age younger than sixty years (49% vs. 35% five-year OS) were also found to be favourable. Results need corroborating in larger contemporary data that includes a range of available radical treatment modalities.

For bone and brain RCC metastases, radical radiotherapy provides a non-invasive potentially curative approach, delivered by SBRT [190, 191]. A 61-case study examined patients with extra-cranial and intracranial metastatic RCC who underwent SBRT [191]. 74% were treated for a solitary metastatic lesion. The median radiotherapy dose was 25 Grey (range 10–52) in 5–10 fractions, which was well tolerated. The pattern of treatment failure was predominantly outside the SBRT field (one year out-field progression-free-survival (PFS) of 39% versus one year in-field PFS of 70%). The systemic treatment-free-rate was 70% and 50% at one and two years, respectively. Another retrospective study of 47 oligometastatic treatment naïve patients treated with SABR showed a 15.2 month delay to starting subsequent therapies. 38.2% (18/47) of patients did not receive subsequent treatment at a median follow-up of thirty months [190]. Although SABR and surgery for selected oligometastatic disease show promising results, without comparison arms these studies cannot definitively determine whether these approaches improve the disease course or whether the delay in starting systemic treatments affects survival or improves quality of life.

4.1.5 Risk prediction in patient who relapse after curative nephrectomy

The Leuven-Undine (LU) metastatectomy risk score [192] was the first to be developed to predict CSS for patients with RCC who underwent partial or radical nephrectomy and at least one metastatectomy between 1988 and 2011. On multivariable regression analysis [193] including data from 109 consecutive patients, a DFI between nephrectomy and first metastases of less than twelve months (HR 2.3 $p< 0.058$) featured as a poor prognostic factor for CSS along with T stage ≥3 (HR 2.8 $p< 0.01$), primary tumour Fuhrman grade ≥3 (HR 2.3, $p< 0.03$), non-pulmonary metastases (HR: 3.1, $p< 0.03$) and multi-organ metastases (HR 2.5, $p< 0.04$). LU prognostic groups A to D were generated based on these co-variates (**Table 25**). The two and five-year CSS were significantly different; group A; 95.8% and 83.1%; group B, 89.9% and 56.4%; group C, 65.6% and 32.6%; and group D, 24.7% and 0% ($p< 0.0001$).

Findings from this study indicate that alongside histological features of the primary tumour, inherent characteristics associated with the first relapse may carry prognostic significance in patients with ccRCC who undergo complete resection of oligometastases. However, without validation in larger contemporary datasets, the practical applicability of these findings remains uncertain.

Table 25 Leuven-Udine prognostic factors and groupings

Factors	Score
T stage ≥3	1
Fuhrman grade ≥ 3	1
Disease free interval ≤ 12m	1
Non-pulmonary metastases	1
Multiple sites	1

Leuven-Udine prognostic groups	
A	0-1
B	2
C	3
D	4-5

Taken from 'Survival and Impact of Clinical Prognostic Factors in Surgically Treated Metastatic Renal Cell Carcinoma' European Urology [192]

4.1.6 Effect of time to recurrence on RCC Survival

Few studies evaluate the prognostic impact of timing of first relapse, or DFI after curative nephrectomy on survival. One study included 259 patients with pT1–4 NX M0 ccRCC treated between 1981 to 2009 and compared those who recurred within and after twelve months of primary surgery to those who remained disease free [194]. The five-year CSS was 98% for those without recurrences, 53% for first recurrences after twelve months and 23% for recurrences occurring within twelve months. T stage and recurrences before twelve months were shown to independently predict CSS, on multivariable analysis ($p < 0.0001$). Another single institution study of 747 patients undergoing curative nephrectomy between 1989 and 2008, compared patients with synchronous metastases, recurrence occurring less than five years and more than five years after surgery and found five-year CSS to be 27%, 41% and 73% respectively ($p<0.001$) [195]. A comprehensive contemporary survival analysis of patients relapsing after nephrectomy with oligometastatic RCC, including a comparison of narrow relapse time points ranging from six months, twelve months, twenty-four months to beyond five years after surgery, accounting for the number of sites/metastases and assessing the treatment modalities used upon relapse has not yet been conducted. An analysis of this sort would enhance the clinical application of time to relapse specifically for patients with resected oligometastatic disease.

4.1.7 The RECUR[8] Consortium Study characterising curability of metastatic RCC

A large retrospective study [22] by the RECUR[9] consortium, utilised multi-institutional European data from 1265 patients with metastatic ccRCC treated with curative intent between 2006 and 2011 [196]. They postulated a link between the extent and characteristics of the first recurrence of disease after curative nephrectomy and OS. Their aim was to examine survival outcomes of an oligometastatic cohort defined by the number of anatomical sites of first relapse. They focused on patients with ccRCC. Recurrences were identified as ‘potentially curable’ or ‘probably incurable’ based not on their final management outcome or whether they were even selected for radical

[1] RECUR consortium, ‘the euRopEan association of urology renal cell carcinoma guidelines panel Collaborative multicenter consortium for the studies of follow-Up and recur- rence patterns in Radically treated renal cell carcinoma patients’

resection, but on the number of sites of initial metastases. ‘Potentially curable’ recurrences were defined as those confined to a single anatomical site with three or fewer metastases. ‘probably incurable’ recurrences were those with multiple sites of first relapses or more than three metastases at a single site.

Of 1265 patients with ccRCC, 286 had a recurrence (23%). 155/286 (54%) of relapses were ‘probably incurable’ and 131/286 (45%) were ‘potentially curable’. The median time-to-relapse (TTR) for ‘potentially curable’ recurrences was 25 months (Interquartile range (IQR) 11.6, 47.4) compared with 17.3 months (IQR 6.2, 40.3, p =0.004) for the ‘probably incurable’ group. The median RCC-specific-survival-after-recurrence was longer in ‘potentially curable’ patients (27.4 months, IQR 11.1, 48.3) compared to ‘probably incurable’ patients (15.2 months, IQR 5.5, 33.4, $p < 0.001$). Although the definitions of ‘potentially curable’ may have over or underestimated the proportion of patients who went on to have curative treatments after relapse, by defining groups based on disease extent at first relapse, patients undergoing a range of radical treatments were included not just those that were surgically resected. The authors showed that the ‘potentially curable’/ ‘probably incurable’ delineation provided useful prognostic information for patients at the point of first relapse. Critically, it underscored the significance of the nature of the first relapse in predicting subsequent disease trajectory.

In summary, prior studies examining prognostic factors in RCC point to TTR and markers of the extent of disease (size, number of deposits and number of organs involved) being informative predictors of survival in the oligometastatic group. However, studies are limited by including outdated and retrospective datasets and by variable quality of data assessment e.g. variable endpoints and selection criteria, inconsistent follow-up and poor standardisation of diagnosis. In addition, few studies compare treatments and outcomes of those with oligometastatic disease to those with advanced unresectable disease. Such an analysis would have significant implications for personalising the surveillance and management of patients after first relapse and for delineating the inclusion criteria for patients with resected oligometastases in adjuvant clinical trials.

4.1.8 Rationale for Analysing Patterns and Timing of Relapse in data from SORCE

Whether there is an oligometastatic phenotype in RCC with clinical characteristics and outcomes distinct from patients with advanced RCC continues to be debated. The management of patients with oligometastatic disease given the broad clinical guidance for this group is imprecisely delineated. In addition, a cohort of patients with limited metastatic disease, suitable for selection to adjuvant RCC trials is yet to be standardised. This study analysed data from patients with fully resected RCC enrolled into the SORCE trial (**Chapter 1, Figure 2**) [5]. SORCE provided prospectively collected data from patients who underwent radical nephrectomy with intermediate and high 2003 Leibovich risk of relapse (**Chapter 1, Table 10**), standardised eligibility criteria and consistent long-term follow-up. Data on anatomical sites and timing of initial relapse, performance status and first treatment on relapse were collected for participants. As such this dataset represented an opportunity to investigate the 'potentially curable' (single site) vs 'probably incurable' (multiple sites) dichotomy proposed by the RECUR consortium [22] albeit that within SORCE the number of individual metastases at each anatomical site was not recorded.

4.1.9 Research question and aims

In this chapter, I addressed the following research question and aims.

Research question

Can the timing and extent of the first recurrence in RCC predict outcomes for patients with initially fully resected (i.e. radical nephrectomy) intermediate and high risk RCC?

Research aims

- To determine the prognostic utility of classifying patients according to the extent of initial relapse i.e. those relapsing in single local or distant anatomical sites compared to those relapsing in multiple anatomical sites.

- To explore differences in the clinical characteristics at relapse of patients developing single and multiple anatomical sites of metastases. Including a comparison of performance status upon relapse, timing and sites of initial recurrence and the treatments offered upon relapse.

- To evaluate whether the extent and timing of initial relapse can be used to define prognostic risk groups in patients with oligometastatic RCC.

- To define a cohort of patients with limited metastatic RCC who may be suitable for inclusion onto adjuvant trials after receiving metastasis directed treatment.

4.2 Methods

4.2.1 Design

A retrospective cohort study examining the patterns and timing of relapses after curative nephrectomy using prospectively collected data from participants who relapsed during SORCE trial follow-up.

4.2.2 Data

SORCE enrolled patients with RCC treated surgically with curative intent and randomly assigned patients between 2007 and 2013. The SORCE trial was a multi-centre European and Australian, randomised phase three double-blind placebo-controlled study (see **Chapter 1.5.5** for trial details). It examined the efficacy and tolerability of sorafenib versus placebo in 1711 patients with resected (total or partial) RCC recruited from 147 centres. As SORCE showed no benefit for sorafenib after nephrectomy in both experimental arms (one year/ three years sorafenib), I decided to include patients from both experimental and placebo arms, as this would retain as much data as possible for this analysis. A data release request for this study was approved by the SORCE Trial Management Group (TMG) in February 2020, allowing access to the SORCE dataset.

Baseline patient and tumour characteristics, site and timing of recurrence events, subsequent treatments and survival outcomes were captured prospectively for participants via the SORCE Case Report Forms (CRFs). Values for components of the 2003 Leibovich score (**Table 10**) were also prospectively collected for each participant. Patients whose tumours had an intermediate (3-5) or high (≥6) 2003 Leibovich score at randomisation, were eligible for SORCE [5, 23]. See **Appendix Table E** for SORCE trial eligibility criteria.

4.2.3 SORCE Participants – Follow-up

Participants were eligible for randomisation within ninety days of surgery. Those who relapsed within this period as evidenced on a baseline computerized tomography (CT)

chest, abdomen and pelvis, (CT CAP) were not eligible for SORCE as they were considered likely to be metastatic at the time of primary surgery. It should be acknowledged that this study evaluates time points relative to random allocation and therefore excludes poor prognosis patients who relapsed within ninety days of surgery. The time points of follow-up imaging collected were pre-specified in the SORCE trial protocol and assessments were conducted by local investigating teams. Participants were clinically and radiologically assessed every three months until the end of year three on study, six monthly until year five, then annually until year ten. A CT scan of chest and abdomen was required prior to randomisation and then every 6 months during the first three years of study. A chest x-ray (CXR) was required 6 monthly in years one to five and annually thereafter to ten years. Clinical suspicion of recurrence was confirmed via CT, CXR, or a positive biopsy earlier than scheduled if clinically indicated.

4.2.4 Analyses Cohorts

The primary analysis cohort was generated from the subset of SORCE patients who relapsed with ccRCC, for ease of comparison with other published analyses. Patients with non-ccRCCs; papillary RCC (pRCC), chromophobe RCC (chRCC) and sarcomatoid RCC (sRCC) were included in a secondary analysis. Non-ccRCC histology's were not examined individually. Patients who remained disease free over the follow-up period and patients who died before relapsing or had a first event as a second primary were excluded.

Participants with 'single site' recurrences were defined as recurrences at a single local or distant anatomical site. Participants with 'multiple sites' first recurrences were disseminated to two or more anatomical sites. SORCE CRFs did not collect data on the number or size of metastatic lesions at a particular site, therefore it was not possible to further delineate the oligometastatic group.

4.2.5 Outcome Measures

DFS for the purposes of this analysis was defined, as the interval from randomisation to first evidence of local recurrence, distant metastasis or death from RCC. RCC-specific survival denoted 'RCC-survival' was the time from randomisation to death from RCC. An additional analysis; RCC-survival from recurrence (where the first recurrence was time point zero as opposed to the date of randomisation) was reported to directly assess

the impact of first recurrence on survival. TTR was the median and IQR time from randomisation to local or distant recurrence.

4.2.6 Statistical Methods

I wrote the statistical analysis plan for this analysis and worked alongside Matthew Burnell, a senior statistician at the MRC CTU at UCL who wrote the statical code using STATA 16.1 (StataCorp LP, College Station, TX, USA). I analysed the STATA transcripts directly for this study.

4.2.7 Missing data

For these analyses only complete observations have been included. Missing data has been reported in the tables.

4.2.8 Statistical Analysis

Descriptive statistics; categorical variables with percentages and continuous variables with median and IQRs were used to present baseline data from the 'single sites' and 'multiple sites' cohorts. TTR and RCC-survival were presented graphically assessing the degree of separation between the Kaplan-Meier curves alongside HRs comparing the survival distributions between the two groups. Percentages were presented to 0 decimal places. P-values were given to 2 significant figures. All statistical tests were 2-sided. The Chi ^{2} test paired t-test or Mann-Whitney test were implemented for categorical data comparisons, as appropriate.

Logistic regression was used to evaluate if any baseline patient and tumour characteristics were associated with developing curable metastatic disease. Patient clinical characteristics and tumour histological features collected at the point of random allocation to SORCE were included in the model with 'single site' status as the dependent variable. Odds ratios were considered significant at the 5% level.

The association between TTR and RCC-survival-from-recurrence was examined using Cox regression models. HRs comparing various time points (6m-1yr, 1-2yrs, 2-3yrs, 3-5yrs, >5yrs) after randomisation to a baseline category of >6m were presented. Leibovich score, age, gender and baseline performance status were accounted for in the models.

Factors contributing to variation in outcomes between those recurring at 'single sites' and 'multiple sites' were explored using the chi^2 test. The factors of interest were:

- performance status at relapse
- treatment upon relapse
- location of first relapse

Multivariable Cox regression models considering the same baseline factors were built for the outcome of RCC-survival-from-recurrence. Prognostic significance was assessed using hazard ratios and results considered statistically significant at the 5% level.

All measures were reported with 95% CIs. p-values were two sided.

4.3 Results

4.3.1 Baseline Characteristics of patient with ‘single sites’ and ‘multiple sites’ of recurrence

Overall, 37% (634/1711) of SORCE participants developed recurrent disease after radical nephrectomy for localised RCC. Participants were followed up for a median of 5.5 years (IQR 2.3, 6.9 years). 56% (965/1711) were excluded because they remained disease free. 7% of participants were excluded with incomplete data, a diagnosis of second cancer or a death without an event **(Figure 23)**. The analysis cohort comprised 634 participants who relapsed; 539/634 (85%) had ccRCC and 95/634 (13%) had non-ccRCCs. Of the 539 patients with ccRCC, 354/539 (66%) developed ’single sites’ of disease and 185/539 (34%) developed ‘multiple sites’ of disease (**Figure 23**). In the ’multiple sites’ group, 35% (65/185) were Leibovich intermediate risk and 65% (120/185) were high risk at the point of random allocation to the trial (i.e. post nephrectomy). In the ‘single sites’ group, 35% (122/354) were intermediate risk and 65% (232/354) were high-risk. The highest Leibovich score was 6 in both groups. Baseline tumour and patient characteristics were well-balanced between the two cohorts (**Table 26)**.

4.3.2 Baseline clinical and histological features associated with ‘single site’ status

The prognostic significance of baseline tumour and patient characteristics on ‘single site’ status were evaluated using logistic regression. Variables included age, performance status, gender, T stage, tumour size, nodal status, nuclear grade, tumour necrosis and Leibovich score. These were selected based on the published literature demonstrating their prognostic significance in patients with locally advanced RCC after undergoing curative nephrectomy. Models with and without each variable were compared using the likelihood ratio test. No variable was found to be significantly associated with ‘single site’ status upon relapse, with all confidence intervals crossing one (**Table 27)**.

Figure 23: Consort diagram

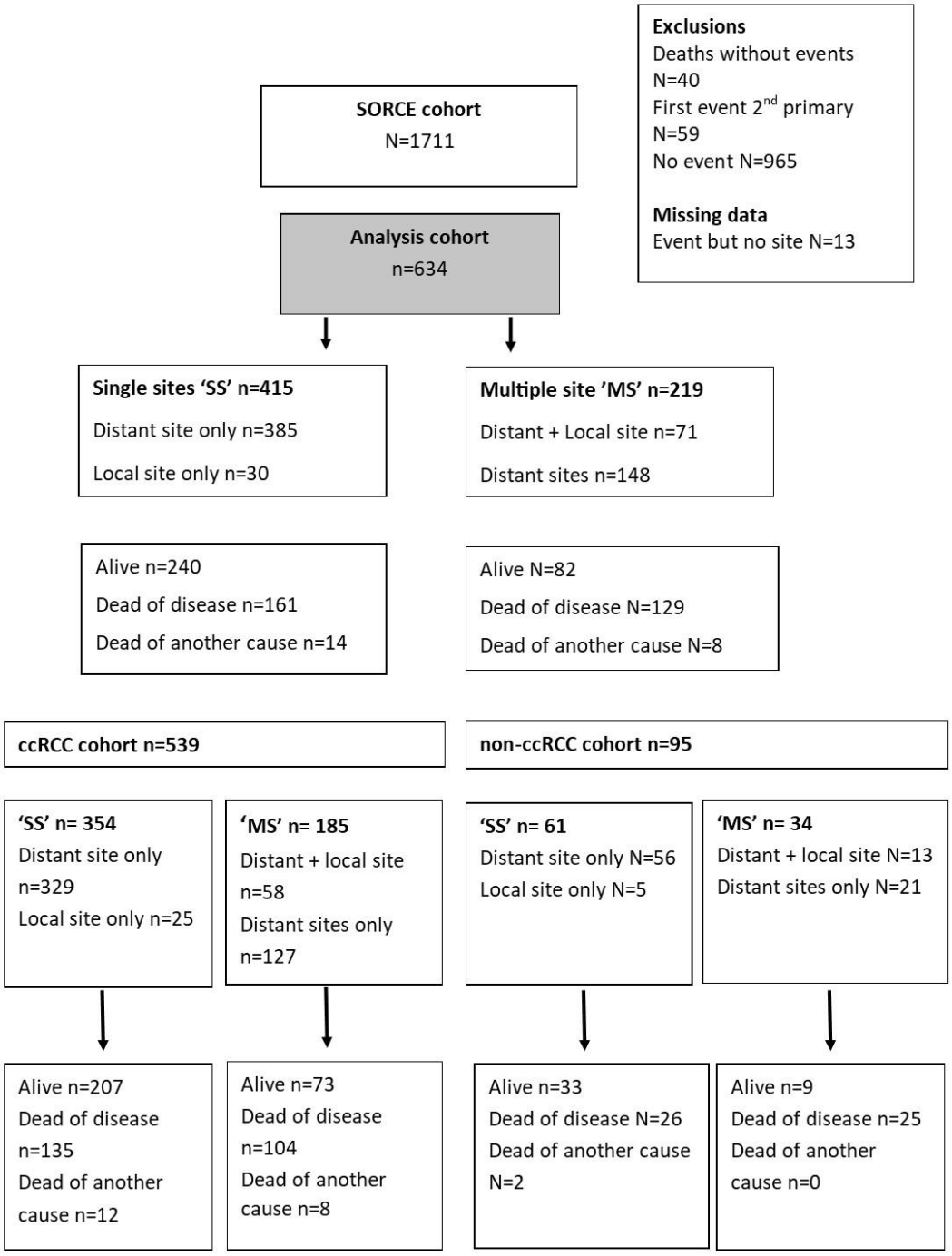


Table 26 Baseline characteristics in the 'single site' and 'multiple site' cohorts

	ccRCC (Primary cohort)			Non-ccRCC (secondary cohort)		
Variable at Baseline	**Single site N (%)**	**Multiple site**	**p-value**	**Single site**	**Multiple site**	**p-value**
Patient number	354 (65.6)	185 (34.4)		61 (64)	34 (36)	
Age at Randomisation	59.1 (52.7-80.2)	60.1 (54.3-82.9)		58.1 (54.3-61.9)	57.7 (54.3-70.0)	
Sex						
Male	248 (70)	133 (72)		49 (80)	22 (65)	
Female	106 (30)	52 (28)	0.657	12 (20)	12 (35)	0.093
Missing	0	0		0	0	
Performance Status						
0	281 (80)	138 (75)		46 (75)	30 (88)	
1	71 (20)	46 (25)	0.438	15 (25)	4 (12)	0.134
Missing	2 (<1)	1 (<1)		0	0	
Tumour stage						
pT1a	1 (<1)	2 (1)		0	0	
pT1b	25 (7)	13 (7)		3 (5)	2 (6)	
pT2	65 (18)	27 (15)		19 (31)	11 (32)	
pT3a-4	263 (74)	143 (77)	0.467	39 (64)	21 (62)	0.968
Missing	0	0		0	0	
Nodal Status						
pNx/ pN0	341 (96)	172 (93)		46 (75)	26 (76)	
pN1/ pN2	13 (4)	13 (7)	0.084	15 (25)	8 (24)	0.908
Missing	0	0		0	0	
Tumour Size						
<10	199 (56)	111 (60)		38 (62)	19 (56)	
>10	155 (44)	74 (40)	0.399	23 (38)	15 (44)	0.541
Missing	0	0		0	0	
Nuclear Grade						
1	11 (4)	8 (4)		6 (10)	1 (3)	
2	63 (18)	31 (17)		10 (16)	7 (21)	
3	188 (53)	98 (53)		24 (39)	20 (59)	
4	92 (26)	48 (26)	0.900	21 (35)	6 (18)	0.134
	0	0		0	0	
Histological Tumour Necrosis						
No	108 (31)	63 (34)		17 (29)	10 (29)	
Yes	246 (70)	122 (66)	0.401	44 (72)	24 (71)	
Missing	0	0		0	0	0.873
Leibovich Risk Group						
Intermediate	122 (35)	65 (35)		16 (27)	10 (29)	
High	232 (65)	120 (65)	0.876	45 (74)	24 (71)	0.739
Missing						
Type of Nephrectomy						
Total (radical)	329 (93)	179 (97)	0.136	57 (94)	32 (94)	
Partial	15 (3)	0		3 (5)	1 (3)	
Missing	14 (4)	6 (3)		1 (2)	1 (3)	0.827
Histology						
Papillary	-	-		28 (46)	22 (65)	
Sarcomatoid	-	-		16 (26)	6 (18)	
Chromophobe	-	-	-	7 (11)	3 (9)	
Missing/other				10 (16)	3 (9)	0.359

Primary analysis; ccRCC cohort

Table 27 baseline multivariable model for 'single site' status

Characteristic		Single site status		
	Category description	**Odds ratio**	**P>|z|**	**95% confidence interval**
Cumulative age		0.99	0.47	0.97- 1.01
Performance status	0	ref		
	1	0.69	0.09	0.44-1.06
Gender	Female	ref		
	Male	0.78	0.22	0.51- 1.17
Tumour stage	1a	0.09	0.19	0.01- 4.58
	1b	ref		
	2	1.78	0.44	0.42- 7.62
	3a-4	1.67	0.40	0.13- 20.7
Tumour size	<10cm >10cm	ref 1.43	 0.59	 0.39- 5.18
Nodal status	N0	ref		
	N+	0.72	0.80	0.05- 10.18
Nuclear grade *	1	ref		
	2	1.26	0.66	0.45- 3.51
	3	1.52	0.60	0.32- 7.19
	4	2.35	0.65	0.06- 97.22
Histological tumour necrosis**	negative	ref		
	positive	1.35	0.65	0.37- 4.92
Type of nephrectomy				
	Radical	ref		
	Partial	23.37	0.07	0.81-677.76
Cumulative 2003 Leibovich score		0.81	0.74	0.25- 2.71
Baseline odds		3.47	0.40	0.19-63.59

Observations= 516

* The SORCE grading system used simplifies and selects the worst ISUP features at each grade. For details of grading components, see supplementary material, Table D.

** For the definition of histological tumour necrosis outlined in SORCE trial protocol, see supplementary material; Figure 1

Comparison of patients who developed ‘single site’ first relapses to those with ‘multiple sites’ first relapses

4.3.3 Time to relapse

TTR was not significantly different between those who relapsed with ‘single sites’ and ‘multiple sites’ of disease (HR 0.94, 95% CI, 0.79-1.13, p =0.523) (**Figure 24).** Median TTR for ‘single sites’ first recurrences was 1.72 years (IQR 0.75, 3.93) and for ‘multiple site’ recurrences was 1.52 years (IQR 0.67, 3.66).

4.3.4 RCC-survival

In contrast, relapsing with ‘single site’ disease was associated with significantly improved RCC-survival compared to those with ‘multiple sites’ disease (HR 0.56 95% CI, 0.43-0.72, p<0.001) (**Figure 25**)**.** Median RCC-survival was 9.91 years (IQR 4.47, NR) in the ‘single sites’ group compared to 6.17 years (IQR 2.48, 10.26) in those relapsing with ‘multiple sites’ disease.

4.3.5 RCC-survival-after-recurrence

For this analysis the point of first recurrence was time point zero. Significantly improved RCC-survival-after-recurrence for those with ‘single sites’ of first relapses compared to those with ‘multiple sites’ of relapses (HR 0.51 95% CI, 0.39-0.66, p<0.001) was shown. The median RCC-survival-after-recurrence for patients in the ‘single sites’ group was 5.62 years (95% CI, 2,49-NR) compared to 2.65 years (95% CI, 1.17-6.72) for those in the ‘multiple sites’ group (**Figure 26**).

Figure 24 Cumulative time to recurrence comparing the ‘single sites’ and ‘multiple sites’ cohorts

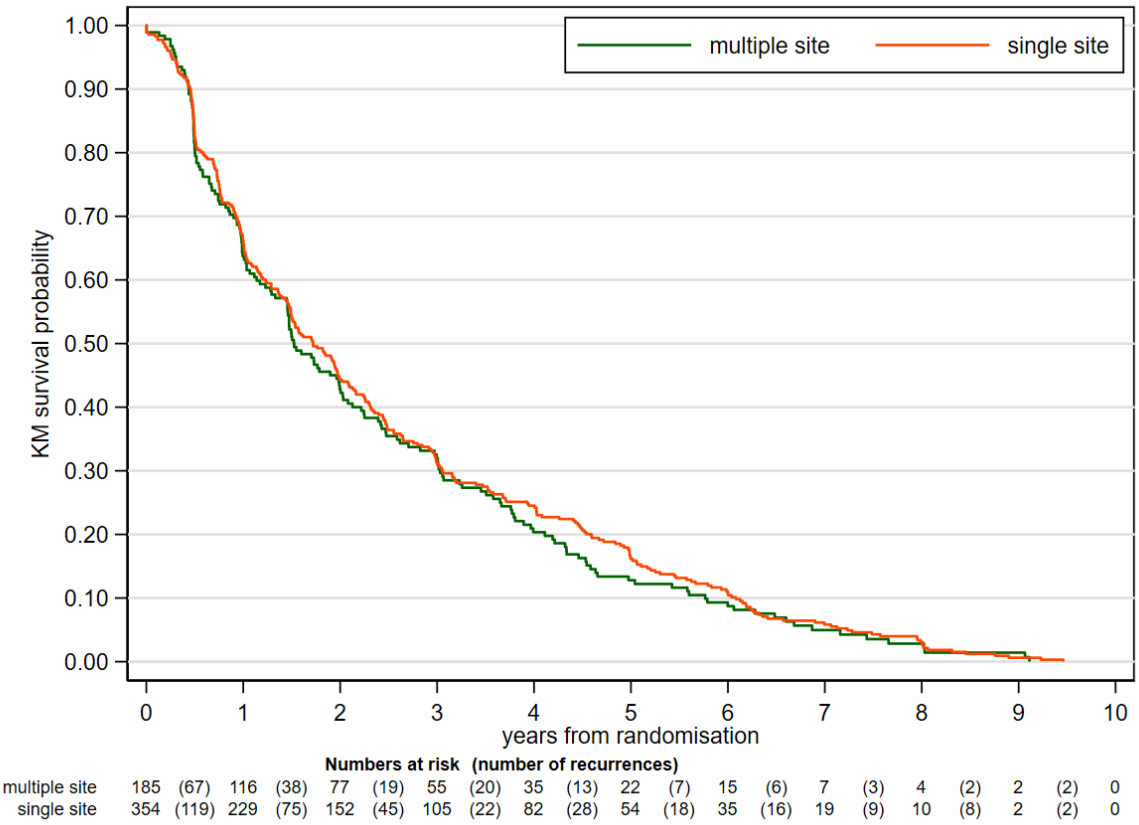
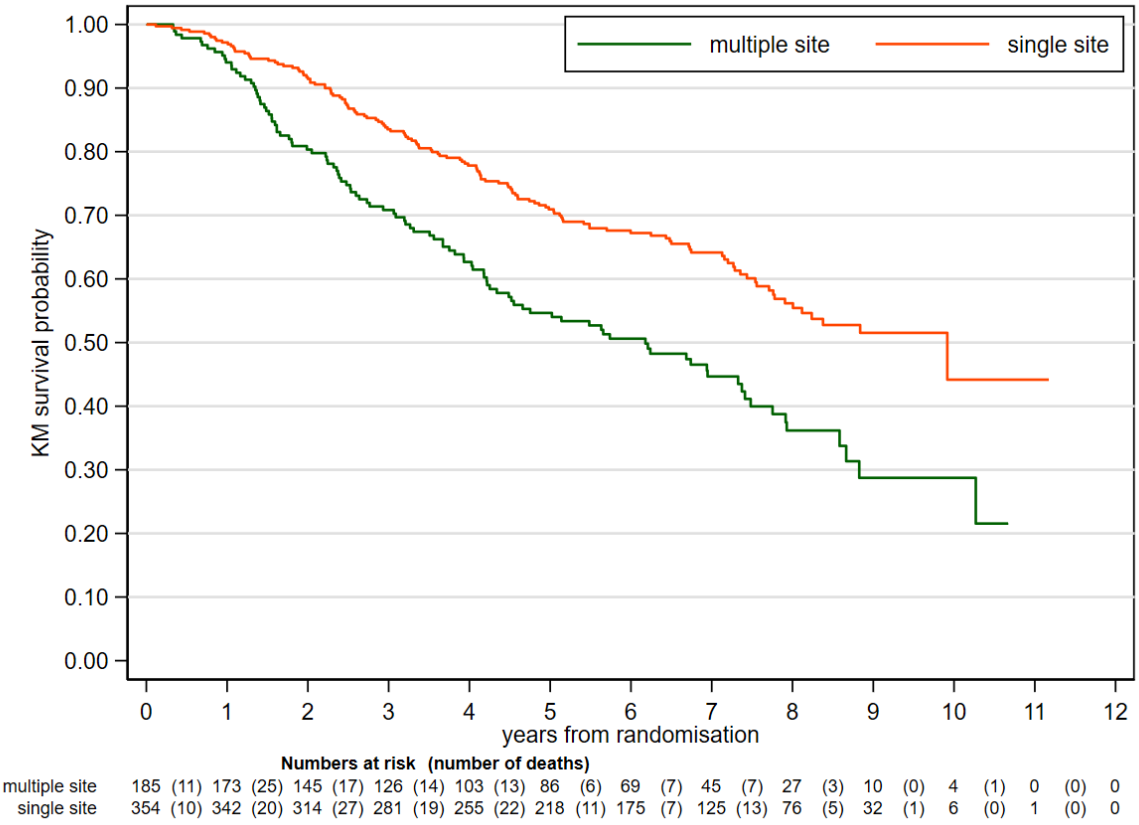
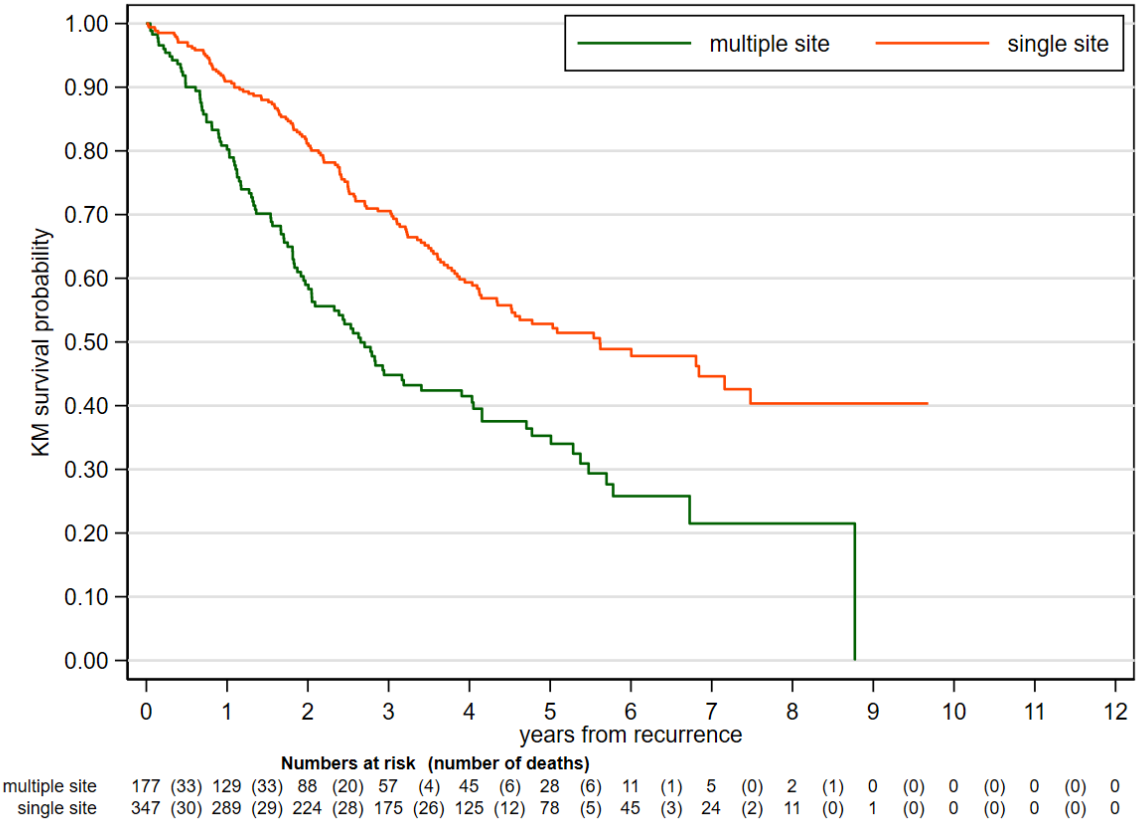


Figure 25 Kaplan-Meier curves for RCC-survival comparing 'single sites' and 'multiple sites' cohorts



Time point zero was set at the point of randomisation to SORCE.

Figure 26 Kaplan-Meier curves for RCC-survival-from-recurrence comparing the 'single sites' and 'multiple sites' cohorts



Time point zero was set at the point of first recurrence

4.3.6 Factors associated with differences in RCC-survival-from-recurrence in patients who initially relapsed with ‘single sites’ and ‘multiple sites’ of disease

Performance status at relapse

Proportionally more patients who developed ‘single site’ first recurrences had performance status zero compared to those with ’multiple sites’ of initial recurrence (45% vs 33%, and fewer patients with ‘single sites’ of relapse were of performance status three or four (11% vs 24%) (**Table 28**). However, an overall comparison of patient performance status between the two cohorts exhibited non-significant difference (p= 0.187) suggesting that patient fitness at relapse cannot fully explain the favourable survival in the ‘single sites’ cohort. Of note, this analysis is limited by incomplete data for performance status on relapse.

Site of relapse

Table 29 shows the locations of first relapses comparing those with ‘single sites’ versus ‘multiple sites’ of disease. Relapses to the lung and the lymph-nodes were commonest in both groups. Significantly fewer relapses to the lymph-nodes, lung, bone, liver, abdominal deposits, pleural lining and skin were reported in the ‘single sites’ group compared to the ‘multiple sites’ group ($p<0.005$). Notably, 52% of first relapses in the ‘multiple sites’ group (compared to 18% in the ‘single sites’ group) involved a lymph-node and 69% (compared to 40%), involved the lung.

Treatment after first relapse

77% of patients relapsing at a single site and 73% of patients at multiple sites had a single treatment upon first relapse (**Table 30**). 68% (88/130) in the ‘multiple sites’ group and 48% (116/246) of patients in the ‘single sites’ groups had treatment that included systemic therapy. Predictably, over double the patients with ’single site’ first relapses received only local treatments compared to those in the ‘multiple sites’ group (39% vs 18%) whereas approximately double the patients in the ’multiple sites’ group compared to the ‘single sites’ group (11% vs 6%) had combined modality (local and systemic) treatment **(Table 31).**

Having only local treatment was associated with markedly favourable RCC-survival-from-recurrence compared to those having systemic treatment only, palliative surgery/ radiotherapy or having a combination of local and systemic treatment. This was shown in patients relapsing at ‘single sites’ and ‘multiple sites’, (median RCC-survival-from-

recurrence for ‘local treatment only’; not reached in both groups). Kaplan-Meier curves **(Figures 27 and 28)** and HRs for RCC-survival-from-recurrence **(Tables 31 and 32)** are shown. Despite small numbers leading to imprecise estimates for this analysis, it suggests that relapses amenable to local treatments have distinctly favourable outcomes, regardless of the number of anatomical sites involved.

Table 28 WHO performance status at the time of 'single site' and 'multiple site' first relapses

	Group				
WHO PS at relapse	**MS (N)**	**MS (%)**	**SS (N)**	**SS (%)**	**Total (N)**
0	39	33	123	45	162
1	46	32	93	34	139
2	16	11	29	11	45
3	14	12	24	9	38
4	5	12	6	2	11
Total	120		275		395
Missing	61		94		

MS; multiple sites, SS; single site

Table 29 Anatomical site of first relapse in the 'single site' and 'multiple site' cohorts

Relapse site	**Group**				**p>chi2**
	MS (N)	**MS (%)**	**SS (N)**	**SS (%)**	**p-value**
Contralateral kidney	21	10	21	5	0.029
Nodal	113	52	76	18	<0.001
Lung	150	69	166	40	<0.001
Bone	39	18	37	9	0.001
Liver	47	22	24	6	<0.001
Abdominal deposits	24	11	7	2	<0.001
Adrenal	17	8	19	5	0.099
Bowel	4	2	2	<1	0.096
Brain	5	2	9	2	0.926
Pancreas	7	3	7	2	0.219
Pleural	5	2	0	0	0.002
Skin	5	2	2	<1	0.039
Local	71	32	30	7	<0.001
Other	19	9	15	4	0.007
Missing	0		25		

MS; multiple sites, SS; single site

Table 30 Number of treatments upon first relapse comparing the 'single site' and 'multiple sites' cohorts

	Group			
Number of treatments	**MS (N)**	**MS (%)**	**SS (N)**	**SS (%)**
1	97	73	197	77
2	15	11	33	13
3	15	12	19	7
4	5	4	6	2
Total	132		255	
Missing data	55		107	

MS; multiple sites, SS; single site

Table 30 Types of treatment received by patients comparing the 'single site' and 'multiple sites' cohorts

Treatment on first relapse	Group				p-value
	MS (N)	MS (%)	SS (N)	SS (%)	
Local only	23	18	95	39	
Systemic only	74	57	102	42	
Pall surgery/RT	17	13	33	14	
Local and systemic	14	11	14	6	
	130		246		<0.001
Missing	55		108		

Treatment on first relapse	Group			
	MS (N)	MS (%)	SS (N)	SS (%)
Systemic therapy*	100	61	134	47
Radiotherapy	22	14	39	13
Ablative therapy	6	3.1	9	3
Brain radiosurgery	3	2	1	<1
Metastasectomy	22	14	86	27
Lymph-node removal	4	2	11	4
Partial nephrectomy	2	2	11	4
Palliative surgery	2	<1	0	0
Palliative procedure	1	<1	3	<1
Total	188		343	

*Includes TKIs, Nivolumab monotherapy and cytokines

MS; multiple sites, SS; single site

Figure 27 Kaplan-Meier curves for RCC-survival-from-recurrence in patients with 'single sites' of first relapses stratified by treatment upon relapse

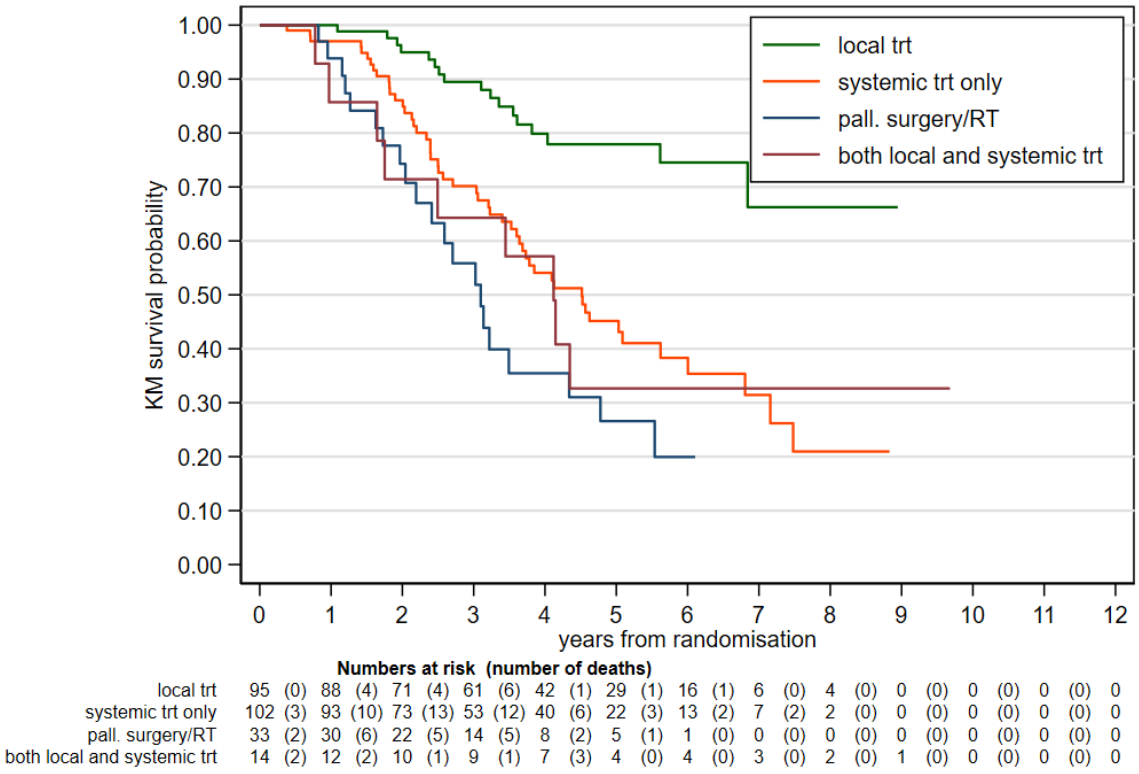


Table 31 Cox model for RCC-survival-from-recurrence by treatment upon recurrence in the 'single site' group

Number of obs = 244

Treatment on relapse	HR	95% CI	p-value	Median RCCS (years)
Local	1.0 (reference)			
Systemic treatment only	3.19	1.84-5,53	<0.001	4.52
Palliative surgery/RT	5.36	2.82-10.21	<0.001	3.10
Both local and systemic treatment	3.22	1.42-7.27	0.005	4.12

HR; hazard ratio, CI; confidence interval, RCCS; RCC-specific survival-from-recurrence

Figure 28 Kaplan-Meier curves for RCC-survival-from-recurrence in patients with 'multiple sites' of first relapses stratified by treatment upon relapse

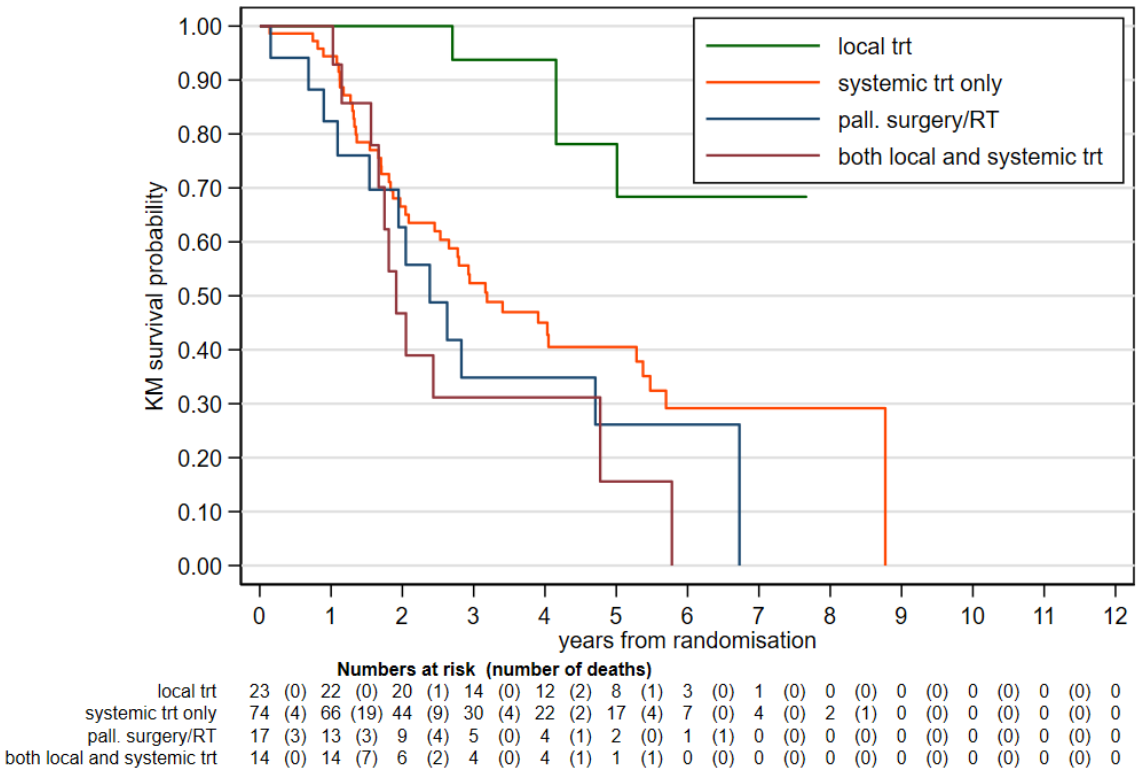


Table 32 Cox model of RCC-survival-from-recurrence by treatment upon recurrence in the 'multiple sites' group

Number of obs = 128

Treatment on relapse	HR	95% CI	p-value	Median (RCCS)
Local	1.0 (reference)			
Systemic treatment only	4.38	1.57-12.23	0.005	3.18
Palliative. surgery/RT	6.70	2.15-20.80	0.001	2.38
Both local and systemic treatment	7.63	2.42-24.08	0.001	1.91

HR; hazard ratio, CI; confidence interval, RCCS- RCC-specific survival-from-recurrence

4.3.7 Can the difference in survival in those with ‘single sites’ and ‘multiple sites’ of first relapses be explained?

Performance status at relapse, organ site and treatments upon relapse exhibit variability between the two cohorts. Their individual effect on RCC-survival-from-recurrence were examined by adjusting for each component separately and in combination in multivariable regression models.

‘Single sites’ status was significantly associated with improved RCC-survival-from-recurrence compared to having ‘multiple sites’ of relapse (HR 0.51 95% CI 0.39-0.66, p<0.001) **(Table 33)**. The prognostic effect of ‘single sites’ status remained statistically significant when adjusting separately for organ site, (HR 0.54 95% CI 0.35-0.82 p=0.004) **(Table 34)** and WHO PS (HR 0.60 95% CI 0.42-0.86, p=0.006) **(Table 35).** Adjusting the model for treatment after relapse **(Table 36)** resulted in an insignificant difference in RCC-survival-from-recurrence between the two groups (HR 0.71 (95% CI, 0.52- 0.98 p=0.037). When relapse site, WHO PS and treatment upon relapse were simultaneously adjusted for, **(Table 37)**, the prognostic benefit associated with ‘single sites’ status was lost (HR 1.06, 95% CI, 0.43-2.61 p= 0.898), implicating an input from the three variables in combination to the survival difference of patients relapsing with ‘single sites’ and ‘multiple sites’ of disease shown.

Table 33 Cox model for RCC-survival evaluating 'single sites' status

Number of observations=524

		HR	95% CI	p-value
Site status	MS	1.0 (reference)		
	SS	0.51	0.39-0.67	<0.001

Table 34 Cox model for RCC-survival evaluating 'single sites' status adjusted for relapse site

Number of obs = 524

		HR	95% CI	p-value
Site status	MS	1.0 (reference)		
	SS	0.54	0.35-0.82	0.004
Organ site	Contralateral kidney	0.52	0.25-0.97	0.075
	Distant nodal	0.97	0.68-1.41	0.913
	Lung	0.87	0.61-1.24	0.44
	Bone	1.48	0.99-2.20	0.056
	Liver	1.38	1.09-2.20	0.015
	Abdominal deposit	1.61	0.86-3.01	0.133
	Adrenal	0.61	0.32-1.16	0.142
	Bowel	0.68	0.17-2.77	0.611
	Brain	3.09	1.48-6.45	0.010
	Pancreas	0.52	0.13-2.15	0.213
	Pleural	1.98	0.71- 5.52	0.227
	Skin	1.12	0.35-3.62	0.825
	Other	1.27	0.77-2.11	0.511

MS; multiple sites, SS; single site

Table 35 Cox model for RCC-survival evaluating 'single sites' status adjusted for performance status

Number of obs = 327

		HR	95% CI	p-value
Site status	MS	1.0 (reference)		
	SS	0.60	0.42-0.86	0.006
WHO PS at relapse				
	0	1.0 (reference)		
	1	3.68	2.07-6.54	0.000
	2	7.23	3.85-12.56	0.000
	3	9.59	5.03-18.26	0.000
	4	6.63	2.75-16.08	0.000

Table 36 Cox model for RCC-survival evaluating 'single sites' status adjusted for treatment after first relapse

Number of obs = 372

		HR	95% CI	p-value
Site status	MS	1.0 (reference)		
	SS	0.71	0.52-0.98	0.037
Treatment on relapse				
	Local	1.0 (reference)		
	Systemic treatment only	3.42	2.12-5.52	0.000
	Palliative. surgery/RT	5.43	3.14-9.42	0.000
	Both local and systemic treatment	4.57	2.47-8.47	0.000

MS; multiple sites, SS; single site

Table 37 Evaluating 'single sites' status for RCC-survival adjusting for a. first relapse site b. performance status c. treatment after first relapse

Number of obs = 253

		HR	95% CI	p-value
Site status	MS	1.0 (reference)		
	SS	1.06	0.43-2.61	0.989
Organ site	Contralateral kidney	1.62	0.41-6.37	0.485
	Distant nodal	1.73	0.81-3.72	0.159
	Lung	1.22	0.57-2.63	0.608
	Bone	1.68	0.76-3.72	0.203
	Liver	1.13	0.47-2.80	0.779
	Abdominal deposit	4.78	1.58-14.46	0.006
	Adrenal	1.22	0.36-4.32	0.744
	Bowel	2.88	0.60-14.82	0.181
	Brain	8.37	0.99-70.44	0.051
	Pancreas	<0.001	-	1.000
	Pleural	<0.001	-	1.000
	Skin	12.72	1.40-115.37	0.024
	Other	2.36	0.95-5.86	0.064
WHO PS at relapse	0	(reference)		
	1	4.16	1.92-9.03	0.000
	2	8.51	3.74-19.39	0.000
	3	8.77	3.6021.33	0.000
	4	8.82	3.01-25.93	0.000
Treatment on relapse	Local	(reference)		
	Systemic treatment only	2.29	1.26-4.18	0.007
	Palliative. surgery/RT	2.38	1.10-5.16	0.027
	Both local and systemic treatment	2.00	0.78-5.15	0.147

4.3.8 Association between RCC-survival-after-recurrence and TTR

Cox models and Kaplan-Meier curves were used to assess TTR in all patients who relapsed and then in those relapsing at ‘single sites’ and at ‘multiple sites’. Assessing TTR as a continuous function in the entire cohort found that every additional relapse free year was associated with a 21% reduced risk of death (HR 0.79, 95% CI, 0.70-0.90).

Comparing dichotomised TTR groups in the entire cohort; patients relapsing between 6 -12 months showed improved survival compared to those relapsing <6 months (HR 0.79 (95% CI, 0.56-1.14 p=0.215) however the difference was statistically insignificant. Recurrences within 12-24 months exhibited significantly improved survival (HR 0.51, 95% CI, 0.35-0.73, p<0.001) as did those relapsing between 24-36 months, (HR 0.31, 95% CI, 0.19-0.50, p<0.001), 36-60 months, (HR 0.33, 95% CI, 0.20-0.80, p<0.001) and >60 months (HR 0.37, 95% CI, 0.18-0.76, p=0.007) (**Table 38**). Separately examining the ‘single sites’ and ‘multiple sites’ cohorts confirmed incrementally improved survival as TTR increased. In both groups a TTR beyond 12 months was associated with statistically significant survival improvement compared to relapses before twelve months **(Tables 39 and 40).**

Finally, the data was used to explore dichotomising patients in the ‘single site’ cohort into three clinically useful TTR based prognostic groups; a. TTR <12 months (high risk) b. TTR <12- 36 months (intermediate risk) c. TTR >36 months (low risk). Patients relapsing <12-36 months showed 44% improved survival (HR 0.65, CI, 0.39-0.82, p=0.003) compared to those relapsing <12 months, and those relapsing >36 months showed 63% improved survival (HR 0.37, CI, 0.21-0.68, p=0.001) **(Figure 29).** Median survival was 3.13 years (high-risk), 5.62 years (intermediate-risk) and not reached (low-risk group).

Table 38 Cox model evaluating dichotomised TTR variables for RCC-survival-from-recurrence in the entire ccRCC group Number of obs = 505

TTR	Events	*Adjusted HR	95% CI	p-value	Median RCCS (years)
<6	98	1.0 (ref)			2.04
6-12	87	0.79	0.56-1.14	0.215	2.70
>12-24	110	0.51	0.35-0.73	<0.001	5.25
>24-36	67	0.31	0.19- 0.50	<0.001	NS
>36-60	79	0.33	0.20- 0.80	<0.001	NS
>60	67	0.37	0.18-0.76	0.007	NS

*Adjusted for Leibovich score, age, gender and baseline PS. TTR- time to relapse, CI; confidence interval, RCCS-RCC-specific survival, NS; not specified

Table 39 Cox model evaluating dichotomised TTR variables for RCC-Survival-from-recurrence in 'single sites' group Number of obs = 336

TTR	Events	*Adjusted HR	95% CI	p-value	Median RCCS (years)
<6	62	1.0 (ref)			3.02
6-12	55	0.72	0.46- 1.12	0.147	3.60
>12-24	74	0.61	0.39- 0.98	0.039	5.52
>24-36	46	0.33	0.18- 0.63	0.001	NS
>36-60	48	0.41	0.20-0.84	0.014	NS
>60	51	0.41	0.15- 1.08	0.071	NS

*Adjusted for Leibovich score, age, gender and baseline PS. TTR- time to relapse, CI; confidence interval, RCCS-RCC specific-survival, NS; not specified

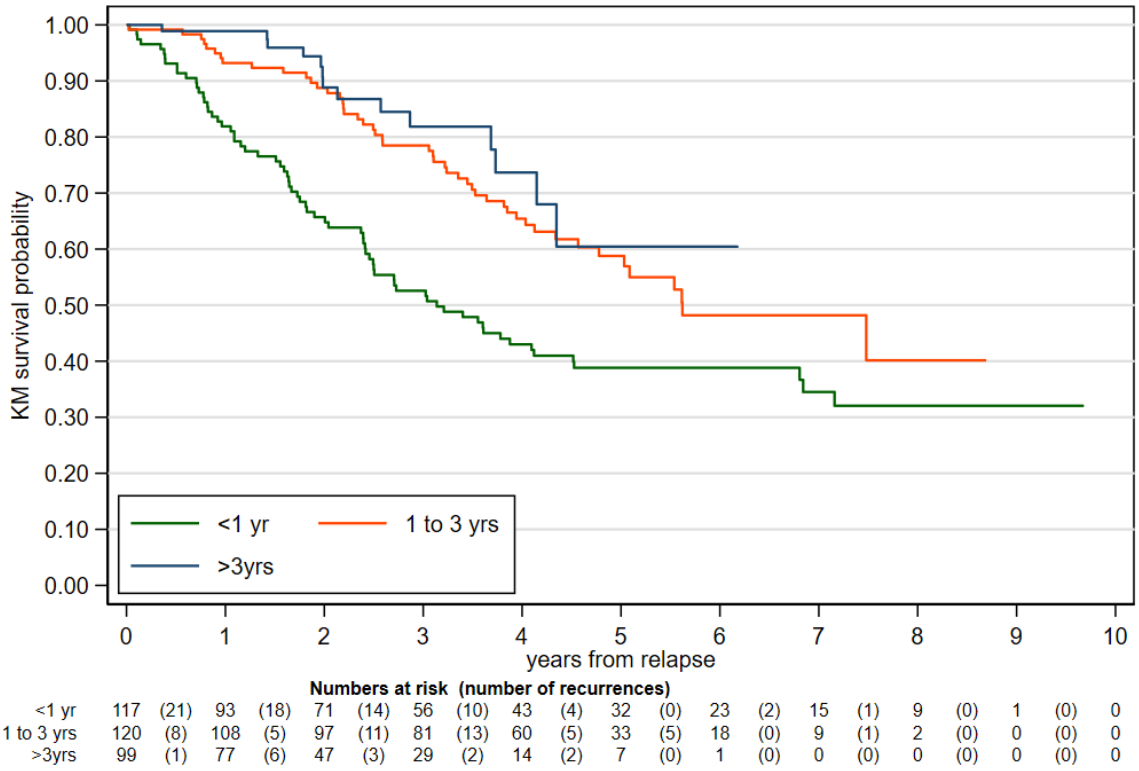
Table 40 Cox model evaluating dichotomised TTR variables for RCC-survival-from-recurrence in the 'multiple sites' group Number of obs = 171

TTR	Events	*Adjusted HR	95% CI	p-value	Median RCCS (years)
<6	36	1.0 (ref)			2.05
6-12	32	0.70	0.41- 1.22	0.212	2.94
>12-24	36	0.23	0.12- 0.44	<0.001	6.73
>24-36	21	0.21	0.10- 0.48	<0.001	NS
>36-60	31	0.22	0.10- 0.47	<0.001	NS
>60	16	0.38	0.12- 1.18	0.095	NS

*Adjusted for Leibovich score, age, gender and baseline PS TTR- time to relapse, CI; confidence interval, RCCS-RCC-specific survival, NS; not specified

Figure 29 Generating clinical risk groups

Kaplan-Meier survival curves showing RCC-Survival-from-recurrence in dichotomised TTR groups (entire ccRCC cohort)



4.3.9 Secondary analysis: non-ccRCC cohort

Time to recurrence

Patients with non-ccRCCs relapsed quicker than those with ccRCC regardless of number of sites status, shown by the Kaplan-Meier curves in **Figure 30**. TTR for patients with non-ccRCC with 'single sites' first recurrences was favourable (HR 0.75, 95% CI, 0.49-1.14, p= 0.180) compared to those with 'multiple sites' recurrences, (**Table 41**) although the difference was not statistically significant.

RCC-survival-from-recurrence

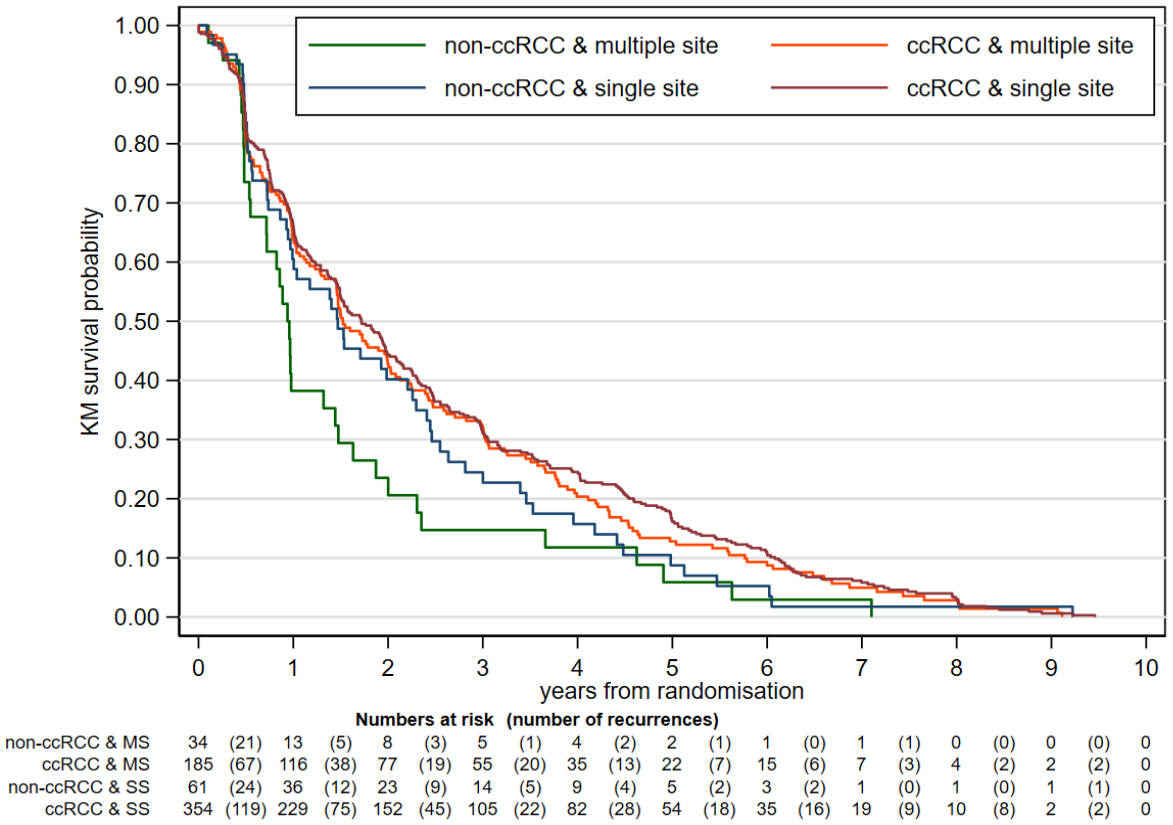
Similar survival outcomes were exhibited by patients with ccRCC and non-ccRCCs developing 'single site' first relapses, shown by aligned Kaplan-Meier curves and overlapping HRs for RCC-survival-from-recurrence **(Figure 31 and Table 42).** In contrast, patients with non-ccRCCs who developed 'multiple sites' of first relapses had markedly worse survival compared to those with non-ccRCCs developing 'single site' relapses (HR 0.33, CI, 0.19-0.58, p<0.001).

Table 41 Cox model showing HRs for TTR comparing groups by number of sites of relapse status and histology

Cohorts	HR	P value	95% CI
ccRCC + MS	0.66	0.026	0.46-0.95
ccRCC +SS	0.62	0.007	0.43-0.88
non-ccRCC +MS	1.0 (reference)		
non-ccRCC +SS	0.75	0.180	0.49-1.14

DFS; disease-free-survival, HR; hazard ratio, CI; confidence interval, ccRCC; clear cell renal cell carcinoma, non-ccRCC; non-clear cell renal cell carcinoma, MS; multiple sites, SS; single site

Figure 30 Kaplan-Meier curves of TTR comparing groups by number of sites of relapse status and histology



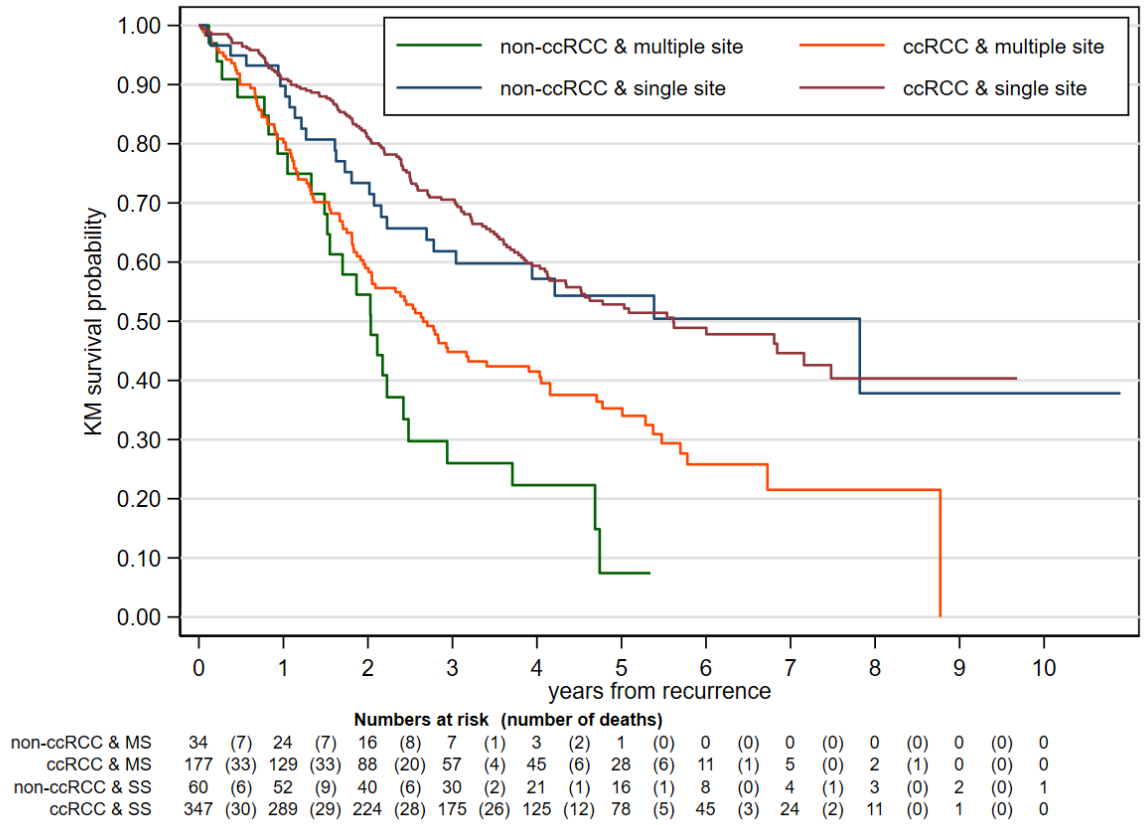
ccRCC; clear cell renal cell carcinoma, non-ccRCC; non-clear cell renal cell carcinoma, MS; multiple sites, SS; single site

Table 42 Cox model showing HRs for RCC-survival-from-recurrence comparing groups by number of sites of relapse status and histology

Cohorts	HR	P value	95% CI
ccRCC + MS	0.59	0.021	0.38-0.92
ccRCC +SS	0.32	<0.001	0.19-0.47
non-ccRCC +MS	1.0 (reference)		
non-ccRCC +SS	0.33	<0.001	0.19-0.58

HR; hazard ratio, CI; confidence interval, ccRCC; clear cell renal cell carcinoma, non-ccRCC; non-clear cell renal cell carcinoma, MS; multiple sites, SS; single site

Figure 31: Kaplan-Meier curves of RCC-survival-from-recurrence comparing groups by number of sites of relapse status and histology



ccRCC; clear cell renal cell carcinoma, non-ccRCC; non-clear cell renal cell carcinoma, MS; multiple sites, SS; single site

4.4 Discussion

The oligometastatic phenotype in RCC, its clinical distinction and optimal management have not been precisely defined. Although adjuvant trials in RCC have expanded eligibility to include participants with fully resected oligo-recurrence, the inclusion criteria vary widely between trials in terms of number, site and timing of relapse. This variability is underscored by low level evidence owing to a lack of comparative studies. To delineate the precise timing and patten of oligometastases in RCC, I analysed outcomes of an international cohort of intermediate and high Lebovich risk patients who relapsed after initially undergoing radical nephrectomy for localised RCC. Patients with single anatomical sites of initial relapse were compared to those with multiple sites of relapse. This was a simple definition, based on previous studies in patients with RCC showing an association with number of initial metastases and metastatic sites with survival [197, 198]. Furthermore, it was a definition easily standardised across the SORCE dataset.

The first important finding was of relapse rates of 37%, notably higher than 17-23% which has been previously reported [158, 196, 199, 200]. This difference may in part reflect the potential for enhanced/early detection of relapses in this study owing to frequent re-imaging mandated (particularly in the first three years) within the SORCE trial protocol. Of note, the percentage of ‘single site’ recurrences (65%) compared to ‘multiple sites’ (35%) was higher than the RECUR consortium ‘potentially curable’ group (45%) compared to the ‘probably incurable’ group (65%) [158, 196]. The RECUR ‘potentially curable’ group excluded patients with more than three metastases at single sites. Arguably, with advances in surgical and non-surgical techniques that enable removal of more extensive disease and multiple sites, the RECUR ‘potentially curable’ definition is too limited for clinical use. Indeed, in the setting of colorectal cancer, curability of liver metastases is not limited by number, size or bi-lobar metastatic involvement and instead is based on an assessment of the expected likelihood of complete resection after multidisciplinary team discussion [201].

This study confirmed that in patients with ccRCC, those with single anatomical sites of first recurrence exhibited vastly improved survival compared to those developing recurrences at multiple sites (HR 0.56 95% CI, 0.43-0.72, $p<0.001$). A statistically significant survival improvement was also observed in patients with non-ccRCCs with single sites of recurrence compared to those with multiple sites, (HR 0.33, CI 0.19-

0.58, P<0.001). Therefore findings support the prognostic utility of classifying patients with clear-cell and non-clear RCCs according to the number of sites of first relapse.

The second aim was to explore differences in performance status upon relapse, timing and sites of recurrence and the treatments offered upon relapse between the two cohorts. This study showed that TTR was not statistically different in patients recurring at single sites compared to those recurring at multiple sites. TTR was however prognostic of RCC-survival-from-recurrence in all patients and in those recurring first at single and multiple sites, which confirms previous data (see section **4.1.6**). In line with the third study aim, I was able to defines three clinically distinct prognostic categories based on the extent and timing of initial relapses. For the cohort recurring at single anatomical sites, relapses within 36 months were 'low risk' of death, relapses between 24-36 months were 'intermediate risk' and relapses within 12 months were 'high risk'. The median survival-from-recurrence in each group was 'not reached', 5.6 years and 3.1 years respectively (**Figure 29**). These preliminary findings require formal validation in other contemporary datasets before that can be recommended in follow-up guidance.

Organ site of recurrence and clinical performance status at relapse were shown to have limited prognostic value. In contrast, treatment modality on first relapse was a useful prognostic surrogate. Patients receiving local treatments including surgical metastasectomy (27%) and non-surgical (radiotherapy 13%, ablation 1%) were associated with statistically improved survival compared those receiving other modalities; systemic treatments, palliative or mixed systemic and local treatments (**Figures 27 and 28**). This pattern was consistent regardless of the number of sites of first relapse. Therefore, findings from this study clearly support the pursuit of local treatments for patients with oligometastases where possible, regardless of the number of sites of disease. A future study confirming the outcome of local treatments, for example confirmation of R0 resection (leaving no tumour at the margin), would add value.

Whether radical resection of metastases alters the disease biology by preventing further tumour seeding and metastatic progression or whether improved survival in patients receiving local treatments is due to their inherently favourable tumour biology, remains uncertain. In this study there were no histopathological characteristics of the primary tumour or clinical characteristics at the time of nephrectomy (**Table 27)**, that

were predictive of developing first relapses at single sites. It is possible to hypothesise that the ability to form metastases may be linked to other tumour or tumour micro-environment related factors pertinent to the metastatic process. Werfel *et al.* conducted a phenome-wide association study (Phe-WAS) using genomic data from 29,000 patients. They focused on the gene TBXA2R encoding thromboxane A2-prostanoid (TPr) a molecule involved in platelet activation via the arachidonic acid signalling pathway [202]. They uncovered a TBXA2R single-nucleotide-polymorphism that correlated with cancer metastasis across several cancer types. In vitro studies have shown that platelet-tumour cell aggregation enhance metastasis by providing survival signals to the tumour cell, shielding it from immune surveillance, and promoting adhesion to the vascular endothelium [203]. Therefore, it is possible that patients developing oligorecurrence may via pathways such as TPr regulation, be worse genetically at metastatic dissemination.

The additional question of interest in the oligometastatic setting is whether the combination of upfront local and systemic treatments provides superior outcomes to local treatments followed by surveillance, (delaying initiation of systemic treatments until further progression). Results from current randomised controlled trials that assess ICIs after radical nephrectomy and include patients with fully resected metastatic disease will be informative. Currently each adjuvant ICI trial provides a different criteria for defining the 'M1NED' group **(Table 43).** The phase three, Keynote-564 trial examining pembrolizumab (anti-PD-1) or placebo after curative nephrectomy was the first to report interim results [55]. Patients with M1NED status were eligible if they had fully resected oligometastatic disease at any site, within twelve months of the primary nephrectomy. At two years of follow-up, the M1NED subgroup (N=58/994) showed a striking HR for DFS of 0.29 (95% CI, 0.12-0.69), favouring pembrolizumab, (**Figure 32)**. The potential for a strong treatment effect in the (albeit small) M1NED cohort to skew results in favour of ICIs in patients with non-metastatic disease highlights the importance of carefully considering the composition and analysis of the M1NED group. Interestingly, recent findings from the phase three adjuvant trial, ImMOTION-101 showed no significant difference in two year DFS between those receiving adjuvant atezolizumab (anti-PD-L1) compared to those receiving standard of care, HR 0·93, 95% CI 0·75-1·15, p=0·50) [56]. ImMOTION101 excluded patients recurring within twelve months unless they were synchronous lung and adrenal metastases and

included those metastasising to the lung, soft tissue and lymph-nodes beyond twelve months after nephrectomy.

Findings from this current study can directly inform on adjuvant trial inclusion criteria. Long-term non progressors (relapsing >36 months) exhibited favourable survival after recurrence and could arguably be managed with surveillance/ local therapies alone, justifying their exclusion from adjuvant trial entry to reserve systemic treatment until the point of further progression. The earliest progressors (relapsing <12 months), given their poor survival after recurrence are likely to include patients with sub-clinically metastatic disease at the time of nephrectomy, providing a justification for their exclusion from adjuvant trial entry. Put together, findings from this study provide support for including patients with single or multiple sites of resectable metastases relapsing between twelve to thirty-six months after nephrectomy into adjuvant clinical trials in RCC.

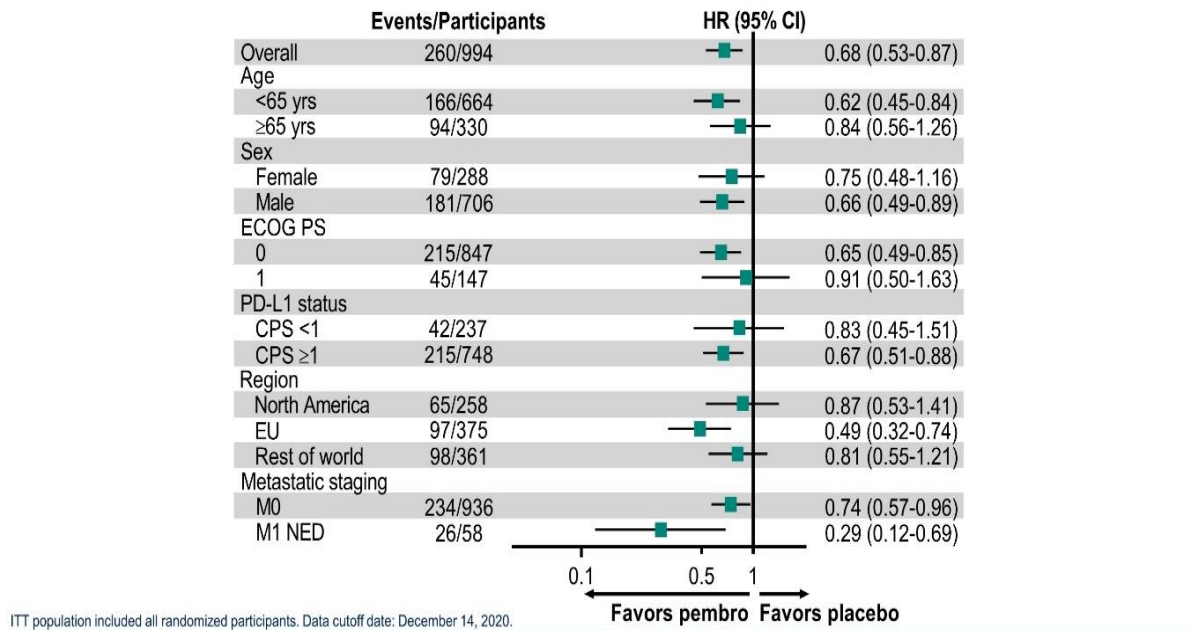
There are notable limitations to this study. Firstly, patients with baseline poor performance status, advanced age and low Leibovich risk were not recruited to SORCE and therefore not accounted for in this analysis. Secondly, as aforementioned there was significant missing data on performance status at recurrence and onward treatment data. In addition, lack of access to post resection histology and imaging reports, meant that it was not possible to conclude with certainty whether patients receiving local treatments achieved complete resection. Instead the study relies on an assumption that patients treated with local approaches were intended for radical resection. In order to extend the evidence base for non-surgical local treatments, for example SABR, in the setting of oligometastatic RCC, a randomised controlled trial is required, one that compares patients receiving radical non-surgical versus surgical approaches to observation or systemic therapy. Finally, this study does not address the optimal management of patients who develop potentially resectable second or third relapses.

4.5 Conclusion

In summary, this study advances understanding of the oligometastatic phenotype in RCC. Although the definition of oligometastatic and oligorecurrent RCC based on number of sites of metastases has scope for refinement, it confirms that patients developing metastases first in one organ have improved survival compared to those metastasising first in multiple organs. This applies to patients with clear-cell and non-

clear cell RCCs. In addition, TTR stratifies patients into clinically useful risk groups and can be used to define an oligometastatic cohort suitable for adjuvant trial entry. This study provides support for preferentially considering local resection for patients with amenable single and multiple sites of first relapses. Finally, it highlights a need to better understanding the underlying biology of oligometastatic RCC and on optimising surgical and non-surgical metastasectomy strategies.

Figure 32 Forrest plot showing the analysis of DFS within key Keynote-564 subgroups, at 24 months of follow-up



Taken from Adjuvant Pembrolizumab after Nephrectomy in Renal-Cell Carcinoma, Toni K Choueri *et al*, NEJM [55].

Programmed death ligand (PD-L1) combined positive score (CPS); the number of PD-L1 staining cells (tumour cells, macrophages and lymphocytes) divided by the total number of viable tumour cells, multiplied by 100. European Union (EU) included the UK. M0; absence of metastases, M1NED; no evidence of disease after resection of primary tumour and solid, isolated soft tissue metastases, CI; confidence interval

Table 43 Risk group stratification and definition of M1NED status in adjuvant checkpoint inhibitor trials

	KEYNOTE-564 [55]	RAMPART [1]	IMMOTION-101 [204]	PROSPER [205]	CHECKMATE-914 [57]
	n=994	n=1750	n=778	n=766	N=1600
Trial design	Double blinded, P3 randomised	Unblinded, P3 randomised	Double blinded, P3 randomised	Unblinded, P3 randomised	Double blinded, P3 randomised
Treatment arm(s)	Pembrolizumab 200mg Q3W	Durvalumab 1500mg +/- tremelimumab 75mg	Atezolizumab 1200mg Q3W	Nivolumab(X1) → nephrectomy → nivolumab (X9)	Nivolumab (PD-1) +/- ipilimumab (CTLA-4)
Comparator arm	Placebo	Active monitoring	Placebo	Nephrectomy alone	Placebo
Duration	12m/ 17 cycles	12m/17 cycles	12m/16 cycles	10 cycles total	6 months
Performance status	ECOG 0-1	ECOG 0-1	ECOG 0-1	ECOG 0-1	ECOG 0-1
Risk group	Intermediate high M1 NED	Intermediate (LS 3-5) High (LS 6-11) M1NED	T2NxM0 TanyN+ M1NED	T2NXM0 TanyN+ M1 NED	T2aG3/4N0M0 T2b/3/4GanyN0M0 TanyGanyN+M0
M1 NED definition	**Synchronou**s or **Metachronous:** ≤12 months: oligometastatic site	**Synchronous:** adrenal	**Synchronou**s: adrenal, lung **Metachronous:** >12 months: pulmonary, soft tissue, LN,	**Synchronous;** ≤12 weeks before or after nephrectomy: oligometastatic sites	N/A

Chapter 5: Discussion and future direction of research

Active surveillance for patients with locally advanced RCC after curative nephrectomy remains a global standard of care. This may change following findings from Keynote-564 [55] which has led to the recent European Medicines Agency and Food Drug Agency approval of adjuvant pembrolizumab (anti-PD-1) for selected high risk patients after nephrectomy [206]. With other adjuvant ICI trials due to publish in the next years, our understanding of the precise effect of upfront immune-checkpoint inhibitors (ICIs) following resection of locally advanced RCC will be consolidated. Given the likely selective response to ICIs in the adjuvant setting and the potential for significant toxicities, the rational selection of patients suitable for these treatments is of upmost importance and should consider their diverse biological categorisation and heterogeneity of relapse behaviours after nephrectomy. I aimed to evaluate whether recent (pre-immunotherapy) phase three trial data can be used to refine the clinical characterisation of patients with higher-risk RCC and to therefore guide their optimal surveillance after nephrectomy and investigation in future adjuvant trials. Findings were pertinent to the design and conduct of the RAMPART trial [1], a phase three multi-arm multistage ICI trial currently recruiting participants after nephrectomy.

The first thesis aim was to externally validate the 2003 Leibovich risk prediction score [3]. I used novel methodology that compared historical data from patients used to derive the score to 'matched' contemporary data from patients enrolled on the SORCE trial. The second aim was to generate a large dataset to explore the characteristics, clinical behaviours and relapse patterns of patients with non-clear-cell RCCs by combining data from SORCE and ASSURE, two international phase three adjuvant trials [5, 97]. The final aim was to refine a clinically relevant classification of oligorecurrent disease in RCC by comparing the outcomes of patients who initially develop single anatomical sites of relapse to those who develop multiple sites after initially undergoing radical nephrectomy. I also explored whether the timing and extent of first recurrences may be used as prognostic tools for patients with higher-risk RCCs.

The 2003 Leibovich score validation is the first study to indicate the score's ability to discriminate between intermediate and high-risk of relapse in patients with locally advanced clear cell RCCs and those with non-clear cell RCCs. As such, it represents a pragmatic tool for the selection of multi-subtype patients for inclusion onto adjuvant trials and hence for adjuvant treatments as they are approved in different global

territories. Several of the recent ICI trials have used TNM and Fuhrman grading for risk stratification ((NCT03024996), (NCT03138512), (NCT03142334), (NCT03055013)). On direct comparison, I showed that the 2003 Leibovich score exhibited discriminative superiority over TNM for patients with intermediate and high risk RCC. RAMPART [1], is currently recruiting participants according to the 2003 Leibovich criteria, and the external validation presented provides confidence in its use.

The methodology developed for this study provides an approach for future validations exercises by enabling a direct comparison of c-statistics between like-for-like patient populations. Previously, prognostic scores have been favoured for clinical application based on studies comparing c-indexes generated from populations of patients with very different risk scores and clinical characteristics.

Ultimately, this study shows that the 2003 Leibovich score remains simple, very easy to apply as well as retaining good discrimination in this contemporary dataset. It is therefore possible to support its use over newer more complicated scores like Leibovich 2018. Future work will be in determining whether trade-off parameters for risk score selection can be quantified, for example, how the added discriminative accuracy of the 2018 score might play out in clinical application and whether a minimal clinically relevant difference in c-index can be determined. Future work may involve developing a decision curve analysis in RCC to determine the net benefit (similar to the idea of net profit in business) of a particular score [207], with benefits and harms (including aspects such as discrimination, implementation costs, accessibility and time requirements), being put on the same scale so they can be compared directly.

The future of risk prediction in RCC may incorporate molecularly-based prognostic tools. Rini *et al.* have shown that it is possible to enhance outcome prediction in RCC by adding a transcript-based recurrence score to the 2003 Leibovich model [208]. However, the recurrence score has not been routinely endorsed for several reasons including significant cost and resource implications globally and its failure to incorporate a prediction of clinical response and toxicity in patients with RCC receiving ICIs. In time, it may be possible to improve upon outcome prediction by adapting the 2003 Leibovich score to include immunological or genetic biomarkers that show additional prognostic and predictive benefit. An overarching aim will be to retain as much of the usefulness, cost-effectiveness and simplicity of the original Leibovich score as possible.

In this thesis, I showed that histological subtype remains an important predictor of relapse behaviour, and of survival outcomes for patients with RCC. Poorer outcomes for patients with higher risk pRCC and sRCC compared to those with ccRCC provides an argument for standards of care for surgical resection, systemic therapy and surveillance to be histology specific. Recent FDA guidance on conducting adjuvant genitourinary studies highlights that a survival advantage specifically in patients with non-clear cell RCCs may be needed for drug approval for these patients [209]. Therefore, it is more important than ever that trials are designed to include non-ccRCC subtypes, accounting for their rarity and clinical heterogeneity, ensuring that those of high enough risk are correctly identified, randomly assigned and appropriately included in primary analyses. In RAMPART [1], although patients with non-ccRCC histology's are being recruited, there are no pre-planned analyses specifically for them. Homogenising the study cohort in this way, although inclusive, fails to detect any variability in clinical behaviour or treatment effect between histology's. In Keynote-546 [55], Immotion-101 [56] and Checkmate-914 [57], all patients in the intention-to-treat population had ccRCC with or without sarcomatoid features. Therefore, the specific impact of ICIs on patients with pure non-ccRCCs remains uncertain.

Trial designs that account for the heterogeneity amongst individual RCCs are of critical need. The National Lung Matrix Trial [210] was a UK-wide multi-arm 'umbrella' trial in which participants were allocated to targeted therapy according to the molecular genotype of their lung cancer rather than being randomly assigned to treatment or placebo arms or stratifying patients by histology. A challenge with using this approach in RCC is that there remains a lag in development of suitable molecular markers and poor availability of molecularly targeted treatments for patients with RCC. To this end, longitudinal sampling of clear cell and non-clear cell tumour and metastases collected as part of TRANSORCE (sample collection and translational study which ran alongside the SORCE trial) will be informative. Findings from my thesis have directly informed on cohorts of interest; notably patients with pRCC relapsing first in the abdomen (conferring poor prognosis) and those relapsing in distant nodes (common site). A digital review of TRANSORCE tumour samples from patients with pRCC, guided by findings from this thesis and led by Professor Harrison at St Andrews University, is underway. Rapid whole slide scanning technology and high quality 2D and 3D digital images from conventional glass slides are being conducted alongside

immune-histological and gene transcription panelling. It aims to unpick the molecular and histological characteristics driving worse outcomes and to explore potential molecular markers for further testing.

In the absence of biomarker stratified trials in RCC, an alternative way of testing treatments across a number of RCC subtypes may be a Bayesian hierarchical basket trial design [211] . The premise of the Bayesian approach are two assumptions. First is an overall treatment effect when grouping all subtypes together and second is an assumption of treatment effect in each subtype separately. This allows results from the overarching cohort to support the effect seen in the rarer cohorts. Applying this to RCC, the first step would be to elicit a consensus as to the expected treatment effect in overall cohort and then separately in each subtype. The second step would be to apply the scientific consensus to determine the statistical power that could be 'borrowed' from the dominant group, (in this case, ccRCC), to support the effect size seen in the rarer types. A Bayesian hierarchical framework would, therefore, allow each subtype to be assessed individually whilst also contributing events towards an overarching analysis. Drug approval for a particular RCC subtype would be contingent on a positive treatment effect in the individual cohorts and on this result being consistent with the overall result.

Patients with resected oligometastatic disease are being evaluated in adjuvant ICI trials (**Table 44**). A parallel focus must therefore be on optimising their risk stratification and participant selection. In this thesis, I show that developing single sites of initial relapse is a clinically favourable oligometastatic situation in RCC. In addition, I show that patients undergoing local treatments upon relapse are associated with favourable survival, regardless of number of sites of relapse. Within the group of patients relapsing after nephrectomy, further stratification based on timings of first relapse was possible. Put together, I was able to recommend inclusion of patients developing fully resected single metastases between twelve to thirty-six months to the RAMPART TMG. In order to cap the number of patients in this category, the TMG decided to exclude those relapsing more than twenty-four months after nephrectomy. If RAMPART yields positive results, the provision of ICIs may therefore be extended to patients with fully resected single metastases at any sites between twelve to twenty-four months, thereby radically re-defining the clinical pathway for this group of patients. Whether patients with multiple sites of radically treated metastases may benefit from

adjuvant systemic treatments is an important additional clinical question. Currently, neither the European Society of Medical Oncology [14] nor European Association of Urology (EAU) [159] offer definitive guidance on the appropriate selection of patients for local treatments of metastases. They state that metastasectomy (surgical or stereotactic radiotherapy) for single or oligometastatic relapses could be considered on an individual bases following multidisciplinary review. Somewhat contradicting this guidance, EAU also references a single-arm prospective and retrospective study supporting the observation of oligometastases for up to 16 months, without resection, before initiating systemic therapy. Findings from my thesis provide a rationale for considering upfront radical resection where possible for all patients relapsing at single or multiple sites and certainly for those relapsing within 12-36 months.

Future work must focus on expanding the evidence base for non-surgical radical treatment options for patients with oligometastatic RCC, for example investigating the radical removal of more extensive metastases or evaluating outcomes in less fit patients receiving radical non-surgical techniques who may not be suitable for surgery. Growing evidence supports the use of stereotactic ablative radiation therapy (SABR) in oligometastatic RCC. A recent meta-analysis of twenty-eight studies examining SABR for patients with RCC oligometastases, found 90% one-year local control rates intra- and extracranially and minimal toxicity [212]. In addition, early pre-clinical and clinical studies indicate the immunomodulatory effects of delivering ablative doses of radiation in limited fractions [213]. Therefore, another promising direction for SABR is in its combination with ICIs. A recent systematic review of eighteen non-randomised studies evaluating ICI-SABR combinations in non-small-cell lung cancer reported local control rates of 71%. They also reported abscopal responses, where focal radiotherapy resulted in systemic anti-tumoral action at distant sites of 30-50% [214]. Eleven studies reported progression-free-survival (PFS) and overall survival (OS), with a mean PFS of 4.6 months (Interquartile range-IQR 2.3, 7 months) and 12.4 months (IQR 9.0, 24.7 months), respectively. Although not directly comparable, OS benefit from ICI-SABR combinations was similar to that observed in phase two and three trials of ICIs without radiotherapy [215, 216]. Toxicity rates were also consistent with those attributable to ICI treatment alone [215, 216]. Of note however most ICI-SABR studies included unselected patients who had been heavily treated beforehand, limiting the interpretation of OS. Ultimately, randomised trials are required to confirm

the synergy between ICIs and SABR in the context of previously untreated oligometastatic RCC. Current phase three trials that are evaluating ICIs in the adjuvant setting and are including patients with highest risk oligometastatic RCCs, will to this end, provide valuable randomised data (**Table 44**).

Given the differential outcomes seen in patients developing single anatomical sites compared to multiple sites of first relapse, another avenue for further work is in investigating the biological processes that underpin low metastatic burden patterns of spread. There may be a role for pharmacological regulation or blockade of cellular metastatic signalling. Studies of thromboxane A2 and prostanoid receptor (TPr) signalling in animal models have shown disruption of the endothelial barrier leading to metastasis via platelet activation and upregulation of platelet-tumour cell aggregation [217]. When tested in mouse models of metastasis, TPr inhibition potently blocked spontaneous metastasis from primary tumours, without affecting tumour cell proliferation, motility, or tumour growth [202]. Mice were randomised to receive either treatment with placebo or CPI211, a potent a selective small molecule of TPr, following surgical resection of primary tumour. Although 100% of placebo treated mice and 90% of CPI211 treated mice developed lung metastases, the number of lung metastases per mouse was reduced in CPI211-treated mice to half of that seen in controls, suggesting a transformation to a lower metastatic burden phenotype. In another single randomised, dose escalating placebo-controlled study, CPI211 has been shown to be safe and well tolerated in humans [202]. Put together, this early work highlights a potential role for pharmacological manipulation of the metastatic process, to promote long-term cytostasis in the post nephrectomy setting, and warrants further investigation.

In summary, I have shown how datasets from international phase three renal trials can usefully inform on the biological heterogeneity and relapse patterns of contemporary patients with RCC. Of note, caution must be applied when generalising clinical trial findings to all real-world patients particularly older, frailer cohorts who are less likely to be fit for radical treatment options and who are excluded from clinical trial inclusion purely based on age. Of added utility is the parallel collection of longitudinal histology samples from trial participants for example TRANSRAMPART. This will allow clinical findings to be correlated to immunohistochemical and molecular data.

Given the rarity of some of the RCC subgroups and the long follow up required to acquire survival data, generating large enough datasets for meaningful results will require pre-specified meta-analyses and sharing of data between trials. Ultimately this requires early data-sharing agreements between institutions (at the protocol development stage) and collaborative analysis plans to ensure timely reporting and optimal quality of results. My collaborations with Professor Hass and the ECOG-ACRIN team at Dana Farber Cancer Institute and with Professor Leibovich and colleagues at the Mayo Clinic Cleveland US, pave the way for future collaborations between the MRC CTU at UCL and these international institutions, with the ultimate aim of improving outcomes for patients with higher risk RCCs.

References

1. Oza, B., et al., *RAMPART: A phase III multi-arm multi-stage trial of adjuvant checkpoint inhibitors in patients with resected primary renal cell carcinoma (RCC) at high or intermediate risk of relapse.* Contemp Clin Trials, 2021. **108**: p. 106482.
2. Meade, A., et al., *RAMPART: A model for a regulatory-ready academic-led phase III trial in the adjuvant renal cell carcinoma setting.* Contemp Clin Trials, 2021. **108**: p. 106481.
3. Oza, B., et al., *External Validation of the 2003 Leibovich Prognostic Score in Patients Randomly Assigned to SORCE, an International Phase III Trial of Adjuvant Sorafenib in Renal Cell Cancer.* J Clin Oncol, 2022. **40**(16): p. 1772-1782.
4. Oza, B., et al., *Reply to U. Capitanio et al.* Journal of Clinical Oncology, 2022. **41**(3): p. 704-706.
5. Eisen, T., et al., *Adjuvant Sorafenib for Renal Cell Carcinoma at Intermediate or High Risk of Relapse: Results From the SORCE Randomized Phase III Intergroup Trial.* J Clin Oncol, 2020. **38**(34): p. 4064-4075.
6. Cancer., W.H.O.I.A.f.R.o., *Global cancer observatory.*
7. International Agency for Research on Cancer, W.H.O.; Available from: https://gco.iarc.fr/.
8. Moch, H., et al., *The 2022 World Health Organization Classification of Tumours of the Urinary System and Male Genital Organs—Part A: Renal, Penile, and Testicular Tumours.* European Urology, 2022. **82**(5): p. 458-468.
9. Sun, M., et al., *Age-adjusted incidence, mortality, and survival rates of stage-specific renal cell carcinoma in North America: a trend analysis.* Eur Urol, 2011. **59**(1): p. 135-41.
10. Mitchell, T.J., et al., *Timing the Landmark Events in the Evolution of Clear Cell Renal Cell Cancer: TRACERx Renal.* Cell, 2018. **173**(3): p. 611-623.e17.
11. Kim, W.Y. and W.G. Kaelin, Jr., *Molecular pathways in renal cell carcinoma--rationale for targeted treatment.* Semin Oncol, 2006. **33**(5): p. 588-95.
12. Network, C.G.A.R., *Comprehensive molecular characterization of clear cell renal cell carcinoma.* Nature, 2013. **499**(7456): p. 43-9.
13. Xu, J., et al., *Mechanistically distinct cancer-associated mTOR activation clusters predict sensitivity to rapamycin.* J Clin Invest, 2016. **126**(9): p. 3526-40.
14. Escudier, B., et al., *Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†.* Ann Oncol, 2019. **30**(5): p. 706-720.
15. Piva, F., et al., *BAP1, PBRM1 and SETD2 in clear-cell renal cell carcinoma: molecular diagnostics and possible targets for personalized therapies.* Expert Review of Molecular Diagnostics, 2015. **15**(9): p. 1201-1210.
16. Kotecha, R.R., R.J. Motzer, and M.H. Voss, *Towards individualized therapy for metastatic renal cell carcinoma.* Nature Reviews Clinical Oncology, 2019. **16**(10): p. 621-633.
17. Vergho, D.C., et al., *Tumor thrombus of inferior vena cava in patients with renal cell carcinoma - clinical and oncological outcome of 50 patients after surgery.* BMC Res Notes, 2012. **5**: p. 5.
18. MacLennan, S., et al., *Systematic review of oncological outcomes following surgical management of localised renal cancer.* European urology, 2012. **61**(5): p. 972-993.
19. Ljungberg, B., et al., *European Association of Urology Guidelines on Renal Cell Carcinoma 2021.* 2021.
20. Ljungberg, B., et al., *European Association of Urology Guidelines on Renal Cell Carcinoma: The 2022 Update.* Eur Urol, 2022. **82**(4): p. 399-410.
21. Lee, Z., et al., *Local Recurrence Following Resection of Intermediate-High Risk Nonmetastatic Renal Cell Carcinoma: An Anatomical Classification and Analysis of the ASSURE (ECOG-ACRIN E2805) Adjuvant Trial.* J Urol, 2020. **203**(4): p. 684-689.
22. Dabestani, S., et al., *Long-term Outcomes of Follow-up for Initially Localised Clear Cell Renal Cell Carcinoma: RECUR Database Analysis.* Eur Urol Focus, 2019. **5**(5): p. 857-866.

23. Leibovich, B.C., et al., *Prediction of progression after radical nephrectomy for patients with clear cell renal cell carcinoma: a stratification tool for prospective clinical trials.* Cancer, 2003. **97**(7): p. 1663-71.
24. Zisman, A., et al., *Improved prognostication of renal cell carcinoma using an integrated staging system.* J Clin Oncol, 2001. **19**(6): p. 1649-57.
25. Zisman, A., et al., *Risk group assessment and clinical outcome algorithm to predict the natural history of patients with surgically resected renal cell carcinoma.* J Clin Oncol, 2002. **20**(23): p. 4559-66.
26. Kattan, M.W., et al., *A postoperative prognostic nomogram for renal cell carcinoma.* J Urol, 2001. **166**(1): p. 63-7.
27. Hartmann, J.T., et al., *Tyrosine kinase inhibitors - a review on pharmacology, metabolism and side effects.* Curr Drug Metab, 2009. **10**(5): p. 470-81.
28. Heng, D.Y., 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, 2009. **27**(34): p. 5794-5799.
29. Rini, B.I., et al., *Pembrolizumab plus Axitinib versus Sunitinib for Advanced Renal-Cell Carcinoma.* N Engl J Med, 2019. **380**(12): p. 1116-1127.
30. Motzer, R.J., et al., *Avelumab plus Axitinib versus Sunitinib for Advanced Renal-Cell Carcinoma.* N Engl J Med, 2019. **380**(12): p. 1103-1115.
31. Choueiri, T.K., et al., *Nivolumab plus Cabozantinib versus Sunitinib for Advanced Renal-Cell Carcinoma.* N Engl J Med, 2021. **384**(9): p. 829-841.
32. Motzer, R., et al., *Lenvatinib plus Pembrolizumab or Everolimus for Advanced Renal Cell Carcinoma.* N Engl J Med, 2021. **384**(14): p. 1289-1300.
33. Haas, N.B., et al., *Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinoma (ECOG-ACRIN E2805): a double-blind, placebo-controlled, randomised, phase 3 trial.* Lancet, 2016. **387**(10032): p. 2008-16.
34. Motzer, R.J., et al., *Randomized Phase III Trial of Adjuvant Pazopanib Versus Placebo After Nephrectomy in Patients With Localized or Locally Advanced Renal Cell Carcinoma.* J Clin Oncol, 2017. **35**(35): p. 3916-3923.
35. Ravaud, A., et al., *Adjuvant Sunitinib in High-Risk Renal-Cell Carcinoma after Nephrectomy.* N Engl J Med, 2016. **375**(23): p. 2246-2254.
36. Gross-Goupil, M., et al., *Axitinib versus placebo as an adjuvant treatment of renal cell carcinoma: results from the phase III, randomized ATLAS trial.* Ann Oncol, 2018. **29**(12): p. 2371-2378.
37. Escudier, B., et al., *Sorafenib in advanced clear-cell renal-cell carcinoma.* N Engl J Med, 2007. **356**(2): p. 125-34.
38. Motzer, R.J., et al., *Sunitinib versus interferon alfa in metastatic renal-cell carcinoma.* N Engl J Med, 2007. **356**(2): p. 115-24.
39. Motzer, R.J., et al., *Pazopanib versus sunitinib in metastatic renal-cell carcinoma.* N Engl J Med, 2013. **369**(8): p. 722-31.
40. Motzer, R.J., et al., *Axitinib versus sorafenib as second-line treatment for advanced renal cell carcinoma: overall survival analysis and updated results from a randomised phase 3 trial.* Lancet Oncol, 2013. **14**(6): p. 552-62.
41. Choueiri, T.K., et al., *Cabozantinib Versus Sunitinib As Initial Targeted Therapy for Patients With Metastatic Renal Cell Carcinoma of Poor or Intermediate Risk: The Alliance A031203 CABOSUN Trial.* J Clin Oncol, 2017. **35**(6): p. 591-597.
42. Mehta, A., G. Sonpavde, and B. Escudier, *Tivozanib for the treatment of renal cell carcinoma: results and implications of the TIVO-1 trial.* Future Oncol, 2014. **10**(11): p. 1819-26.
43. Motzer, R.J., et al., *Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma.* N Engl J Med, 2015. **373**(19): p. 1803-13.

44. Motzer, R.J., et al., *Nivolumab plus Ipilimumab versus Sunitinib in Advanced Renal-Cell Carcinoma.* N Engl J Med, 2018. **378**(14): p. 1277-1290.
45. Rini, B.I., et al., *Atezolizumab plus bevacizumab versus sunitinib in patients with previously untreated metastatic renal cell carcinoma (IMmotion151): a multicentre, open-label, phase 3, randomised controlled trial.* Lancet, 2019. **393**(10189): p. 2404-2415.
46. Rini, B.I., et al., *Pembrolizumab plus Axitinib versus Sunitinib for Advanced Renal-Cell Carcinoma.* N Engl J Med, 2019. **380**(12): p. 1116-1127.
47. Motzer, R.J., et al., *Adjuvant Pazopanib Versus Placebo After Nephrectomy in Patients With Localized or Locally Advanced Renal Cell Carcinoma: Final Overall Survival Analysis of the Phase 3 PROTECT Trial.* Eur Urol, 2021. **79**(3): p. 334-338.
48. Saltz, L.B., et al., *Bevacizumab in combination with oxaliplatin-based chemotherapy as first-line therapy in metastatic colorectal cancer: a randomized phase III study.* Journal of clinical oncology, 2008. **26**(12): p. 2013-2019.
49. Miller, K., et al., *Paclitaxel plus bevacizumab versus paclitaxel alone for metastatic breast cancer.* New England journal of medicine, 2007. **357**(26): p. 2666-2676.
50. Gridelli, C., et al., *The role of bevacizumab in the treatment of non-small cell lung cancer: current indications and future developments.* Oncologist, 2007. **12**(10): p. 1183-93.
51. Cameron, D., et al., *Adjuvant bevacizumab-containing therapy in triple-negative breast cancer (BEATRICE): Primary results of a randomised, phase 3 trial.* The Lancet Oncology, 2013. **14**(10): p. 933-942.
52. Dickson, P.V., et al., *Bevacizumab-induced transient remodeling of the vasculature in neuroblastoma xenografts results in improved delivery and efficacy of systemically administered chemotherapy.* Clin Cancer Res, 2007. **13**(13): p. 3942-50.
53. Motzer, R.J., et al., *Nivolumab plus Ipilimumab versus Sunitinib in Advanced Renal-Cell Carcinoma.* N Engl J Med, 2018. **378**(14): p. 1277-1290.
54. Msaouel, P., *Less is More? First Impressions From COSMIC-313.* Cancer Invest, 2023. **41**(1): p. 101-106.
55. Choueiri, T.K., et al., *Adjuvant Pembrolizumab after Nephrectomy in Renal-Cell Carcinoma.* N Engl J Med, 2021. **385**(8): p. 683-694.
56. Pal, S.K., et al., *Adjuvant atezolizumab versus placebo for patients with renal cell carcinoma at increased risk of recurrence following resection (IMmotion010): a multicentre, randomised, double-blind, phase 3 trial.* Lancet, 2022. **400**(10358): p. 1103-1116.
57. Motzer, R.J., et al., *Adjuvant nivolumab plus ipilimumab versus placebo for localised renal cell carcinoma after nephrectomy (CheckMate 914): a double-blind, randomised, phase 3 trial.* Lancet, 2023. **401**(10379): p. 821-832.
58. Antonia, S.J., et al., *Durvalumab after Chemoradiotherapy in Stage III Non-Small-Cell Lung Cancer.* N Engl J Med, 2017. **377**(20): p. 1919-1929.
59. Choodari-Oskooei, B., et al., *Adding new experimental arms to randomised clinical trials: Impact on error rates.* Clin Trials, 2020. **17**(3): p. 273-284.
60. Choodari-Oskooei, B., et al., *Adding new experimental arms to randomised clinical trials: Impact on error rates.* Clinical trials (London, England), 2020. **17**(3): p. 273-284.
61. Vasudev, N.S., et al., *UK Multicenter Prospective Evaluation of the Leibovich Score in Localized Renal Cell Carcinoma: Performance has Altered Over Time.* Urology, 2020. **136**: p. 162-168.
62. Tan, M.H., et al., *Comparison of the UCLA Integrated Staging System and the Leibovich score in survival prediction for patients with nonmetastatic clear cell renal cell carcinoma.* Urology, 2010. **75**(6): p. 1365-1370; 1370.e1-3.
63. Pichler, M., et al., *External validation of the Leibovich prognosis score for nonmetastatic clear cell renal cell carcinoma at a single European center applying routine pathology.* J Urol, 2011. **186**(5): p. 1773-7.

64. Correa, A.F., et al., *Predicting Renal Cancer Recurrence: Defining Limitations of Existing Prognostic Models With Prospective Trial-Based Validation.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2019. **37**(23): p. 2062-2071.
65. Ljungberg, B., et al., *European Association of Urology Guidelines on Renal Cell Carcinoma: The 2019 Update.* Eur Urol, 2019. **75**(5): p. 799-810.
66. Swami, U., et al., *Revisiting AJCC TNM staging for renal cell carcinoma: quest for improvement.* Ann Transl Med, 2019. **7**(Suppl 1): p. S18.
67. Shao, N., et al., *Modification of American Joint Committee on cancer prognostic groups for renal cell carcinoma.* Cancer Med, 2018. **7**(11): p. 5431-5438.
68. Altman, D.G. and P. Royston, *What do we mean by validating a prognostic model?* Stat Med, 2000. **19**(4): p. 453-73.
69. Royston, P. and D.G. Altman, *External validation of a Cox prognostic model: principles and methods.* BMC Med Res Methodol, 2013. **13**: p. 33.
70. Kattan, M.W., et al., *American Joint Committee on Cancer acceptance criteria for inclusion of risk models for individualized prognosis in the practice of precision medicine.* CA Cancer J Clin, 2016. **66**(5): p. 370-4.
71. Harrell, F.E., Jr., et al., *Evaluating the yield of medical tests.* Jama, 1982. **247**(18): p. 2543-6.
72. Leibovich, B.C., et al., *Predicting Oncologic Outcomes in Renal Cell Carcinoma After Surgery.* Eur Urol, 2018. **73**(5): p. 772-780.
73. Karakiewicz, P.I., et al., *Multi-institutional validation of a new renal cancer-specific survival nomogram.* J Clin Oncol, 2007. **25**(11): p. 1316-22.
74. Frank, I., et al., *An outcome prediction model for patients with clear cell renal cell carcinoma treated with radical nephrectomy based on tumor stage, size, grade and necrosis: the SSIGN score.* J Urol, 2002. **168**(6): p. 2395-400.
75. Yaycioglu, O., et al., *A Preoperative Prognostic Model Predicting Recurrence-free Survival for Patients With Kidney Cancer.* Japanese Journal of Clinical Oncology, 2012. **43**(1): p. 63-68.
76. Cindolo, L., et al., *A preoperative clinical prognostic model for non-metastatic renal cell carcinoma.* BJU Int, 2003. **92**(9): p. 901-5.
77. Burns, P.B., R.J. Rohrich, and K.C. Chung, *The levels of evidence and their role in evidence-based medicine.* Plastic and reconstructive surgery, 2011. **128**(1): p. 305-310.
78. Sackett, D.L., *Rules of evidence and clinical recommendations on the use of antithrombotic agents.* Chest, 1989. **95**(2 Suppl): p. 2s-4s.
79. Karakiewicz, P.I., et al., *A preoperative prognostic model for patients treated with nephrectomy for renal cell carcinoma.* Eur Urol, 2009. **55**(2): p. 287-95.
80. Tan, M.-H., et al., *The Karakiewicz nomogram is the most useful clinical predictor for survival outcomes in patients with localized renal cell carcinoma.* Cancer, 2011. **117**(23): p. 5314-5324.
81. Warren, A.Y. and D. Harrison, *WHO/ISUP classification, grading and pathological staging of renal cell carcinoma: standards and controversies.* World J Urol, 2018. **36**(12): p. 1913-1926.
82. Correa, A.F., et al., *Predicting Renal Cancer Recurrence: Defining Limitations of Existing Prognostic Models With Prospective Trial-Based Validation.* J Clin Oncol, 2019. **37**(23): p. 2062-2071.
83. Klatte, T., et al., *The VENUSS prognostic model to predict disease recurrence following surgery for non-metastatic papillary renal cell carcinoma: development and evaluation using the ASSURE prospective clinical trial cohort.* BMC medicine, 2019. **17**(1): p. 182-182.
84. Lee, H.J., et al., *External validation of the updated Leibovich prognostic models for clear cell and papillary renal cell carcinoma in an Asian population.* Urol Oncol, 2019. **37**(6): p. 356.e9-356.e18.
85. Klatte, T., et al., *Development and external validation of a nomogram predicting disease specific survival after nephrectomy for papillary renal cell carcinoma.* J Urol, 2010. **184**(1): p. 53-8.

86. Buti, S., et al., *Validation of a new prognostic model to easily predict outcome in renal cell carcinoma: the GRANT score applied to the ASSURE trial population.* Ann Oncol, 2017. **28**(11): p. 2747-2753.
87. Erdem, S., et al., *External validation of the VENUSS prognostic model to predict recurrence after surgery in non-metastatic papillary renal cell carcinoma: A multi-institutional analysis.* Urol Oncol, 2022. **40**(5): p. 198.e9-198.e17.
88. Eisen, T., et al., *Adjuvant Sorafenib for Renal Cell Carcinoma at Intermediate or High Risk of Relapse: Results From the SORCE Randomized Phase III Intergroup Trial.* Journal of Clinical Oncology, 2020. **38**(34): p. 4064-4075.
89. Amin, M.B., et al., *Prognostic impact of histologic subtyping of adult renal epithelial neoplasms: an experience of 405 cases.* Am J Surg Pathol, 2002. **26**(3): p. 281-91.
90. Patzer, R.E., A.H. Kaji, and Y. Fong, *TRIPOD Reporting Guidelines for Diagnostic and Prognostic Studies.* JAMA Surg, 2021. **156**(7): p. 675-676.
91. Ashrafi, A.N. and I.S. Gill, *Minimally invasive radical nephrectomy: a contemporary review.* Transl Androl Urol, 2020. **9**(6): p. 3112-3122.
92. Cunningham, D., et al., *LBA-03 Neoadjuvant chemotherapy for resectable oesophageal and junctional adenocarcinoma: results from the UK Medical Research Council randomised OEO5 trial (ISRCTN 01852072).* Annals of Oncology, 2015. **26**: p. iv117.
93. Group, M.R.C.O.C.W., *Surgical resection with or without preoperative chemotherapy in oesophageal cancer: a randomised controlled trial.* Lancet, 2002. **359**(9319): p. 1727-33.
94. Luzzago, S., et al., *Effect of stage and grade migration on cancer specific mortality in renal cell carcinoma patients, according to clear cell vs. non-clear cell histology: A contemporary population-based analysis.* Urol Oncol, 2020. **38**(5): p. 506-514.
95. Moch, H., et al., *The 2022 World Health Organization Classification of Tumours of the Urinary System and Male Genital Organs-Part A: Renal, Penile, and Testicular Tumours.* Eur Urol, 2022. **82**(5): p. 458-468.
96. Wei, S. and T. Al-Saleem, *The Pathology and Molecular Genetics of Sarcomatoid Renal Cell Carcinoma: A Mini-Review.* Journal of kidney cancer and VHL, 2017. **4**(2): p. 19-23.
97. Haas, N.B., et al., *Adjuvant Treatment for High-Risk Clear Cell Renal Cancer: Updated Results of a High-Risk Subset of the ASSURE Randomized Trial.* JAMA Oncol, 2017. **3**(9): p. 1249-1252.
98. Muglia, V.F. and A. Prando, *Renal cell carcinoma: histological classification and correlation with imaging findings.* Radiologia brasileira, 2015. **48**(3): p. 166-174.
99. Brugarolas, J., *Molecular genetics of clear-cell renal cell carcinoma.* J Clin Oncol, 2014. **32**(18): p. 1968-76.
100. Linehan, W.M., et al., *Comprehensive Molecular Characterization of Papillary Renal-Cell Carcinoma.* N Engl J Med, 2016. **374**(2): p. 135-45.
101. Garje, R., et al., *Comprehensive review of chromophobe renal cell carcinoma.* Critical Reviews in Oncology/Hematology, 2021. **160**: p. 103287.
102. Blum, K.A., et al., *Sarcomatoid renal cell carcinoma: biology, natural history and management.* Nat Rev Urol, 2020. **17**(12): p. 659-678.
103. Pedrosa, I., et al., *MR imaging of renal masses: correlation with findings at surgery and pathologic analysis.* Radiographics, 2008. **28**(4): p. 985-1003.
104. Arjumand, W. and S. Sultana, *Role of VHL gene mutation in human renal cell carcinoma.* Tumour Biol, 2012. **33**(1): p. 9-16.
105. UK, C.R. Accessed 9/1/23; Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/kidney-cancer#heading-Two.
106. Padala, S.A. and A. Kallam, *Clear Cell Renal Carcinoma*, in *StatPearls*. 2022, StatPearls Publishing

107. Motzer, R.J., et al., *Conditional survival and long-term efficacy with nivolumab plus ipilimumab versus sunitinib in patients with advanced renal cell carcinoma.* Cancer, 2022. **128**(11): p. 2085-2097.
108. Rao, A., C. Wiggins, and R.C. Lauer, *Survival outcomes for advanced kidney cancer patients in the era of targeted therapies.* Annals of translational medicine, 2018. **6**(9): p. 165-165.
109. Mendhiratta, N., et al., *Papillary renal cell carcinoma: Review.* Urologic Oncology: Seminars and Original Investigations, 2021. **39**(6): p. 327-337.
110. Mendhiratta, N., et al., *Papillary renal cell carcinoma: Review.* Urol Oncol, 2021. **39**(6): p. 327-337.
111. Attalla, K., M.H. Voss, and A.A. Hakimi, *Prognostic models in papillary renal cell carcinoma.* Annals of translational medicine, 2020. **8**(21): p. 1334-1334.
112. Ljungberg, B.J., et al., *Different vascular endothelial growth factor (VEGF), VEGF-receptor 1 and -2 mRNA expression profiles between clear cell and papillary renal cell carcinoma.* BJU Int, 2006. **98**(3): p. 661-7.
113. Rhoades Smith, K.E. and M.A. Bilen, *A Review of Papillary Renal Cell Carcinoma and MET Inhibitors.* Kidney Cancer, 2019. **3**(3): p. 151-161.
114. Ren, W., et al., *Prognostic factors for the survival of patients with papillary renal cell carcinoma after surgical management.* Clin Transl Oncol, 2020. **22**(5): p. 725-733.
115. Ledezma, R.A., et al., *Clinically localized type 1 and 2 papillary renal cell carcinomas have similar survival outcomes following surgery.* World J Urol, 2016. **34**(5): p. 687-93.
116. Herrmann, E., et al., *Prognostic factors of papillary renal cell carcinoma: results from a multi-institutional series after pathological review.* J Urol, 2010. **183**(2): p. 460-6.
117. Przybycin, C.G., et al., *Chromophobe renal cell carcinoma: a clinicopathologic study of 203 tumors in 200 patients with primary resection at a single institution.* Am J Surg Pathol, 2011. **35**(7): p. 962-70.
118. Büscheck, F., et al., *Prevalence and clinical significance of VHL mutations and 3p25 deletions in renal tumor subtypes.* Oncotarget, 2020. **11**(3): p. 237-249.
119. Yuan, X., C. Larsson, and D. Xu, *Mechanisms underlying the activation of TERT transcription and telomerase activity in human cancer: old actors and new players.* Oncogene, 2019. **38**(34): p. 6172-6183.
120. Gay-Bellile, M., et al., *TERT promoter status and gene copy number gains: effect on TERT expression and association with prognosis in breast cancer.* Oncotarget, 2017. **8**(44): p. 77540-77551.
121. Roldan-Romero, J.M., et al., *Molecular characterization of chromophobe renal cell carcinoma reveals mTOR pathway alterations in patients with poor outcome.* Mod Pathol, 2020. **33**(12): p. 2580-2590.
122. Amin, M.B., et al., *Chromophobe Renal Cell Carcinoma: Histomorphologic Characteristics and Evaluation of Conventional Pathologic Prognostic Parameters in 145 Cases.* The American Journal of Surgical Pathology, 2008. **32**(12).
123. de Peralta-Venturina, M., et al., *Sarcomatoid differentiation in renal cell carcinoma: a study of 101 cases.* Am J Surg Pathol, 2001. **25**(3): p. 275-84.
124. Chang, A., et al., *Use of PAX8 and GATA3 in diagnosing sarcomatoid renal cell carcinoma and sarcomatoid urothelial carcinoma.* Hum Pathol, 2013. **44**(8): p. 1563-8.
125. Mian, B.M., et al., *Prognostic factors and survival of patients with sarcomatoid renal cell carcinoma.* J Urol, 2002. **167**(1): p. 65-70.
126. Shuch, B., et al., *Sarcomatoid renal cell carcinoma: a comprehensive review of the biology and current treatment strategies.* The oncologist, 2012. **17**(1): p. 46-54.
127. Sircar, K., et al., *Biphasic components of sarcomatoid clear cell renal cell carcinomas are molecularly similar to each other, but distinct from, non-sarcomatoid renal carcinomas.* J Pathol Clin Res, 2015. **1**(4): p. 212-24.

128. Pal, S.K., et al., *RNA-seq reveals aurora kinase-driven mTOR pathway activation in patients with sarcomatoid metastatic renal cell carcinoma.* Mol Cancer Res, 2015. **13**(1): p. 130-7.
129. Tickoo, S.K., et al., *Immunohistochemical expression of hypoxia inducible factor-1alpha and its downstream molecules in sarcomatoid renal cell carcinoma.* J Urol, 2007. **177**(4): p. 1258-63.
130. Kyriakopoulos, C.E., et al., *Outcome of patients with metastatic sarcomatoid renal cell carcinoma: results from the International Metastatic Renal Cell Carcinoma Database Consortium.* Clin Genitourin Cancer, 2015. **13**(2): p. e79-85.
131. Vera-Badillo, F.E., et al., *Systemic Therapy for Non–clear Cell Renal Cell Carcinomas: A Systematic Review and Meta-analysis.* European Urology, 2015. **67**(4): p. 740-749.
132. Motzer, R.J., et al., *Phase II randomized trial comparing sequential first-line everolimus and second-line sunitinib versus first-line sunitinib and second-line everolimus in patients with metastatic renal cell carcinoma.* J Clin Oncol, 2014. **32**(25): p. 2765-72.
133. Armstrong, A.J., et al., *Everolimus versus sunitinib for patients with metastatic non-clear cell renal cell carcinoma (ASPEN): a multicentre, open-label, randomised phase 2 trial.* The Lancet Oncology, 2016. **17**(3): p. 378-388.
134. Tannir, N.M., et al., *Everolimus versus sunitinib prospective evaluation in metastatic non–clear cell renal cell carcinoma (ESPN): a randomized multicenter phase 2 trial.* European urology, 2016. **69**(5): p. 866-874.
135. Knox, J., et al., *Final overall survival analysis for the phase II RECORD-3 study of first-line everolimus followed by sunitinib versus first-line sunitinib followed by everolimus in metastatic RCC.* Annals of Oncology, 2017. **28**(6): p. 1339-1345.
136. Choueiri, T.K., et al., *Phase II and biomarker study of the dual MET/VEGFR2 inhibitor foretinib in patients with papillary renal cell carcinoma.* Journal of Clinical Oncology, 2013. **31**(2): p. 181.
137. Albiges, L., et al., *Non–Clear Cell Renal Cell Carcinomas: From Shadow to Light.* Journal of Clinical Oncology, 2018. **36**(36): p. 3624-3631.
138. Choueiri, T.K., et al., *Efficacy of sunitinib and sorafenib in metastatic papillary and chromophobe renal cell carcinoma.* J Clin Oncol, 2008. **26**(1): p. 127-31.
139. Borchiellini, D. and P. Barthélémy, *Cabozantinib: a new first-line option for papillary renal cell carcinoma?* Lancet, 2021. **397**(10275): p. 645-647.
140. Choueiri, T.K., et al., *Cabozantinib Versus Sunitinib As Initial Targeted Therapy for Patients With Metastatic Renal Cell Carcinoma of Poor or Intermediate Risk: The Alliance A031203 CABOSUN Trial.* Journal of Clinical Oncology, 2016. **35**(6): p. 591-597.
141. Golshayan, A.R., et al., *Metastatic sarcomatoid renal cell carcinoma treated with vascular endothelial growth factor-targeted therapy.* J Clin Oncol, 2009. **27**(2): p. 235-41.
142. Choueiri, T.K., et al., *PD-L1 expression in nonclear-cell renal cell carcinoma.* Ann Oncol, 2014. **25**(11): p. 2178-2184.
143. Motzer, R.J., et al., *Nivolumab plus Ipilimumab versus Sunitinib in Advanced Renal-Cell Carcinoma.* New England Journal of Medicine, 2018. **378**(14): p. 1277-1290.
144. Iacovelli, R., et al., *Patients with sarcomatoid renal cell carcinoma - re-defining the first-line of treatment: A meta-analysis of randomised clinical trials with immune checkpoint inhibitors.* Eur J Cancer, 2020. **136**: p. 195-203.
145. Vogelzang, N.J., et al., *Safety and Efficacy of Nivolumab in Patients With Advanced Non-Clear Cell Renal Cell Carcinoma: Results From the Phase IIIb/IV CheckMate 374 Study.* Clin Genitourin Cancer, 2020. **18**(6): p. 461-468.e3.
146. McDermott, D.F., et al., *Atezolizumab, an Anti-Programmed Death-Ligand 1 Antibody, in Metastatic Renal Cell Carcinoma: Long-Term Safety, Clinical Activity, and Immune Correlates From a Phase Ia Study.* J Clin Oncol, 2016. **34**(8): p. 833-42.

147. Gupta, R., et al., *Clinical activity of ipilimumab plus nivolumab (Ipi/Nivo) in patients (pts) with metastatic non-clear cell renal cell carcinoma (nccRCC).* Journal of Clinical Oncology, 2019. **37**(15_suppl): p. e16084-e16084.
148. Ross, K. and R.J. Jones, *Immune checkpoint inhibitors in renal cell carcinoma.* Clinical science (London, England : 1979), 2017. **131**(21): p. 2627-2642.
149. Pal, S.K., et al., *Cabozantinib in Combination With Atezolizumab for Advanced Renal Cell Carcinoma: Results From the COSMIC-021 Study.* J Clin Oncol, 2021. **39**(33): p. 3725-3736.
150. Campbell, S.C., et al., *Renal Mass and Localized Renal Cancer: Evaluation, Management, and Follow-Up: AUA Guideline: Part I.* J Urol, 2021. **206**(2): p. 199-208.
151. Narayan, V., et al., *Patterns of Relapse and Implications for Post-Nephrectomy Surveillance in Patients with High Risk Nonclear Cell Renal Cell Carcinoma: Subgroup Analysis of the Phase 3 ECOG-ACRIN E2805 Trial.* The Journal of urology, 2019. **201**(1): p. 62-68.
152. Abu-Ghanem, Y., et al., *The Impact of Histological Subtype on the Incidence, Timing, and Patterns of Recurrence in Patients with Renal Cell Carcinoma After Surgery-Results from RECUR Consortium.* Eur Urol Oncol, 2021. **4**(3): p. 473-482.
153. Dudani, S., et al., *Evaluation of Clear Cell, Papillary, and Chromophobe Renal Cell Carcinoma Metastasis Sites and Association With Survival.* JAMA Netw Open, 2021. **4**(1): p. e2021869.
154. Neves, J.B., et al., *Pattern, timing and predictors of recurrence after surgical resection of chromophobe renal cell carcinoma.* World J Urol, 2021. **39**(10): p. 3823-3831.
155. Thomas, A.Z., et al., *The Role of Metastasectomy in Patients with Renal Cell Carcinoma with Sarcomatoid Dedifferentiation: A Matched Controlled Analysis.* J Urol, 2016. **196**(3): p. 678-84.
156. van Wanrooij, L.L., et al., *Pooling individual participant data from randomized controlled trials: Exploring potential loss of information.* PloS one, 2020. **15**(5): p. e0232970-e0232970.
157. Serour, D.K., et al., *Lymph nodes in the head and neck cancer: would diffusion-weighted magnetic resonance imaging solve the diagnostic dilemma?* Egyptian Journal of Radiology and Nuclear Medicine, 2020. **51**(1): p. 190.
158. Dabestani, S., et al., *Renal cell carcinoma recurrences and metastases in primary non-metastatic patients: a population-based study.* World Journal of Urology, 2016. **34**(8): p. 1081-1086.
159. Ljungberg, B., et al., *EAU guidelines on renal cell carcinoma: 2014 update.* Eur Urol, 2015. **67**(5): p. 913-24.
160. Hellman, S. and R.R. Weichselbaum, *Oligometastases.* J Clin Oncol, 1995. **13**(1): p. 8-10.
161. Niibe, Y. and K. Hayakawa, *Oligometastases and oligo-recurrence: the new era of cancer therapy.* Jpn J Clin Oncol, 2010. **40**(2): p. 107-11.
162. Niibe, Y. and J.Y. Chang, *Novel insights of oligometastases and oligo-recurrence and review of the literature.* Pulm Med, 2012. **2012**: p. 261096.
163. Alt, A.L., et al., *Survival after complete surgical resection of multiple metastases from renal cell carcinoma.* Cancer, 2011. **117**(13): p. 2873-2882.
164. Donini, M., et al., *Management of oligometastatic and oligoprogressive renal cell carcinoma: state of the art and future directions.* Expert Rev Anticancer Ther, 2020. **20**(6): p. 491-501.
165. Magné, N. and I. Latorzeff, *Oligometastatic renal cell carcinoma: radiotherapy as a new standard of care?* Lancet Oncol, 2021. **22**(12): p. 1644-1645.
166. Van Cutsem, E., et al., *ESMO consensus guidelines for the management of patients with metastatic colorectal cancer.* Ann Oncol, 2016. **27**(8): p. 1386-422.
167. Argilés, G., et al., *Localised colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up.* Ann Oncol, 2020. **31**(10): p. 1291-1305.
168. Onaitis, M.W., et al., *Prognostic factors for recurrence after pulmonary resection of colorectal cancer metastases.* Ann Thorac Surg, 2009. **87**(6): p. 1684-8.
169. Vatandoust, S., T.J. Price, and C.S. Karapetis, *Colorectal cancer: Metastases to a single organ.* World J Gastroenterol, 2015. **21**(41): p. 11767-76.

170. Gonzalez, M., et al., *Risk factors for survival after lung metastasectomy in colorectal cancer patients: a systematic review and meta-analysis.* Ann Surg Oncol, 2013. **20**(2): p. 572-9.
171. Treasure, T., et al., *Pulmonary Metastasectomy versus Continued Active Monitoring in Colorectal Cancer (PulMiCC): a multicentre randomised clinical trial.* Trials, 2019. **20**(1): p. 718.
172. Nardo, B., et al., *Liver resection for metastases from colorectal cancer in very elderly patients: New surgical horizons.* Int J Surg, 2016. **33 Suppl 1**: p. S135-41.
173. Biller, L.H. and D. Schrag, *Diagnosis and Treatment of Metastatic Colorectal Cancer: A Review.* Jama, 2021. **325**(7): p. 669-685.
174. Coyle, C., et al., *ADD-ASPIRIN: A phase III, double-blind, placebo controlled, randomised trial assessing the effects of aspirin on disease recurrence and survival after primary therapy in common non-metastatic solid tumours.* Contemp Clin Trials, 2016. **51**: p. 56-64.
175. Ost, P., et al., *Surveillance or Metastasis-Directed Therapy for Oligometastatic Prostate Cancer Recurrence: A Prospective, Randomized, Multicenter Phase II Trial.* J Clin Oncol, 2018. **36**(5): p. 446-453.
176. Phillips, R., et al., *Outcomes of Observation vs Stereotactic Ablative Radiation for Oligometastatic Prostate Cancer: The ORIOLE Phase 2 Randomized Clinical Trial.* JAMA Oncol, 2020. **6**(5): p. 650-659.
177. Palma, D.A., et al., *Stereotactic Ablative Radiotherapy for the Comprehensive Treatment of Oligometastatic Cancers: Long-Term Results of the SABR-COMET Phase II Randomized Trial.* J Clin Oncol, 2020. **38**(25): p. 2830-2838.
178. Gillessen, S., et al., *Management of Patients with Advanced Prostate Cancer. Part I: Intermediate-/High-risk and Locally Advanced Disease, Biochemical Relapse, and Side Effects of Hormonal Treatment: Report of the Advanced Prostate Cancer Consensus Conference 2022.* Eur Urol, 2023. **83**(3): p. 267-293.
179. Wuttig, D., et al., *Gene signatures of pulmonary metastases of renal cell carcinoma reflect the disease-free interval and the number of metastases per patient.* Int J Cancer, 2009. **125**(2): p. 474-82.
180. Turajlic, S., et al., *Tracking Cancer Evolution Reveals Constrained Routes to Metastases: TRACERx Renal.* Cell, 2018. **173**(3): p. 581-594.e12.
181. Mehta, M.P., et al., *The American Society for Therapeutic Radiology and Oncology (ASTRO) evidence-based review of the role of radiosurgery for brain metastases.* Int J Radiat Oncol Biol Phys, 2005. **63**(1): p. 37-46.
182. Wang, X.S., et al., *Stereotactic body radiation therapy for management of spinal metastases in patients without spinal cord compression: a phase 1-2 trial.* Lancet Oncol, 2012. **13**(4): p. 395-402.
183. Goering, J.D., et al., *Cryoablation and liver resection for noncolorectal liver metastases.* American journal of surgery, 2002. **183 4**: p. 384-9.
184. Kato, S., et al., *Surgical metastasectomy for renal cell carcinoma: which patients are the real candidates for surgery?* Ann Transl Med, 2019. **7**(Suppl 8): p. S273.
185. Lyon, T.D., et al., *Complete Surgical Metastasectomy of Renal Cell Carcinoma in the Post-Cytokine Era.* J Urol, 2020. **203**(2): p. 275-282.
186. Dabestani, S., et al., *Local treatments for metastases of renal cell carcinoma: a systematic review.* Lancet Oncol, 2014. **15**(12): p. e549-61.
187. Pogrebniak, H.W., et al., *Renal cell carcinoma: resection of solitary and multiple metastases.* Ann Thorac Surg, 1992. **54**(1): p. 33-8.
188. Jakubowski, C.D., et al., *Complete metastasectomy for renal cell carcinoma: Comparison of five solid organ sites.* J Surg Oncol, 2016. **114**(3): p. 375-9.
189. Kavolius, J.P., et al., *Resection of metastatic renal cell carcinoma.* J Clin Oncol, 1998. **16**(6): p. 2261-6.

190. Zhang, Y., et al., *Stereotactic Ablative Radiation Therapy (SAbR) Used to Defer Systemic Therapy in Oligometastatic Renal Cell Cancer.* Int J Radiat Oncol Biol Phys, 2019. **105**(2): p. 367-375.
191. Marvaso, G., et al., *Oligo metastatic renal cell carcinoma: stereotactic body radiation therapy, if, when and how?* Clin Transl Oncol, 2021. **23**(8): p. 1717-1726.
192. Tosco, L., et al., *Survival and impact of clinical prognostic factors in surgically treated metastatic renal cell carcinoma.* Eur Urol, 2013. **63**(4): p. 646-52.
193. Tosco, L., et al., *Survival and Impact of Clinical Prognostic Factors in Surgically Treated Metastatic Renal Cell Carcinoma.* European Urology, 2013. **63**(4): p. 646-652.
194. Rodriguez-Covarrubias, F., et al., *Time to recurrence after nephrectomy as a predictor of cancer-specific survival in localized clear-cell renal cell carcinoma.* Urol Int, 2011. **86**(1): p. 47-52.
195. Park, Y.H., et al., *Late recurrence of renal cell carcinoma >5 years after surgery: clinicopathological characteristics and prognosis.* BJU Int, 2012. **110**(11 Pt B): p. E553-8.
196. Dabestani, S., et al., *Renal cell carcinoma recurrences and metastases in primary non-metastatic patients: a population-based study.* World J Urol, 2016. **34**(8): p. 1081-6.
197. Miyao, N., et al., *Late recurrence of renal cell carcinoma: retrospective and collaborative study of the Japanese Society of Renal Cancer.* Urology, 2011. **77**(2): p. 379-84.
198. Abreu, D., et al., *Prognostic Factors in De Novo Metastatic Renal Cell Carcinoma: A Report From the Latin American Renal Cancer Group.* JCO Glob Oncol, 2021. **7**: p. 671-685.
199. Noguchi, G., et al., *Time-dependent change in relapse sites of renal cell carcinoma after curative surgery.* Clin Exp Metastasis, 2018. **35**(1-2): p. 69-75.
200. Crispen, P.L., et al., *Lymph node dissection at the time of radical nephrectomy for high-risk clear cell renal cell carcinoma: indications and recommendations for surgical templates.* Eur Urol, 2011. **59**(1): p. 18-23.
201. Chow, F.C. and K.S. Chok, *Colorectal liver metastases: An update on multidisciplinary approach.* World J Hepatol, 2019. **11**(2): p. 150-172.
202. Werfel, T.A., et al., *Repurposing of a Thromboxane Receptor Inhibitor Based on a Novel Role in Metastasis Identified by Phenome-Wide Association Study.* Mol Cancer Ther, 2020. **19**(12): p. 2454-2464.
203. Lucotti, S., et al., *Aspirin blocks formation of metastatic intravascular niches by inhibiting platelet-derived COX-1/thromboxane A2.* J Clin Invest, 2019. **129**(5): p. 1845-1862.
204. Khene, Z.E., D. Borchiellini, and K. Bensalah, *Re: Adjuvant Atezolizumab Versus Placebo for Patients with Renal Cell Carcinoma at Increased Risk of Recurrence Following Resection (IMmotion010): A Multicentre, Randomised, Double-blind, Phase 3 Trial.* Eur Urol, 2023.
205. Patel, H.D., et al., *The future of perioperative therapy in advanced renal cell carcinoma: how can we PROSPER?* Future Oncol, 2019. **15**(15): p. 1683-1695.
206. Agency, E.M., *Summary of product characteristics*. 2022.
207. Vickers, A.J., B. Van Calster, and E.W. Steyerberg, *Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests.* Bmj, 2016. **352**: p. i6.
208. Rini, B., et al., *A 16-gene assay to predict recurrence after surgery in localised renal cell carcinoma: development and validation studies.* Lancet Oncol, 2015. **16**(6): p. 676-85.
209. Services, U.S.D.o.H.a.H. *Renal Cell Carcinoma: Developing Drugs and Biologics for Adjuvant Treatment Guidance for Industry*. 2020; Available from: https://www.fda.gov/media/142553/download.
210. Middleton, G., et al., *The National Lung Matrix Trial of personalized therapy in lung cancer.* Nature, 2020. **583**(7818): p. 807-812.
211. Chu, Y. and Y. Yuan, *A Bayesian basket trial design using a calibrated Bayesian hierarchical model.* Clin Trials, 2018. **15**(2): p. 149-158.

212. Zaorsky, N.G., et al., *Stereotactic ablative radiation therapy for oligometastatic renal cell carcinoma (SABR ORCA): a meta-analysis of 28 studies.* Eur Urol Oncol, 2019. **2**(5): p. 515-523.
213. Bernstein, M.B., et al., *Immunotherapy and stereotactic ablative radiotherapy (ISABR): a curative approach?* Nat Rev Clin Oncol, 2016. **13**(8): p. 516-24.
214. Chicas-Sett, R., et al., *Stereotactic Ablative Radiotherapy Combined with Immune Checkpoint Inhibitors Reboots the Immune Response Assisted by Immunotherapy in Metastatic Lung Cancer: A Systematic Review.* Int J Mol Sci, 2019. **20**(9).
215. Brahmer, J., et al., *Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer.* N Engl J Med, 2015. **373**(2): p. 123-35.
216. Herbst, R.S., et al., *Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial.* Lancet, 2016. **387**(10027): p. 1540-1550.
217. Schumacher, W.A., et al., *Antiplatelet activity of the long-acting thromboxane receptor antagonist BMS 180,291 in monkeys.* Prostaglandins, 1992. **44**(5): p. 389-97.

Appendix

SORCE trial protocol can be found here:

https://www.ctu.mrc.ac.uk/media/1297/sorce-protocol-v70-jul-2017.pdf

Table A: Comparison of systems used to define nuclear grade; Fuhrman grading, WHO/ISUP and the modified system used in SORCE and by Leibovich and colleagues to determine the 2003 Leibovich score.

GRADE	FUHRMAN	WHO/ISUP	2003 LEIBOVICH AND SORCE GRADING SYSTEM
1	Small nuclei (10 microns) with round uniform nuclei, inconspicuous nucleoli	Inconspicuous or absent nucleoli at x400 magnification	Small, round Inconspicuous, visible only at x400 magnification
2	Larger nuclei (15 microns) slightly irregular and with small nucleoli visible at x400.	Nucleoli should be distinctly visible at x400, but inconspicuous or invisible at x100 magnification;	Round to slightly irregular Mildly enlarged, visible at x200 magnification
3	Larger nuclei (20 microns) with irregular outlines and large prominent nucleoli at x100	Nucleoli should be distinctly visible at x100 magnification.	Round to irregular Prominent, visible at x100 magnification
4	Pleomorphic nuclei, e.g., polylobated	Extreme nuclear pleomorphism and those showing rhabdoid or sarcomatoid differentiation	Enlarged, pleomorphic or giant cells

Table B: Definition of histological tumour necrosis in the SORCE protocol

Histological tumour necrosis: The presence of any microscopic, coagulative tumour necrosis and is distinguished from degenerative changes such as hyalinization, haemorrhage and fibrosis.

Table C: Comparison of scales of performance status used to assess patients in SORCE and ASSURE

WHO (SORCE)	ECOG (ASSURE)
0: able to carry out all normal activity without restriction	0: Fully active; no performance restrictions.
1: restricted in strenuous activity but ambulatory and able to carry out light work	1: Strenuous physical activity restricted; fully ambulatory and able to carry out light work.
2: ambulatory and capable of all self-care but unable to carry out any work activities; up and about more than 50% of waking hours	2: Capable of all self-care but unable to carry out any work activities. Up and about >50% of waking hours.
3: symptomatic and in a chair or in bed for greater than 50% of the day but not bedridden	Capable of only limited self-care; confined to bed or chair >50% of waking hours.
4: completely disabled; cannot carry out any self-care; totally confined to bed or chair.	Completely disabled; cannot carry out any self-care; totally confined to bed or chair.

Table E Comparison of participant eligibility criteria SORCE and ASSURE

	ASSURE	SORCE
Disease characteristics	Primary-intact RCC treated with curative intent M0 disease; fully resected nodes, renal vein thrombosis, IVC thrombus No residual macroscopic disease on post operative CT (<4 weeks post-surgery) All surgical specimens must have negative margins >= pT1bNany (resectable) Randomization <12 weeks after surgery	All RCCs except pure oncocytoma No residual macroscopic disease on post operative CT (<28 days post-surgery) No pulmonary nodules ≥ 5mm diameter or multiple pulmonary nodules. Intermediate- or high-risk disease (Leibovich score 3 to 11) Surgery for RCC at least 4 weeks but no more than 3 months prior to randomisation
Patient characteristics	ECOG PS 0-1 Platelet count >= 100,000/mm^3 Serum creatinine =< 2.0 x ULN Total bilirubin =< 1.5 x ULN SGOT and SGPT =< 2.5 x ULN AGC>= 1,500/mm^3 QTc < 500 msec on baseline ECG No prior anti-cancer therapy for RCC No other current malignancies, except those specified in protocol No serious intercurrent illness including; Clinically significant cardiovascular disease New York Heart Association grade II or greater congestive heart failure Psychiatric illness Myocardial infarction within the last 6 months. Uncontrolled HTN No HIV Not pregnant or nursing	WHO PS 0-1 Platelet count > 99,000/mm³ Serum Creatinine < 2.5 x ULN LFTs< 1.5 x ULN serum amylase < 1.5 x ULN WBC > 3,400/mm³ PT/INR < 1.5 times ULN PTT < 1.5 times ULN No prior anti-cancer therapy for RCC No prior malignancy (except those specified in protocol No cardiac arrhythmias or uncontrolled hypertension No congestive cardiac failure > NYHA Class II. · Active No clinically serious bacterial or fungal infections. Not pregnant or nursing No known history of HIV infection No chronic hepatitis B or C

KEY: ECOG; Eastern Cooperative Oncology Group. PS; performance status, WHO; world health organisation, RCC; renal cell carcinoma, SGOT; Serum glutamic oxaloacetic transaminase SGPT; serum glutamate pyruvate transaminase, cQT; Corrected QT, ECG; electrocardiogram, AGC; Absolute granulocyte count, M0; non-metastatic; HTN; hypertension, HIV; human immunodeficiency virus, NYHA; New York Heart Association PT; prothrombin time PTT; partial thromboplastic time INR' International normalised ratio WBC; white blood cell, LFT; liver function test, IVC; inferior vena-cava CT; computerised tomography

Table F: ASSURE Baseline Characteristics by Subtype

Variable at Baseline	Clear cell Total (N=1436)	Papillary Total (N=141)	Chromophobe (N=105)	Sarcomatoid (N=118)	Total (N=1800)
Age (years)	56.3 (20 - 84)	57 (28 - 80)	49.8 (24 - 84)	56.5 (27 - 81)	56 (20 - 84)
Sex					
Males	972 (68%)	108 (77%)	54 (51%)	85 (72%)	1219 (68%)
Females	464 (32%)	33 (23%)	51 (49%)	33 (28%)	581 (32%)
Performance Status					
0	1150 (80%)	114 (81%)	88 (84%)	88 (75%)	1440 (80%)
1	251 (17%)	25 (18%)	15 (14%)	28 (24%)	319 (18%)
2	1 (0%)				1 (0%)
3	2 (0%)	1 (1%)			3 (1%)
Missing	32 (2%)	1 (1%)	2 (2%)	2 (2%)	37 (2%)
Risk Category					
Intermediate High	716 (50%)	79 (56%)	75 (71%)	49 (42%)	919 (51%)
Very High	720 (50%)	62 (44%)	30 (29%)	69 (58%)	881 (49%)
Pathological T stage					
T1	147 (10%)	17 (12%)	7 (7%)	11 (9%)	182 (10%)
T2	363 (25%)	49 (35%)	50 (48%)	27 (23%)	489 (27%)
T3	916 (64%)	73 (52%)	47 (45%)	78 (66%)	1114 (62%)
T4	10 (1%)	2 (1%)	1 (1%)	2 (2%)	15 (1%)
Pathological N stage					
N0	553 (39%)	34 (24%)	40 (38%)	43 (36%)	670 (37%)
N1	41 (3%)	15 (11%)	3 (3%)	16 (14%)	75 (4%)
N2	33 (2%)	18 (13%)	3 (3%)	11 (9%)	65 (4%)
NX	809 (56%)	74 (52%)	59 (56%)	48 (41%)	990 (55%)
Fuhrman Grade					
Grade 1	39 (3%)	5 (4%)	2 (2%)	1 (1%)	47 (3%)
Grade 2	486 (34%)	46 (33%)	40 (38%)	10 (8%)	582 (32%)
Grade 3	697 (49%)	75 (53%)	51 (49%)	25 (21%)	848 (47%)
Grade 4	212 (15%)	11 (8%)	6 (6%)	81 (69%)	310 (17%)
Missing	2 (0%)	4 (3%)	6 (6%)	1 (1%)	13 (1%)
Type of Surgery					
Radical	1363 (95%)	122 (87%)	104 (99%)	111 (94%)	1700 (94%)
Partial	73 (5%)	19 (13%)	1 (1%)	7 (6%)	100 (6%)
Surgical Approach					
Open	808 (56%)	84 (60%)	55 (52%)	71 (60%)	1018 (57%)
Laparoscopic	628 (44%)	57 (40%)	50 (48%)	47 (40%)	782 (43%)
DFS Event					
No	810 (56%)	88 (62%)	90 (86%)	38 (32%)	1026 (57%)

Yes	626 (44%)	53 (38%)	15 (14%)	80 (68%)	774 (43%)
Death					
Alive/Censored	1138 (79%)	113 (80%)	104 (99%)	67 (57%)	1422 (79%)
Dead	298 (21%)	28 (20%)	1 (1%)	51 (43%)	378 (21%)

Table G: SORCE Baseline Characteristic by Subtype

Variable at Baseline	Clear Cell (N=1445)	Papillary (N=128)	Chromophobe (N=96)	Sarcomatoid (N=20)	Total (N=1689)
Age at Randomisation	58.5 (20.1 - 85.6)	58.2 (18.7 - 78.6)	54.9 (27.8 - 82.3)	56.2 (41.1 - 75.8)	58.2 (18.7 - 85.6)
Sex					
Female	426 (29%)	26 (20%)	33 (34%)	6 (30%)	491 (29%)
Male	1019 (71%)	102 (80%)	63 (66%)	14 (70%)	1198 (71%)
Performance Status					
0	1155 (80%)	99 (77%)	79 (82%)	13 (65%)	1346 (80%)
1	278 (19%)	29 (23%)	16 (17%)	7 (35%)	330 (20%)
2	1 (0%)				
Missing	11 (1%)		1 (1%)		12 (1%)
Pathological T cat. of Primary Tumour					
pT1a	5 (0%)		1 (1%)	1 (5%)	7 (0%)
pT1b	170 (12%)	17 (13%)	7 (7%)		194 (11%)
pT2	298 (21%)	43 (34%)	50 (52%)	5 (25%)	396 (23%)
pT3a-4	972 (67%)	68 (53%)	38 (40%)	14 (70%)	1092 (65%)
Regional Lymph Node Status					
pNx/ pN0	1405 (97%)	107 (84%)	90 (94%)	16 (80%)	1618 (96%)
pN1/ pN2	40 (3%)	21 (16%)	6 (6%)	4 (20%)	71 (4%)
Tumour Size					
<10	993 (69%)	78 (61%)	47 (49%)	15 (75%)	1133 (67%)
>10	452 (31%)	50 (39%)	49 (51%)	5 (25%)	556 (33%)
Nuclear Grade					
1	68 (5%)	16 (13%)	5 (5%)		89 (5%)
2	374 (26%)	32 (25%)	30 (31%)		436 (26%)
3	735 (51%)	68 (53%)	45 (47%)	3 (15%)	851 (50%)
4	268 (19%)	12 (9%)	15 (16%)	17 (85%)	312 (18%)
Missing			1 (1%)		1 (0%)
Histological Tumour Necrosis					
No	671 (46%)	43 (34%)	47 (49%)	6 (30%)	767 (45%)
Yes	774 (54%)	85 (66%)	49 (51%)	14 (70%)	922 (55%)
Leibovich Score Group					
Intermediate	776 (54%)	66 (52%)	61 (64%)	1 (5%)	904 (54%)
High	669 (46%)	62 (48%)	35 (36%)	19 (95%)	785 (46%)
Type of Nephrectomy					
Radical	1349 (93%)	112 (88%)	90 (94%)	18 (90%)	1569 (93%)
Partial	43 (3%)	10 (8%)	4 (4%)	2 (10%)	59 (3%)
Missing	53 (4%)	6 (5%)	2 (2%)		61 (4%)

Type of Operation					
Open	761 (53%)	70 (55%)	58 (60%)	9 (45%)	898 (53%)
Laparoscopic	606 (42%)	52 (41%)	35 (36%)	11 (55%)	704 (42%)
Missing	78 (5%)	6 (5%)	3 (3%)		87 (5%)
DFS Event					
No	812 (56%)	65 (51%)	71 (74%)	9 (45%)	957 (57%)
Yes	633 (44%)	63 (49%)	25 (26%)	11 (55%)	732 (43%)
OS Event					
No	1139 (79%)	90 (70%)	85 (89%)	14 (70%)	1328 (79%)
Yes	306 (21%)	38 (30%)	11 (11%)	6 (30%)	361 (21%)
RCC Survival Event					
No	1196 (83%)	97 (76%)	88 (92%)	14 (70%)	1395 (83%)
Yes	249 (17%)	31 (24%)	8 (8%)	6 (30%)	294 (17%)
RFS Event					
No	892 (62%)	74 (58%)	74 (77%)	9 (45%)	1049 (62%)
Yes	553 (38%)	54 (42%)	22 (23%)	11 (55%)	640 (38%)

Table H: Cause-Specific Adjusted Cox Models assessing the effect of histological subtype on A. risk of abdominal or B. chest relapse

A. Abdominal relapse

Variables	Hazard Ratio	CI	P value
T stage cont.	1.21	(1.05,1.39)	<0.01
N stage: pN1/pN2	1.45	(1.11,1.91)	<0.01
PS: >= 1	1.25	(1.01,1.55)	0.041
Study ID	1.22	(1,1.48)	0.044
Papillary	1.33	(.98,1.81)	0.063
Chromophobe	.663	(.38,1.15)	0.142
Sarcomatoid	2.13	(1.59,2.86)	<0.01

* T stage, nodal involvement, PS and study were adjusted for

B. Chest relapse

Variables	Hazard Ratio	CI	P value
T stage cont.	1.08	(0.92,1.26)	0.368
N stage: pN1/pN2	0.909	(0.60,1.37)	0.648
PS: >= 1	0.883	(0.70,1.11)	0.284
Study ID	1.18	(0.97,1.43)	0.091
Papillary	1.1	(0.71,1.71)	0.681
Chromophobe	2.72	(0.81,9.1)	0.105
Sarcomatoid	2.65	(1.81,3.88)	<0.01

* T stage, nodal involvement, PS and study were adjusted for

Table I: Cause-Specific Adjusted Cox Models assessing the effect of histological subtype on OS in patients who A. relapsed to the abdomen or B. to the chest.

A. OS – Abdominal relapses

Variables	Hazard Ratio	CI	P value
T stage cont.	1.63	(1.29,2.06)	<0.01
N stage: pN1/pN2	1.37	(0.97,1.94)	0.07
PS: >= 1	1.42	(1.07,1.89)	0.015
Study ID	1.12	(0.84,1.48)	0.438
Papillary	1.7	(1.15,2.5)	<0.01
Chromophobe	.388	(0.12,1.22)	0.105
Sarcomatoid	3.61	(2.53,5.15)	<0.01

* T stage, nodal involvement, PS and study were adjusted for

B. OS – Chest relapses

Variables	Hazard Ratio	CI	P value
T stage cont.	1.11	(0.87,1.4)	0.398
N stage: pN1/pN2	1.21	(0.66,2.2)	0.527
PS: >= 1	1.32	(0.97,1.8)	0.076
Study ID	1.01	(0.75,1.34)	0.973
Papillary	1.21	(0.65,2.24)	0.539
Chromophobe	1.85	(0.42,8.21)	0.421
Sarcomatoid	2.13	(1.33,3.42)	<0.01

* T stage, nodal involvement, PS and study were adjusted for

Table J: Distribution of local recurrences and distant LNs in SORCE and ASSURE

SORCE defined ‘local’ relapses as in either remnant kidney, local node, renal bed. ASSURE did not delineate between local vs distant nodal relapse

SORCE	Clear-cell	Papillary	Chromophobe	Sarcomatoid
Renal bed	56	8	2	1
Remnant kidney	6	1	0	0
Local node	11	2	1	0
Distant node*	138	18	7	4
Local and distant	8	3	0	0

Inclusive of relapses at multiple sites

ASSURE	Clear-cell	Papillary	Chromophobe	Sarcomatoid
Renal bed	58	9	2	15
Any node	64	13	4	20

Table K: Nodal anatomy at baseline in patients who relapsed to lymph-nodes in SORCE participants only

Table 3

Nodal status at baseline	Local/distant nodal relapse	Clear-cell	Papillary	Chromophobe	Sarcomatoid
NX/N0	Local	19	3	1	-
NX/N0	Distant*	174	19	7	12
N1/N2	Local	1	2	0	-
N1/N2	Distant*	31	13	5	12

*any non-local lymph-node within chest/abdomen/pelvis

Figure A: Kaplan-Meier plots of MFS in the derivation and validation cohorts, split by Leibovich score; scores of 9 and above were grouped together due to small numbers.

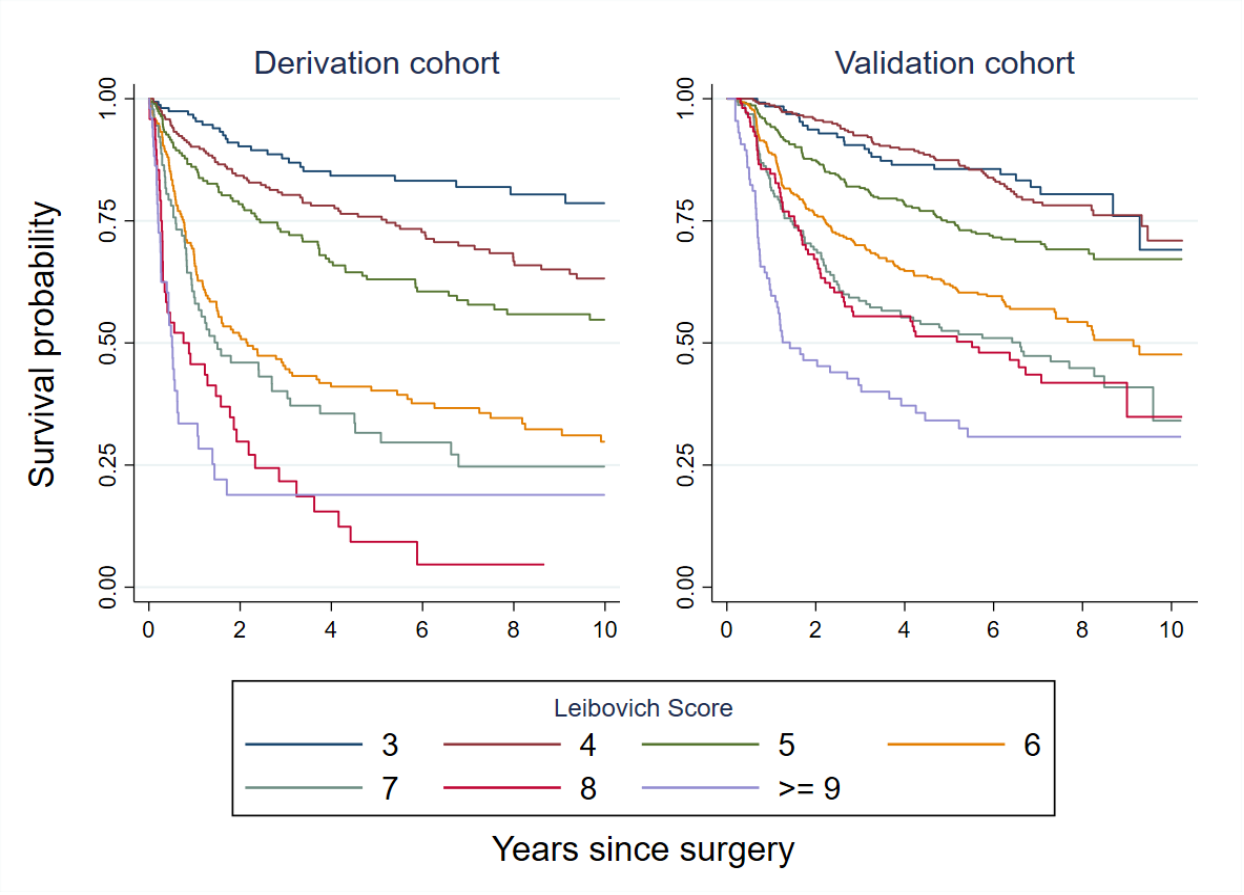
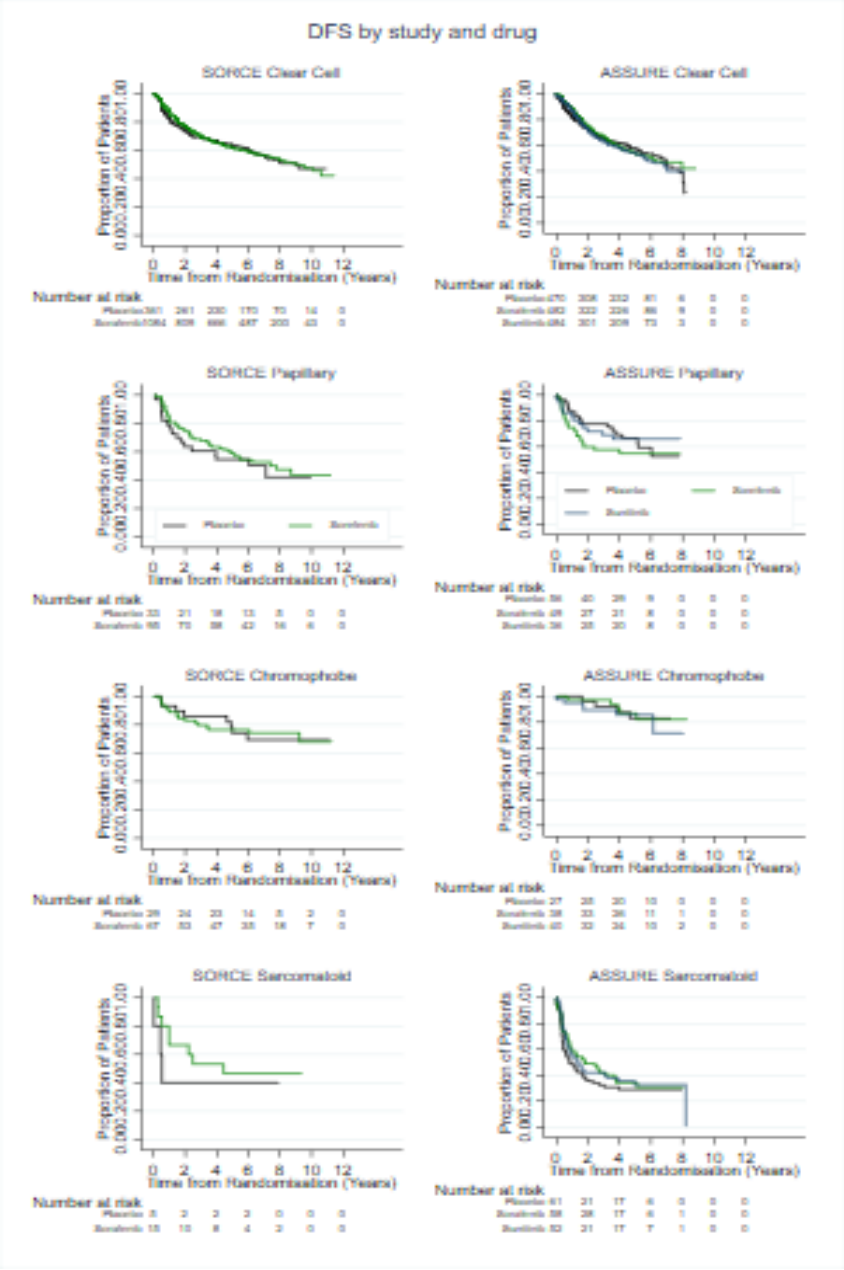


Figure B: The effect of TKI treatment on i. DFS and ii. OS in patients stratified by histological subtype for the SORCE and ASSURE trials separately.

i. DFS by study and TKI treatment



Histology	SORCE Log-rank p-values	ASSURE Log-rank p-values	
Clear Cell	0.930	0.703	
Papillary	0.595	0.423	
Chromophobe	0.887	0.931	
Sarcomatoid	0.373	0.566	

ii. OS by study and TKI treatment

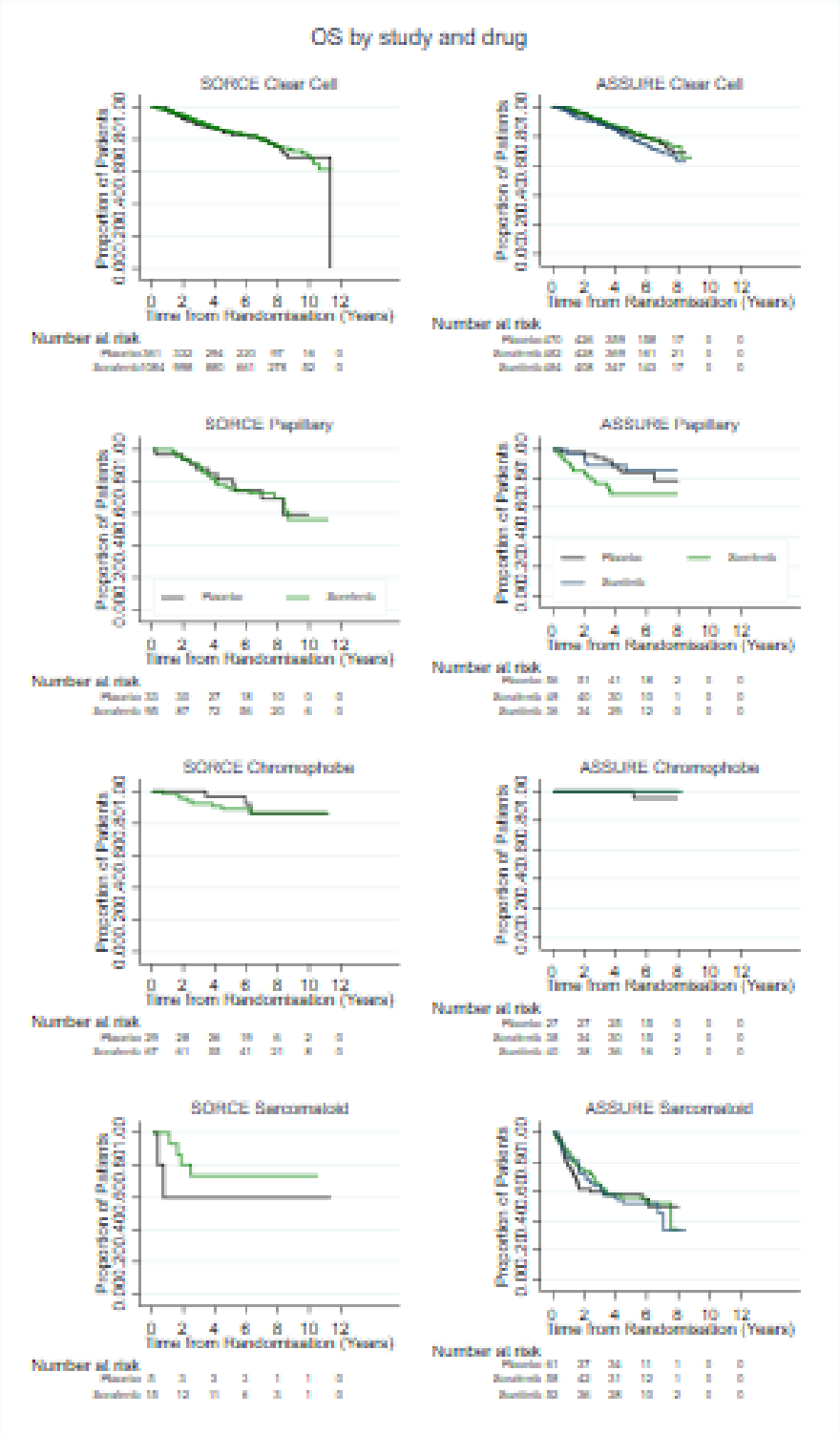
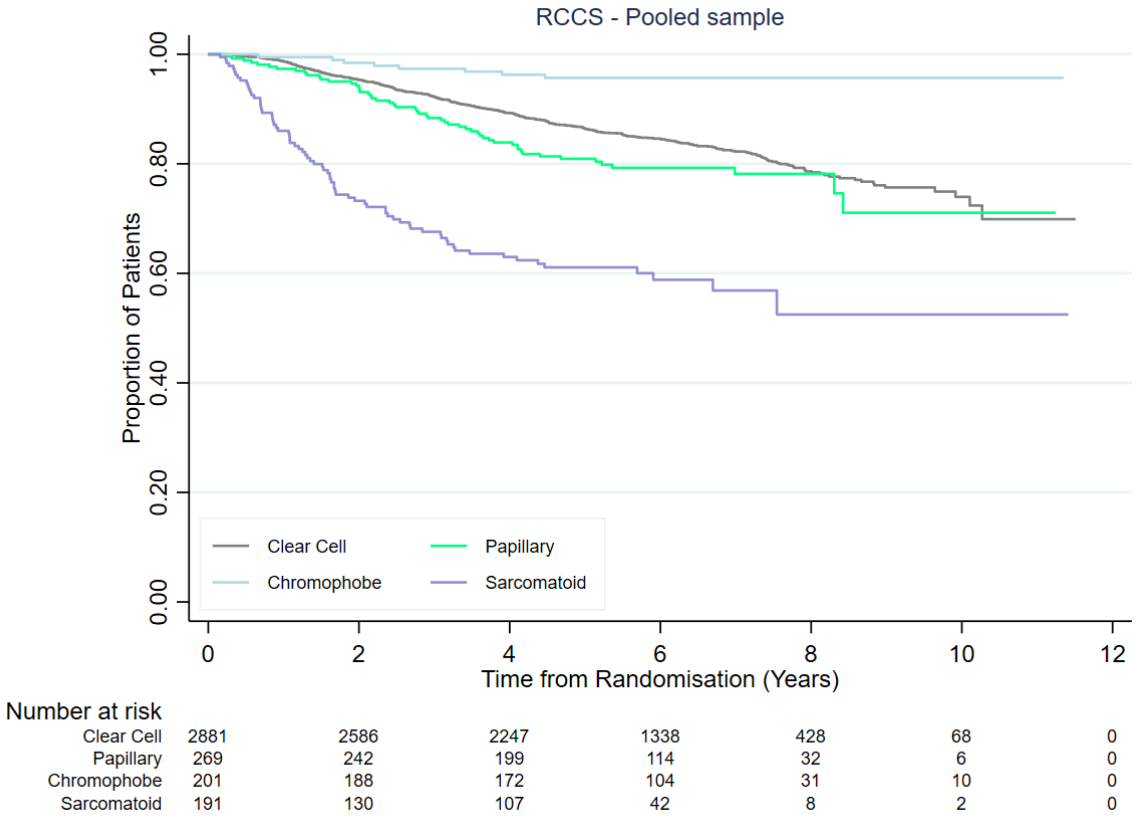


Figure C: RCCS for patients with pRCC, chRCC and sRCC compared to those with ccRCC



Number at risk

Clear Cell	2881	2586	2247	1338	428	68	0
Papillary	269	242	199	114	32	6	0
Chromophobe	201	188	172	104	31	10	0
Sarcomatoid	191	130	107	42	8	2	0

Chapter 1 Supplementary Analysis

Analysis of time-dependence of c-index

To assess whether the 2003 Leibovich score reduced its predictive ability at longer follow-ups, I conducted a longitudinal analysis of time dependence of c-index in three instances:

Analysis 1. 2003 Leibovich dataset, full range of risk groups, ungrouped risk score

Analysis 2. Updated Leibovich dataset, restricted range of risk groups, grouped risk score (scores 3-5 and scores > 6).

Analysis 3. SORCE dataset, grouped risk score

Motivation for analysis 1 was to correspond with what is reported in the 2003 Leibovich paper [23].

Motivation for analyses 2 and 3 was to work with the matched datasets and risk groups used in the primary analyses of this study.

To assess time dependence, maximum follow-up time, t*, was progressively restricted and c-indexes were evaluated (with 95% CIs). For each t* = 1 and 10 years, t* was censored for all observation times with more than t* years of follow-up. Bootstrap was used, with 1000 replicates, to estimate the standard error of the differences, confidence intervals, and the p-values for testing the differences against 0. The bootstrap distribution of the differences was close to normal in each case.

C-index	Analysis 1	Analysis 2	Analysis 3
t= 1 year	0.850	0.698	0.687
t= 10 years	0.823	0.672	0.633
difference	0.027	0.026	0.054
p-value (95% CI)	p<0.001 (0.013 - 0.042)	p = 0.006 (0.007 - 0.044)	p < 0.001 (0.025 to 0.082)

In analysis 1, values of c-indexes decrease steadily from 0.850 after 1 year follow-up to 0.823 after 10 yr. The difference is statistically significant, but modest: 0.027 (95% CI: 0.013 to 0.042; p < 0.001). The second two analyses show similar patterns to the first.

ASCO

Journal of Clinical Oncology®

External Validation of the 2003 Leibovich Prognostic Score in Patients Randomly Assigned to SORCE, an International Phase III Trial of Adjuvant Sorafenib in Renal Cell Cancer

Bhavna Oza, MBBS[1]; Tim Eisen, PhD[2]; Eleni Frangou, MSc (Res)[1]; Grant D. Stewart, MBChB[3]; Axel Bex, MD[4,5]; Alastair W.S. Ritchie, MD[6]; Rick Kaplan, MD[1]; Benjamin Smith, BSc[1]; Ian D. Davis, MBBS, PhD[7,8,9]; Martin R. Stockler, MBBS[9,10]; Laurence Albiges, MD[11]; Bernard Escudier, MD[11]; James Larkin, PhD[12]; Steven Joniau, MD[13]; Barry Hancock, MD[14]; Gregers G. Hermann, DMSc[15]; Joaquim Bellmunt, MD[16]; Mahesh K.B. Parmar, DPhil[1]; Patrick Royston, DSc[1]; and Angela Meade, DPhil[1]

PURPOSE The 2003 Leibovich score guides prognostication and selection to adjuvant clinical trials for patients with locally advanced renal cell carcinoma (RCC) after nephrectomy. We provide a robust external validation of the 2003 Leibovich score using contemporary data from SORCE, an international, randomized trial of sorafenib after excision of primary RCC.

METHODS Data used to derive the 2003 Leibovich score were compared with contemporary data from SORCE. Discrimination and calibration of the metastasis-free survival outcome were assessed in data from patients with clear-cell RCC, using Cox proportional hazards regression, Kaplan-Meier curves, and calculation of Harrell's c indexes. Secondary analyses involved three important SORCE groups: patients with any non–clear-cell subtype, papillary, and chromophobe carcinomas.

RESULTS Four hundred seven recurrences occurred in 982 patients in the Leibovich cohort and 520 recurrences were recorded in 1,445 patients in the primary SORCE cohort. Clear discrimination between intermediate-risk and high-risk SORCE cohorts was shown; hazard ratio 2.74 (95% CI, 2.29 to 3.28), c-index 0.63 (95% CI, 0.61 to 0.65). A hazard ratio of 0.61 (95% CI, 0.53 to 0.70) confirmed poor calibration of the two cohorts. Discrimination was observed in secondary populations, with c-indexes of 0.64 (95% CI, 0.59 to 0.69) for non–clear-cell RCC, 0.63 (95% CI, 0.56 to 0.69) for papillary RCC, and 0.65 (95% CI, 0.55 to 0.76) for chromophobe RCC.

CONCLUSION The 2003 Leibovich score discriminates between intermediate-risk and high-risk clear-cell and non–clear-cell RCC groups in contemporary data, supporting its use for risk stratification in adjuvant clinical trials. Over time, metastasis-free survival for patients with locally advanced RCC has improved. Contemporary data from adjuvant RCC trials should be used to improve prognostication for patients with RCC. J Clin Oncol 40:1772-1782.

ASSOCIATED CONTENT
Data Sharing Statement
Data Supplement
Protocol
Author affiliations and support information (if applicable) appear at the end of this article.
Accepted on January 20, 2022 and published at ascopubs.org/journal/jco on February 25, 2022: DOI https://doi.org/10.1200/JCO.21.01090

INTRODUCTION

The Leibovich score,[1] published in 2003, is widely used to guide postnephrectomy prognostication for patients with locally advanced renal cell carcinoma (RCC)[2] and for risk-stratifying patients into adjuvant clinical trials.[2]

Leibovich et al developed the score using retrospective data from patients with clear-cell RCC who underwent radical nephrectomy at the US Mayo Clinic between 1970 and 2000. Five features that were significantly associated with time-to-distant metastases (P, .001) comprised the final multivariable model: tumor category (6th TNM 2002), regional lymph-node status, maximum tumor diameter, nuclear grade, and presence of tumor necrosis. For clinical application, risk groups were defined as low (scores 0-2), intermediate (3-5), and high (6 or higher). Five-year metastasis-free probabilities were reported as 97.1%, 73.8%, and 31.2% respectively.[1]

The SORCE trial (ClinicalTrials.gov identifier: NCT00492258), evaluated the effect of sorafenib after nephrectomy and is one of the largest internationally recruiting randomized controlled trial in patients with locally advanced RCC, to date.[3] In SORCE, and now RAMPART (ClinicalTrials.gov identifier: NCT03288532),[4] the 2003 Leibovich score determines participant eligibility and guides their random allocation to trial arms.

Selection of the 2003 Leibovich score for this purpose is supported by its superior discriminative accuracy on direct comparison with several other prognostic scores.[5] Furthermore, the 2003 Leibovich score is

ASCO
Journal of Clinical Oncology®
1772 Volume 40, Issue 16

CONTEXT

Key Objective

The 2003 Leibovich score guides the prognostication and the selection of clear-cell and non–clear-cell patients with locally advanced renal cell carcinoma (RCC) into clinical trials. Its up-to-date validation in contemporary data is necessary to support its continued use. To our knowledge, an evaluation of the 2003 Leibovich score's discrimination between risk groups for non–clear-cell RCCs has not previously been demonstrated.

Knowledge Generated

The 2003 Leibovich score demonstrated discriminative accuracy in contemporary clear-cell and non–clear-cell groups, supporting its use for recruiting and guiding the random assignment of participants to adjuvant RCC trials. Outcomes for patients with RCC have improved over time, rendering the 2003 Leibovich score poorly calibrated to contemporary outcomes.

Relevance

We support the use of the 2003 Leibovich score to risk-stratify patients with RCC suspected of being at intermediate or high

risk of relapse.

simple to calculate. All score components are tumorderived and routinely reported on RCC pathology, negating the need for additional expertise or training. Clinical markers such as patient's performance status are not included in the score, reducing the chance of subjective bias.

An external validation of the 2003 Leibovich score, using data from SORCE participants, was prespecified within the SORCE Protocol (online only). We focused on the intermediate-risk and high-risk patients as they are of specific interest for recruitment to adjuvant clinical trials. SORCE provided a large contemporary data set of individual participant data (IPD) with detailed and long follow-up. Unusually for a validation study, we accessed IPD used to derive the 2003 Leibovich score.[1] By creating closely matched data sets, we were able to compute measures of discrimination and calibration,[6-8] to directly compare the performance of the 2003 Leibovich score in the historical and contemporary cohorts. Accordingly, we provide a high-quality evaluation of the Leibovich score's ability to discriminate between patients at intermediate risk and high risk of relapse.

Although the 2003 Leibovich score is used in clinical trials that recruit patients with non–clear-cell RCC, its ability to stratify risk in this group has not been evaluated. Newer prognostic scores[9-11] have been developed (including some specifically for non–clear-cell subtypes), but none are commonly used in clinical trials, where straightforward application is key. We present the first exploration of the 2003 Leibovich score's discriminative accuracy within important histologic SORCE subpopulations: any non–clear-cell, papillary-only, and chromophobe-only carcinomas.

METHODS

Participants: Leibovich Score Calculation

SORCE participants were recruited from July 2007 to April 2013 from 147 centers in seven countries: United Kingdom, Australia, France, Belgium, the Netherlands, Spain, and Denmark, and followed up until July 2019.[3] Only patients with intermediate (3-5) or high ($ 6) Leibovich scores were included in SORCE.[1,3] Participants with any histology except pure oncocytoma were eligible. Values for components of the 2003 Leibovich score (Data Supplement, online only) were prospectively collected for each participant on random assignment to SORCE.

The 6th TNM 2002 system was used by Leibovich et al and in SORCE. The same nuclear grading system that selects the worst WHO/International Society of Urological Pathology[12] features at each grade was used in both data sets (Data Supplement). In SORCE, this system pragmatically standardized grading across international trial sites and was used for all histologic subtypes including chromophobe and other non–clear-cell RCCs.

The SORCE trial was approved by national regulatory and ethical committees in each participating country and was conducted in accordance with the principles of Good Clinical Practice, the Declaration of Helsinki, and all applicable regulatory requirements and laws. All participants signed an informed consent form before entry into the study.

Participants and Outcomes: Leibovich and SORCE Populations

Two matched cohorts were analyzed. A derivation cohort derived from the 2003 Leibovich data set included patients with clear-cell RCC only and excluded the low-risk group (Leibovich scores 0-2). A validation cohort was derived from intermediate-risk and high-risk clear-cell RCC participants in SORCE. All patients in the 2003 Leibovich data set underwent radical nephrectomy. Partial or radical nephrectomy was permitted in SORCE, reflecting contemporary surgical practice.

The primary outcome was time to metastasis-free survival (MFS), defined as the interval between nephrectomy and the date of distant metastases. In the study by Leibovich et al,[1] deaths preceding presumed metastasis were treated as censored observations (C. Lohse, personal communication, April 2020). We defined MFS in the same way for this analysis. We censored time to MFS at 10 years in both cohorts to reflect available follow-up data in SORCE.

Secondary exploratory analyses were conducted in the three SORCE subpopulations: patients with any non–clearcell histology, papillary-only, and chromophobe-only carcinomas.

Statistical Methods

Model validation was performed adhering to transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) guidelines[13] (Data Supplement). The time origin used for both cohorts was date of surgery. A survival analysis allowing for late entry was used,[14] capturing the post hoc nonexposure of a SORCE participant to the risk of an MFS event between surgery and random assignment. In SORCE, 56/1711 (3%) dates of surgery were missing; they were estimated by taking a random selection of 56 values from the distribution of observed intervals between surgery and random assignment.

The performance of the 2003 Leibovich score was assessed using discrimination and calibration.[6,7] Discrimination denotes the ability of a model to distinguish between patients who have and have not experienced an event. Calibration relates to a model's predictive accuracy.

Discrimination was assessed graphically by observing the degree of separation between the Kaplan-Meier curves and by the hazard ratio (HR) between intermediate-risk and high-risk Leibovich risk groups in each cohort. We quantified discrimination according to Harrell's c-index,[8] which denotes the proportion of all usable patient pairs in whom the observed and predicted survival times are concordant. The c-index ranges from 0.5 (performance no better than chance) through to 1 (perfect discrimination).

Calibration measures agreement between predicted and observed outcomes. Good calibration is inferred if KaplanMeier curves for risk groups in the derivation and validation cohorts are similar. We quantified calibration through the HR of the indicator variable for the two cohorts (0 5 derivation cohort, 1 5 validation cohort) separately for the two risk groups (0 5 intermediate risk, 1 5 high risk). An HR around 1 suggests accurate calibration.

We also analyzed the ungrouped Leibovich scores 3,4, …,11, to compare the HRs between the individual scores and a base category (taken as score 5 3). We fitted Cox models separately for the derivation and the validation data sets, with each individual score as the explanatory variable and graphed the results.

Furthermore, we analyzed the ungrouped scores as a single entity to compare the discrimination (c-index) of the Leibovich score with that of 2002 TNM staging.

The secondary (exploratory) analyses were conducted with the three SORCE subpopulations using same procedures as with the primary analysis.

All measures were reported with 95% CIs. P values were two-sided.

All analyses were performed in STATA (16.1; StataCorp LLC, College Station, TX).

RESULTS

The 2003 Leibovich data included 479 MFS events in 1671 US-based patients who had radical nephrectomies between 1970 and 2000. The SORCE data had 614 MFS events in 1711 patients enrolled between 2007 and 2013 (Fig 1). The derivation cohort included 407 MFS events in 982 patients with a median follow-up of 7.3 years
(interquartile range, 3-10 years), whereas the validation cohort included 520 MFS events in 1,445 patients with a median follow-up of 7.2 years (interquartile range, 6.1-8.4 years; Fig 1).

Table 1 describes the demographic, clinical, and histologic characteristics of patients in the 2003 Leibovich data, the derivation cohort, the SORCE data set, and the validation cohort. The validation cohort included more high-risk than intermediate-risk patients (46% v 38%). The validation cohort included 652 (45%) patients who had a laparoscopic nephrectomy and 43 (3%) patients who had a partial nephrectomy, whereas all patients in the derivation cohort underwent radical open nephrectomy (Data Supplement). The median time to MFS in the derivation cohort was 9.2 years, whereas in SORCE, this was not reached within 10 years of follow-up.

Primary Analysis Population: Discrimination and Calibration

Discrimination. Figure 2 presents the results of the validation exercise graphically, showing Kaplan-Meier curves of MFS in the intermediate-risk and high-risk groups for each cohort. Figure 2 shows that discrimination between intermediate-risk and high-risk groups in the derivation cohort is substantial but not entirely maintained in the validation cohort. The c-index in the derivation cohort is
0.67 (95% CI, 0.65 to 0.69) compared with 0.63 (95% CI, 0.61 to 0.65) in the validation cohort (P5 .01, chi square test).

Discrimination between high-risk and intermediate-risk groups in the derivation cohort, with intermediate-risk as the baseline category, is further indicated by an HR of 3.88 (95% CI, 3.18 to 4.74), compared with 2.74 (95% CI, 2.29 to 3.28) in the validation cohort. Thus, discrimination is maintained in the validation cohort, albeit significantly reduced (P5 .003, interaction analysis) compared with the derivation cohort.

To assess whether the discrimination of the 2003 Leibovich score degrades over follow-up time, c-indexes at one and

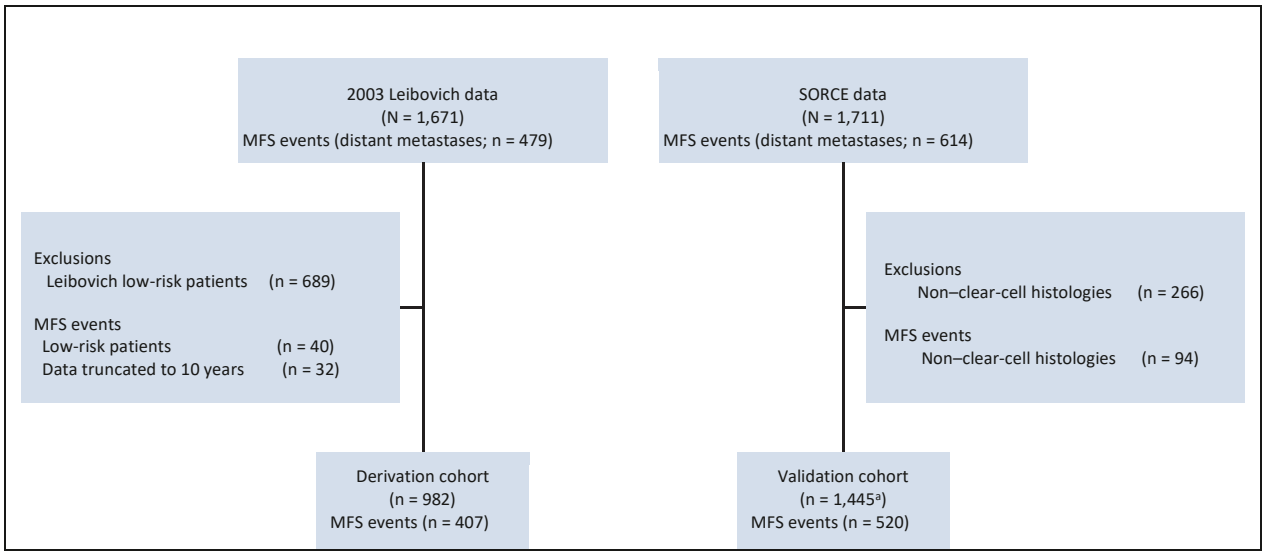


FIG 1. The primary analysis cohorts. MFS: time from nephrectomy to the date of distant metastases; deaths preceding metastasis were censored.
[a] A nuclear-grade assignment was missing for one participant, which we imputed singly by substituting the most common nuclear grade value (3), to ensure completeness of the validation data set. MFS, metastasis-free survival.

10 years after nephrectomy were compared in both cohorts (Data Supplement). We show that although c-index values for the Leibovich score reduce over time, the difference is small in both data sets.

Calibration. The validation and derivation survival curves for the intermediate-risk and high-risk groups are not aligned (Fig 2), suggesting poor calibration.

Overall, the MFS rate was 26% lower in the validation than in the derivation cohort (HR 5 0.74; 95% CI, 0.65 to 0.85). For the intermediate-risk group, the reduction in MFS rate was 24% (HR 5 0.76; 95% CI, 0.61 to 0.94), compared with 46% (HR 5 0.54; 95% CI, 0.45 to 0.64) in the high-risk group. The results confirm a distinct lack of calibration between data sets.

Analysis of Ungrouped Leibovich Scores

Figure 3 shows that the HRs comparing individual scores with the reference category (Leibovich score 3) increase markedly as the score increases in both the derivation and validation cohorts, reflecting consistently higher discrimination with increasing Leibovich score. We combined groups with scores of 9 and above because very few patients had score 10 or 11, giving unreliable estimates. See the Data Supplement for Kaplan-Meier curves (Fig 1A), HR values (Data Supplement), and c-indexes for each score (Data Supplement). Lower values of c-index and HR for each score group in the validation cohort confirm in detail that discrimination is maintained, albeit attenuated, in the contemporary cohort.

In both the validation and the derivation cohorts, the collapse of scores 3-5 and 6-11 into two larger prognostic groups (intermediate-risk and high-risk) results in reduced discrimination compared with the original Leibovich score. This is a compromise to achieve a clinically more useful risk stratification tool.

To compare the discrimination of the ungrouped 2003 Leibovich score with that of 2002 TNM, we calculated c-indexes using the primary analysis data sets. The Leibovich score outperformed the 2002 TNM system in the derivation cohort (c-indexes of 0.72 [SE 0.01] v 0.56 [SE 0.01]) and in the validation cohort (c-indexes of 0.67 [SE 0.01] v 0.56 [SE 0.01]).

Secondary Analyses: Discrimination

Three cohorts were included in the secondary analysis: those with any non–clear-cell RCC (N 5 266; MFS events, n 5 94), papillary RCC (N 5 128; MFS events, n 5 49), and chromophobe RCC (N 5 96; MFS events, n 5 21). Discrimination between intermediate-risk and high-risk groups within each SORCE subcohort was compared with that of the derivation cohort.

Figure 4 shows Kaplan-Meier estimates of MFS in the intermediate-risk and high-risk groups for SORCE non– clear-cell, papillary, and chromophobe populations. The maintained separation between the curves beyond 6 months indicates that the 2003 Leibovich score retains long-term discriminative capability in these SORCE subpopulations. Compared with the derivation cohort (c-index 0.67), we obtained c-indexes of 0.64 (95% CI, 0.59 to 0.69) for the SORCE non–clear-cell cohort, 0.63 (95% CI, 0.56 to 0.69) for SORCE papillary, and 0.65 (95% CI, 0.55 to 0.76) for the SORCE chromophobe group.

An HR of 3.88 (95% CI, 3.18 to 4.74) between risk groups was observed in the derivation cohort, compared with 3.21 (CI,

2.05 to 5.03) for SORCE non–clear-cell patients, 2.61 (95% CI, 1.44 to 4.70) for papillary, and 3.88 (95% CI, 1.56 to 9.61) for the chromophobe cohort. Despite smaller cohort sizes with correspondingly larger imprecision, these results highlight the Leibovich score's preserved discrimination in these SORCE subpopulations.

Secondary Analyses: Calibration

Attenuated calibration between the SORCE subpopulations and the derivation cohort for each risk group is shown by

TABLE 1. Histopathologic Characteristics, Leibovich Score Components, and Median Follow-Up in Leibovich and SORCE Data

Variable at Baseline	2003 Leibovich Data (N 5 1,671)	Derivation Cohort (N 5 982)	SORCE Data (N 5 1711)	Validation Cohort (N 5 1,445)
Histologic characteristics, No. (%)				
Histology				
Clear-cell	1,671 (100)	982 (100)	1,445 (84)	1,445 (100)
Papillary	—	—	128 (7)	—
Chromophobe	—	—	96 (6)	—
Collecting duct	—	—	4 (, 1)	—
Other	—	—	38 (2)	—
Other histologies				
Mixed	—	—	8 (21)	—
Sarcomatoid	—	—	23 (61)	—
Unclassified	—	—	5 (13)	—
Translocation	—	—	2 (5)	—
Tumor stage				
pT1a	384 (23)	6 (, 1)	7 (, 1)	5 (1)
pT1b	440 (26)	129 (13)	197 (12)	170 (12)
pT2	335 (20)	335 (34)	400 (23)	298 (10)
pT3a-4	512 (31)	512 (52)	1,107 (65)	972 (67)
Regional lymph node status				
pNx/pN0	1,605 (96)	916 (93)	1,637 (96)	1,405 (97)
pN1/pN2	66 (4)	66 (7)	74 (4)	40 (3)
Tumor size, cm				
, 10	1,312 (79)	623 (63)	1,152 (67)	996 (69)
$ 10	359 (21)	359 (37)	559 (33)	452 (31)
Nuclear grade[a]				
1	182 (11)	47 (5)	89 (5)	68 (5)
2	786 (47)	284 (29)	440 (26)	374 (26)
3	600 (36)	548 (56)	859 (50)	735 (51)
4	103 (6)	103 (10)	322 (19)	268 (18)

Histologic tumor necrosis[b]				
No	1,232 (74)	561 (57)	774 (45)	671 (46)
Yes	439 (26)	421 (43)	937 (55)	774 (54)
Leibovich score groups in the Leibovich and SORCE data				
Leibovich score group, No. (%)				
Low risk	689 (41)	—	—	—
Intermediate risk	608 (36)	608 (62)	910 (53)	776 (54)
High risk	374 (22)	374 (38)	801 (47)	669 (46)
Median follow-up				
Years (IQR)	7.5 (3.2-10)	7.3 (3.0-10)	7.3 (6.1-8.4)	7.2 (6.1-8.4)

NOTE. Categorical data are presented as No. (%).

Abbreviation: IQR, interquartile range. [a]The Leibovich and SORCE grading system used simplifies and selects the worst International Society of Urological Pathology features at each grade. For details of grading components, see the Data Supplement.

[b]For the definition of histologic tumor necrosis outlined in SORCE trial protocol, see the Data Supplement.

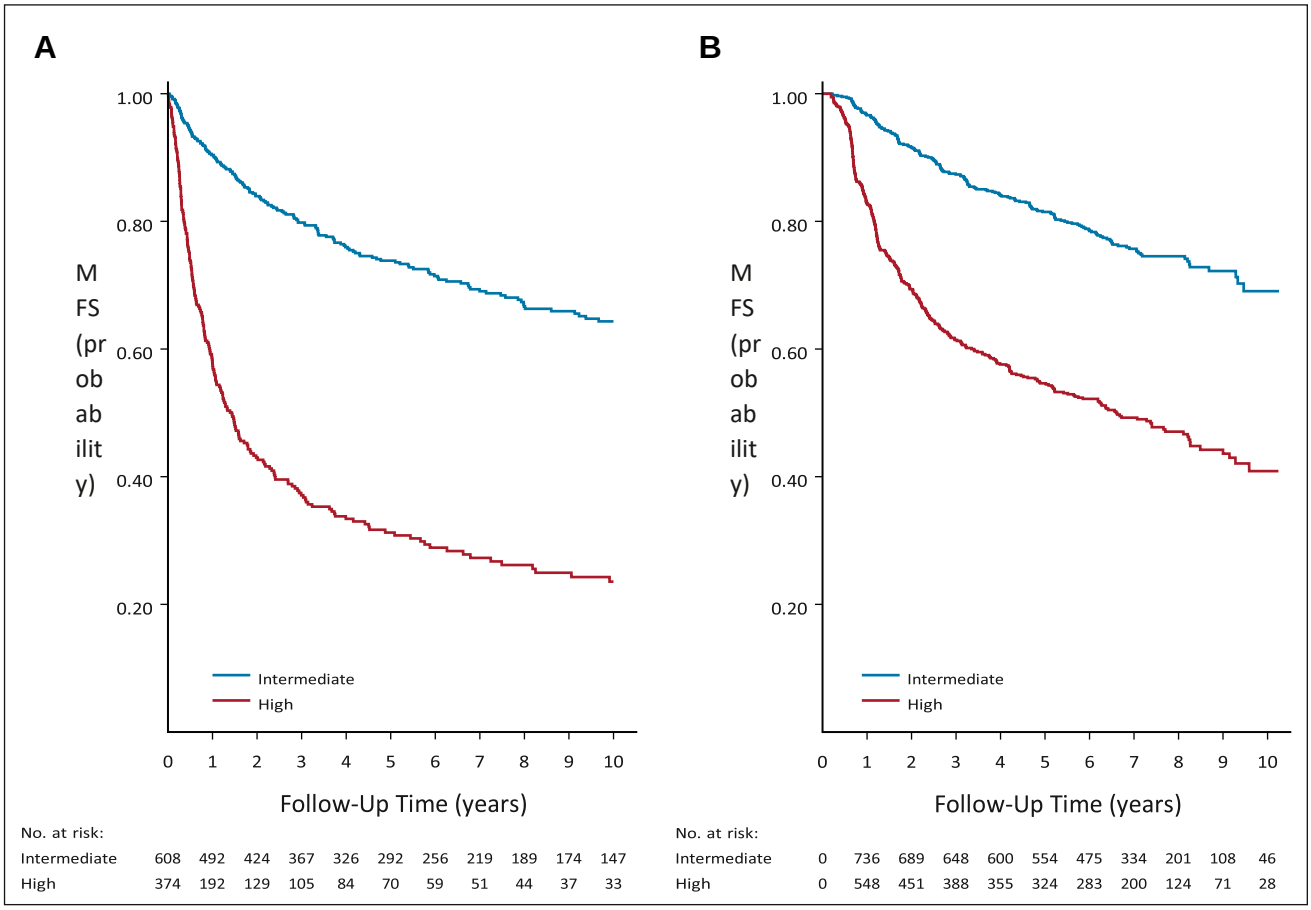


FIG 2. Kaplan-Meier curves for MFS by Leibovich risk group in the (A) derivation and (B) validation cohorts. In the validation cohort Kaplan-Meier plot, the number of patients entering at time 0 is given as 0 in the at-risk tables. It is a consequence of the late entry character of the follow-up data. Patients were not deemed at risk until they were randomly assigned into SORCE, which occurs after t 5 0. MFS, metastasis-free survival.

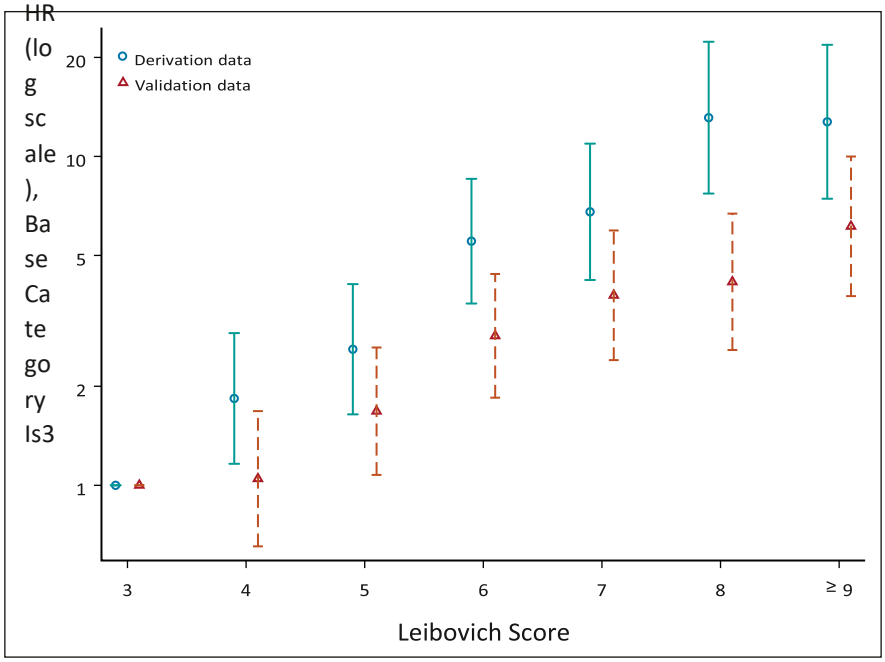


FIG 3. HRs estimated for ungrouped 2003 Leibovich scores in the derivation data set and in the validation data set. Values are presented with 95% CIs. The lowest score (3) in the validation data set is the reference category. HR, hazard ratio.

observing the misalignment of the corresponding survival curves (Fig 4). This is quantified by HRs for MFS after fitting a Cox regression model to each risk group separately (Table 2). Five-year relapse probabilities (Table 3) show improved MFS compared with the corresponding derivation cohort in all SORCE subgroups. The difference is most marked between the high-risk groups.

DISCUSSION

Validation of the 2003 Leibovich score using contemporary IPD from a large international trial represents the highest quality of validation, according to the American Joint Committee on Cancer criteria for model selection.[15] We focused on the intermediate-risk and high-risk clear-cell patients, a group commonly recruited to adjuvant clinical trials. This study confirms that the grouped 2003 Leibovich score, although developed two decades ago, largely retains discrimination in the SORCE validation cohort (c-index 0.63; 95% CI, 0.61 to 0.65) when compared with the derivation cohort (c-index 0.67; 95% CI, 0.65 to 0.69). We therefore support its ongoing use for risk stratification in this setting.

Uniquely, we show that the 2003 Leibovich score discriminates comparably between intermediate-risk and high-risk patients in the non–clear-cell SORCE cohort (cindex 0.64; 95% CI, 0.59 to 0.69). Since the non–clear-cell cohort is limited by inherent variability in clinical trajectories, we explored the two largest non–clear-cell subtypes separately: papillary (c-index 0.63; 95% CI, 0.56 to 0.69) and chromophobe groups (c-index 0.65; 95% CI, 0.55 to 0.76). Although the latter analyses are limited by smaller patient numbers, they indicate negligibly attenuated discrimination compared with the derivation cohort.

Some of the immune-oncology–focused adjuvant RCC trials, including IMMOTION010 (ClinicalTrials.gov identifier: NCT03024996) and KEYNOTE-564,[16] use the TNM staging system for patient random assignment to trial arms. We show that discrimination of the 2003 Leibovich score exceeds that of 2002 TNM in the derivation cohort (c-indexes of 0.72 [SE 0.01] v 0.56 [SE 0.01]) and in the validation cohort (c-indexes of 0.67 [SE 0.01] v 0.56 [SE 0.01]). The improvement is noteworthy, considering that a c-index of 0.5 represents a performance that is no better than chance. This finding has implications for using TNM for participant selection to clinical trials. We also show that the 2003 Leibovich score loses discrimination over followup time, with c-indexes of 0.63 at 10 years compared with 0.69 at 1 year after surgery in SORCE (Data Supplement). We suggest that this small difference over long follow-up should not impact on the 2003 Leibovich score's use.

In 2018, Leibovich et al[10] published five scoring systems, modeling progression-free survival (PFS) and cancer-specific survival individually for clear-cell and papillary RCC and PFS for chromophobe carcinomas. A major tradeoff for histologic specificity is added complexity in terms of the number of scoring systems for different subtypes and

models that comprise many more components for clear-cell RCC. This is important when considering trial practicalities including standardization and limiting the workload associated with assigning prognostic risk for eligibility purposes. In addition, the 2018 scores offers only minor improvement in discrimination for PFS and cancer-specific survival in clear-cell patients, with internally validated c-indexes of 0.83 and 0.86, respectively, versus 0.82 for MFS for the 2003 score.[1,10]

Overall, the simplicity, practical utility, and maintained discrimination in a multisubtype population shown by the 2003 Leibovich score support its standardized use for risk stratification in adjuvant RCC trials in preference to recently published, yet to be widely externally validated, subtypespecific scores.[9,10]

We were able to perform a robust calibration analysis using IPD from the original Leibovich study, matching risk groups and unifying the MFS definition across cohorts. We clearly demonstrate longer MFS in patients with intermediate-risk and high-risk clear-cell RCC in the validation cohort (5-year MFS; 78% [CI, 75 to 81] and 52% [CI, 48 to 56], respectively), compared with the corresponding derivation cohorts (5-year MFS; 72% [CI, 67 to 75] and 30% [CI, 25 to 35], respectively). Comparatively longer MFS for contemporary non–clear-cell, papillary, and chromophobe cohorts are also shown (Table 3). On the basis of this, it may be necessary to reconsider trial eligibility for patients with long-term low relapse risk, for example, those with intermediate-risk chromophobe RCC where 5 year MFS approaches 87% (CI, 75-94; Table 3). Overall, better outcomes for patients with locally advanced RCC over time corroborate findings in contemporary literature.[17,18] Improved MFS may be linked to factors such as improved radiologic and pathologic practices over time and the introduction of minimally invasive surgical techniques such as laparoscopic nephrectomy.[19] Differences may additionally reflect an evolution in renal tumor biology over time, driven by changing rates of modifiable risk factors such as obesity and smoking.

Our analysis is not without limitations. First, pathology samples were not centrally reviewed. However, strict guidance for their assessment was provided in the SORCE protocol. Second, patients with low Leibovich risk (score 02) were not included this validation, because they are usually cured by surgery or ablation and not usually considered for recruitment to adjuvant trials. We acknowledge that excluding the low-risk group is likely to have resulted in loss of some discrimination compared with that achieved by the complete Leibovich data. Third, our validation was performed using the whole SORCE cohort rather than being restricted to the placebo group. As SORCE showed a clear lack of benefit of sorafenib as an adjuvant strategy after nephrectomy, we considered that including patients from the experimental arms would have no detrimental impact on this analysis.

Finally, patient and tumor characteristics differed between the Leibovich and SORCE cohorts. The median age of SORCE patients was 5 years younger and included higher rates of T3a-4 tumors compared with the Leibovich cohort (67% v 52%). Other differences included higher rates of histologic tumor necrosis in the SORCE cohort (54% v 43%) and more nuclear grade 4 cases (18% v 10%) were present.

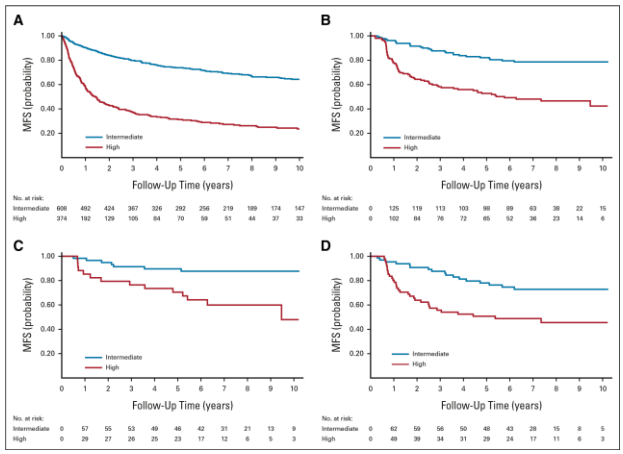


FIG 4. Kaplan-Meier curves for MFS in (A) the derivation cohort and in the SORCE (B) non–clear-cell, (C) chromophobe, and (D) papillary subcohorts stratified by 2003 Leibovich risk group. Derivation cohort included for reference. In the validation cohort Kaplan-Meier plot, the number of patients entering at time 0 is given as 0 in the at-risk tables. It is a consequence of the late entry character of the follow-up data. Patients were not deemed at risk until they were randomly assigned into SORCE, which occurs after t 5 0. MFS, metastasis-free survival.

TABLE 2. HRs Comparing MFS for Each SORCE Subpopulation to the Derivation Cohort Separately in Intermediate-Risk and High-Risk Groups

Patient Population	SORCE Non–Clear-Cell	95% CI	SORCE Papillary	95% CI	SORCE Chromophobe	95% CI
No. of patients	266		128		96	
No. of MFS events	94		49		21	
Median follow-up, years (IQR)	7.3 (6.1-8.3)		7.3 (6.1-8.4)		7.3 (6.3-8.3)	
Intermediate-risk HRs	0.65	0.43 to 0.97	0.84	0.51 to 1.38	0.36	0.17 to 0.77

High-risk HRs	0.57	0.44 to 0.76	0.59	0.40 to 0.85	0.40	0.23 to 0.69

NOTE. Presented with 95% CI.

Abbreviations: HR, hazard ratio; IQR, interquartile range; MFS, metastasis-free survival.

TABLE 3. Five-Year Survival Probabilities for Metastasis-Free Survival in the Derivation Cohort, the Validation Cohort, and Each SORCE Subcohort

Patient Population	SORCE Validation Cohort (clear-cell), % (95% CI)	SORCE Non–Clear-Cell, % (95% CI)	SORCE Papillary, % (95% CI)	SORCE Chromophobe, % (95% CI)
Derivation, intermediate risk		72 (67 to 75)		
Intermediate risk	78 (75 to 81)	79 (71 to 85)	75 (62 to 84)	87 (75 to 94)
Derivation, high risk		30 (25 to 35)		
High risk	52 (48 to 56)	50 (41 to 58)	49 (36 to 60)	64.9 (45 to 78)

NOTE. Presented with 95% CI.

In time, it may be possible to improve upon outcome prediction in RCC by adapting prognostic scores to include immunologic or genetic biomarkers that show both prognostic and predictive benefit. An example is the transcriptbased recurrence score,[20] which adds prognostic information when included with the 2003 Leibovich score. However, as it does not predict response to adjuvant treatment and is expensive and complex, it has not been routinely used.

Alongside prognostic and predictive biomarker studies, a pragmatic step will be to refine the 2003 Leibovich score by further unpicking the characteristics known to drive worse outcomes in RCC. A digital pathology review of SORCE tumor samples is underway. This will allow a comprehensive analysis of the heterogeneity among RCC tumor specimens.[21] It may also reveal further granularity within current 2003 Leibovich score features, to enhance the prognostication and the prediction of recurrence for patients with RCC. A practical goal will be to retain as much of the usefulness and simplicity of the original Leibovich score as possible.

In conclusion, the 2003 Leibovich score is a validated prognostic score which, in contemporary data, discriminates between patients with clear-cell RCC at intermediate risk and high risk of disease recurrence. In addition, it comparably discriminates relapse risk in patients with non–clear-cell, papillary, and chromophobe RCCs in our data set. Over time, MFS rates among patients have improved; therefore, clinicopathologic prognostic scores need to be regularly reviewed. With the wealth of data available from recent RCC trials, there is an opportunity to build upon the 2003 Leibovich score to better reflect the changing landscape of RCC.

AFFILIATIONS [1] MRC Clinical Trials Unit at UCL, Institute of Clinical Trials and Methodology, London, United Kingdom [2] Department of Oncology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, United Kingdom [3] Department of Surgery, University of Cambridge, Cambridge, United Kingdom [4] Royal Free London NHS Foundation Trust UCL Division of Surgery and Interventional Science, London, United Kingdom [5] Netherlands Cancer Institute, Amsterdam, the Netherlands [6] MRC Clinical Trials Unit Trial, London, United Kingdom [7] Department of Medical Oncology, Eastern Health, Melbourne, Australia [8] Eastern Health Clinical School, Monash University, Melbourne, Australia [9] ANZUP Cancer Trials Group, Sydney, Australia [10] NHMRC Clinical Trials Centre, Medicine, Central Clinical School, Sydney, NSW, Australia [11] Institut Gustave Roussy, Villejuif, France [12] The Royal Marsden Hospital, NHS Foundation Trust, London, United Kingdom [13] Department of Development and Regeneration—Urogenital, Abdominal and Plastic Surgery, Leuven, Belgium [14] University of Sheffield, Sheffield, United Kingdom [15] Urology Research Unit, Department of Urology, Herlev, Denmark [16] Beth Israel Deaconess Medical Center-IMIM Research Lab, Boston, MA

CORRESPONDING AUTHOR

Bhavna Oza, MBBS, MRC Clinical Trials Unit at UCL, Institute of Clinical Trials and Methodology, 2nd Floor, 90 High Holborn, London WC1V 6LJ, United Kingdom; e-mail: b.oza@ucl.ac.uk.

DISCLAIMER

Cancer Research UK and Bayer had no role in the study design; in the collection,analysis,andinterpretation ofdata;inthe writing ofthereport; or in the decision to submit the article for publication. Staff funded by the Medical Research Council and University College London contributed to study design, data collection, data analysis, data interpretation, and writing of this report. The views expressed are those of theauthors and not necessarily those of the NIHR or the Department of Health and Social Care.

SUPPORT

Supported by Cancer Research UK (C2845/A6858), Bayer, UK Medical ResearchCouncil, University College London. Also supported by the Mark Foundation for Cancer Research (G.D.S.), the Cancer Research UK Cambridge Centre (C9685/A25177; G.D.S.), Cancer Research UK (grant no.: A28690; G.D.S.) and the NIHR Cambridge Biomedical Research Centre (BRC-1215-20014; G.D.S.).

CLINICAL TRIAL INFORMATION

The SORCE trial (NCT00492258) is closed to recruitment.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at DOI https://doi.org/10.1200/JCO.21.01090.

AUTHOR CONTRIBUTIONS

Conception and design: Tim Eisen, Alastair W.S. Ritchie, Ian D. Davis, Martin R. Stockler, Bernard Escudier, Barry Hancock, Mahesh K.B. Parmar, Patrick Royston, Angela Meade

Administrative support: Benjamin Smith, Angela Meade

Provision of study material or patients: Tim Eisen, Grant D. Stewart, Axel Bex, Ian D. Davis, Martin R. Stockler, Laurence Albiges, Bernard Escudier, James Larkin, Steven Joniau, Barry Hancock, Gregers G. Hermann, Joaquim Bellmunt

Collection and assembly of data: Eleni Frangou, Benjamin Smith, Angela Meade

Data analysis and interpretation: Tim Eisen, Bhavna Oza, Eleni Frangou, Alastair W.S. Ritchie, Rick Kaplan, Ian D. Davis, Martin R. Stockler, Laurence Albiges, Bernard Escudier, James Larkin, Axel Bex, Steven Joniau, Joaquim Bellmunt, Grant D. Stewart, Mahesh K.B. Parmar, Patrick Royston, Angela Meade

Manuscript writing: All authors

Final approval of manuscript: All authors

Accountable for all aspects of the work: All authors

ACKNOWLEDGMENT

The authors thank the 1,711 participants for their willingness to take part in research to inform future generations. They also thank all of the staff at each of our 147 sites, represented here by the lead investigator at each one: Velindre Hospital, Dr Jim Barber; Addenbrooke's Hospital, Prof Tim Eisen; St James University Hospital, Dr Naveen Vasudev, Weston Park Hospital; Dr Janet Brown; Freeman Hospital, Dr Rhona McMenemin; Mount Vernon Hospital, Dr Paul Nathan; Weston Park Hospital, Dr Omar Din; Cheltenham General Hospital, Dr David Farrugia; Royal Marsden Hospital, Prof Martin Gore; James Cook University Hospital, Dr Alison Humphreys; Leicester General Hospital, Dr Subramanian Vasanthan; Beatson West of Scotland Cancer Centre, Dr Rob Jones; Clatterbridge Centre for Oncology, Dr Richard Griffiths;QueenAlexandraHospital,DrJoannaGale;RoyalDerbyHospital,Dr PrabirChakraborti;RoyalFreeHospital,DrEkateriniBoleti;ChristieHospital, Dr Tom Waddell; Kent and Canterbury Hospital, Dr Natasha Mithal; Royal PerthHospital,DrSimonTroon;BristolHaematologyandOncologyCentre,Dr Amit Bahl; Norfolk and Norwich University Hospital, Dr Rob Wade; Birmingham City Hospital, Dr Daniel Ford; Lincoln County Hospital, Dr Miguel Panades; Castle Hill Hospital, Dr Anthony Maraveyas; St Bartholomews Hospital, Dr Jonathon Shamash; Rigshospitalet University Hospital, Dr Gregers Hermann; University Hospital Coventry and Warwickshire, Dr Jane Worlding; Universitaire Ziekenhuizen Leuven, Prof. HeinvanPoppel;DerrifordHospital,DrMartinHighley;AustinHospital,Prof. Ian Davis; Centre Hospitalier Universitaire de Besançon, Dr Antoine ThieryVuillemin;InstitutGustave-Roussy,Prof.BernardEscudier;IpswichHospital, DrChristopher Scrase; RoyalStoke UniversityHospital,DrFawziAdab; Royal ShrewsburyHospital,DrNarayananSrihari;RoyalDevonandExeterHospital, Dr Rajaguru Srinivasan; Southend University Hospital, Dr Imtiaz Ahmed; Northampton General Hospital, Dr Mario Uccello; Royal Bournemouth Hospital, Dr Tom Geldart; Scunthorpe General Hospital, Dr Sanjay Dixit; Erasmus MC, Dr Willem Harm Jan Kruit; Centre Alexis Vautrin, Dr Lionnel Geoffrois;InstitutPaoliCalmettes,DrGwenaelleGravis;GlanClwydHospital, Dr Carey Macdonald-Smith; Prince of Wales Hospital, Dr Elizabeth Hovey; Southampton General Hospital, Dr Matthew Wheater; St George's Hospital, Dr Lisa Pickering; Maidstone Hospital, Dr Sharon Beesley; Wexham Park Hospital, Dr Nicola Dallas; Westmead Hospital, Prof. Howard Gurney; Churchill Hospital, Prof Andrew Protheroe; Dorset County Hospital, Mr Stephen Andrews; Nottingham University Hospital, Prof Poulam Patel;

REFERENCES

Weston General Hospital, Dr Serena Hilman; Royal Brisbane & Women's Hospital, Dr Jeffrey Goh; Darent Valley Hospital, Mr Sanjeev Madaan; Diana Princess of Wales Hospital, Dr Sanjay Dixit; King's Mill Hospital, Dr Santhanam Sundar; Fremantle Hospital, Dr Phillip Claringbold; Canberra Hospital, Dr Desmond Yip; Royal Prince Alfred Hospital, Dr Michelle Harrison; Basildon Hospital, Dr Awais Jalil; Royal Berkshire Hospital, Dr HelenO'Donnell;TorbayDistrictGeneralHospital,DrAnnaLydon;Institutde Cancerologie de la Loire, Dr Aline Guillot; Aberdeen Royal Infirmary, Dr Graham MacDonald; Belfast City Hospital, Dr Seamus McAleer; Charing Cross Hospital, Dr Naveed Sarwar; Poole Hospital, Dr Tom Geldart; Queen's Hospital, Dr Mike Smith-Howell; Southmead Hospital, Mr Raj Persad; Worcestershire Royal Hospital, Dr Lisa Capaldi; The Alfred, Dr Sanjeev Gill; Onze Lieve Vrouwziekenhuis, Dr Paul Carpentier; Hopital Saint-Andre, Drˆ Alain Ravaud; Centre Hospitalier Universitaire de Rouen, Prof Christian Pfister; Great Western Hospital, Dr Omar Khan; Guy's Hospital, Dr Simon Chowdhury; Pilgrim Hospital, Dr Miguel Panades; Queen Elizabeth Queen Mother Hospital, Dr Natasha Mithal; Salisbury District Hospital, Dr Adityanarayan Bhatnagar; Warwick Hospital, Dr Andrew Chan; Flinders Medical Centre, Dr Ganessan Kichenadasse; Royal Adelaide Hospital, Dr Nimit Singhal; Concord Hospital, Prof Martin Stockler; University Hospital Gent, Prof Sylvie Rottey; Institut Claudius Regaud, Dr Christine Chevreau; SalfordRoyalHospital,ProfNoelClarke;WilliamHarveyHospital,DrNatasha Mithal; Box Hill, Dr Joseph McKendrick; Royal North Shore Hospital, Dr Nick Pavlakis; Hopital Nord, Dr Marjorie Baciuchka; Broomˆ field Hospital, Dr Gopalakrishnan Srinivasan; South Tyneside District Hospital, Dr Ashraf Azzabi; Western General Hospital, Dr Duncan McLaren; Wycombe Hospital, Dr Prabir Chakraborti; Ysbyty Gwynedd, Dr Anna Mullard; Hospital Ramon y Cajal, Dr Maria Lopez; James Paget Hospital, Dr Rob Wade; Stafford General Hospital,DrRajaneeBhana;SunderlandRoyalHospital,DrAshrafAzzabi;Sir Charles Gairdner Hospital, Dr Siobhan Ng; Port Macquarie Base Hospital, Dr Stephen Begbie; Radboud University Nijmegen Medical Centre, Prof. Peter Mulders; Centre Jean Perrin, Dr Hakim Mahammedi; Centre Hospitalier Universitaire de Limoges, Dr Aurelien Descazeaud; Instituto Valenciano Oncologı´a,Dr Miguel Climent; Alexandra Hospital,DrLisaCapaldi;Conquest Hospital, Dr Kathryn Lees; Musgrove Park Hospital, Dr John Graham; Queen Elizabeth Hospital, Dr Anjali Zarkar; Royal United Hospital, Dr Mark Beresford; Border, Dr Craig Underhill; Peninsula Oncology Centre, Dr Vinod Ganju; Centre Leon B´ erard, Prof. Sylvie Negrier; Centre Paul Papin, Dr Remy´
Delva; Hospital del Mar, Dr Joaquim Bellmunt; Essex County Hospital, Dr Dakshinamoorthy Muthukumar; Scarborough General Hospital, Mr Simon Hawkyard; St Richard's Hospital, Mr James Hicks; Royal Hobart Hospital, Dr Louise Nott;Jules BordetInstitute,DrThierry Gil;Fundacio Puigvert,DrJoan´ Palou; Whipps Cross University Hospital, Dr Thomas Powles; Yeovil District Hospital, Mr Tim Porter; Centre George-Francois Leclerc, Dr Sylvain Ladoire; Clinique Valdegour, Dr Eric Legouffe; Val d'Aurelle, Dr Diego Tosi; Centre Hospitalier Aix-en-Provence, Dr Sophie Nahon; Clinica Universitaria Navarra, Dr Jose Perez-Gracia. We would also like to thank our colleagues at Bayer; Centre Leon Berard, France; European Organisation for Research and

Treatment of Cancer (EORTC); ADKNOMA Research, Spain; Australian and New Zealand Urogenital and Prostate (ANZUP) Cancer Trials Group, University of Sydney/NHMRC (National Health and Medical Research CouncilofAustralia);InstituteofCancerResearch,UniversityofLeeds,Leeds Teaching Hospitals NHS Trust; INC Research; Catalent and Rigshospitalet, Denmark, for their collaboration and support. They also warmly acknowledge the support of Bradley Leibovich et al at the US Mayo Clinic who provided them with original data used to derive the 2003 Leibovich Score. They are particularly grateful for the expedient and collaborative manner in which the Mayo Clinic Team answered their data queries.

1. Leibovich BC, Blute ML, Cheville JC, et al: Prediction of progression after radical nephrectomy for patients with clear cell renal cell carcinoma: A stratification tool for prospective clinical trials. Cancer 97:1663-1671, 2003
2. Escudier B, Porta C, Schmidinger M, et al: Renal cell carcinoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 30: 706-720, 2019
3. Eisen T, Frangou E, Oza B, et al: Adjuvant sorafenib for renal cell carcinoma at intermediate or high risk of relapse: Results from the SORCE randomized phase III intergroup trial. J Clin Oncol 38:4064-4075, 2020
4. Oza B, Frangou E, Smith B, et al: RAMPART: A phase III multi-arm multi-stage trial of adjuvant checkpoint inhibitors in patients with resected primary renal cell carcinoma (RCC) at high or intermediate risk of relapse. Contemp Clin Trials 108:106482, 2021
5. Correa AF, Jegede O, Haas NB, et al: Predicting renal cancer recurrence: Defining limitations of existing prognostic models with prospective trial-based validation. J Clin Oncol 37:2062-2071, 2019
6. Mackillop WJ, Quirt CF: Measuring the accuracy of prognostic judgments in oncology. J Clin Epidemiol 50:21-29, 1997
7. Harrell FE Jr, Lee KL, Mark DB: Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 15:361-387, 1996
8. Harrell FE Jr, Califf RM, Pryor DB, et al: Evaluating the yield of medical tests. JAMA 247:2543-2546, 1982
9. Klatte T, Gallagher KM, Afferi L, et al: The VENUSS prognostic model to predict disease recurrence following surgery for non-metastatic papillary renal cell carcinoma: Development and evaluation using the ASSURE prospective clinical trial cohort. BMC Med 17:182, 2019
10. Leibovich BC, Lohse CM, Cheville JC, et al: Predicting oncologic outcomes in renal cell carcinoma after surgery. Eur Urol 73:772-780, 2018
11. Buti S, Puligandla M, Bersanelli M, et al: Validation of a new prognostic model to easily predict outcome in renal cell carcinoma: The GRANT score applied to the ASSURE trial population. Ann Oncol 28:2747-2753, 2017
12. Amin MB, Epstein JI, Ulbright TM, et al: Best practices recommendations in the application of immunohistochemistry in urologic pathology: Report from the International Society of Urological Pathology consensus conference. Am J Surg Pathol 38:1017-1022, 2014
13. Moons KG, Altman DG, Reitsma JB, et al: Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. Ann Intern Med 162:W1-W73, 2015
14. Cleves M, Gould WW, Marchenko YV: An Introduction to Survival Analysis Using Stata, Revised 3rd Edition. College Station, TX, Stata Press, 2016
15. Kattan MW, Hess KR, Amin MB, et al: American Joint Committee on Cancer acceptance criteria for inclusion of risk models for individualized prognosis in the practice of precision medicine. CA Cancer J Clin 66:370-374, 2016
16. Choueiri TK, Tomczak P, Park SH, et al: Adjuvant pembrolizumab after nephrectomy in renal-cell carcinoma. N Engl J Med 385:683-694, 2021
17. Haas NB, Manola J, Uzzo RG, et al: Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinoma (ECOG-ACRIN E2805): A double-blind, placebo-controlled, randomised, phase 3 trial. Lancet 387:2008-2016, 2016
18. Vasudev NS, Hutchinson M, Trainor S, et al: UK multicenter prospective evaluation of the Leibovich score in localized renal cell carcinoma: Performance has altered over time. Urology 136:162-168, 2020
19. Ljungberg B, Albiges L, Abu-Ghanem Y, et al: European Association of Urology guidelines on renal cell carcinoma: The 2019 update. Eur Urol 75:799-810, 2019
20. Rini B, Goddard A, Knezevic D, et al: A 16-gene assay to predict recurrence after surgery in localised renal cell carcinoma: Development and validation studies. Lancet Oncol 16:676-685, 2015
21. Um IH, Scott-Hayward L, Mackenzie M, et al: Computerized image analysis of tumor cell nuclear morphology can improve patient selection for clinical trials in localized clear cell renal cell carcinoma. J Pathol Inform 11:35, 2020

ASCO Journal of Clinical Oncology

Reply to U. Capitanio et al

We thank Capitanio et al[1] for their comments on our publication.[2]

In our paper, we describe a prespecified per-protocol validation of the 2003 Leibovich score, which we used to select patients for the SORCE trial of adjuvant sorafenib in renal cell carcinoma (RCC).[2,3] We provide a robust approach to validating the 2003 Leibovich score using a prospectively collected contemporary data set, which is far superior to that possible using retrospective data. We showed that the 2003 Leibovich score, although published 17 years before completing SORCE, is still able to discriminate between intermediate-risk and high-risk patients with RCC and therefore remains pertinent in a contemporary context.

We thank the authors for highlighting their recent headto-head comparison and external validation of available prognostic models in RCC.[4] We note that discriminative accuracy of progression-free survival and cancerspecific survival achieved with the Leibovich 2018 score was only slightly improved compared with Leibovich 2003 (0.839 and 0.810 for Leibovich 2003 v 0.881 and 0.868 for Leibovich 2018). Another study by Blackmur et al[5] showed even less difference between the two scores (Fig 1).

The challenge is in judging the practical clinical importance of a small difference in c-index and in assessing whether the gain in discriminative accuracy provided by the 2018 score justifies the added complexity in its calculation. The 2018 score requires nine components for progression-free survival and 13 for cancer-specific survival, and both clinical and pathologic information is required. This complexity may explain why the 2018 score is rarely used in routine practice or in clinical trials. The five components comprising the Leibovich 2003 score are easy to derive and are routinely reported on RCC pathology, negating the need for additional expertise or training. Furthermore, unlike the 2018 score, constitutional symptoms and Eastern Cooperative Oncology Group performance status, two inherently subjective factors, are not required, rendering the 2003 score less prone to interuser variability.

The relevance of the Rosiello group findings to adjuvant trial populations must be considered with caution given the inclusion of predominantly low-risk patients in the study, who in practice are unlikely to be considered for adjuvant treatments after nephrectomy. We evaluated the Leibovich 2003 score specifically in intermediate-risk and high-risk patients with both clear cell and nonclear cell RCCs. We found that the 2003 score provides acceptable discrimination of risk within a multisubtype population and therefore represents a pragmatic tool for selection of patients suitable for inclusion in adjuvant trials and hence for adjuvant treatments as they are approved in different global territories.

Keynote-564[6] is the first of several adjuvant checkpoint inhibitor trials to report results and indicates that adjuvant immunotherapy is likely to form the new standard of care over the coming years. Several of these trials have used TNM and Fuhrman grading for risk stratification (ClinicalTrials.gov identifiers: NCT03024996, NCT03138512, NCT03142334, and NCT03055013). On direct comparison, we were able to show that the 2003 Leibovich score showed discriminative superiority over TNM. RAMPART[7] is a phase III multiarm multistage checkpoint inhibitor trial, which is currently recruiting participants according to the 2003 Leibovich criteria, and our external validation gives us confidence in its use.

Ultimately, the 2003 Leibovich score remains simple, very easy to apply, and of comparable accuracy with other risk tools. Until the development of molecularly based prognostic tools that markedly improve predictive accuracy, we believe that the 2003 score should remain the clinical standard. In the meantime, an ongoing digital pathology review of SORCE tumor samples aims to unpick the heterogeneity among RCC tumor specimens and may enhance histopathologically based prognostication. Given the current interest in evaluating adjuvant checkpoint inhibitors in patients with resectable metastatic disease, we suggest that a parallel focus should be

on optimizing risk stratification and patient selection within this particular group. To this end, longitudinal sampling of tumor and metastases collected as part of TransRAMPART (sample collection and translational study that run alongside the RAMPART trial) will be informative.

Journal of Clinical Oncology

Bhavna Oza, MBBS
MRC Clinical Trials Unit, UCL Institute of Clinical Trials and Methodology, London, United Kingdom

Tim Eisen, PhD
Department of Oncology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, United Kingdom

Grant D. Stewart, MBChB
Department of Surgery, University of Cambridge, Cambridge, United Kingdom

Axel Bex, MD
Division of Surgery and Interventional Science, Royal Free London NHS Foundation Trust UCL, London, United Kingdom
Netherlands Cancer Institute, Amsterdam, the Netherlands

Patrick Royston, DSc
MRC Clinical Trials Unit, UCL Institute of Clinical Trials and Methodology, London, United Kingdom

Angela Meade, DPhil
MRC Clinical Trials Unit, UCL Institute of Clinical Trials and Methodology, London, United Kingdom

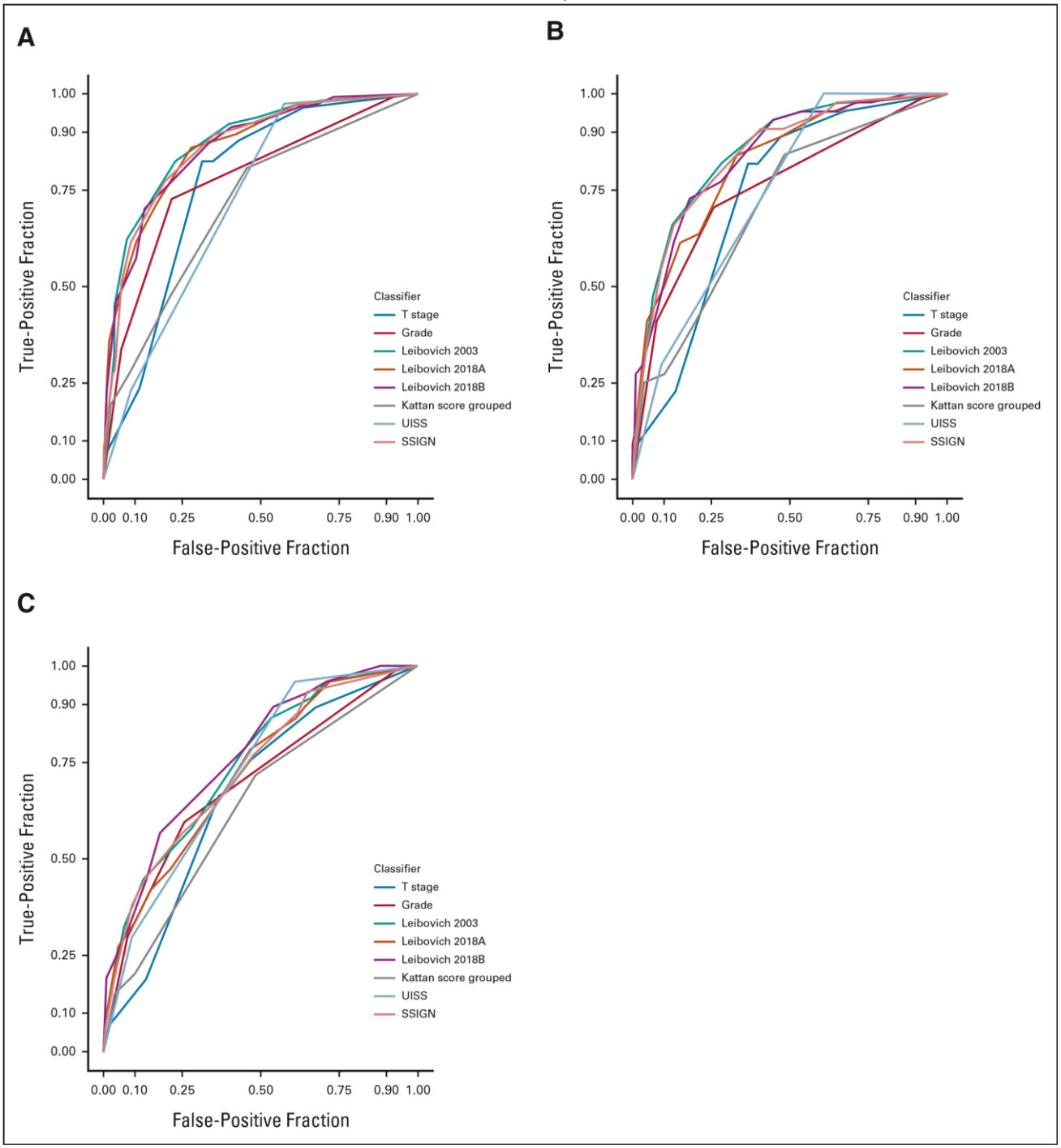


FIG 1. Receiver operating characteristic curves of (A) 5-year RFS, (B) 5-year CSS, and (C) 5-year OS stratified by prognostic scoring systems, adapted from Blackmur et al.[5] This article was published in Urologic Oncology: Seminars and Original Investigations, vol. 39, James P. Blackmur, Fortis Gaba, Dilini Fernando et al, "Leibovich score is the optimal clinico-pathological system associated with recurrence of non-metastatic clear cell renal cell carcinoma," 438.e11-438.e21, © Elsevier (2021). CSS, cancer-specific survival; OS, overall survival; RFS, relapse-free survival; SSIGN, Stage, Size, Grade, and Necrosis score; UISS, University of California Los Angeles Integrated Staging System.

REFERENCES

1. Capitanio U, Rosiello G, Necchi A, et al: Predicting the risk of progression in patients with kidney cancer: Aiming at the optimal tool in the optimal setting. J Clin Oncol 41:703, 2023
2. Oza B, Eisen T, Frangou E, et al: External validation of the 2003 Leibovich prognostic score in patients randomly assigned to SORCE, an international phase III trial of adjuvant sorafenib in renal cell cancer. J Clin Oncol 40: 1772-1782, 2022
3. Eisen T, Frangou E, Oza B, et al: Adjuvant sorafenib for renal cell carcinoma at intermediate or high risk of relapse: Results from the SORCE randomized phase III intergroup trial. J Clin Oncol 38:4064-4075, 2020
4. Rosiello G, Larcher A, Fallara G, et al: Head-to-head comparison of all the prognostic models recommended by the European Association of Urology Guidelines to predict oncologic outcomes in patients with renal cell carcinoma. Urol Oncol 40:271.e19-271.e27, 2022
5. Blackmur JP, Gaba F, Fernando D, et al: Leibovich score is the optimal clinico-pathological system associated with recurrence of nonmetastatic clear cell renal cell carcinoma. Urol Oncol 39: 438.e11-438.e21, 2021
6. Choueiri TK, Tomczak P, Park SH, et al: Adjuvant pembrolizumab after nephrectomy in renal-cell carcinoma. N Engl J Med 385:683-694, 2021
7. Oza B, Frangou E, Smith B, et al: RAMPART: A phase III multi-arm multistage trial of adjuvant checkpoint inhibitors in patients with resected primary renal cell carcinoma (RCC) at high or intermediate risk of relapse. Contemp Clin Trials 108:106482, 2021

DOI: https://doi.org/10.1200/JCO.22.01124; Published at ascopubs.org/journal/jco on September 27, 2022.

Contemporary Clinical Trials 108 (2021) 106482

Contents lists available ScienceDirec

Contemporary Clinical Trials

journal www.elsevier.com/locate/conclintr

RAMPART: A phase III multi-arm multi-stage trial of adjuvant checkpoint inhibitors in patients with resected primary renal cell carcinoma (RCC) at high or intermediate risk of relapse

Bhavna Oza [a,*], Eleni Frangou [a], Ben Smith [a], Hanna Bryant [a], Rick Kaplan [a], Babak Choodari-Oskooei [a], Tom Powles [b], Grant D. Stewart [c], Laurence Albiges [d], Axel Bex [e,p], Toni K. Choueiri [f], Ian D. Davis [g,h,i], Tim Eisen [j], Alison Fielding [a], David Harrison [k], Anita McWhirter [l], Salena Mulhere [a], Paul Nathan [m], Brian Rini [n], Alastair Ritchie [a], Sarah Scovell [b], Clare Shakeshaft [a], Martin R. Stockler [i,o], Nat Thorogood [a], Mahesh K. B. Parmar [a], James Larkin [l], Angela Meade [a]

[a] MRC Clinical Trials Unit at UCL, Institute of Clinical Trials & Methodology, 2nd Floor 90 High Holborn, London WC1V 6LJ, UK [b] St Bartholomew's Hospital, W Smithfield, London EC1A 7B, UK

1. Introduction

Patients with intermediate or high-risk locally advanced renal cell cancer (RCC), as stratified by RCC specific prognostic scores, are at significant risk of relapse after surgical tumour resection. 20–30% of patients with intermediate risk and 40–60% of patients with high risk RCC develop metastatic disease following nephrectomy [1–3]. Between 1 and 4% of patients with RCC present with synchronous ipsilateral adrenal metastases that can be resected at the time of nephrectomy [4]. These patients are treated adjuvantly, and they are considered to behave similarly to patients with high risk RCC. An effective strategy for reducing the risk of recurrence or death for patients with locally advanced, fully resected RCC remains an unmet clinical need.

TKIs targeting the vascular endothelial growth-factor receptor are established in treating metastatic RCC and have been extensively tested in the adjuvant setting. Five TKI trials have produced results: ASSURE, PROTECT, S-TRAC, ATLAS and now SORCE [5–10]. These studies evaluated the effect of oral TKIs compared to placebo and none have shown a benefit of TKI on overall survival (OS) [10,11]. Only the S- TRAC trial showed a modest DFS benefit with 1 year of sunitinib (HR 0.76; 95% CI 0.59–0.98; $p = 0.03$) compared to placebo on blinded independent central review [7]. 63% of sunitinib treated patients experienced Grade ≥ 3 toxicities, with many patients unwilling or unable to complete treatment. On this basis, the Food and Drug Administration (FDA) approved sunitinib for the adjuvant treatment of patients with high risk RCC. However, given the toxicity and cost associated with sunitinib in this setting, the results have not been universally practice changing. Therefore, nephrectomy followed by active surveillance for relapse, remains the predominant standard of care globally.

Treatment with ICIs, either as a dual combination or in combination with TKIs have revolutionised the management of patients with advanced RCC. The combination of ipilimumab, a monoclonal antibody against human cytotoxic T-lymphocyte-associated antigen-4 (CTLA-4) with nivolumab, a monoclonal antibody against programmed cell death protein-1(PD-1) is now a first line treatment for patients with intermediate or poor risk advanced RCC, classified according to the International Metastatic Renal-Cell

Carcinoma Database Consortium (IMDC) model [12]). This followed findings from CHECKMATE-214 [13], a phase III study showing an 18 month OS of 75% (95% CI; 70–78) for the combination ICI group versus 60% (95% CI; 55–65) for sunitinib [14]. In addition, single agent nivolumab, is routinely available in the second line setting for patients who have progressed on TKI therapy regardless of IMDC risk, based on a 5.6 month OS benefit of nivolumab over everolimus [15]. There are now a number of ICI and TKI combination strategies which have shown efficacy benefit over sunitinib in the first line setting, including axitinib (TKI) with avelumab (anti-programmed death-ligand 1 (PD-L1)), axitinib with pembrolizumab (anti-PD-1), pembrolizumab with lenvatinib (TKI) and nivolumab with cabozantinib (TKI) [16–19].

Durvalumab (anti PD-L1) and tremelimumab (anti-CTLA-4) are agents of the same class as other ICIs. Durvalumab is efficacious for patients with non-small-cell lung cancer who have completed definitive chemo-radiotherapy, showning a progression free survival (PFS) advantage of 11.2 months in patients receiving durvalumab compared with placebo [20]. Durvalumab in combination with tremelimumab has shown benefit to OS and PFS in the third-line treatment of patients with metastatic NSCLC, (in those with PD-L1 tumour cells ≥25%) and is being evaluated in the advanced setting in patients with various tumour types (NCT03298451), (NCT03994393) and (NCT02516241). Therefore, the RAMPART trial, led by the MRC CTU at University College London (UCL) and working in partnership with AstraZeneca, is investigating the activity of durvalumab alone and in combination with tremelimumab after nephrectomy for patients with locally advanced RCC. There is also potential for adding additional research arms as the trial progresses.

2. Methods

2.1. Overview of design

RAMPART is a phase III, multi-arm multi-stage (MAMS) trial, initiated with a control (Arm A; active monitoring) and two research arms, (Arm B; durvalumab and Arm C; durvaumab with tremilimumab)

(Fig. 1**)**. Initially, participants with Leibovich scores 3 to 11 (intermediate and high risk) [21] are eligible to be randomised. Intermediate-risk participants (Leibovich scores 3–5) will be capped at 25% of the total accrual target or after four years of recruitment, whichever is earlier. Inclusion of intermediate risk participants acknowledges that they are at a substantial risk of relapse, although they tend to occur later than those at high risk. By including the intermediate-risk participants during the early years of the trial, there will be sufficient numbers, followed for long enough, to contribute to the disease-free survival (DFS) analysis. Recruitment of participants with Leibovich scores 6 to 11 will continue to the accrual target of 1750 participants. Participants with ipsilateral adrenal metastases that are completely resected at the time of nephrectomy are eligible for RAMPART, which was a mid-trial protocol change implemented in July 2021.

2.2. Outcome measures

There are two co-primary endpoints in RAMPART, DFS and OS. DFS is defined as the interval from randomisation to first evidence of local recurrence, new primary renal cell carcinoma (RCC), distant metastases, or death from any cause, whichever occurs first. OS is defined as all- cause mortality; the time from randomisation to death from any cause. An adjuvant trial focusing on DFS and OS independently would take up to fifteen to twenty years to report its results and would potentially deny many thousands of patients the opportunity to benefit from promising new treatments. Therefore, both DFS and OS were accepted, after regulatory and scientific review, as co-primary endpoints. OS will be examined conditional on seeing improvements in DFS. However, DFS is the primary outcome on which regulatory approval will be sought for durvalumab monotherapy and/or the combination of durvalumab and tremelimumab. Reporting both DFS and OS as co-primary endpoints will provide the complete picture and allow clinicians and regulators to make fully informed treatment decisions.

The following secondary outcome measures will be analysed:

- Safety
- Metastasis-free survival (MFS), defined as the interval from randomisation to first evidence of metastases or death from RCC;

- RCC specific survival time, defined as the time from randomisation to death from RCC;
- Quality of Life (EQ-5D, EORTC QLQ-C30)
- Preferences for Adjuvant Immunotherapy in RAMPART (PAIR) sub- study questionnaire at baseline, week 16 and after completing treatment at month 15.

2.3. Sample size

RAMPART is powered for both the DFS and OS outcomes. The sample size calculations and design characteristics for RAMPART were obtained using nstage (version 3.0.1, 10-Sep-2014). Specifically, the nstage program was used to obtain the 'ideal' target number of control arm events needed at each stage for each comparison and an approximate idea for the timing of the stages. Artpep (version 1.0.4 PR 05-Jul-2013) was then used to project a more realistic analysis timeline using accrual and time- to-event patterns based on the SORCE trial (ISRCTN: ISRCTN38934710, EUDRACT: 2006–006079-19; NCT00492258). ART (version 1.1.0, 10- Dec-2013) was used to determine the absolute differences in DFS and OS at relevant time points. All calculations were performed in Stata 14.1.

Using control arm data from the SORCE trial [9] we anticipate a 3- year DFS rate of 65% for the control arm of RAMPART. We plan to recruit 1750 participants (750 to Arm A, 500 to Arm B and 500 to Arm C) over approximately 5.5 years but will continue until the accrual target is reached.

2.4. Adjusting sample size estimates for multiple comparisons

We have adjusted the RAMPART sample size to allow for multiple comparisons. The overall type I error rate – i.e. the family-wise type I error rate (FWER) [19] - is strongly controlled at 2.5% for all the pairwise comparisons whether or not a new research arm is added. Simulations were used to find the final stage significance level that control the FWER across the three pairwise comparisons. We considered two different scenarios: 1) the trial starts (and possibly concludes) with two research arms B and C, 2) or a new research arm (Arm D) is added before accrual to the current 3-arm trial completes. We applied Dunnett's approach to calculate the FWER in both scenarios 1 and 2. The results showed that the final stage significance level of 0.0097 in all pairwise comparisons controls the overall FWER at 2.5% when Arm D is added later on. Our simulations also showed that the final stage significance level of the two original pairwise comparisons can be increased to 0.015 if the deferred arm is not added, to buy back the unspent type I error of the third pairwise comparison. The relevant methods to calculate the correlation structure are described in Choodari-Oskooei et al. (2020) [22].

2.5. Eligibility and participant recruitment

Participants entering the RAMPART trial have undergone potentially curative nephrectomy for RCC and must satisfy the eligibility criteria, summarised in Tables 1 and 2. The time window for entry (up to 12 weeks post nephrectomy) allows treatment to be started at the earliest opportunity to maximise the potential benefits, whilst also considering safety from a post-surgical perspective.

2.6. Site recruitment

RAMPART is open to hospitals throughout the United Kingdom (UK). Recruitment will commence in Australia, New Zealand, France and Spain in mid 2021 (other countries may join subsequently).

UK site recruitment has been organised in 'waves' of hospitals, grouped by geography and also by their accrual to SORCE. Doing this has enabled more individualised site support. Lessons learnt from opening successive waves helps with optimising and refining activation processes, training of hospital staff and the provision of more useful guidance documents. Given the novel nature of the treatments and concerns around potential toxicity, the trial management team have been able to keep a closer eye on safety with a smaller initial group of sites.

2.7. RAMPART and the COVID-19 pandemic

Recruitment of participants into RAMPART and the treatment of existing RAMPART participants was suspended on 23-Mar-2020 due to the COVID-19 pandemic as it was unknown whether durvalumab and/or tremelimumab would lead to an increased risk or severity of COVID-19 for RAMPART participants. AstraZeneca subsequently advised that there

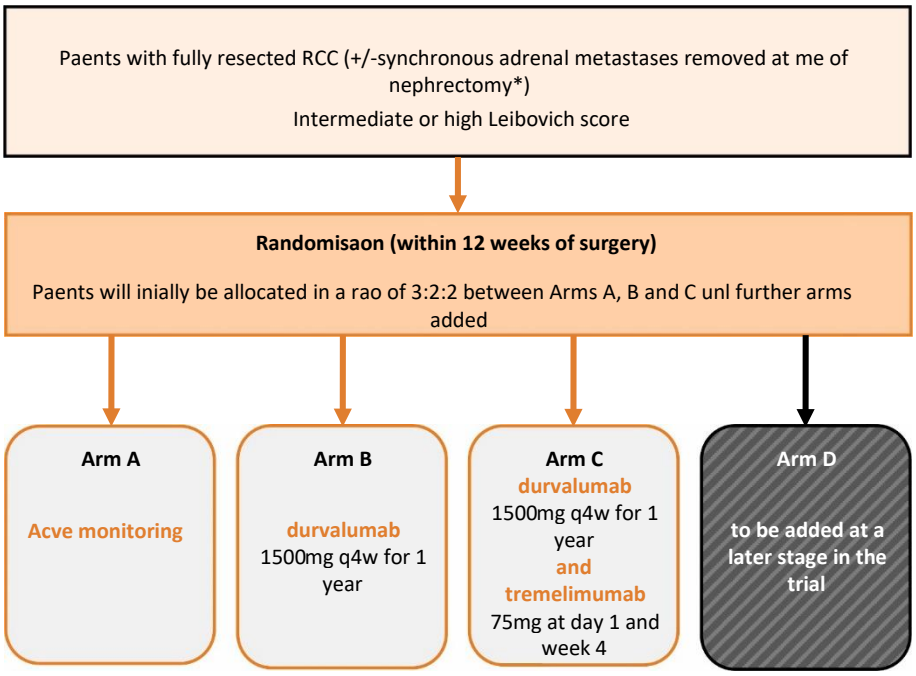


Fig. 1. Trial schema.

Table 1

Inclusion Criteria
Histologically proven RCC all cell types except for pure oncocytoma, collecting duct, medullary and transitional cell cancer
Microscopically positive resection margins after **radical** nephrectomy at the nephrectomy bed, renal vein or inferior vena cava
Synchronous ipsilateral adrenal metastases, provided they are fully resected at the time of nephrectomy
No residual macroscopic disease on post-operative CT scan after resection of RCC
Leibovich Score 3-11
Nephrectomy 28-91 days prior randomisation
WHO PS 0-1
FFPE tissue available
Haemoglobin ≥9.0g/dL
Neutrophil count ≥1.5 x 10^9/L
Platelet count ≥100 x 10^9
Bilirubin ≤1.5 ULN
AST/ALT ≤2.5 ULN
Creatinine Clearance > 40mL/min

Inclusion criteria (see protocol for more detailed list).	≥ 18 years of age
	Following trial's contraception policy (see RAMPART protocol V5.0 for details)

WHO World Health Organisation; PS Performance status; FFPE Formalin-fixed paraffin embedded; AST aspartate aminotransferase; ALT alanine aminotransferase; ULN upper limit normal.

was no evidence linking participants treated with either drug or the combination to a higher risk or severity of COVID-19 infection. As the potential benefits of treatment in the RAMPART trial outweigh the risks overall, RAMPART was re-opened on 07-Jul-2020. Sites were advised to restart treatment of existing participants and recruit new participants as and when they could.

Significant changes have been made to the RAMPART protocol during the pandemic to optimise safety for participants in terms of minimising time spent in hospital. Sites are now permitted to complete the participant consent process via either video or phone, where it forms part of local policy to reduce patient exposure to COVID-19. For more details on the Remote Consent Policy see RAMPART protocol version 5.0 [23]. In addition, the COVID-19 pandemic has significantly impacted the way in which clinical assessments can be conducted by sites and is highlighted below. These changes have enabled the RAMPART trial to remain active through subsequent waves of the pandemic.

2.8. Randomisation

Participants are randomised centrally between Arm A and the research Arms B and C using stratified block randomisation in the ratio 3:2:2. Treatment allocation is not blinded. To decrease determinability, the stratification factors are not listed here but are described in the RAMPART Statistical Analysis Plan, which will be finalized and published prior to the first interim analysis. Participants randomised to Arms B and C start treatment within 14 days of randomisation.

2.9. Treatment schedule and assessments

Participants in Arm B receive a fixed dose of 1500 mg durvalumab via IV infusion every four weeks for up to 13 cycles. Participants in Arm C receive four weekly 1500 mg durvalumab IV for a total of 13 cycles and two doses of 75 mg tremelimumab IV with the first and second cycles of durvalumab.

Participants in the active monitoring arm (Arm A) receive no drug; however they are radiologically assessed at the same frequency as participants on the active treatment arms. Arm A participants are clinically assessed at weeks 16, 32 and 52. Participants in arms B and C are assessed at day 1, then on a 2 weekly basis until week 8, and then every 4 weeks until week 52.

Since the COVID-19 pandemic there is now a greater emphasis on remote clinical assessments and this approach is supported by patient groups. Therefore, where it is deemed appropriate by the investigator, the pre-treatment clinical assessments can be carried out remotely (via telephone or video). The laboratory tests may be completed at the participants GP or a local hospital. Once the assessments have been completed at week 52 all participants will move into follow-up phase.

Table 2
Exclusion criteria.

Exclusion Criteria
Previous diagnosis of RCC
Metastatic or residual macroscopic disease (synchronous adrenal metastases which are fully resected at the time of nephrectomy are permitted).
Single pulmonary nodule ≥5mm (unless benign). multiple small, less than 5 mm nodules may be eligible if nodules are radiologically stable for at least 8 weeks
Prior anti-cancer treatment (other than nephrectomy) for RCC
Unresolved toxicity CTCAE v4.03 Grade ≥2 from previous anticancer therapy
Major surgical procedure within 28 days prior randomisation
Clinically significant pneumonitis or fibrosis
Concurrent enrolment in other RCT unless observational (non-interventional) or during follow-up period of an interventional study
Current or prior use of immunosuppressive therapy within 14 days prior to first dose of trial IMP
Active or prior autoimmune or inflammatory disorder
History of immunodeficiency syndrome
History of allogeneic organ transplant
Uncontrolled inter-current illness; congestive heart failure, unstable angina, uncontrolled cardiac arrhythmia, acute peptic ulcer/gastritis, acute bleeding, psychiatric illness that would limit study compliance
Active infection (participants who are exhibiting symptoms consistent with COVID-19, or who have tested positive, should not be randomised into the study until they are asymptomatic and at least 14 days after a positive test)
Live attenuated vaccine within 30 days prior to randomisation
Pregnant or breastfeeding
Patient has archival FFPE pathology tissue available, and agrees to provide at least one sample, as well as baseline CPDA and PAXgene blood samples for future translational research
Clinically significant pneumonitis or fibrosis

CTCAE Common Terminology Criteria for Adverse Events; RCT randomised controlled trial, FFPE Formalin-fixed paraffin embedded; CPDA citrate-phosphate-dextrose solution with adenine.

2.10. Criteria for discontinuing allocated interventions

An individual participant may stop treatment early for any of the following reasons.

- Disease progression
- Unacceptable toxicity, Inter-current illness or change in patient's condition that justifies discontinuation
- Any change in the patient's condition that in the clinician's opinion makes continuing investigational medicinal product a safety risk.
- Pregnancy or intent to become pregnant
- Grade ≥ 3 infusion reaction
- Initiation of alternative anticancer therapy including another investigational agent

- Withdrawal of consent for treatment by the patient
  Missing consecutive treatment visits[1].

[1] For treatment breaks due to COVID-19 infection, there is no specified maximum time between durvalumab infusions prior to restarting treatment. The second dose of tremelimumab for Arm C participants however, must be given within 12 weeks of starting trial treatment. Regardless of any administration delays, the maximum duration of durvalumab treatment must not exceed 1 year.

2.11. Participant follow-up

Participants on all arms who have not progressed are seen at week 52 (end of year one), and are followed-up on a three monthly basis until the end of year three. Initially, they have CT scans (with contrast) of the chest, abdomen and pelvis every six months. After year three, participants have annual clinical follow-up with a CT until year five. For the purposes of the trial, participants have a CT scan at either years seven or eight and a further CT scan at ten years. Follow-up assessments may be conducted in person or remotely (via telephone or video) if deemed appropriate. The trial follow-up schedule is described in Table 3.

2.12. Tumour assessment

The DFS outcome will be investigator-reported and must be based on thorough investigation and evidence such unequivocal radiological progression or biopsy. Site specific combined imaging and surgical review, followed by histopathological discussion (if relevant) at multi- disciplinary team discussion is the routine way in which DFS events are classified.

2.13. Safety/toxicity management

ICIs aim to boost endogenous immune responses directed against tumour cells. This augmented immune response may lead to activation of autoreactive T-cells that damage host tissues. Immune-related adverse events (ir-AEs) can affect any body system. Combined PDL-1 plus CTLA- 4 blockade is likely to trigger more ir-AEs than anti PD-1 alone. A comprehensive toxicity management guide is provided to all investigators to manage potential ir-AEs based on the severity of treatment-emergent toxicities graded per NCI CTCAE v4.03 [23]. It includes criteria for permanent discontinuation of study drug/study regimen based on CTCAE grade. In addition, study drug should be permanently discontinue for the following conditions:

- If an ir-AE results in missing consecutive treatment visits. Inability to reduce corticosteroid to a dose of ≤10 mg of prednisone per day (or equivalent) within 8 weeks after last dose of study drug/study regimen
- Grade ≥ 3 recurrence of an ir-AE following resumption of dosing
- Grade ≥ 3 infusion reaction

Table schedule assessment follow-up Routine clinically required As DFS Event Month 120 Month 84 Month 60 Month 54 Month 48 Month 42 Month 36 Month 33 Month 30 Month 27 Month 24 Month 21 Month 18 Month 15 Week 52	Assessments Clinical Physical	Examination	Concomitant Medications	Events Adverse WHO	Performance Status	Radiology Scans CT	Tests Laboratory	Haematology	Chemistry Clinical	Questionnaires EQ-5D	QLQ-C30	PAIR	Questionnaire Life of Quality Cancer Treatment and Medications Research Concomitant Radiology omitant the Organisation European QLQ-C30 Questionnaire; Dimension Euro 5 Qol- EQ-5D Organisation; Health World 3.0 version WHO
		X				X		X	X				
						X				X	X		
				X									
				X		X							
				X		X							
				X		X							
				X		X							
				X		X				X	X		
				X									
				X									
				X		X							
				X									
				X		X							
		X	X	X		X		X	X			X	
				X						X	X		
		X	X	X		X		X	X				

2.14. Monitoring

The monitoring plan for RAMPART is based on a formal risk assessment, initially undertaken during trial development. The plan is reviewed and updated as appropriate on at least an annual basis.

The RAMPART team conduct central and on-site monitoring checks to identify potential issues with consent, eligibility, treatment administration, drug supply and safety monitoring. Any issues identified will be raised and discussed with the local team.

Each participating site will have their first on-site monitoring visit within one year of randomising a patient to active treatment, (Arms B or C). The frequency of subsequent visits to sites will be determined by the outcomes of central monitoring. At times when site visits are difficult to conduct (e.g. COVID-19 pandemic), more central monitoring checks are performed to check site compliance.

2.15. Data management

Paper CRFs are used in RAMPART. Original copies of CRFs are retained at individual sites whilst copies are sent via secure email to the MRC CTU at UCL where they are stored securely. Key variable checks are performed on receipt of all RAMPART CRFs to ensure any potential safety issues are rapidly identified. Data is single entered onto a customised in-house database by trained staff. The database has many built

in validations (value ranges, date inconstancies, treatment administration) to help ensure the data is both accurate and correct. The RAMPART Data Management Plan provides a comprehensive breakdown of how data is to be acquired, handled and secured.

2.16. Statistical analysis plan

All efficacy analyses will be performed in the intention-to-treat population. The treatment effect for each of the initial two treatment comparisons and each of the two co-primary outcomes will be assessed using a stratified Cox model, stratifying for the factors used in randomisation. Data will be presented graphically using Kaplan-Meier plots. The Chi-squared test or Mann-Whitney test will be implemented for categorical data comparisons, including toxicity, as appropriate. Subgroup analyses will be conducted to assess consistency of treatment effect across potential or expected prognostic factors, including subgroup analyses by PD-L1 baseline expression status. Full details of all planned analyses are documented in the RAMPART Statistical Analysis Plan.

2.17. Interim analyses for disease-free survival (DFS)

As part of the MAMS design, one interim analysis is planned for the comparison of Arm C vs Arm A (combination vs control) and three interim analyses are planned for the comparison of Arm B vs Arm A (monotherapy vs control). Each one will consider both a lack-of-benefit and overwhelming benefit of treatment on DFS. Both sets of stopping boundaries are included in the RAMPART protocol [23].

Based on assumptions for accrual and survival distribution for the control arm, at the time of the trial design the first interim analyses were planned 4.75 years after the trial started. However, the exact timing of the interim analyses will be subject to change as they are time-to-event analyses. The most up-to-date information on the timeline for interim analyses can be found in the RAMPART protocol [23].

2.18. Primary DFS analysis

The primary DFS analysis of Arm C vs Arm A is planned when 276 control arm event events have been observed, The target HR for Arm C versus Arm A is 0.70, which translates to an absolute improvement in 3- year DFS of 9%, from 65% to 74%. This design gives 87.3% power to detect this difference at the 0.0097 one-sided significance level. If the DFS result at this time point is positive, OS will also be analysed using a closed test, even though the data will be not be fully mature, allowing for a more complete assessment of the DFS results.

2.19. Overall survival

The primary OS analysis is planned in high risk participants only (with Leibovich Score 6–11). With approximately 940 high-risk participants in the Arm C vs Arm A comparison, we will have 80% power to detect a HR of 0.7. This HR translates to an absolute difference in OS at 5 years of 6.5%, increasing survival from 76% to 82.5%.

With approximately 940 high-risk participants in the Arm B vs Arm A comparison, we will have 80% power to detect a HR of 0.75. This HR translates to an absolute difference in OS at 5 years of 5.4%, increasing survival from 76% to 81.4%. *2.20. Translational studies/sub-studies*

2.20.1. TransRAMPART

TransRAMPART is the Cancer Research UK (CRUK) funded translational study linked to RAMPART. TransRAMPART is an expanded sample collection, building on the samples already obtained through

RAMPART and supplementing them with additional sample types and collection time points. To ensure that all sites can contribute to the study, we have defined three participation levels (Bronze, Silver and Gold). For more details see TransRAMPART protocol [24] and the sample collection manual [25].

2.21. Safety and efficacy of the COVID-19 vaccines

There is little published prospective data on the safety of COVID-19 vaccines for participants receiving ICI therapy. In RAMPART we will be publishing the vaccinations our participants receive, any adverse events that they experience following administration of the vaccine as well as any subsequent COVID-19 infections. Relevant information for trial participants should be submitted on the COVID-19 case report form (CRF).

2.22. Regulatory and ethical considerations

The trial will be conducted in compliance with the approved protocol of the Declaration of Helsinki 2008, the principles of Good Clinical Practice (GCP), Commission Directive 2005/28/EC with implementation in national legislation in the UK by Statutory Instrument 2004/1031 and subsequent amendments, the EU General Data Protection Regulation (GDPR), and the National Health Service (NHS) Research Governance Framework for Health and Social Care (RGF). International centres will comply with the principles of GCP as laid down by the ICH topic E6 R2 (Note for Guidance on GCP) and applicable national regulations.

Scientific advice on the clinical trial protocol has been obtained from both the European Medicines

Agency (EMA) and FDA. Sufficient elements are in place to enable compliance with ICH GCP (both retrospectively and prospectively) if it is decided at a later stage that trial data are to be submitted to regulatory authorities as part of a licensing application.

The Medicines and Healthcare Products Regulatory Agency (MHRA) granted the Clinical Trials Authorisation on the 24th November 2018. The London Riverside Research Ethics Committee granted ethical approval on 8th January 2018. *2.23. Patient and public involvement*

The RAMPART Trial Management Group (TMG) is committed to engaging with the public and involving patient representatives in all aspects of the trial. Patient Public Involvement (PPI) activities include commenting on grant applications, promotion of the trial at start-up, advising on strategies to aid patient recruitment and ongoing engagement with relevant patient groups and charities. Patient newsletters, information sheets and trial promotional videos have been developed by the trial management team with the support of PPI delegates.

Patient representatives are members of the RAMPART TMG while other patient representatives are members of the MRC CTU Genitourinary Trial Steering Committee (TSC) and therefore are actively involved in discussions on trial progress including IDMC recommendations.

2.24. Trial oversight

RAMPART is sponsored by UCL. The MRC CTU at UCL has overall responsibility for the study working closely with the Chief Investigator, all members of the TMG and all collaborators. The Trial Management Team (TMT) meet on a weekly basis to discuss all aspects of trial conduct, including for example site set-up, participant accrual and safety management. The TMT report to the TMG. The TMG is responsible for the running of the trial and meets at least six times a year. An Independent Data Monitoring Committee (IDMC) meet approximately annually to review safety, compliance with

treatment and efficacy data (at pre-planned interim analyses). They are the only group who see the confidential, accumulating efficacy data for the trial. The IDMC advise the TSC. The TSC provides overall supervision of the trial and provides advice and recommendations through its independent Chair. The ultimate decision for the continuation of the trial lies with the TSC.

2.25. Use of electronic health records (EHR)

We plan to incorporate the use of EHR to improve on follow-up of participants enrolled to RAMPART by charting their health status and survival, for example, from records maintained by NHS Central Registry or any applicable national registry.

3. Conclusion

RAMPART is an international, UK-led trial that will assess the benefit of ICIs in participants at intermediate and high risk of recurrence after surgical resection of locally advanced RCC. Participants with synchronous ipsilateral adrenal metastases, removed at the time of nephrectomy are included in RAMPART (implemented in July 2021).

The first patient was randomely assigned in October 2018. By the end of June 2021, 259 of the target of 1750 participants from 34 UK sites were recruited. Recruitment of intermediate risk participants continues, (25% cap has not been reached). After a short pause at the start of the COVID-19 pandemic, RAMPART has re-established recruitment in the UK and in mid 2021 opened in France, Australia and Spain. There have been no safety concerns highlighted in IDMC safety reviews to date.

RAMPART is a three arm adaptive MAMS platform trial upon which at least one new research arm (Arm D) can be added over the coming years. Importantly, the RAMPART trial design allows the control arm to be amended should the standard of care change. This will allow progress to be made in the treatment of locally advanced RCC within the framework of one trial rather than starting a competing trial or waiting a number of years, until the first trial reports. This is critical in the adjuvant setting in RCC where it takes international collaborations many years to develop, launch and deliver a trial. ‘RAMPART: A Model for a regulatory ready academic led phase III trial in the adjuvant RCC setting’ (Contemporary Clinical Trials [26]), outlines the pertinent lessons we have learnt during the process of trial design, development and conduct.

TransRAMPART is a unique scientific collaboration that will provide an opportunity to address unanswered issues for patients with locally advanced RCC including which patients are most in need of adjuvant ICIs. Research will also explore biomarkers that predict treatment response and those that might pre-empt the onset of significant toxicity. It is likely that the research conducted on the TransRAMPART samples will lead to tangible benefits for patients.

The full support of our UK and international collaborators is essential to meet our ambitious accrual target of 1750 participants in order to complete this important trial aimed at improving the adjuvant treatment of renal cell carcinoma.

An up-to-date version of the RAMPART protocol can be found at htt ps://www.rampart-trial.org/.

Credit

Conceptualisation; TMG clinicians, Mahesh K B Parmar, Rick Kaplan, Angela Meade. Input from all members of the RAMPART TMG.

Writing - Original Draft and data curation; Bhavna Oza, Elena

Frangou, Ben Smith, Hanna Bryant, Angela Meade

Visualization; Bhavna Oza, Angela Meade

Supervision; Angela Meade, Clare Shakeshaft

Project administration/Funding acquisition; Mahesh K B Parmar,

Rick Kaplan, Angela Meade

Passive PSO reviewer- Eric Goluboff

Trial status

The trial is currently open to recruitment.

Abbreviations

Noted in text

Funding

AstraZeneca LP have provided an educational grant for the trial and free of charge durvalumab and tremelimumab. A small grant is also provided by Kidney Cancer UK. MRC CTU at UCL also provides funding for staff working on the trial. The TransRAMPART sample collection is being funded by a Prospective Sample Collection award from Cancer Research UK.

Availability of data and materials N/A.

Author's contributors N/A.

Ethics approval and consent to participate

The RAMPART trial was approved by the Riverside Research Ethics Committee and the Health Research Authority (HRA) and is part of the UK National Cancer Research Network (NCRN) portfolio. Reference number 17/LO/1875. The RAMPART trial is an investigator-led academic trial sponsored by UCL and co-ordinated by the MRC CTU at UCL. All participants signed an Informed Consent Form prior to entry into the study. A separate consent process will be employed for TransRAMPART sub-study.

Consent for publication N/A.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to acknowledge the efforts and contributions of all of our patient representatives, the trial investigators and research teams at sites. We would also like to show our gratitude to all the trial participants, without whom RAMPART would not be possible. All authors approved of the final version of the article, including the authorship list.

References

[1] S. Dabestani, et al., Long-term outcomes of follow-up for initially localised clear cell renal cell carcinoma: RECUR database analysis, Eur. Urol. Focus 5 (5) (2019) 857–866.
[2] S. Dabestani, et al., Renal cell carcinoma recurrences and metastases in primary non-metastatic patients: a population-based study, World J. Urol. 34 (8) (2016) 1081–1086.
[3] J.S. Lam, et al., Postoperative surveillance protocol for patients with localized and locally advanced renal cell carcinoma based on a validated prognostic nomogram and risk group stratification system, J. Urol. 174 (2) (2005) 466–472, discussion 472. (quiz 801).
[4] C.J. Weight, et al., The role of Adrenalectomy in renal Cancer, Eur. Urol. Focus 1 (3) (2016) 251–257.
[5] N.B. Haas, et al., Adjuvant treatment for high-risk clear cell renal Cancer: updated results of a high-risk subset of the ASSURE randomized trial, JAMA Oncol. 3 (9) (2017) 1249–1252.
[6] R.J. Motzer, et al., Randomized phase III trial of adjuvant Pazopanib versus placebo after nephrectomy in patients with localized or locally advanced renal cell carcinoma, J. Clin. Oncol. 35 (35) (2017) 3916–3923.
[7] A. Ravaud, et al., Adjuvant Sunitinib in high-risk renal-cell carcinoma after nephrectomy, N. Engl. J. Med. 375 (23) (2016) 2246–2254.
[8] M. Gross-Goupil, et al., Axitinib versus placebo as an adjuvant treatment of renal cell carcinoma: results from the phase III, randomized ATLAS trial, Ann. Oncol. 29 (12) (2018) 2371–2378.
[9] T. Eisen, et al., Adjuvant Sorafenib for renal cell carcinoma at intermediate or high risk of relapse: results from the SORCE randomized phase III intergroup trial, J. Clin. Oncol. 38 (34) (2020) 4064–4075.
[10] R.J. Motzer, et al., Adjuvant Pazopanib versus placebo after nephrectomy in patients with localized or locally advanced renal cell carcinoma: final overall survival analysis of the phase 3 PROTECT trial, Eur. Urol. 79 (3) (2021) 334–338.
[11] R.J. Motzer, et al., Adjuvant Sunitinib for high-risk renal cell carcinoma after nephrectomy: subgroup analyses and updated overall survival results, Eur. Urol. 73 (1) (2018) 62–68.
[12] D.Y. Heng, 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, J. Clin. Oncol. 27 (34) (2009) 5794–5799.
[13] R.J. Motzer, et al., Nivolumab plus ipilimumab versus sunitinib in advanced renal- cell carcinoma, N. Engl. J. Med. 378 (14) (2018) 1277–1290.
[14] R.J. Motzer, et al., Nivolumab plus ipilimumab versus sunitinib in first-line treatment for advanced renal cell carcinoma: extended follow-up of efficacy and safety results from a randomised, controlled, phase 3 trial, Lancet Oncol. 20 (10) (2019) 1370–1385.
[15] R.J. Motzer, et al., Nivolumab versus everolimus in advanced renal-cell carcinoma, N. Engl. J. Med. 373 (19) (2015) 1803–1813.
[16] B.I. Rini, et al., Pembrolizumab plus axitinib versus sunitinib for advanced renal- cell carcinoma, N. Engl. J. Med. 380 (12) (2019) 1116–1127.
[17] T.K. Choueiri, et al., Nivolumab plus cabozantinib versus sunitinib for advanced renal-cell carcinoma, N. Engl. J. Med. 384 (9) (2021) 829–841.
[18] R.J. Motzer, et al., Avelumab plus axitinib versus sunitinib for advanced renal-cell carcinoma, N. Engl. J. Med. 380 (12) (2019) 1103–1115.
[19] R. Motzer, et al., Lenvatinib plus pembrolizumab or everolimus for advanced renal cell carcinoma, N. Engl. J. Med. 384 (14) (2021) 1289–1300.
[20] S.J. Antonia, et al., Durvalumab after chemoradiotherapy in stage III non–small- cell lung cancer, N. Engl. J. Med. 377 (20) (2017) 1919–1929.
[21] B.C. Leibovich, et al., Prediction of progression after radical nephrectomy for patients with clear cell renal cell carcinoma: a stratification tool for prospective clinical trials, Cancer 97 (7) (2003) 1663–1671.
[22] B. Choodari-Oskooei, et al., Adding new experimental arms to randomised clinical trials: impact on error rates, Clin. Trials 17 (3) (2020) 273–284.
[23] MRC CTU at UCL, Clinical Trial Protocol V5.0;Renal Adjuvant MultiPle Arm Randomised Trial (RAMPART): An International Investigator-Led Phase III Multi Arm Multi Stage Multi-Centre Randomised Controlled Platform Trial of Adjuvant Therapy in Patients with Resected Primary Renal Cell Carcinoma (RCC) at High or Intermediate Risk of Relapse, RE06; ISRCTN53348826; 2017–002329-39; NCT03288532, 2021.
[24] UCL, M.C.A, TransRAMPART protocol v1.0 24.08.2020, 2020.
[25] UCL, M.C.A, RAMPART Sample Collection Manual v3.0 17.08.2020, 2020.
[26] RAMPART: A Model for a regulatory ready academic led phase III trial in the adjuvant RCC setting, Contemporary Clinical Trials.