THE MOLECULAR EPIDEMIOLOGY OF RENAL CELL CARCINOMA : SUBTYPES AND PROGNOSIS

TAN MIN-HAN
(M.B.,B.S., M.R.C.P.)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF EPIDEMIOLOGY AND PUBLIC HEALTH
NATIONAL UNIVERSITY OF SINGAPORE

2011

ACKNOWLEDGEMENTS

During the five years of my PhD studies, many people at the National Cancer Centre Singapore, Van Andel Research Institute and the National University of Singapore contributed either directly or indirectly to my work, and to whom I owe a great debt of gratitude. Specifically, I would like to thank :

Bin Tean Teh, my main supervisor, who has over the years, provided both guidance and freedom, without whom this would not be possible;

Chia Kee Seng, my co-supervisor, for all the guidance, support, encouragement and direction;

Koo Wen Hsin, overseeing the medical oncology service at the National Cancer Centre Singapore, for his patience and unyielding support;

Rajasoorya, for setting me down this path at a fateful lunch with my bosses of past, present and future some ten years ago;

Colleagues and friends at the Van Andel Research Institute, including but not restricted to Jonathon Ditlev, Mark Betten, Masayuki Takahashi, Khoo Sok Kean, David Petillo, Julie Koeman, Daisuke Matsuda, Miles Qian, Eric Kort, Kyle Furge, Jacob Zhang and James Resau;

Colleagues and friends at the National Cancer Centre Singapore and Singapore General Hospital, including but not restricted to Tan Hwei Ling, Li Huihua, Tan Puay Hoon, Wong Chin Fong, Ooi Aik Seng, Nay Min Htun, Eileen Poon;

The Singapore Millennium Foundation, the National Kidney Foundation, the Singapore Cancer Society and Singapore Health Services for their support;

My parents, Tan Kim Lee and Low Ken Yin, for the love and blessings lavished on to me over all my life;

My wife and best friend, Carolina Png – thank you for all the years, the laughter, the tears and the patience.

LIST OF PUBLICATIONS

This thesis is based on the following manuscripts:

1. **Tan MH,** Ravindran K, Li H, Tan HL, Tan PH, Wong CF, Chia KS, Teh BT, Yuen J, Chong TW. A comparison of the UCLA Integrated Staging System (UISS) and the Leibovich scores in survival prediction for patients with non-metastatic clear cell renal cell carcinoma. Urology 2010 June: 75(6) 1365-70.

2. **Tan MH,** Choong C, Tang T, Chia KS, Chong TW, Li H, Tan PH. The Karakiewicz nomogram is optimal in post-operative prediction of survival outcomes in nonmetastatic renal cell carcinoma. Cancer 2011 (in press).

3. **Tan MH**, Takahashi M, Ditlev JA, Kim HL, Rogers CG, Kort EJ, Zhang J, Furge KA, Kanayama H, Belldegrun A, Teh BT. Gene expression profiling identifies a prognostic signature in both primary and metastatic renal cell carcinoma (manuscript in preparation)

4. Yang XJ*, **Tan MH***, Kim HL, Ditlev JA, Betten MW, Png CE, Kort EJ, Futami K, Furge KA, Takahashi M, Kanayama H, Tan PH, The BS, Luan C, Wang K, Pins M, Tretiakova M, Anema J, Kahnoski R, Nicol T, Stadler W, Vogelzang NG, Amato R, Seligson D, Figlin R, Belldegrun A, Rogers CG, Teh BT. A molecular classification of papillary renal cell carcinoma. Cancer Res. 2005; 65(13): 5628-37. ***Co-first authors**

5. **Tan MH**, Wong CF, Tan HL, Yang XJ, Ditlev JA, Matsuda D, Khoo SK, Sugimura J, Furge KA, Kort E, Giraud S, Ferlicot S, Vielh P, Amsellem-Ouazana D, Debre B, Flam T, Thiounn N, Zerbib M, Benoit G, Droupy S, Molinie V, Vieillefond A, Tan PH, Richard S, Teh BT. Genomic expression and single nucleotide polymorphism profiling discriminates chromophobe renal cell carcinoma and renal oncocytoma. BMC Cancer 2010 May 12:10:196

6. Koeman JM*, Russell RC*, **Tan MH***, Petillo D, Westphal M, Koelzer K, Metcalf JL, Zhang ZF, Matsuda D, Dykema KJ, Houseman HL, Kort EJ, Furge LL, Kahnoski RJ, French Kidney Cancer Consortium, Swiatek PJ, Teh BT , Ohh M, Furge KA. Somatic pairing of chromosome 19 in renal oncocytoma is associated with deregulated EGLN2-mediated oxygen-sensing response. PLoS Genet. 2008 Sep 5; 4(9). e1000176*** Co-first authors**

TABLE OF CONTENTS

SUMMARY

The field of renal cell carcinoma (RCC) has evolved rapidly over the last five years, with the advent of novel therapies targeting specific molecular pathways dysregulated in RCC. The development of these drugs was via a classic bench-to-bedside fashion, where an understanding of the underlying biology in RCC permitted relevant drug development. The foundation of these biological insights was the careful pathologic subtyping of RCC, supported by advances in familial cancer genetics. These subtypes have tremendous clinical and biologic relevance, further illustrated by the clinical observation that survival outcomes in RCC may diverge more dramatically than almost any other cancer.

The work presented here is divided into **two areas** – the first being the evaluation of existing clinical models for outcome predictions in RCC, and the second being the evaluation of molecular models in RCC, and corresponding molecular insights. For the **first area**, we focused on the clinical models where epidemiologists and clinicians are actively seeking an optimal combination of clinico-pathologic variables for subtyping patients with RCC and predicting survival outcomes. Indeed, the literature is replete with a variety of proposed pre-operative and post-operative models. However, much less work has been invested in comparing these multiple models to choose one that is performing optimally. The work presented here compares multiple algorithms and nomograms to select an optimal and practical predictor in

localized RCC that may be recommended for use internationally for individual prognostication and in clinical trials of adjuvant therapy. We compare several clinical post-operative models including the Leibovich model, the UCLA Integrated Staging System (UISS), the Karakiewicz nomogram, the Kattan nomogram and the Sorbellini nomogram, and conclude that the best performing model is the Karakiewicz nomogram. This finding is of relevance in individual patient counseling, biomarker research and pharmaceutical trial design for adjuvant therapy.

For **the second area** on molecular models in RCC, I derive and evaluate useful molecular predictors in the various subtypes of RCC in terms of pathology and prognosis. Thus, various hitherto undescribed subtypes of RCC with distinct molecular and clinical profiles may be defined here. We have generated novel expression predictors of prognosis in clear cell RCC as well as papillary RCC, while concurrently generating insights into the molecular mechanisms underpinning these prognostic differences. For the rarer chromophobe RCC, we have reported a novel expression predictor discriminating chromophobe RCC from its close benign counterpart, renal oncocytoma, which was externally validated. We also found that somatic pairing of chromosome 19q, an unusual cytogenetic finding, was found in renal oncocytoma but not in chromophobe RCC, and was associated with deregulated oxygen-sensing response. Overall, our findings provide not only a comprehensive analysis of gene expression in the various molecular

subtypes of RCC, but has also provided multiple insights into the potential pathogenesis of each RCC subtype.

Finally, I hope that this work embodied in this thesis allows the scientific community investigating RCC to prepare its labours with a firm foundation from a clear understanding of the molecular epidemiology and pathology of RCC.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AUC	Area Under the Curve
BHD	Birt Hogg Dube
cDNA	Complementary Deoxyribonucleic Acid
CGMA	Comparative Genomic Microarray Analysis
CSS	Cancer Specific Survival
DFS	Disease Free Survival
DNA	DeoxyriboNucleic Acid
ECOG	Eastern Cooperative Oncology Group
FDR	False Discovery Rate
FISH	Fluorescent In Situ Hybridization
GEO	Gene Expression Omnibus
GO	Gene Ontology
HIF	Hypoxia-Inducible Factor
IHC	Immunohistochemistry
LIMMA	Linear Models for Microarray Data
LR	Likelihood Ratio
mRNA	Messenger RiboNucleic Acid
MSKCC	Memorial Sloan-Kettering Cancer Centre
MVD	Microvessel Density
OS	Overall Survival
PAM	Prediction Analysis of Microarrays
PBMC	Peripheral Blood Mononuclear Cells
RCC	Renal Cell Carcinoma
RMA	Robust Multichip Average
RNA	RiboNucleic Acid
ROC	Receiver Operating Characteristic
SAM	Significance Analysis of Microarrays
SSIGN	Stage, SIze, Grade and Necrosis Score
UISS	UCLA Integrated Staging System
VEGF	Vascular Endothelial Growth Factor
VHL	Von Hippel-Lindau
WCP	Whole Chromosome Painting

OVERALL BACKGROUND

Renal cell carcinoma (RCC) is the 7th most common cancer in males with an estimated 131,010 new cases diagnosed with 28,100 deaths from the disease in the USA in 2009(*Jemal et al.* 2009). In the USA, it is currently the 10th most common cancer overall. In Singapore, it is currently the 10th most common cancer in males. Over the last four decades, RCC is one of the few cancers to see a continued rise in incidence(*Chow et al.* 1999). This has been attributed to both a true increase, based on autopsy studies of individuals dying of unrelated causes, as well as ascertainment bias as a result of increased screening(*Chow et al.* 1999). Approximately 70% of the patients with RCC presents with localized disease, which is usually curative with nephrectomy. However, about a third of these patients eventually develop metastases during the course of follow-up(*Mejean et al.* 2003). The overall prognosis for metastatic RCC is poor, and even the development of novel agents used in systemic therapy has yielded only modest benefits(*Rini et al.* 2009). However, there is substantial heterogeneity in survival within clinical staging groups, and individual outcome remains difficult to predict. Risk factors for renal cell carcinoma include male sex, obesity, smoking, dialysis, hypertension and underlying germline mutations of specific tumour suppressor genes that result in hereditary RCC syndromes(*Chow et al.* 2000).

PATHOLOGY

Renal cell carcinoma is usually classified into several distinct histologic subtypes. Based on morphologic features first proposed in 1986(*Bostwick and Eble* 1999; *Thoenes et al.* 1986), RCC can be divided into clear cell (conventional), papillary (chromophil), chromophobe, collecting duct, and unclassified subtypes. Clear cell RCC constitutes more than 80% of all kidney cancers(*Cheville et al.* 2003), with papillary RCC, the second most common subtype comprising 10% to 15% of kidney cancers (*Bostwick and Eble* 1999). Rarer histologies include chromophobe RCC (approximately 5%) and collecting duct carcinoma (<1%).

While a variety of clinical models have been used for prognosticating RCC(*Cindolo et al.* 2005; *Galfano et al.* 2008), the understanding of genetic mechanisms underlying the variability in RCC behavior is more limited. Research in this area has been recognized as a key priority for oncology. The potential for identification of prognostic subgroups of patients for adjuvant treatment has been reinforced recently by the establishing of multi-targeted kinase inhibitors as a treatment modality in clear cell RCC (*Hutson et al.* 2008; *Motzer et al.* 2007). These findings have resulted in considerable excitement in the oncology community.

CLEAR CELL RCC

Clear cell RCC is the most common subtype of RCC (>70%), and also has the poorest overall prognosis(*Bostwick and Eble* 1999). Morphologically, clear cell RCC has a characteristic gross appearance, usually golden-brown as a result of lipid-rich cells. The tumour usually presents as a well defined mass, that can be heterogenous as a result of necrosis or haemorrhage. On microscopic appearance, it is typically characterized by malignant epithelial cells with clear cytoplasm and a compact-alveolar (nested) or acinar growth pattern interspersed with intricate, arborizing vasculature. The most common underlying mutation is the *VHL* gene mutation, which occurs as a somatic mutation in up to 90% of all patients with sporadic clear cell RCC(*Nickerson et al.* 2008), where there is evidence of somatic biallelic inactivation of the *VHL* gene (*Chen et al.* 1995) (*Iliopoulos et al.* 1995). It is recognized that clear cell RCC is also a common manifestation of the VHL syndrome(*Kaelin* 2004), where germline mutations of the *VHL* gene predispose to the development of multiple tumours, including clear cell RCC, cranial and spinal haemangioblastoma, phaeochromocytoma and multiple visceral cysts. Most recently, somatic mutations of PBRM1, a component of the SWI/SNF chromatin remodeling complex, have been identified in approximately 41% of clear cell RCC samples examined(*Varela et al.* 2011).

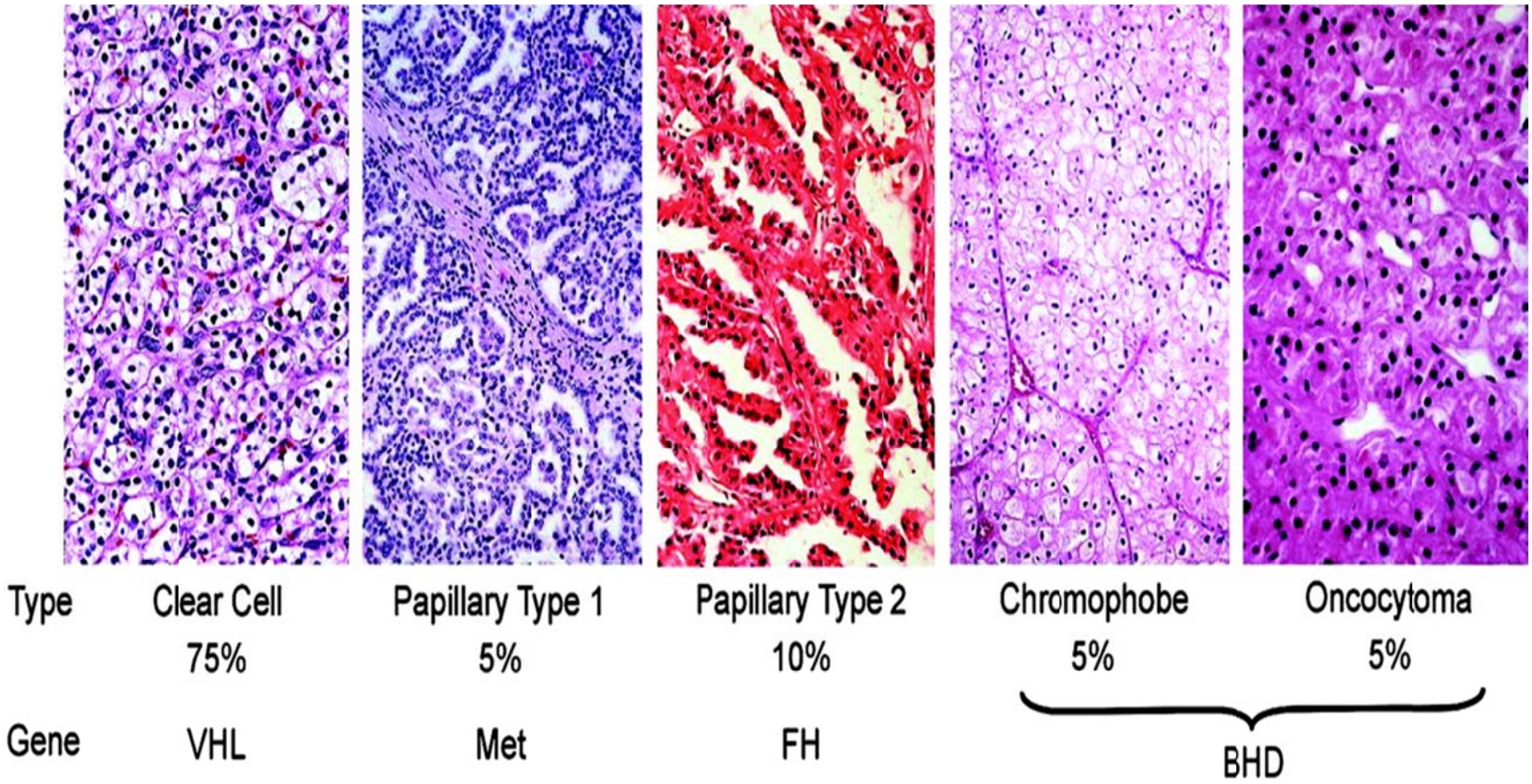

Figure 1 : Histologic subtypes of epithelial renal tumours,
The corresponding germline gene mutations in familial cancers is reported below together with the relative somatic mutation prevalence in sporadic cases of these tumours. For example, fumarate hydratase (*FH*) mutations are found in approximately 10% of papillary type 2 tumours.

PAPILLARY RCC

For papillary RCC, it is similarly true that the majority of these tumours show indolent behavior and have a limited risk of progression and mortality, but a distinct subset displays highly aggressive behavior. It is the second most common subtype comprising 10% to 15% of kidney cancers with an estimated annual incidence of between 3,500 and 5,000 cases in the United States (*Jemal et al.* 2009). Delahunt and Eble have proposed that papillary RCC can be morphologically classified into two subtypes (Figure 1, preceding page) (*Delahunt and Eble* 1997). Type 1 is characterized by the presence of small cuboidal cells covering thin papillae, with a single line of small uniform nuclei and basophilic cytoplasm. Type 2 is characterized by the presence of large tumour cells with eosinophilic cytoplasm and pseudostratification. Generally, type 2 tumours have a poorer prognosis than type 1 tumours (*Waldert et al.* 2008). However, the morphologic classification remains controversial, and there is limited molecular and biochemical evidence to support this morphologic classification. The relatively high incidence of mixed type 1 and 2 tumours poses additional difficulties for such a method of classification. As a result, some studies of papillary RCC do not stratify papillary RCC into type 1 and 2 tumours (*Cheville et al.* 2003). Despite the moderate incidence of PRCC, comparable to that of chronic myeloid leukemia, there is a disproportionately limited knowledge about the underlying molecular basis for development and progression of papillary RCC.

CHROMOPHOBE RCC

Chromophobe RCCs account for about 4-8% of all renal tumours, with a more favorable prognosis relative to clear cell renal cell carcinoma, which comprises the majority of all RCCs (*Cheville et al.* 2003). On the other hand, oncocytoma is the most common benign renal tumour, comprising 5-8% of resected renal masses. The overlapping characteristics of these entities may be explained by a possible common origin from the intercalated cells of the distal tubule (*Storkel et al.* 1989). Patients with Birt-Hogg-Dubé (BHD) syndrome, a familial multi-tumour syndrome linked to mutation of the *BHD* gene, exhibit bilateral oncocytomas, chRCC and hybrid tumours (*Khoo et al.* 2001; *Nickerson et al.* 2002).

DIAGNOSIS

Patients present to clinicians either in the asymptomatic setting (screening) or with a variety of symptoms that may be suggestive of either the local extension of the tumour, or the systemic spread of the cancer to distant sites. Local symptoms may include haematuria, loin pain or abdominal mass. Systemic symptoms may include fever, loss of appetite or weight, organ compromise or paraneoplastic symptoms. Patient evaluation for primary RCC involves usually radiologic imaging of the abdomen, using modalities such as ultrasound, computed tomographic scanning or magnetic resonance imaging. The use of urine cytology for histological confirmation of RCC is usually of low yield. The demonstration of a renal mass is followed by a clinical decision as

to the appropriate intervention: while renal masses may be biopsied to determine its nature, it is most often that clinicians will decide based on radiologic characteristics to intervene directly with the use of surgical treatment.

THERAPY

The treatment of renal cell carcinoma depends on the final pathologic and radiologic staging of the patient. Essentially, in the localized setting, a complete resection of the tumour is regarded as the standard of care. The current approach involves a nephrectomy (or removal of the kidney), with or without radical lymph node dissection. It should be noted that in the elderly and asymptomatic, or in patients with multiple comorbidities, a decision for surveillance may be undertaken(*Chen and Uzzo* 2009), to evaluate if the disease is indolent. In the metastatic setting, the standard of care involves the consideration of nephrectomy, with the first-line use of targeted therapies. Unusually, removal of the primary tumour has been demonstrated to confer a low, but definite survival benefit in patients with metastatic disease(*Flanigan et al.* 2001; *Flanigan* 2004). In the selection of targeted therapy for patients, tumour histology and risk stratification of patients are regarded as the primary factors of importance. The most widely used model for risk stratification currently is the MSKCC model(*Motzer et al.* 1999; *Motzer et al.* 2004), which classifies patients according to the presence of several adverse prognostic factors: Karnofsky performance scale of 70 or less, the presence of anaemia,

corrected serum calcium above the upper limit of normal, time from diagnosis/nephrectomy to therapy of less than one year, serum lactate dehydrogenase levels greater than 1.5 times the upper limit of normal. Patients with none of these factors are regarded as good prognosis; those with 1 or 2 factors considered as intermediate risk; patients with 3 or more factors considered as poor-risk. Currently, in patients who are categorized as good- or intermediate- prognosis by the Memorial Sloan-Kettering Cancer Centre (MSKCC) criteria, the standard of care for first-line treatment is a targeted therapy utilizing tyrosine kinase inhibitors, most commonly sunitinib, but which include agents such as sorafenib and bevacizumab in combination with interferon(*Rini* 2009). For patients with poor-prognosis MSKCC, the current standard of care is an mTOR inhibitor administered intravenously (temsirolimus) (*Hudes et al.* 2007). The current second-line standard of care following failure of first-line VEGF-targeted therapy is everolimus(*Motzer et al.* 2008). Adjuvant therapy using antiangiogenic therapy is currently under active research with several ongoing clinical trials recruiting patients. Based on the success of sunitinib and sorafenib, the UK Medical Research Council (MRC) SORCE and the Sunitinib Treatment of Renal Adjuvant Cancer (STAR) multi-centre Phase III trials are ongoing. Respectively, these trials are testing placebo versus sorafenib versus sunitinib, as well as sunitinib versus placebo in the adjuvant setting for high-risk patients following surgery.

AIMS

OVERALL AIMS

We aim to evaluate both clinical and molecular parameters of RCC, a heterogenous disease, with a view to determining underlying mechanisms of disease and developing useful models for predicting survival outcomes. We aimed to evaluate how molecular profiling may improve or complement these survival predictions, and how these studies may provide biologic insight on the clinical heterogeneity observed.

SPECIFIC AIMS (CLINICAL MODELS)

To evaluate clinical models in predicting survival outcomes in patients with renal cell carcinoma;

SPECIFIC AIMS (MOLECULAR MODELS)

To evaluate the molecular profiles of three primary subtypes of RCC clear cell RCC, papillary RCC and chromophobe RCC using high-throughput gene expression profiling technology. This would be in the context of clinical outcomes, specifically survival, regional gene expression biases, high throughput single-nucleotide polymorphism profiling and protein expression, using immunohistochemistry.

CLINICAL MODELS IN RENAL CELL CARCINOMA

BACKGROUND

CLINICAL PROGNOSTIC MODELS

From a clinical viewpoint, there are several prognostic models and nomograms developed to estimate survival outcomes of patients with localized RCC(*Cindolo et al.* 2005). These models are used in clinical practice to aid in counseling, follow-up planning and most recently, patient classification into groups for trials of adjuvant therapy(*Haas and Uzzo* 2008). However, there is substantial heterogeneity in survival within clinical staging groups, and individual outcomes remain difficult to predict. These prognostic models and nomograms incorporate multiple clinical and pathologic variables in their scoring. There are several prognostic risk groups and nomograms developed to estimate outcomes of patients with localized renal cell carcinoma (RCC), for whom there is an overall relapse risk of between 20-30%(*Cindolo et al.* 2005). We present a summary table describing the key similarities and differences between risk grouping models and nomograms (Table 1, following page). It should be noted in particular that risk grouping models are far more widespread in clinical acceptance than nomograms, primarily due to their simplicity.

Table 1: Comparison of algorithms and nomograms in predicting survival outcomes

	Algorithms	Nomograms
Objective and Approach	To predict outcomes for the individual patient by classifying similar, but not identical, patients into risk groups, where all patients in the risk group have the same predicted outcome	To predict outcomes for patients using formulae that computes predictions for the individualized patient, rather than for risk groups. Points on a semicontinuous scale are assigned to individual variables.
Scales of Measurement	Relatively fewer groups on a semiquantitative, or ordinal scale (e.g. classification into groups of low-, medium- and high- risk patients)	A directly calculated quantitative outcome on a numerical scale (e.g. predicted overall survival of 69% at 5 years for a patient)
Clinical use	Widespread e.g. International Prognostic Index in lymphoma	Limited, most commonly used in prostate cancer
Examples in RCC	UCLA Integrated Staging System, Leibovich score	Karakiewicz nomogram, Kattan nomogram, Sorbellini nomogram

A recent systematic review(*Galfano et al.* 2008) found 11 different mathematical models proposed for this purpose, including both models describing risk groups and nomograms. Unlike risk groupings (the most common approach), nomograms use continuous scales and thus are able to calculate the continuous probability of a particular outcome. This maximizes the predictive power of the nomogram as it eliminates the spectrum bias that occurs when predictors are stratified(*Karakiewicz and Hutterer* 2007). However, these nomograms have not been used within trial design simply because nomograms do not have established cut-offs for decision-making in

current practical use in these trials. Even in publications where nomograms were compared with other prognostic models, relatively arbitrary cut-offs were employed for the nomograms with multiple strata(*Cindolo et al.* 2005)(*Liu et al.* 2009).

Table 2 : UCLA Integrated Staging System (UISS) for Non-Metastatic RCC

T stage	1				2	3				4
Grade	1-2		3-4		↓	1		>1		↓
ECOG PS	0	≥1	0	≥1		0	≥1	0	≥1	
Risk group	Low	Intermediate							High	

Commonly evaluated risk grouping models are the UCLA Integrated Staging System (UISS) model(*Zisman et al.* 2001; *Zisman et al.* 2002), the Mayo stage, size, grade, necrosis (SSIGN) model (*Frank et al.* 2002; *Leibovich et al.* 2003), and the Leibovich prognostic score. The UISS model (Table 2) was developed to predict overall survival in patients with kidney cancer regardless of histological subtype, while the SSIGN model was developed separately to predict cancer-specific survival in patients with clear cell RCC only. The Mayo group further developed a model distinct from the SSIGN to predict disease-free survival in patients with non-metastatic clear cell RCC, referred to as the Leibovich prognostic score(*Leibovich et al.* 2003). For non-metastatic tumours, the SSIGN and the Leibovich models use the same tumour features, but differ in terms of scoring weights. While the UISS

and the SSIGN scores have been validated in Western populations of patients with clear cell RCC (*Cindolo et al.* 2005; *Ficarra et al.* 2006; *Han et al.* 2003; *Patard et al.* 2004), and the SSIGN score has been validated in Asia (*Fujii et al.* 2008), the Leibovich score has not been previously externally validated in any population. The first direct comparison of the UISS and the SSIGN performed in Italy (*Ficarra et al.* 2009) reported that the SSIGN score is more accurate than the UISS for predicting cancer specific survival in patients with clear cell RCC, using a comparison of the respective areas under the ROC curve (AUC).

With the introduction of the UISS and the Leibovich scores (Table 3, following page) into inclusion criteria of separate Phase III adjuvant trials, it is urgently required that the utility of these scores be directly compared to ensure uniformity of future trial designs and minimize confusion. In particular, the absence of external validation of the Leibovich score was noted. The UK MRC SORCE trial is currently recruiting patients with intermediate and high-risk Leibovich scores for randomization between placebo and sorafenib. Both the ASSURE (ECOG 2805) trial (comparing placebo versus sorafenib versus sunitinib) and the S-TRAC trial (comparing sunitinib versus placebo) are selecting patients for adjuvant therapy based on UISS scores. To our knowledge, the SSIGN score is not used currently for selection of patients in adjuvant Phase III trials in RCC.

Table 3 : Leibovich Algorithm to predict metastasis after nephrectomy

Feature	Score
Primary tumour status (pathologic T stage)	
pT1a	0
pT1b	2
pT2	3
pT3a	4
pT3b	4
pT3c	4
pT4	4
Regional lymph node status (N stage)	
pNx	0
pN0	0
pN1	2
pN2	2
Tumour size (cm)	
<10	0
>=10	1
Nuclear grade	
1	0
2	0
3	1
4	3
Histologic tumour necrosis	
No	0
Yes	1

In contrast, nomograms have received far less attention as compared to models involving risk groups. The Kattan nomogram (*Kattan et al.* 2001) and Karakiewicz nomogram (*Karakiewicz and Hutterer* 2007) are two such nomogram-based models. The Kattan nomogram (Figure 2, following page) was developed in 2001 to predict 5-year disease-free survival (DFS) in patients undergoing radical nephrectomy for non-metastatic RCC. The variables used in this post-operative nomogram were symptoms, histological subtype, pathological tumour size and T-stage. More recently, the

Karakiewicz nomogram (Figure 3, following page) was developed to predict 1-, 2-, 5-, and 10- year cancer-specific survival (CSS) of patients undergoing nephrectomy for RCC of all stages. The Karakiewicz nomogram was designed as a post-operative nomogram with the variables T, N and M stages, tumour size, Furhman grade, histological type, age and symptom classification.

Overall, there have been more new models than comparative studies for selecting an optimal model. We therefore chose to perform a comparative effectiveness study to clarify the field by externally validating the various models, rather than develop another model from a relatively smaller dataset. The two algorithm-based models, the UISS and the Leibovich models have been incorporated for patient selection in large trials of adjuvant therapy in RCC. The UK Medical Research Council SORCE trial is currently recruiting patients with intermediate and high-risk Leibovich scores for randomization between placebo and sorafenib. Both the adjuvant sorafenib or sunitinib for unfavorable renal carcinoma (Eastern Cooperative Oncology Group [ECOG] 2805) trial (comparing placebo vs sorafenib vs sunitinib) and the Sunitinib Treatment of Renal Adjuvant Cancer trial (comparing sunitinib vs placebo) are selecting patients for adjuvant therapy based on UISS scores.

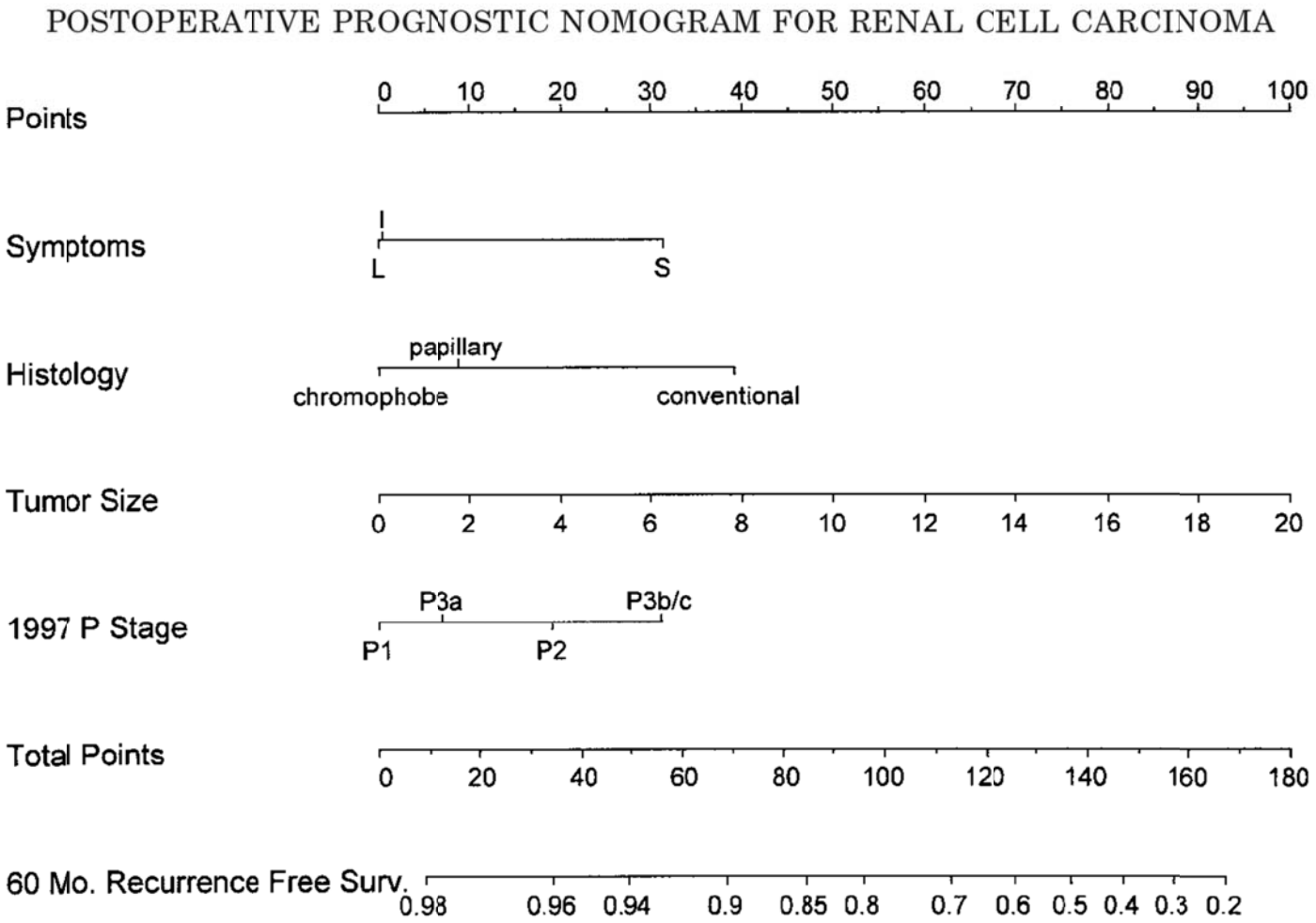

Figure 2: The Kattan nomogram for obtaining a corresponding individual point 5-year recurrence free survival (RFS) prediction. Each feature (symptoms, histology, tumour size, pathologic T-stage) is allocated a weight in terms of points using the top bar. After totaling the number of points, the corresponding 5-year recurrence-free survival can be calculated by a vertical line from the bar corresponding to 'Total points'. For example, 70 points corresponds to approximately less than 0.9 5 year RFS. (Reprinted with permission © 2001 Elsevier)

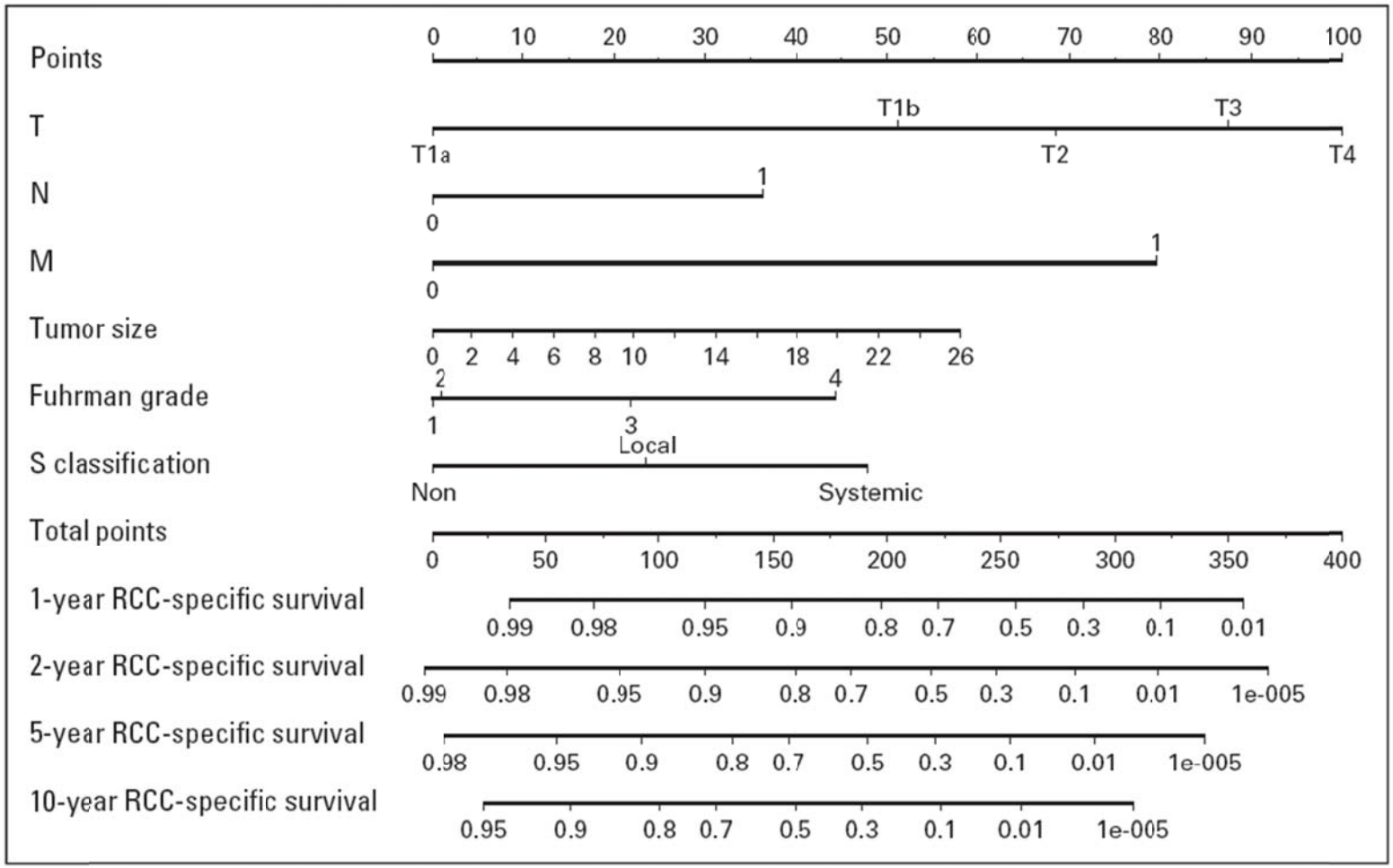

Figure 3: The Karakiewicz nomogram for obtaining a corresponding individual point survival probability.
Each feature (T stage, N nodal stage, M metastatic stage, tumour size, Fuhrman grade, symptom classification corresponding to none, local only and systemic symptoms) is allocated a weight in terms of points using the top bar. After totaling the number of points, the corresponding 1-year, 2-year, 5-year and 10-year RCC-specific survivals can be calculated by a vertical line from the bar corresponding to 'Total points'. For example, 200 points corresponds to approximately 1-year RCC-specific survival of 0.8. (Reprinted with permission **)**

While both the above-discussed UISS and the Leibovich models are risk grouping methods to predict outcomes for patients(*Kattan* 2008), nomograms differ from risk groups in several relevant ways, as described earlier. These models are used in clinical practice to aid in counseling, follow-up planning and most recently, patient classification into groups for trials of adjuvant therapy(*Borowiak et al.* 2004). To illustrate the current use of risk grouping, the UK Medical Research Council SORCE trial is recruiting patients with intermediate and high-risk Leibovich scores for randomization between sorafenib and placebo(*Eisen* 2007), whereas the ASSURE and the Sunitinib Treatment of Renal Adjuvant Cancer (S-TRAC) trials are selecting patients based on modified UCLA Integrated Staging System (UISS) criteria. Although it is recognized that nomograms often outperform risk grouping(*Di Blasio et al.* 2003), nomograms have not been integrated into RCC clinical trials, and no study has comprehensively examined nomogram performance relative to that of risk grouping in predicting survival in RCC. Several studies evaluating nomogram performance have imposed multiple thresholds on the nomogram outcomes(*Cindolo et al.* 2005; *Liu et al.* 2009), with no disclaimer of exploratory analysis or expanded justification presented for the selection of these thresholds.

While discretization of a continuous variable inevitably results in a loss of statistical information(*Morgan and Elashoff* 1987), this is practically difficult to avoid in a trial setting, where thresholds for patient recruitment are used.

For implementation of a trial design using nomograms, a single threshold for discretizing risk, so as to divide patients into two risk-groups, is usually most useful for establishing clear inclusion criteria and for decision making by clinicians and patients.

Several post-operative clinical nomograms have been recently developed for RCC(*Galfano et al.* 2008). The Kattan nomogram was developed to predict 5-year freedom-from-recurrence (FFR) in patients undergoing radical nephrectomy for non-metastatic RCC(*Kattan et al.* 2001). The variables used in this post-operative nomogram were symptoms, histological subtype, tumour size and T-classification. More recently, the Karakiewicz nomogram was developed to predict 1-, 2-, 5-, and 10- year cancer-specific survival (CSS) of patients undergoing nephrectomy for RCC of all stages. The Karakiewicz nomogram was designed as a post-operative nomogram with the variables T, N and M classifications, size, Fuhrman grade, histologic subtype, age and symptom classification(*Karakiewicz et al.* 2007). The Sorbellini nomogram was developed to predict 5-year FFR for patients undergoing surgical treatment for localized clear cell RCC, using tumour size, T classification, Fuhrman grade, tumour necrosis, vascular invasion, and symptom presentation(*Sorbellini et al.* 2005).

Despite the multiple RCC nomograms published in the literature, none are currently in widespread clinical use. Instead, ongoing clinical trials employ risk grouping for risk estimation and patient selection. The UK Medical Research Council SORCE trial is recruiting patients with intermediate and

high-risk Leibovich scores for randomization between sorafenib and placebo(*Eisen* 2007), whereas the ASSURE and the Sunitinib Treatment of Renal Adjuvant Cancer (S-TRAC) trials are selecting patients based on modified UCLA Integrated Staging System (UISS) criteria. In evaluating risk group performance, we have shown that the SORCE trial criteria (a discretized Leibovich model) performs better in discrimination than the ASSURE trial criteria (a discretized modified-UISS model). The purpose of the present study was to clarify which clinical model is most useful for survival prediction in localized RCC, so as to establish a standard for guiding trial design and biomarker research.

AIMS (CLINICAL MODELS)

To evaluate clinical models in predicting survival outcomes in patients with renal cell carcinoma, and determining which clinical and pathologic parameters are most useful in determining prognosis.

METHODS

SUBJECTS

We conducted two distinct analyses of the data, in order to account for differences in the selection criteria of each model. Analysis I was conducted for comparing the UISS and the Leibovich score, two risk models in existing use by pharmaceutical companies in recruiting high-risk post-nephrectomy patients for adjuvant trials. Analysis II was conducted for comparing nomograms against the best performing risk model from Analysis I. Due to minor differences in inclusion criteria for each model, the datasets for each analysis differed slightly, details of which are provided below.

For the comparison of the UISS and the Leibovich score, we identified 364 patients with unilateral non-metastatic clear cell RCC and who underwent nephrectomy at the Singapore General Hospital between 1990 and 2006 through a comprehensive search of the Singapore General Hospital Pathology database and the National Cancer Centre Department of Cancer Informatics. ECOG (Eastern Cooperative Oncology Group) (*Oken et al.* 1982) scores exceeding 1 were excluded (n=9), as our study focused on patients

who were candidates for adjuvant trials. Eventually, 355 patients were selected for evaluation of the UISS and Leibovich scores. For the Leibovich model, we categorized the patients into low (0-2), intermediate (3-5) and high risk (≥6) groups (*Leibovich et al.* 2003). In terms of terminology, we refer to this categorization of UISS and Leibovich scores into these three risk groups as the UISS and the Leibovich models respectively. We refer to the modification of these systems into two categories (low risk versus intermediate and high risk groups) as either UISS or Leibovich trial criteria.

For Analysis II, where we compared nomograms and risk models, a different approach to selection was adopted in view of the fact that several of the risk models were constructed with patient sets with different features. 413 patients with unilateral non-metastatic RCC of all subtypes who underwent nephrectomy at the Singapore General Hospital between 1990 and 2006 were identified through a database search (as contrasted to the earlier comparison, where only patients with clear cell histology were selected). Survival status and cause of death, if any, were obtained from a national registry. To ensure the most accurate comparisons between nomograms, we used an approach to select common selection criteria. Broadly, the Karakiewicz nomogram had the least restrictive selection criteria and similar to the Kattan nomogram, was applicable to all RCC subtypes. The Sorbellini nomogram and the Leibovich score were restricted to clear cell RCC. Therefore, in comparing the Karakiewicz nomogram with the Kattan nomogram, we excluded patients with large tumours (pT4), ECOG>1, and

patients with subtypes other than clear cell RCC, papillary RCC or chromophobe RCC (n=33), with a remaining data-set of 390 subjects. The ECOG limitation was imposed to ensure that this comparison would be useful in the appropriate patient set for trial design, and is consistent with our previous approach. All comparisons with the Sorbellini nomogram (n = 329) and the Leibovich score (n = 322) similarly were restricted to clear cell subtype only, and selection mirrored the more restrictive criteria of each score to ensure a fair comparison.

In our studies, all specimens were reviewed by a pathologist for histological subtype, tumour grade, lymphovascular invasion and necrosis. The tumour size of pathological specimens was determined as the greatest dimension in centimeters and the Fuhrman grading scheme was used to determine the nuclear grade of tumours. Pathologic staging was determined in accordance with the AJCC 2002 primary tumour TNM classification,(*Hudes et al.* 2007) except when scoring the Kattan nomogram, where AJCC 1997 stage grouping was used. The symptoms were classified as incidental, local (hematuria, flank pain, palpable mass), or systemic (weight loss, anorexia, asthenia, fever)(*Bugert and Kovacs* 1996).

STATISTICAL ANALYSES

We compared the clinico-pathologic profile of patients in our data-set to that of the Mayo data-set using Fisher's exact test for categorical variables to assess for baseline differences. 5-year cancer specific survival (CSS), overall survival (OS) and disease-free survival outcomes were estimated by

the Kaplan-Meier method. Cox regression was also performed to evaluate the effect of UISS and Leibovich scores on CSS, OS and DFS separately. Proportional hazards assumptions were verified systematically for each score graphically (data not shown). To test for a difference in the predictive value of the UISS and Leibovich models and trial criteria for a variety of survival outcomes, we used the LR χ^2 test for nested models to assess whether the UISS model adds predictive value to a model including the Leibovich model, and vice versa, as well as whether the UISS trial criteria adds predictive value to a model including the Leibovich trial criteria, and vice versa. An adequacy index using likelihood ratio methods was used to quantify the percentage variation explained by a subset of the predictors (UISS or Leibovich scores separately) compared with the information contained in the full set of predictors (both UISS and Leibovich scores) by means of log-likelihood. Harrell's c-index was calculated to evaluate the concordance between predicted and observed responses of individual subjects in terms of UISS and Leibovich scores separately.

We chose three study endpoints to evaluate in common across the different models, these being cancer-specific survival (CSS), freedom-from-recurrence (FFR), and overall survival (OS). CSS was defined as the interval between diagnosis date and cancer-related death date, or last-follow up date for censored patients. FFR was defined as the interval between surgery date and relapse date, or the date of last follow up for censored patients. OS was defined as the interval between diagnosis date and death date or last-follow

up date for censored patients. The Leibovich score was developed on metastasis-free survival (MFS). As there is considerable overlap between the definition of FFR, MFS and disease-free-survival, and hence replicated analyses did not show material differences between these three outcomes, we selected to present FFR here. Outcomes were estimated by the Kaplan-Meier approach, and Cox regression was used to evaluate the effects of covariates. Proportional hazards assumptions were verified graphically (data not shown). We used LR χ^2 of nested models to perform pairwise comparisons of the models involved. An adequacy index using likelihood ratio (LR) methods was used to quantify the percentage of the variation explained by a subset of the individual predictors compared with the information contained in the full set of predictors by means of log-likelihood(*Al-Radi et al.* 2007; *Harrell* 2001). Harrell's c-index was calculated to evaluate the concordance between predicted and observed responses of individual subjects separately. Calibration is useful for evaluating whether actual outcomes approximate predicted outcomes for each model in our dataset. For calibration comparisons, we evaluated each model by its defined 5-year survival outcome (Karakiewicz: CSS; Kattan, Sorbellini: FFR; Leibovich: MFS), collapsing each nomogram to approximate the risk grouping of the Leibovich score for greater comparability (risk thresholds 0.9 and 0.6), with expected outcome in each risk-group determined by the median scorer. The Leibovich score was calibrated by prespecified low, intermediate and high-risk groups(*Leibovich et al.* 2003). Decision curve analyses were performed to

determine the clinical net benefit derived by examining the theoretical relationship between the threshold probability of developing an event and the relative value of false-positive and false-negative results as described by Vickers et al(*Vickers et al.* 2009). To evaluate whether the nomogram has potential to outperform current standards of risk evaluation, we further tested several pre-specified Karakiewicz nomogram thresholds (estimated 5-year CSS of 0.90, 0.85 and 0.80) against the SORCE trial criteria (a discretized Leibovich score) using similar methodology on an exploratory basis. The SORCE trial criteria divided patients into low-risk (0-2) and intermediate-/high-risk (≥3) individuals by Leibovich score. A separate decision analytic approach was also performed to determine estimated cut-off(*Vickers et al.* 2009), based on a threshold benefit of 0.05, similar to considerations of adjuvant therapy in gastric(*Paoletti et al.* 2010), colorectal(*Baddi and Benson* 2005), and breast cancer(*Seruga et al.* 2010), with a 0.5 risk reduction (*Motzer et al.* 2007). STATA 11 and R 2.11.1 were used for analysis, and all tests were two-sided with a significance level of 0.05.

RESULTS

The clinico-pathologic characteristics of the Singapore cohort is reported in Table 4 (following page). A comparison is provided against the Leibovich data-set; no equivalent data is available for the UCLA data-set. Over a median follow-up of 56 months, 78 patients had relapsed, 46 had died

Table 4 : Characteristics of patients for the comparisons between the Leibovich score and the UCLA Integrated Staging System (Analysis I) and between the nomograms and the Leibovich score (Analysis II)

		Analysis I	Analysis II
Total Number of Patients		355	390
Patients relapsing (n)		78	95
Patients dying of RCC		46	63
Patients dying of other causes		26	23
Median Followup (months)		56	65
Gender	Male (%)	228 (64.2)	256 (66)
	Female (%)	127 (35.8)	134 (34)
Age	Median ± SD	57.0 ± 12.4	56.7 ± 12.4
Race	Chinese (%)	289 (81)	318 (82)
	Non-Chinese (%)	66 (19)	72 (18)
ECOG	0 (%)	242 (68.2)	269 (70)
	1 (%)	104 (29.3)	121 (30)
	2 (%)	6 (1.7)	0 (0)
	3 (%)	3 (0.9)	0 (0)
Tumour Size	Mean ± SD	5.9 ± 3.2	6.2 ± 3.5
	≥ 10 cm (%)	44 (12.4)	61 (15.6)
Pathological Stage	pT1 (%)	178 (50.1)	214 (55)
	pT2 (%)	76 (21.4)	65 (17)
	pT3a (%)	57 (16.1)	69 (18)
	pT3b/c (%)	40 (11.3)	42 (11)
	pT4 (%)	4 (1.1)	0 (0)
Fuhrman Grade	1 (%)	61 (17.2)	57 (15)
	2 (%)	186 (52.4)	206 (53)
	3 (%)	83 (23.4)	99 (25)
	4 (%)	25 (7.0)	28 (7)
Coagulative necrosis	Yes (%)	130 (36.6)	152 (39)
	No (%)	225 (63.4)	236 (61)
Histology	Clear cell (%)	355 (100)	334 (86)
	Papillary (%)		44 (11)
	Chromophobe (%)		12 (3)

of disease and 26 had died of causes other than cancer. The survival outcomes are presented as Kaplan Meier survival curves (Figure 4, following page).

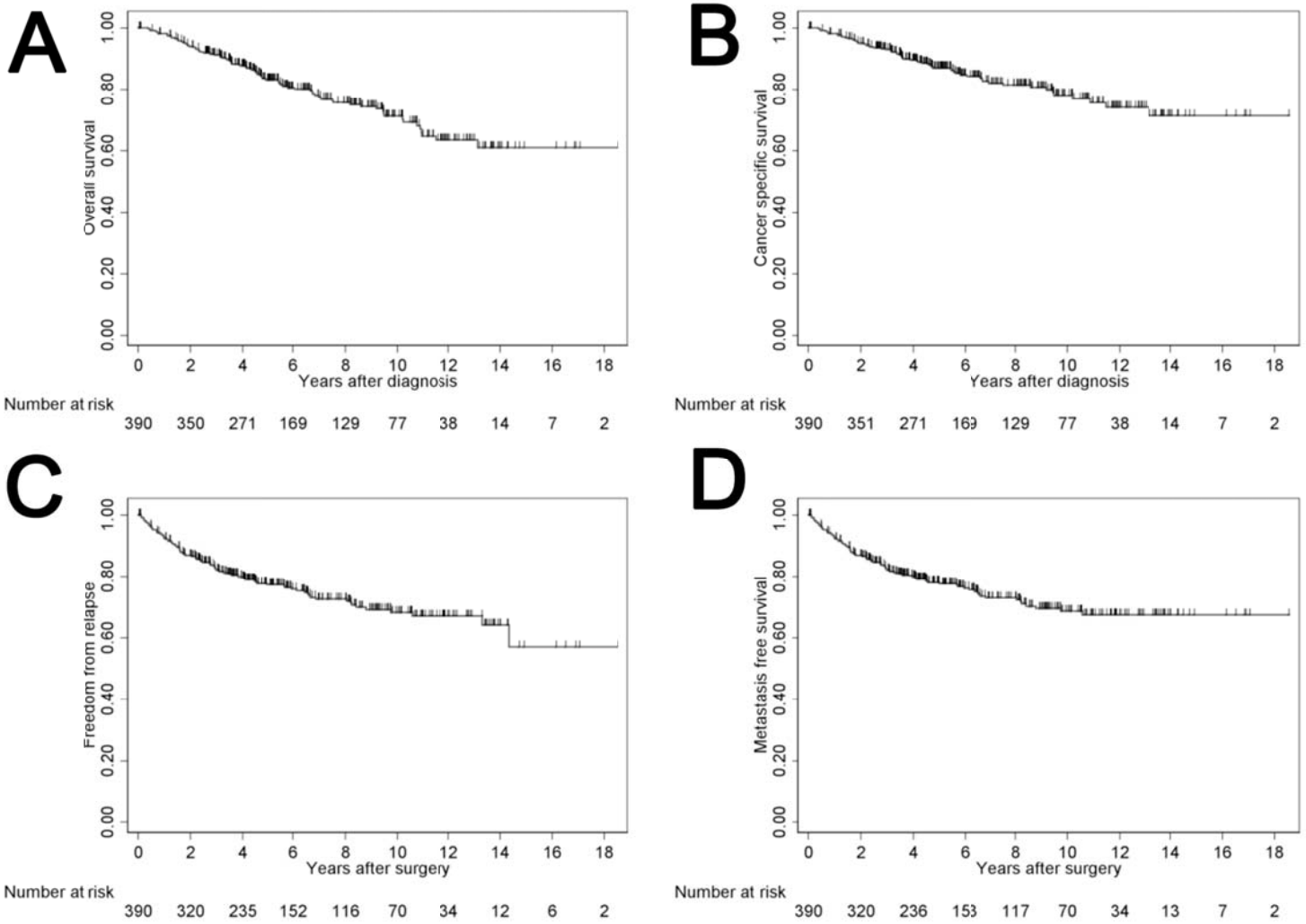

Figure 4: Kaplan-Meier survival curves for Singapore patient cohort with localized renal cell carcinoma.
Censoring is denoted by vertical ticks. Presented here are the outcomes of (A) overall survival; (B) cancer-specific survival; (C) freedom from recurrence; and (D) metastasis-free survival outcomes for the cohort. Numbers at risk row corresponds to the total number of subjects on follow-up at the start of each timepoint.

Cox regression showed that patients with a higher score (either UISS or Leibovich scores) have a higher chance of dying than those with a lower score (Table 5, following page). The concordance indices are reported in Table 5 as well. We show that the addition of the Leibovich model to one containing the UISS model significantly improves the predictive value of the final model, but there is no significant difference in adding the UISS model to the Leibovich model. A similar conclusion for the UISS and Leibovich trial criteria is seen for both cancer specific survival and disease-free survival, but there is no significant difference in terms of overall survival. The higher adequacy and concordance indices of the Leibovich score supports a similar conclusion that the Leibovich score is a superior predictor to the UISS score, both models used in patient recruitment for pharmaceutical trials.

Similarly, we show that the Karakiewicz nomogram is overall the best performing nomogram when individually compared against the other major nomograms (Table 6, 7(*Leibovich et al.* 2003)). The Karakiewicz nomogram had consistently higher adequacy and concordance indices for all tested outcomes. Its inclusion in a full model resulted in highly statistically significant accuracy improvements for all outcomes when tested with LR analysis against the Kattan nomogram ($p<0.001$), the Sorbellini nomogram ($p<0.001$), with marginal improvements over the Leibovich score ($p=0.04$ for CSS, $p=0.03$ for DFS, with equivalent performance for OS). This supports our conclusion that the Karakiewicz nomogram is a superior predictor to the Kattan or Sorbellini nomograms.

Table 5 : Comparison of the various models by survival outcomes and concordance indices

		OS		CSS		DFS	
		HR (95% C.I.)	C-index	HR (95% C.I.)	C-index	HR (95% C.I.)	C-index
Analysis I (n=355)	UISS model		0.64		0.65		0.66
	IR vs. LR	2.63 (1.28, 5.38)		3.94 (1.39, 11.19)		2.72 (1.50, 4.95)	
	HR vs. LR	4.28 (1.92, 9.55)		6.54 (2.10, 20.34)		5.17 (2.64, 10.12)	
	Leibovich model		0.67		0.74		0.7
	IR vs. LR	2.05 (1.09, 3.86)		3.41 (1.30, 8.97)		2.32 (1.34, 4.02)	
	HR vs. LR	5.17 (2.59, 10.32)		10.84 (4.00, 29.41)		7.74 (4.32, 13.87)	
	UISS trial criteria		0.63		0.65		0.63
	HR/IR vs LR	3.02 (1.54, 5.93)		5.63 (2.01, 15.74)		3.27 (1.88, 5.68)	
	Leibovich criteria		0.63		0.66		0.64
	HR/IR vs LR	3.02 (1.58, 5.77)		4.91 (1.93, 12.45)		3.66 (2.11, 6.36)	
Analysis II (n = 390)	Kattan nomogram		0.64		0.7		0.66
	Kattan LR	Reference		Reference		Reference	
	Kattan HR	2.95 (1.84, 4.73		5.37 (2.80, 10.28)		3.54 (2.33, 5.37)	
	Karakiewicz nomogram		0.71		0.75		0.71
	Karakiewicz LR			Reference		Reference	
	Karakiewicz HR	5.89 (3.42, 10.14)		12.69 (5.47, 29.45)		5.64 (3.59, 8.84)	

Table 6 : Likelihood ratio testing comparisons of the Kattan and the Karakiewicz nomograms

	Likelihood			P-value		Adequacy index	
	Kattan+Karakiewicz	Kattan	Karakiewicz	Kattan[a]	Karakiewicz[b]	Kattan[c]	Karakiewicz[d]
Kattan nomogram classification (DFS cutoff of 0.9 at 5 years) and Karakiewicz nomogram classification (CSS cutoff of 0.9 at 5 years)							
CSS	63.79	34.38	63.68	0.9052	<0.0001	53.9%	99.8%
DFS	72.88	41.18	72.66	0.7340	<0.0001	56.5%	99.7%
OS	54.87	22.83	54.36	0.4784	<0.0001	41.6%	99.1%
Kattan nomogram and Karakiewicz nomogram as semi continuous scores							
	Kattan+Karakiewicz	Kattan	Karakiewicz	Kattan[a]	Karakiewicz[b]	Kattan[c]	Karakiewicz[d]
CSS	83.59	42.11	83.16	0.5113	<0.0001	50.4%	99.5%
DFS	97.73	52.69	97.72	0.9106	<0.0001	53.9%	100.0%
OS	65.38	31.81	64.95	0.5128	<0.0001	48.7%	99.3%

[a]: The comparison of the model with the predictor of the Karakiewicz nomogram only with the one with predicators of both the Kattan nomogram and the Karakiewicz nomogram

[b]: The comparison of the model with the predictor of the Kattan nomogram only with the one with predicators of both the Kattan nomogram and the Karakiewicz nomogram

[c]: The proportion of the variation explained by the Kattan nomogram compared to that explained by both the Kattan nomogram and the Karakiewicz nomogram

[d]: The proportion of the variation explained by the Karakiewicz nomogram compared to that explained by both the Kattan nomogram and the Karakiewicz nomogram

Table 7 : Comparison of the Karakiewicz nomogram and the Leibovich score in outcome prediction

	Likelihood			P-value		Adequacy index	
	Karakiewicz+Leibovich	Leibovich	Karakiewicz	Leibovich[a]	Karakiewicz[b]	Leibovich[c]	Karakiewicz[d]
Karakiewicz nomogram classification (CSS cutoff of 0.9 at 5 years) and Leibovich prognostic score							
CSS	56.94	33.54	56.90	0.8456	<0.0001	58.9%	99.9%
DFS	72.77	53.72	69.86	0.0884	<0.0001	73.8%	96.0%
OS	49.75	32.56	49.55	0.6535	<0.0001	65.4%	99.6%
Karakiewicz nomogram and Leibovich scores as semi-continuous scores							
	Karakiewicz+Leibovich	Leibovich	Karakiewicz	Leibovich[a]	Karakiewicz[b]	Leibovich[c]	Karakiewicz[d]
CSS	68.87	61.66	67.97	0.3426	0.0072	89.5%	98.7%
DFS	88.40	82.31	85.73	0.1087	0.0136	93.1%	97.0%
OS	57.71	54.65	55.10	0.1205	0.0912	94.7%	95.5%

[a]: The comparison of the model with the predictor of the Karakiewicz nomogram only with the one with predicators of both the Leibovich prognostic score and the Karakiewicz nomogram

[b]: The comparison of the model with the predictor of the Leibovich prognostic score only with the one with predicators predicators of both the Leibovich prognostic score and the Karakiewicz nomogram

[c]: The proportion of the variation explained by the Leibovich prognostic score compared to that explained by both the Leibovich prognostic score and the Karakiewicz nomogram

[d]: The proportion of the variation explained by the Karakiewicz nomogram compared to that explained by both the Leibovich prognostic score and the Karakiewicz nomogram

On inspection of the calibration plots, all 4 models were reasonably calibrated with best calibration seen in the Karakiewicz nomogram and the Leibovich model (Figure 5). Maximum departure from ideal outcomes was as follows: Karakiewciz (4%), Kattan (11%), Sorbellini (15%), Leibovich score (17%), and the Leibovich model with risk groupings (3%).

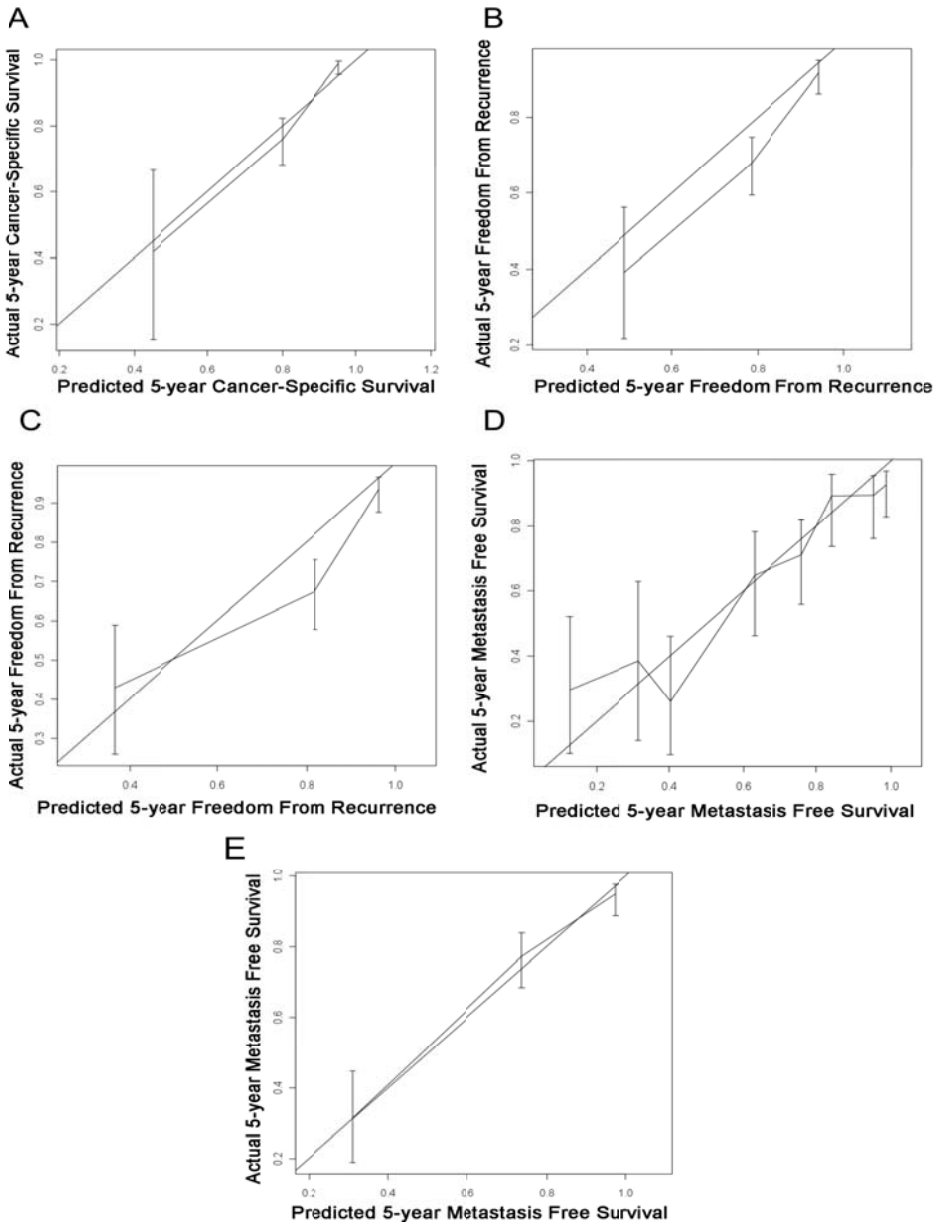

Figure 5: Calibration plots for the Singapore data set. Shown are (A) the Karakiewicz nomogram and cancer-specific survival; (B) the Kattan nomogram and freedom from recurrence; (C) the Sorbellini nomogram and freedom from recurrence; and (D) Leibovich prognostic score and metastasis-specific survival. (E) Due to multiple risk groups of the Leibovich score, increasing perceived variability, we evaluated calibration of corresponding low- (2 or less), intermediate- (3-5), and high-risk (6 or more) groups. Vertical bars are 95% confidence intervals. Model performance in terms of agreement with predicted outcomes is shown by the plot relative to the diagonal line representing perfect prediction.

In exploratory pairwise comparisons, we noted that several discretized Karakiewicz nomograms generally performed better than the ongoing SORCE trial criteria (a discretized Leibovich score) in terms of LR analysis, concordance indices and clinical net benefit in multiple survival outcomes (Table 6 and 7, preceding pages), particularly at the probability threshold of 0.90. Superiority of the Karakiewicz threshold remained, but reduced as the threshold was reduced. Similarly, LR analysis showed consistently higher adequacies for the Karakiewicz nomograms at the 0.85 and 0.80 discretizations for multiple survival outcomes, supporting superiority of the discretized Karakiewicz nomogram. A decision analytic method for threshold derivation yielded a similar approximate threshold (corresponding to predicted 5-year CSS of 0.89, or 99 points). While for all outcomes, the 0.85 and 0.80 discretization thresholds yielded improved net benefit relative to the SORCE criteria over a higher range of threshold probabilities, the SORCE criteria showed benefit over the 0.85 and 0.80 Karakiewicz thresholds in a more limited range of lower threshold probabilities.

The conducted likelihood ratio analyses between the UISS and the Leibovich score may be presented graphically for a clearer depiction of the relative adequacies and the analyses used to derive significance when testing nested models against full models (Figure 6, following page).

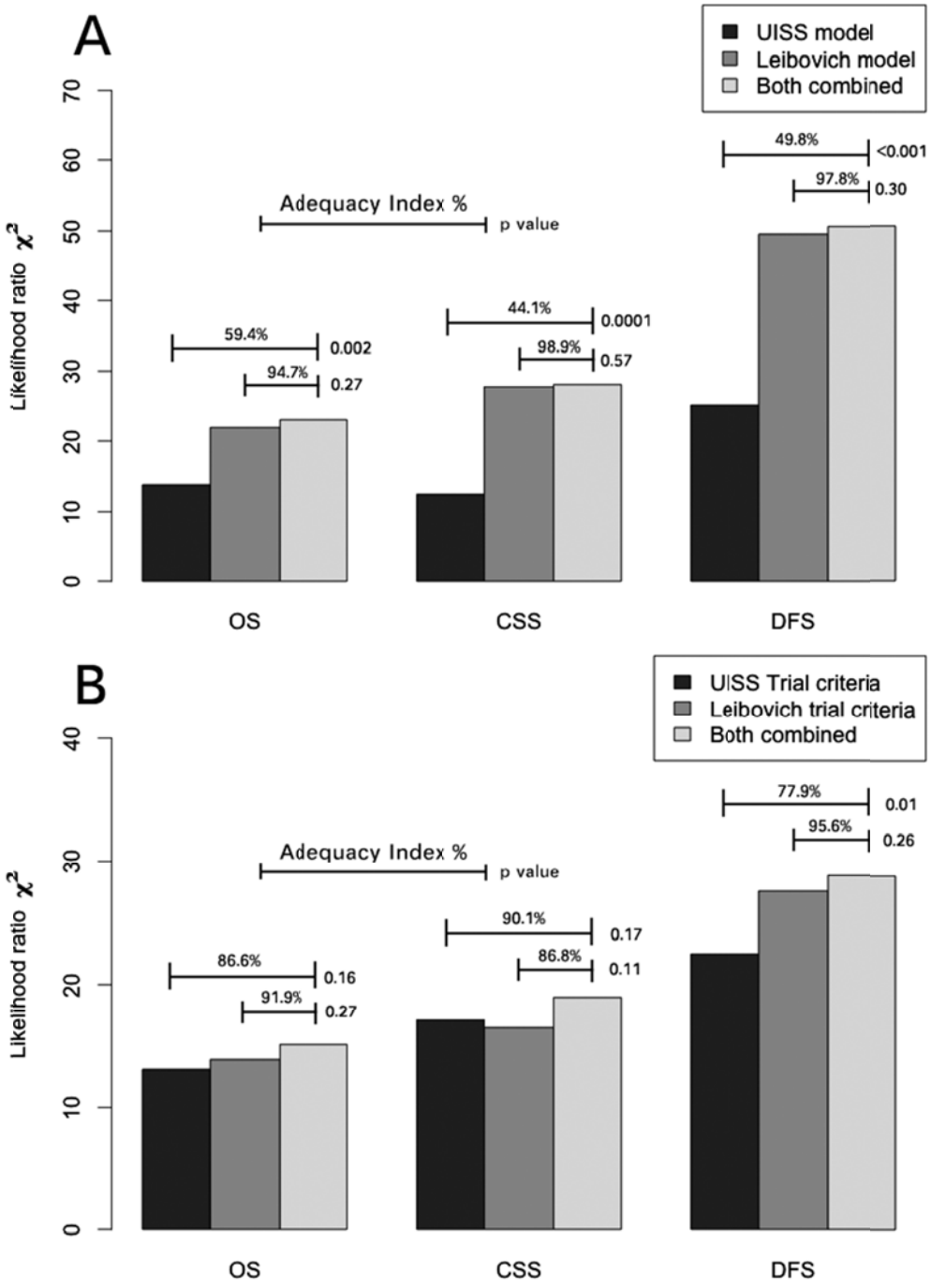

Figure 6: Predictive values for the models and the trial criteria for the clear cell RCC dataset

(A) the UISS and the Leibovich models and (B) the corresponding trial criteria adapted from each model. Models of OS, CSS and DFS are presented here. The predictive value is represented by the model LR and adequacy index (% likelihood ratio χ^2 attributable to the nested model, which does include both UISS and Leibovich scores). For example, in A, the Leibovich model is superior to the UISS model in predicting cancer specific survival in terms of adequacy (98.9% vs 44.1%).

Similarly, a relative comparison between the Karakiewicz nomogram and the Kattan nomograms (Figure 7) can also be presented. We showed that the addition of the Karakiewicz nomogram to the Kattan nomogram improved the model significantly at all selected cut-off thresholds and for all outcomes, but there was no significant difference in adding the Kattan nomogram to the Karakiewicz nomogram (Figure 7). This is true when we tested the models as semi continuous scores as well as a nomogram classification. Analysis II first demonstrated that the Karakiewicz nomogram was superior to the Kattan nomogram at multiple thresholds in making binary predictions of outcome. We report that the use of prespecified cutoffs (0.9, 0.85, and 0.8) for both the Kattan nomogram and Karakiewicz nomogram allow us to report superior outcomes for the Karakiewicz nomogram. Low risk (LR) nomogram patients are reported as patients with scores below the threshold, and high risk (HR) patients are patients with scores equivalent or exceeding the threshold.

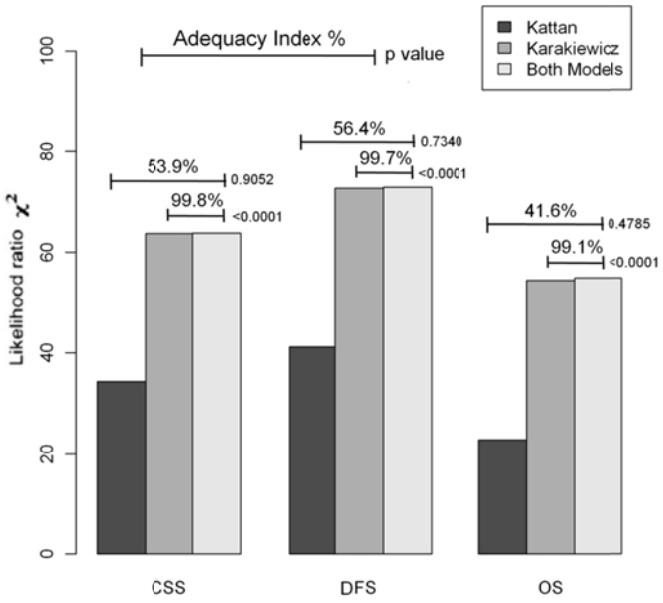

Figure 7: Comparison of the Kattan and Karakiewicz nomograms
Models of OS, CSS and DFS are presented here. The predictive value is represented by the model likelihood ratio and adequacy index (% likelihood ratio χ^2 attributable to the nested model. For example,, the Karakiewicz is superior to the Kattan model in predicting cancer specific survival in terms of adequacy (99.8% vs 53.9%).

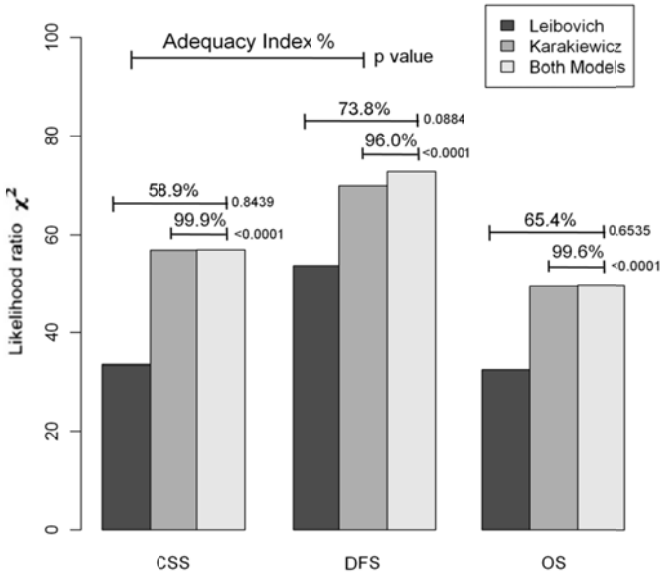

Figure 8 : Comparison of the Leibovich trial criteria and Karakiewicz nomogram
Models of OS, CSS and DFS are presented here. The predictive value is represented by the model likelihood ratio and adequacy index (% likelihood ratio χ^2 attributable to the nested model. For example, in A, the Karakiewciz model is superior in predicting cancer specific survival relative to the Leibovich model in terms of adequacy (99.9% vs 58.9%)

It should be also noted that in addition to the excellent performances on likelihood ratio analysis, the Karakiewicz nomogram had higher concordance indices as well relative to the other models, supporting the idea that the Karakiewicz model had better overall performance.

In order to establish an optimal approach towards a dichotomous classification of RCC, as well as to benchmark the performance of a nomogram against a risk model, we compared a discretized semi-continuous version of the superior nomogram, the Karakiewicz nomogram classification, with the Leibovich trial criteria(*Schuetz et al.* 2005) (Figure 8). We reported in the first part of our study (Analysis I) comparing the Leibovich score and the

UISS criteria that the Leibovich score has superior outcomes relative to the UISS score. It is currently used by the UK MRC SORCE trial for recruiting intermediate and high-risk patients for randomization between placebo and sorafenib. Log likelihood tests showed that the addition of Karakiewicz nomogram classification to the model including Leibovich trial criteria improved the model significantly in terms of all outcomes, while addition of the Leibovich trial criteria to the Karakiewicz nomogram classification did not yield any benefits. This is true for both the semi-continuous scoring as well as the nomogram classification.

To evaluate the question of whether ethnicity may have an impact on performance of the models, we conducted similar analyses within Chinese subjects (comprising approximately 82% of each cohort). The number of Malay and Indian subjects was too small for a meaningful validation. We performed comparisons of the UISS and Leibovich score, as well as comparisons of the Karakiewicz nomogram against the Kattan model and the Leibovich model, and report that the results are consistent with our findings above, where the Leibovich score showed superiority to the UISS model (OS: concordance index 0.66 vs 0.62; CSS: 0.72 vs 0.64; DFS 0.68 vs 0.63) and the Karakiewicz nomogram was superior to the Kattan nomogram (OS: concordance index 0.75 vs 0.66; CSS: 0.82 vs 0.72; FFR: 0.79 vs 0.71). A narrow advantage of the Karakiewicz nomogram over the Leibovich score based on this analysis is also reported (OS: concordance index 0.77 vs 0.76; CSS: 0.82 vs 0.81 and FFR: 0.79 vs 0.79)

DISCUSSION

LEIBOVICH AND UISS MODELS

A number of prognostic models have been developed to improve survival prediction in patients with RCC. A recent systematic review found 11 different models proposed for this purpose(*Galfano et al.* 2008), and the UISS and the Leibovich models are two such risk model based models. Over the last decade, several prognostic models have been developed to predict survival outcomes in RCC patients. Many investigators have generally preferred to develop new models on their own data-sets, rather than compare pre-existing models. There are additional models such as nomograms (the Kattan and the Karakiewicz nomogram), but these are less established in trial design. Both of the risk models, the UISS and Leibovich models, have been incorporated for patient selection in large trials of adjuvant therapy in RCC.

Our results show that both the UISS and the Leibovich scores provide excellent estimates of various survival outcomes in our single institution series of patients with non-metastatic clear cell RCC, and that the Leibovich score is superior to the UISS score in its predictions. Ours represents the first such comparison of these two competing scoring systems, which is of particular relevance to trial design.

A previous European study in 342 patients with non-metastatic clear cell RCC suggested that the SSIGN model was superior to the UISS model

by a comparison of an AUC of 0.830 (SSIGN) against an AUC 0.760 (UISS) in patients with non-metastatic RCC (*Cindolo et al.* 2005). In this study, no direct statistical testing for superiority was conducted, and it was limited to the prediction of cancer-specific survival. While both the SSIGN and the Leibovich models use the same pathologic features for scoring, the scoring system is distinct (SSIGN: 0 – 16, Leibovich score: 0 – 11, tiered into 3 risk strata). The SSIGN is designed for estimation of cancer-specific survival, but the Leibovich model is designed for estimation of disease-free survival. The urgency in direct comparison of these models has arisen from the use of the UISS and the Leibovich scores in the Phase III adjuvant trials.

We have reported here that the Leibovich model is superior to the UISS model in estimation of all tested survival outcomes in our series, a finding of consequence for pharmaceutical trial design. As mentioned, the UISS model was developed to predict overall survival and the Leibovich model was developed to predict cancer-specific survival. This conclusion was derived using a likelihood ratio approach. While rank concordance methods(*Harrell et al.* 1996) have been used to compare the various RCC prognostic models, a direct comparison can be accomplished using a likelihood ratio-based approach, which more powerful than rank-concordance methods (*Harrell* 2001). This approach has not been previously used, and we present it here together with analysis based on rank concordance methods to facilitate comparisons.

The reasons why the Leibovich model may be superior to the UISS model in predicting outcomes in non-metastatic patients are not immediately clear. However, the UISS model was developed using an automated stepwise modeling risk model comprising RCC tumours of all subtypes. It is recognized that certain pathologic variables such as tumour necrosis(*Sengupta et al.* 2005) and lymph node involvement may vary in prognostic importance depending on tumour subtype. Thus, additional variables of prognostic importance may not have been eventually selected in the UISS model through this stepwise approach.

There is ongoing research in the integration of biomarkers into these clinico-pathologic models to enhance prediction of outcomes(*Parker et al.* 2002; *Parker et al.* 2009). However, in the absence of definitive data supporting adjuvant therapy in RCC, these promising results have not entered routine practice in estimating survival outcomes for non-metastatic RCC.

NOMOGRAMS AND RISK MODELS

Although the Leibovich model demonstrated superiority over the UISS model in predicting outcomes, it is reasonable to evaluate whether other clinical models may yield superior outcomes. For example, multiple studies have documented the superior performance of nomograms when compared to risk groups using rank-concordance methods(*Atkin and Jackson* 1996; *Belizaire et al.* 2004; *Brown et al.* 1994; *Yasui et al.* 1999; *Zhao et al.* 2007). Following the establishment of the Leibovich model as the superior risk model

above, we proceeded to compare several nomograms against the Leibovich model, showing that the Karakiewicz nomogram is the most useful clinical predictor of survival outcomes in our data-set relative to other nomograms and the Leibovich model, with superior accuracy in terms of LR analyses and concordance indices. While the Karakiewicz nomogram was designed to test for CSS and the Sorbellini and Kattan nomograms for FFR, the Karakiewicz nomogram discriminated better in predicting CSS, FFR and OS than the other nomograms and the Leibovich model. Thus, in terms of individual counseling, the Karakiewicz nomogram is likely to be more useful that the other models.

For nomogram comparisons, it should be noted that the Kattan nomogram and Karakiewicz nomogram include similar variables like T stage, tumour size and symptoms but they assign different weights to each variable. We report that the Karakiewicz nomogram is superior to the Kattan nomogram in estimation of all test survival outcomes in our series. We would like to point out that while the Karakiewicz nomogram was designed to test for CSS and the Kattan nomogram for RFS, the Karakiewicz nomogram performed better in predicting CSS, RFS and OS.

We have used a wide array of tools for comparing the various models, including comparisons of discrimination (likelihood, adequacy, concordance indices) as well as calibration analyses. In addition to these comparisons, we have used decision curve analysis to show that the Karakiewicz nomogram is consistently superior to the other nomograms across a range of clinical threshold probabilities, with equivalent clinical benefit as the Leibovich score.

Our work provides a framework for considering the choice of model in an adjuvant therapy setting in RCC, best highlighted by the clinical net benefit analysis. It is critical to note that evaluation of clinical net benefit will evolve with information on benefits and toxicities of adjuvant therapy. Practically, the selection of a cut-off in a clinical trial is additionally driven by trade-offs between event rate and accrual rate(*Kattan* 2010). We considered it interesting that the best performing cutoff for our discretization (estimated 5-year CSS 0.90) coincided with one derived by decision analysis. Nonetheless, we caution here that although the Karakiewicz nomogram at all tested discretizations compares favorably to the SORCE trial criteria, prespecified or data-derived thresholds should be considered exploratory, and external validation of such thresholds are required for definitive conclusions on discretization.

Other than improved individualized prediction, the immediate real-world implications of our findings would be a more efficient accrual of patients for the ongoing adjuvant therapy RCC trials. For example, of practical real-world interest is the finding that if the discretized Karakiewicz nomogram, rather than the standard trial criteria using the Leibovich score, had been used to select patients in our centre for the SORCE trial, 151 patients instead of 187 would have been recruited for 47 events in terms of cancer-specific survival. Other than clinical trials, the use of the Karakiewicz nomogram as a standard in biomarker research will allow for the development of more

rigorous biomarkers when the strongest available clinical model is incorporated into multivariable analysis.

Overall, it is important to note that we have focused on post-operative nomograms rather than pre-operative nomograms, as post-operative models have yielded better predictions than pre-operative models(*Cindolo et al.* 2005).

As relatively unusual methods have been used to compare the models, we wish to discuss aspects of the model comparisons here. Although rank concordance methods(*Harrell et al.* 1996) have been used to compare the various RCC prognostic models, a direct comparison can be accomplished using a likelihood ratio-based approach, which is more powerful than rank-concordance methods (*Harrell* 2001). This approach is also useful to address the issue of multicollinearity. In developing models with two highly correlated variables, as clearly occurring in our approach, multicollinearity does affect individual predictors, but does not reduce the predictive power or reliability of the full model, allowing for reliable calculations of likelihood ratios in comparing full versus nested models.

THRESHOLD DETERMINATION

There have been several studies suggesting the superior performance of nomograms (with multiple thresholds) compared to risk models using rank-concordance methods (*Galfano et al.* 2008; *Haas and Uzzo* 2008; *Karakiewicz and Hutterer* 2007; *Liu et al.* 2009; *Shariat et al.* 2008). Nomograms are stronger predictors of survival than risk models because

nomograms assign points to continuous variables on a continuous scale, maximizing their predictive power by reducing spectrum bias. Risk models group patients with similar (though not identical) characteristics into risk groups, resulting in heterogeneity within the risk groups that reduces the predictive power of the risk model (spectrum bias). However, for purposes of a trial, categorization using a threshold is inevitable. A good analogy is that while a speedometer is informative, a limit is still needed whereupon exceeded, a policeman will issue a ticket. For example, the CALGB 90203 trial (*Eastham et al.* 2003) recruited prostate cancer patients with a nomogram-predicted probability of ≤60% of remaining free from disease recurrence.

The choice of an optimal cut off value for the Karakiewicz nomogram is selected at 0.9 at 5 years (corresponding to a nomogram score of 94) in our data set, which is particularly useful for stratifying patients into 2 groups. While a binary classification is susceptible to the usual concerns about categorization of semi-continuous data, it has the advantage of being simple and direct for trial recruitment criteria. The use of nomograms for trial recruitment criteria has not been done before in RCC trials, possibly because there are no established cut offs for practical use, and the degree of handicap in reducing nomograms from a continuous to a binary prediction is uncertain. Our study indicates that with careful selection of a threshold, a nomogram classification can be superior to the conventional risk models in predicting survival outcomes in RCC. We address the issue of loss of information by

categorization by providing a full analysis of Karakiewicz nomogram versus the other semi-continuous scores, showing that it also performs better in the absence of categorization.

Hence, not only have we demonstrated superiority of the Karakiewicz nomogram as a continuous predictor, our results also establish that the Karakiewicz nomogram is able to predict survival better than the usual risk modelic approach even as a binary predictor, outperforming the Leibovich trial criteria that is currently used in clinical trials of adjuvant therapy.

LIMITATIONS

We discuss the epidemiologic issues surrounding our clinical modeling work here. Given that the work is essentially comprised of retrospective, single institution studies drawn from a hospital-based case series, this can raise the usual and valid concerns about selection bias and generalization of results to other populations. Given that the patients in our dataset were selected over a more recent period than either the UISS or the Leibovich score, stage migration may account for some of the improved outcomes in our series relative to the Western centres (*Chow et al.* 1999). However, the excellent calibration demonstrated generally suggests that any issue of generalization is likely of relatively minor concern.

Nevertheless, it is important to note that the settings where these scores were derived varied as well, and with the exception of the Karakiewicz nomogram, were derived from data from single centres of excellence in the

United States. The Karakiewicz nomogram was derived from 2,530 patients in a range of ten academic medical centres in Europe. We agree that our findings need to be replicated in other populations. While population-based data, such as the Surveillance, Epidemiology and End-Results (SEER) registries, would be most appropriate for evaluation and comparison of nomograms, the absence of systematic collection of all evaluated variables precludes this possibility.

The issue of ethnic differences and whether results based on our Asian population can be applicable for the Western population is a valid consideration. This may be partially addressed by my subsequent work in high throughput gene expression profiling (see Molecular Models, pg 62 onwards) indicating that a prognostic genetic signature in clear cell RCC yielded excellent and similar predictions in both Japanese as well as American populations, suggesting that clinical models derived in the West are generally applicable to Asian populations. This is supported by the observation of excellent calibration of the Western models in our Asian dataset. It should be noted that our analyses of the predictions within the Chinese ethnic group alone (comprising approximately 80%) of the dataset demonstrated results similar to that derived from analyzing the full dataset. A future expanded number of members of the other ethnic groups will permit analysis of each racial group, to evaluate whether these models perform in a similar fashion in each racial group. It is not possible to perform similar

analyses in the current dataset due to the relatively small numbers of Malay and Indian patients.

We describe the limitations of our comparative effectiveness study here. The nature of our study as a retrospective, single institution study raises the usual concern about generalization of results to other populations. There are significant baseline differences between our data-set and the Mayo and UCLA data-sets. It is unclear if these differences may have asymmetrically affected predictor generalizability and resulting performance, but individually assessed, both predictors performed well in our data-set. Generally our data-set has the highest proportion of better-risk patients. In particular, the UISS data-set had a remarkably high proportion of high-risk patients (32%) even after having excluded patients with regional lymph node metastasis. The Mayo dataset of 1,671 patients included 66 patients with lymph node-positive disease; this group was excluded in our data-set to allow for direct comparison. Multiple explanations are possible for these baseline differences. Selection bias may account in part for this : the Mayo Clinic and UCLA are major US referral centers, which would be expected to care for a higher proportion of patients with advanced disease. Given that the patients in our data-set being selected over the most recent period, stage-migration may account in part for this difference (*Chow et al.* 1999). However, we do note that the SSIGN, which shares variables with the Leibovich score albeit with a different weighting method, has been shown to predict outcomes better than the UISS score in a European population. We did not study nomogram-

based models, since such models are not currently integrated into trial designs, although they do provide useful predictions for individual patients.

VALIDITY

The validity of studies is usually distinguished in two components – the inferences as they pertain to the members of the source population (internal validity) and people outside the source population (external validity). Our studies provide an overview of the critical nature of validity, with a wide range of techniques being applied to multiple different data-sets and tissue samples from a broad range of centres (*Larkin et al.* 2005; *Quackenbush* 2006b). In terms of validity, the Singapore dataset represents an external test set to evaluate the value of the multiple risk models and nomograms that have been generated to predict a range of outcomes.

It should be noted that the use of adequacy indices may be misleading in the situation when both scores are poor predictors of survival; its use would lead to the potential conclusion that one model is not poorly performing relative to the model to both scores. However, this may not apply in our dataset, as each model individually has been demonstrated to yield excellent predictions of outcomes.

COLINEARITY

Given that we are comparing scores that have a degree of correlation, colinearity is a potential issue. Hence, a full model that incorporates both scores (as we do here) will definitely result in colinearity when estimating the

standard errors of individual beta weights. However, calculation of deviance (which likelihood ratio testing is based on) is less affected by colinearity, and hence permits us to compare full relative to nested models. It should be noted that the error of individual beta weights of individual prognostic scores in the full models will be high, but that we are not utilizing the coefficients.

MOLECULAR MODELS IN RENAL CELL CARCINOMA

BACKGROUND

We have described extensively the use of clinical modeling to estimate outcomes for RCC in our initial analysis. These methods, while useful, may benefit from additional evaluation of additional biological variables, and the same studies have provide additional biological insights. These inevitable limitations result in heterogeneity within risk groups, as well as the provision of probabilistic estimates for individuals by nomograms. Even while accurate, a probabilistic estimate is not completely satisfactory to clinicians and patients, since each individual has a specific individual outcome. Hence, additional biomarkers are of high interest in improving these estimates. These biomarkers may include genetic, epigenetic and molecular variables of interest. Recent advances in materials science, bioengineering and information technology have permitted the development of high throughput methods for determination of these profiles, and we focus here on the use of high throughput expression profiling in selecting candidate signatures and biomarkers for RCC. Generally, the use of these signatures and markers are important for determining cancer classes, using unsupervised methods such as class discovery, and correlating these outcomes with clinically relevant endpoints such as survival, relapse, genetic alterations as well as drug resistance.

HIGH THROUGHPUT EXPRESSION PROFILING

We used high throughput methods, primarily gene expression profiling, a promising adjunct for cancer diagnosis and prognosis(*Tan et al.* 2004b), to identify prognostic predictors and genetic programs in tumour tissue (*Quackenbush* 2006b). These techniques, usually known as microarrays, allow for the concurrent collection of expression data of multiple transcripts from a single sample. These techniques allow for the measurement of expression data in a wide range of biological processes. With the widespread use of this technique in recent years, the explosion of data has resulted in the creation of large data repositories for archival and open access of this data, such as the Gene Expression Omnibus (GEO)(*Sayers et al.* 2010).

Microarray experiments are associated with their own challenges, and these primarily centre on variation (*Abdullah-Sayani et al.* 2006). Data variation may occur as a result of technical (extraction, labelling, hybridization), measurement (fluorescent signal) or biological variation. Biological replication is the most important, but is not always possible due to the high costs of microarrays. Broadly, the general techniques of preparation for microarray analysis involve extraction of RNA (total or mRNA), purification, reverse transcription, labeling and hybridization. There are two major types of expression microarray platforms in use : oligonucleotide based and complementary DNA (cDNA)-based.

The oligonucleotide array class consists of usually smaller probes, ranging from 25 to 80 bp in length. These may be presynthesized and spotted

directly onto the substrate, or synthesized directly on the substrate. This platform is most commonly used in the commercial Affymetrix chips, which depend on the spot intensity measures of oligonucleotide probes hybridized to the array chip. Following hybridization, the array is scanned, producing an image, and the relative fluorescence of each spot is ascertained. Pixel intensity is measured using software, and the intensity of each spot is normalized, allowing for comparison within and between arrays.

cDNA arrays comprise PCR products spotted onto the array, enabling clone banks and DNA from limited templates to be spotted. The size of a PCR amplified fragment is in the order of 400 – 1000 bp. The spotted microarray is hybridised with probes derived from the mRNA of the biological samples being assessed. In the technique known as dye swapping, in which multiple extracts are hybridised to arrays, the mRNA is typically reverse transcribed into cDNA and labelled with a spectrally distinguishable red (Cy5) or green (Cy3) fluorescent dye. Samples are then hybridized on the microarray, allowing labelled cDNA strands complementary to sequences on the microarray to bind. Generally two dyes are used; if only one dye is used there is little measure of the amount of DNA targeted to any particular spot. However, the relative fluorescence of two dyes to each other can be measured, with the sample containing higher levels of transcript producing a greater signal.

After transformation of the raw data into a gene expression matrix, data is analysed using various packages of R and Bioconductor (*Gentleman et al.* 2004). These software packages permit the analysis of microarray experiments using a variety of tools for high dimensional data(*Quackenbush* 2006a).

RCC EXPRESSION PROFILING

Studies have demonstrated that the various subtypes of RCC are readily distinguishable with gene expression profiling (*Higgins et al.* 2003; *Takahashi et al.* 2003; *Yamazaki et al.* 2003; *Young et al.* 2001). These results support the intuitive hypothesis that each subtype has its own individual biological features, clinical behavior, and, by extension, unique sensitivity to therapy. Although the discriminatory ability of gene expression profiling makes it potentially an excellent diagnostic tool, from a practical perspective the technology is currently not readily accessible to many pathologists in clinical practice. In RCC pathology, microarrays have already been instrumental in discovering new immunohistochemical markers for distinguishing the different subtypes. For example, Yamazaki *et al.* (*Yamazaki et al.* 2003) profiled several histologic subtypes of RCC and identified c-kit as being up-regulated in chromophobe RCC, and our group identified the following potential markers for the different histologic subtypes: glutathione *S*-transferase α for clear cell RCC, α-methylacyl racemase for papillary RCC, carbonic anhydrase II for chromophobe RCC, and K19 for transitional cell

carcinoma (*Takahashi et al.* 2003). These studies represent a direct effort to enhance the practice of pathology through the use of microarray technology.

Gene expression profiling is not limited to tumour tissue. Interestingly, Twine *et al.* (*Twine et al.* 2003) profiled peripheral blood mononuclear cells (PBMCs), instead of tumour tissue, from RCC patients and healthy volunteers and demonstrated that these expression profiles could distinguish the PBMCs of RCC patients from those of healthy volunteers. Although the biological significance of the identified discriminatory genes is uncertain, these preliminary findings are of considerable interest, because the authors indicate that their ongoing studies suggest that expression profiling of PBMCs can distinguish RCC patients from patients with other types of solid tumours. It is possible that additional work may establish the presence of disease-specific gene sets in PBMCs.

Refining prognostic systems to more accurately predict patient outcomes and thereby guide more effective treatment decisions is an ongoing process. To date, key prognostic factors identified include TNM staging, tumour grade, functional status, and various biochemical assessments. Integrated prognostic systems have been developed by several groups combining clinical and pathological data to better stratify patients and improve prognostic power(*Gettman et al.* 2001; *Motzer et al.* 1999; *Zisman et al.* 2002). Further integration of molecular markers defined by expression and proteomic profiling into these prognostic systems is likely to further increase prediction accuracy.

Our group initially reported a pilot study in which 29 specimens of clear cell RCC with patient-matched normal tissue were profiled with cDNA microarrays(*Takahashi et al.* 2001). In this study, unsupervised clustering demonstrated two subsets of tumours, with clear segregation by cause-specific survival at 5 years. Approximately 40 genes were identified that discriminated between these two groups. These results suggest that there are two distinct groups of clear cell RCC that vary in aggressiveness. A major limitation of this study was the relatively small number of cases, resulting in the inability to externally validate the results. However, the prediction accuracy of 5-year survival by using microarrays exceeded that of staging, and accurate cross-validated predictions were obtained for patients with clinically indolent metastatic RCCs and clinically aggressive localized RCCs, which suggested that the prognostic signature was not confounded by metastasis. In view of the relatively small numbers of this study, we believe that this study should be considered preliminary, and we have sought to expand this study here. In particular, there is a goal of identifying high-risk patients with localized RCC for early systemic therapy. It is plausible that expression profiling will actually result in individually tailored therapeutic regimens in the near future.

MICROARRAY PLATFORM

The Affymetrix GeneChip platform which is used primarily here is an high-density oligonucleotide expression array platform. A unique gene is

represented by multiple such probe sets on the Affymetrix chip, with probe pairs comprising perfect match (PM) and mismatch (MM) probes. The PM probes are designed to specifically complement a 25 bp sequence in the target gene; the MM probes are identical to the PM probes save for a single base pair change in the centre position. Mismatches are supposed to represent non-specific hybridization. As is evident, to analyze GeneChip data, the probe-level preprocessing needs to be considered carefully. The four steps include background correction (subtraction of nonspecific background), normalization (reduction of non-biological signal and equalization of intensities across chips), PM correction (subtraction of non-specific hybridization) and summarization. Multiple methods exist for summarization, including the Affymetrix Microarray Suite (MAS 5.0) (*Hubbell et al.* 2002). Several approaches have been proposed to improve this conventional method, including Model Based Expression Index (MBEI) (*Li and Wong* 2001) and the Robust Multi-array Average (RMA) (*Irizarry et al.* 2003). Although RMA is used most commonly in this thesis, there are other methods developed such as Factor Analysis for Robust Microarray Summarization (FARMS) (*Hochreiter et al.* 2006). Each method takes a different approach towards summarization, but the goals are similar – to generate improved methods of gene expression evaluation while accounting mathematically for non-biological flaws and bias in data acquisition.

SIGNIFICANCE ANALYSIS OF MICROARRAYS

Advanced statistical methods are commonly used in this thesis. The method best known as SAM, or Significance Analysis of Microarrays (*Tusher et al.* 2001), was developed at Stanford for the identification of genes with statistically significant changes in expression, by evaluating a set of gene-specific *t* tests. It is crucial to note that while cluster, or unsupervised analysis may identify specific classes of gene expression, identification of corresponding discriminating genes may be challenged by the large number of transcripts evaluated, resulting in a distinct false-discovery rate (FDR). SAM estimates FDR by first assigning each gene a score based on its change in gene expression relative to the standard deviation of repeated measurements for that gene. The FDR is estimated by analyzing permutations of the measurements, and the threshold for this can be adjusted to calculate varying FDRs for each set. This has a distinct advantage over the Bonferroni method, which assumes independence of the different tests, which is clearly an invalid assumption resulting in an overly conservative estimate. Other methods outside those discussed here include the step-down correction method, which permits dependent tests (*Westfall et al.* 2002). For this thesis, SAM was selected for supervised analysis primarily because of availability and ease; alternatives such as LIMMA(*Smyth* 2004) (Linear Models for Microarray Data) are certainly possible alternatives. LIMMA is also implemented in R, and is used to create and test linear models for microarray data. A moderated t-statistic is used to evaluate the average difference in log

expression levels for each gene; the t-statistic is generated by the average log ratio divided by a derived standard error. Several multiple comparison procedures may be used to control for the resulting FDR.

SPECIFIC AIMS (MOLECULAR MODELS)

To evaluate the molecular profiles of three primary subtypes of RCC – clear cell RCC, papillary RCC and chromophobe RCC in an unbiased fashion using high-throughput gene expression profiling technology to obtain validated clinical and mechanistic insights into diagnosis, prognosis and pathogenesis. The techniques used here to supplement the high throughput expression profiling include regional gene expression bias determination, high throughput single-nucleotide polymorphism profiling and protein expression, using immunohistochemistry. This would be in the context of correlating a molecular phenotype with clinical outcomes. These outcomes include histological discrimination / recognition, relapse and most importantly survival and prognosis.

CLEAR CELL RENAL CELL CARCINOMA

Considering the pathological and clinical issues for clear cell RCC discussed in Overall Background and Background (Molecular Models), our goal was to use gene expression profiling as a promising adjunct for cancer diagnosis and prognosis to identify prognostic predictors and gene expression programs during both primary and relapsed stages of RCC.

METHODS

SUBJECTS AND STUDY DESIGN

A total of 92 primary tumour and 25 metastatic clear cell RCC tissue specimens from independent patients operated on in 8 centers (6 in USA, 2 in Japan) between 1989 and 2002 were obtained. Japanese centres were University of Tokushima and University of Kitasato. US Centres included University of California Los Angeles, Spectrum Health Grand Rapids, Detroit Wayne State University, Johns Hopkins University, University of Chicago, and Baylor College. Primary tumour tissue samples were acquired from patients diagnosed with sporadic clear cell RCC with at least two years' follow-up following nephrectomy as available. Metastatic tumour tissue samples were acquired from patients with a history of localized clear cell RCC, who had relapsed with metastases after primary nephrectomy as available. These samples represented contributions from institutional tissue banks to the Van Andel Research Institute. 12 non-neoplastic kidney samples from unrelated patients were also obtained as controls for regional expression bias analysis;

these are the same 12 controls for regional expression bias used for the other microarray studies in other subtypes in this thesis.

Written informed consent for analysis of clinical samples was obtained from all patients, and all IRB boards of participating institutions approved the study. Tumour tissue was flash frozen in liquid nitrogen immediately after nephrectomy and stored at –80°C. Portions of the tumours were fixed in buffered formalin. Each sample was confirmed by pathologic analysis and anonymized prior to the study. A portion of the tumour sample was frozen in liquid nitrogen immediately after surgery and stored at –80˚C. Total RNA was isolated from the frozen tissues using Trizol reagent (Invitrogen, Carlsbad, CA). This was subsequently purified with a RNeasy kit (Qiagen, CA), and quality was assessed on denaturing gel electrophoresis. Representative tissue sections of RCC were cut and stained with hematoxylin-eosin for confirmation of histological diagnosis and confirmation of tumour tissue content (>70%). All samples were examined by a central expert uropathologist. Clinicopathologic information was derived by review of pathologic, radiologic, and case notes by individual clinicians. Tumours with sarcomatoid change were classified as grade 4 tumours. Tumour size was defined as the maximum tumour dimension on direct pathological measurement. Patient follow-up status was assessed by directly contacting patients where possible, or at routine clinical follow-up.

An *a priori* external validation approach was used (Figure 4), where samples from Japan were designated as the training set (n=33) and those from the

USA as the test set (n=59). The predictor was derived by 2-means clustering followed by application of nearest shrunken centroids(*Tibshirani et al.* 2002) for optimal derivation. The choice of a 2-means clustering approach was determined by previous work reporting there to be two molecular classes for clear cell RCC (*Takahashi et al.* 2001). For supervised analysis and hypothesis generation, the 92 oligonucleotide profiles were divided into 2 predicted classes based on the prognostic gene predictor. 2260 discriminating transcripts, corresponding to a delta of 3.0 with a FDR of approximately 0.001%, was selected for pathway analysis. In order to test this predictor across array platforms, we employed the following strategy to assess coexpression across different microarray platforms and hybridization designs, since differing data structures preclude direct application of the oligonucleotide predictor to two-channel spotted cDNA data. Accordingly, we tested our predictor in the complete microarray dataset of a Stanford dataset of Swedish patients with clear cell RCC (GEO GSE3538) (n=177) (hereafter referred to as the Stanford dataset). (*Zhao et al.* 2006) We identified well-measured (<50% missing data) expression values of matching cDNA clones. K-nearest neighbour imputation (10 neighbours, largest block of 1,500 transcripts) was used to impute remaining missing values and hierarchical clustering was applied on the data. Imputation was required to avoid omission of samples. Of genes overlapping the oligonucleotide predictor and the Stanford dataset, only 2 of 20 genes in the 177 sample dataset had

sufficiently adequate data (<50% missing data) for analysis, with 55 and 26 missing values respectively

Figure 9: Flowchart for analysis of the gene expression profiles.
The training set, comprising 33 Japanese samples, was analyzed using unsupervised 2-means clustering, and a predictor generated for classification. This generated predictor was applied to external test sets A, B and C to evaluate performance within each dataset.

OLIGONUCLEOTIDE ARRAY PROFILING

For oligonucleotide array gene profiling, we extracted and purified total RNA from homogenized samples using Trizol reageant (Invitrogen, CA) followed by RNeasy columns (Qiagen, CA) according to manufacturers' recommendations. Oligonucleotide array gene profiling was performed using the manufacturer's recommended protocol (GeneChip Expression Analysis Technical Manual, April 2003, Affymetrix, CA). Approximately 5-20

micrograms of total RNA was used to prepare antisense biotinylated RNA. A subset of cases were spiked with external poly(A) RNA controls. Synthesis of single-stranged and double-stranded cDNA was done with the use of T7-oligo(dT) primer (Affymetrix). In vitro transcription was done using Enzo Bioarray Transcript Labeling kit (Enzo, Farmingdale, NY). The biotinylated cRNA was subsequently fragmented and 10 micrograms were hybridized to each array at 45°C ovr 16 hours. We used the HGU133 Plus 2.0 GeneChip, containing 54,675 probe sets, representing approximately 47,000 transcripts and variants. Scanning was done in a GeneChip 3000 scanner. Quality indices reviewed for all samples included mean percentage present, mean background, mean scaling factor and a mean GADPH 3'/5' ratio. All clinical and microarray data for published data has been uploaded to the Gene Expression Omnibus.

STATISTICAL ANALYSES

Statistical analyses were performed with R, utilizing packages from the Bioconductor project(*Gautier et al.* 2004; *Gentleman et al.* 2004) Packages used include *affy* (*Gautier et al.* 2004), *survival*, and *impute*. The default significance threshold was 0.05. For pre-processing of the Affymetrix data, the robust multichip average (RMA) algorithm was the default method used to perform background adjustment, quantile normalization and summarization. All survival analyses were performed by fitting the data to a Cox proportional-hazards model, and the end point of interest was disease-specific mortality. Overall survival time in all patient groups was defined as the time from initial

nephrectomy to the date of death or the last known date of contact with provider. Nagelkerke's R^2 values have been used to report goodness-of-fit, where a value of 1 represents a perfect fit. Where appropriate, bivariate survival analysis was performed, the utility of the predictor being assessed after adjusting for clinicopathologic parameters singly; this approach was selected rather than multivariate analysis based on available event numbers. Case deletion was used for missing data. Adjusted variables included age, gender, tumour stage, metastatic status, tumour size, tumour grade, Eastern Clinical Oncology Group (ECOG) performance status and UCLA Integrated Staging System (UISS) status. Significance levels in all cases were calculated with the likelihood ratio test.

Regional expression biases are genetic intervals where gene expression is coordinately changed, corresponding well with cytogenetic aberrations detected by comparative genomic hybridization. We inferred cytogenetic profiles for the tumours through the use of a refinement of the comparative genomic microarray analysis (CGMA) risk model, which predicts chromosomal alterations based on regional changes in expression. Briefly, for each measured gene, the gene expression value was normalized such that the average gene expression value in the nondiseased samples was subtracted from the tumour-derived gene expression value. Relative expression profiles (R) were generated from the single-channel tumour expression profiles (T) and the mean expression values of the 12 single-channel kidney cortical expression profiles (N), such that $R = \log_2(T) - \log_2(N)$.

The improved resolution-comparative genomic microarray analysis (IR-CGMA) risk model that we had previously reported was applied to achieve this. Chromosomal abnormalities was predicted as implemented in the *reb* package (*Furge et al.* 2005).

PATHWAY ANALYSIS

For pathway analysis, the Ingenuity Pathways database version 3.0 (Ingenuity Systems, Mountain View, CA) was used. A list of transcripts was generated using significance analysis of microarrays (SAM) based on a 2-class unpaired analysis with 1000 permutations.

IMMUNOHISTOCHEMISTRY

Immunohistochemistry (IHC) was performed using CD31. CD31 is a transmembrane protein expressed in vascular endothelium, and CD31 immunostaining is commonly used for assessment of microvessel density (MVD). 45 formalin-fixed paraffin-embedded tissue specimens from patients operated between 1993 and 2002 were obtained from 3 centers. Automated immunostaining was performed with the murine anti-human monoclonal antibody, clone JC/70A (M0823) (30 min incubation at room temperature). 1 slide was stained per specimen. We employed a modified strategy to maximize correlation between mRNA expression in bulk tissue and CD31 immunostaining. 9 fields (each 0.2596 mm^2 in size) at 200X magnification were captured (Spot Insight Camera, Nikon Eclipse E600) in a 3 x 3

arrangement with the widest possible distribution across each slide, avoiding areas of necrosis. MVD was calculated as the mean count over the 9 images. Interactive image analysis was performed using Cytometrix, an in-house program developed for quantitative analysis of histologic images. Immunostaining, image capture and analysis were performed independently by blinded individuals. A countable microvessel was defined as a stained endothelial cell cluster with or without a lumen distinct from other cell clusters. Contiguous or anastomosing structures were considered as one microvessel.

RESULTS AND DISCUSSION

Baseline characteristics of the patients of the Japanese and the US cohort are presented in Table 8 (following page)

Prognostic predictors were identified, externally validated and demonstrated to be independent of standard clinicopathologic parameters. We defined a prognostic oligonucleotide predictor (Table 9, pg 81) using unsupervised (2-means clustering) methods in Japanese patients (R^2 = 0.276, p=0.003) (Figure 10, pg 83), which successfully predicted survival in a separate test set comprising tumours from patients from 6 US centers (R^2 = 0.236, p=0.0001) (Figure 10, pg 83). Respective expression data is shown (Figure 11A-B, pg 85). Bivariate analysis in the test set demonstrated that the predictor remained significant after adjusting for standard clinicopathologic parameters, including the UCLA Integrated Staging System (UISS), an aggregate variable of stage, grade and functional status (Table 10, page 82).

The prognostic predictor demonstrated utility in patients with metastasis at relapse. We demonstrated utility for the predictor from the primary tumour data set in metastatic clear cell RCC samples resected from a variety of distant sites after relapse (R^2 = 0.228, p=0.03) (Figure 10, pg 83).

Table 8 : Individual patient demographic data for the clear cell RCC dataset

Country		Japan	USA
Number of samples		33	59
Age	Range	39 – 80	34 - 82
	Median	62	62
Gender – n (%)	Male	24 (73)	36 (61)
	Female	9 (27)	23 (39)
TNM stage – n (%)	1	15 (45)	20 (34)
	2	4 (12)	5 (8)
	3	7 (21)	19 (32)
	4	7 (21)	15 (25)
Tumour T stage – n (%)	1	17 (52)	22 (37)
	2	8 (24)	8 (14)
	3	8 (24)	28 (4)
	4	0 (0)	1 (2)
Tumour M stage – n (%)	0	26 (79)	44 (75)
	1	7 (21)	15 (25)
Tumour Grade – n (%)	1	13 (39)	2 (3)
	2	13 (39)	24 (41)
	3	7 (21)	24 (41)
	4	0 (0)	9 (15)
Tumour size – cm (%)	Range	2.5 – 17	1.4 - 19.5
	Median	5.2	6
	Unknown	2	3
ECOG PS – n (%)	0	20 (71)	28 (47)
	1	7 (25)	28 (47)
	2	1 (4)	3 (5)
	Unknown	5	0
UISS staging – n (%)	1	11 (39)	9 (15)
	2	11 (39)	26 (44)
	3	2 (7)	7 (12)
	4	4 (14)	16 (27)
	5	0 (0)	1 (2)
	Unknown	5	0
Post-nephrectomy treatment – n (%)	Palliative immunotherapy	5 (15)	12 (20)
Immunotherapy response	Complete response	0	0
	Partial response	0	2
	Stable disease	1	1
	Progressive disease	4	9
Patient status – n (%)	Deaths (cancer related)	9 (27)	22 (37)
	Deaths (other causes)	4 (12)	2 (3)
	Alive with disease	3 (9)	7 (12)
	No evidence disease	17 (51)	28 (47)
Follow-up duration(yr)	Mean	5.5	2.9
	Range	0.25 – 11.7	0.05 – 9.3

ECOG PS: Eastern Cooperative Oncology Group Performance Status
The total percentage count for each category may not total to 100 due to rounding.

Table 9 : Prognostic predictor of transcripts in clear cell RCC

Affymetrix Probe ID	Gene symbol	Gene title	Good prognosis score *	Poor prognosis score *
238178_at	EST	Transcribed locus	0.4703	-0.4997
238066_at	RBP7	retinol binding protein 7, cellular	0.2399	-0.2549
228434_at	BTNL9	butyrophilin-like 9	0.2397	-0.2547
217177_s_at	EST	CDNA FLJ13658 fis, clone PLACE1011567	0.1511	-0.1606
204368_at	SLCO2A1	solute carrier organic anion transporter family, member 2A1	0.147	-0.1561
227874_at	EST	MRNA; cDNA DKFZp586N0121	0.131	-0.1392
219436_s_at	EMCN	endomucin	0.1264	-0.1343
205478_at	PPP1R1A	protein phosphatase 1, regulatory (inhibitor) subunit 1A	-0.124	0.1317
219304_s_at	PDGFD	platelet derived growth factor D	0.1238	-0.1316
228575_at	FNDC6	fibronectin type III domain containing 6	-0.1133	0.1204
230250_at	PTPRB	Protein tyrosine phosphatase, receptor type, B	0.1086	-0.1153
204273_at	EDNRB	endothelin receptor type B	0.1035	-0.1099
238649_at	PITPNC1	phosphatidylinositol transfer protein, cytoplasmic 1	0.0899	-0.0956
205651_x_at	RAPGEF4	Rap guanine nucleotide exchange factor (GEF) 4	0.0896	-0.0952
204271_s_at	EDNRB	endothelin receptor type B	0.0894	-0.095
205357_s_at	AGTR1	angiotensin II receptor, type 1	0.0858	-0.0911
202242_at	TSPAN7	tetraspanin 7	0.0796	-0.0846
209047_at	AQP1	aquaporin 1 (Colton blood group)	0.0724	-0.0769
201150_s_at	TIMP3	TIMP metallopeptidase inhibitor 3	0.0718	-0.0763
203934_at	KDR	kinase insert domain receptor	0.06	-0.0638
221031_s_at	APOLD1	apolipoprotein L domain containing 1	0.0507	-0.0538
219315_s_at	C16orf30	chromosome 16 open reading frame 30	0.0502	-0.0534
230645_at	FRMD3	FERM domain containing 3	0.0449	-0.0477
224215_s_at	DLL1	delta-like 1 (Drosophila)	0.0428	-0.0455
209147_s_at	PPAP2A	phosphatidic acid phosphatase type 2A	0.0389	-0.0413
208096_s_at	COL21A1	collagen, type XXI, alpha 1	0.036	-0.0383
205846_at	PTPRB	protein tyrosine phosphatase, receptor type, B	0.0341	-0.0363
218723_s_at	RGC32	response gene to complement 32	0.0311	-0.0331
207542_s_at	AQP1	aquaporin 1 (Colton blood group)	0.0191	-0.0203
206701_x_at	EDNRB	endothelin receptor type B	0.0187	-0.0199

238169_at	EST	Transcribed locus	0.0141	-0.015
1560359_at	PELO	Pelota homolog (Drosophila)	0.0072	-0.0077
1555725_a_at	RGS5	regulator of G-protein signalling 5	0.0065	-0.0069
222717_at	SDPR	serum deprivation response	0.0053	-0.0056
205150_s_at	KIAA0644	KIAA0644 gene product	0.0045	-0.0048
212230_at	PPAP2B	phosphatidic acid phosphatase type 2B	0.0035	-0.0037
227372_s_at	BAIAP2L1	BAI1-associated protein 2-like 1	-9.00E-04	9.00E-04

* Class discriminant scores for each prognostic group

Table 10 : Univariate adjustment of survival predictor

Adjusted variable	Adjusted hazard ratio of predictor	95% confidence intervals	p-value
Unadjusted	6.25	2.10-18.6	
Age	6.44	2.15-19.3	0.0001
Sex	6.20	2.06-18.6	0.0002
Stage	3.28	1.01-10.6	0.03
Grade	4.81	1.57-14.7	0.002
Metastatic Status	3.95	1.22-12.7	0.004
Size (>4 cm)*	4.70	1.50-14.8	0.005
ECOG PS	6.48	2.15-19.5	0.0001
UISS	4.74	1.51-14.9	0.002

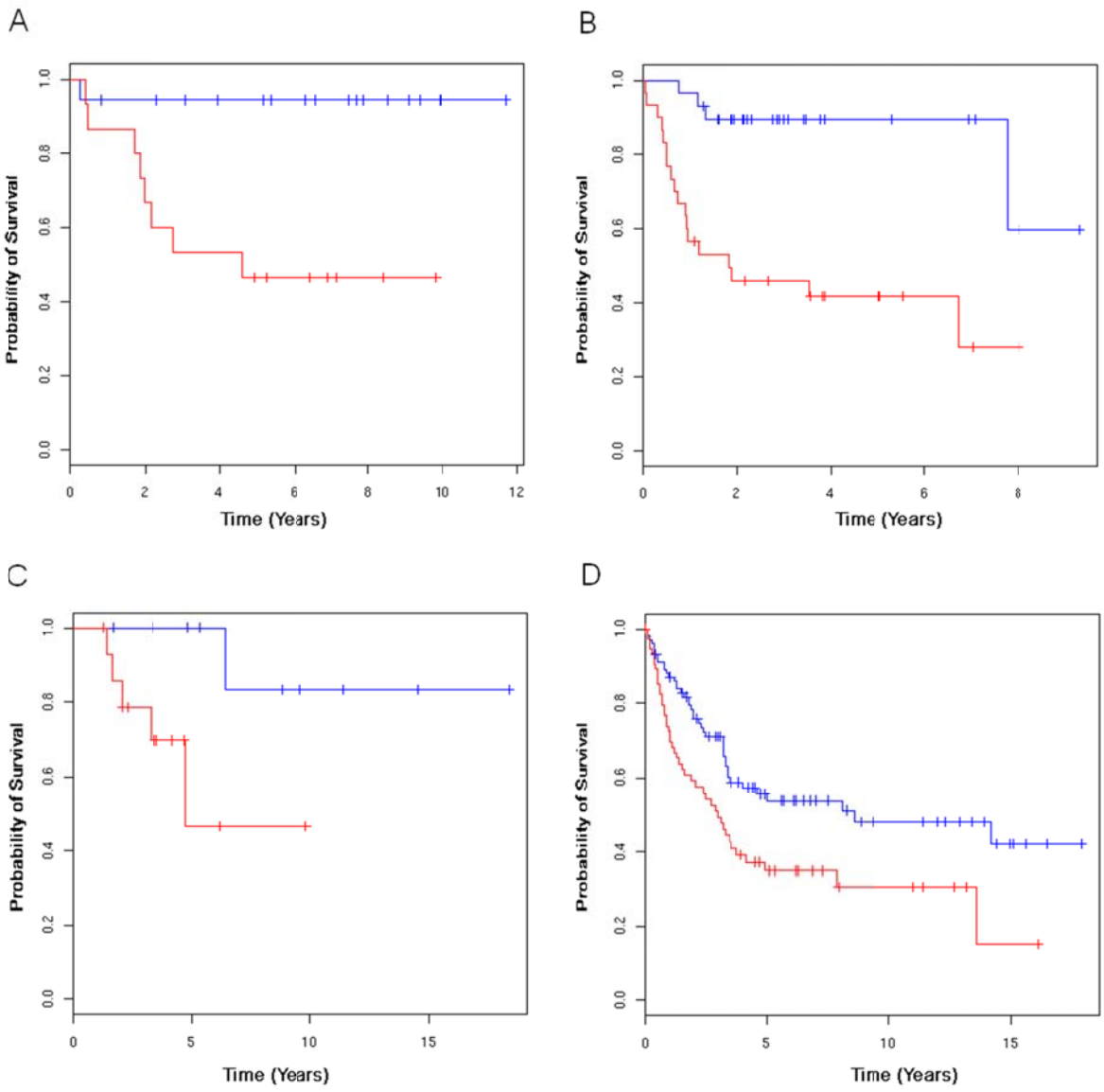

Figure 10: Predicted outcomes in the various training and test sets.
(A) survival of training set, comprising Japanese patients; (B) survival of test set A, comprising US patients; (C) survival of test set B, comprising samples of distant metastatic tumours, and (D) survival of test set C, comprising an external Swedish dataset. Lines in blue correspond to patients with predicted good-risk molecular class, and lines in red correspond to patients with predicted poor-risk molecular class.

Matching genes in the prognostic predictor demonstrated utility in an external dataset. Although only 2 of 37 transcripts in our oligonucleotide predictor were *a priori* determined as being of sufficient quality in the Stanford spotted cDNA dataset, being PPAP2B (T71976) and EDNRB (H28710) (Figure 11, following page), hierarchical clustering of the sample data with these 2 genes generated 2 distinct prognostic clusters (R^2 = 0.05, p = 0.005) (Figures 11, following page). Nonetheless, it is clear from the much reduced R^2 that using these 2 genes only would certainly result in a poorer prediction relative to a potential full transcript predictor.

Prognostic chromosomal aberrations were inferred from the expression profiles We inferred chromosomal aberrations associated with these distinct molecular subtypes of clear cell RCC with CGMA (Figure 11E, following page). This technique identified distinct regional expression biases corresponding to common cytogenetic abnormalities in clear cell RCC(*Kovacs and Brusa* 1988), such as deletions of chromosomes *3p*, *6q*, *9pq* and *14q*, as well as amplifications of chromosomes *1*, *3q*, *5q*, *8q* and *12*. These profiles were consistent with previous reports of associations between deletions of *9pq* and *14q* and poor prognosis(*Meloni-Ehrig* 2002). Other known cytogenetic aberrations not previously linked to survival, such as amplification of *1q*, *3q* and deletion of *6q*, were highlighted in our study as being of potential prognostic value.

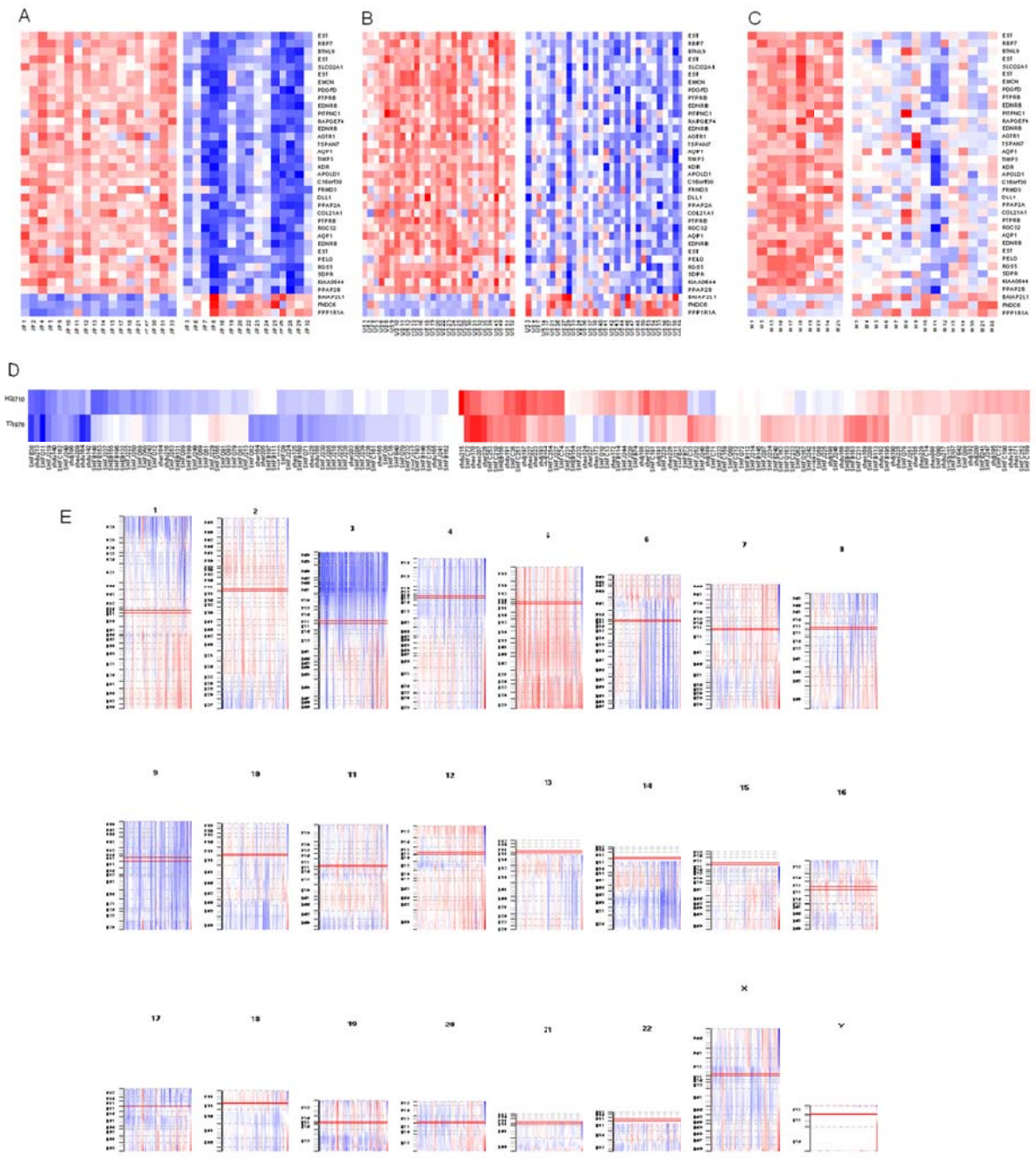

Figure 11 : Expression of gene predictor in the various data-sets by heatmaps. Red cells represent relatively upregulated transcripts, and blue cells represent relatively downregulated transcripts. Distinct expression profiles evident in (A) training set (B) test set and (C) metastatic tumours. (D) shows expression of 2 matched transcripts in the Swedish dataset of adequate quality, with imputation of missing data, similarly classified into good-prognosis (right) and poor-prognosis tumours (left). (E) Regional expression biases in the data-set (in this virtual karyotype, red cells indicates inferred amplified regions and blue cells indicates inferred deleted regions). Centromeres are marked as horizontal red lines on the cytoband maps.

High expression of VEGF-signaling pathway and increased MVD counts are associated with good outcomes. We noted during pathway analysis that high expression of many genes involved in angiogenesis, hypoxia-response and endothelial cell formation optimally predicted longer patient survival. To further confirm this, we examined expression of the specific genes involved in the canonical VEGF signaling pathway within the 2260 differentially expressed transcripts found using significance analysis of microarrays, which confirmed this observation (Figure 12, following page). We chose to evaluate this hypothesis using IHC to examine microvessel counts. Conventional CD31 immunostaining yielded a median MVD of 65 vessels / field for 45 tumour samples. Using this as a cutoff, univariate analysis demonstrated that increased MVD was significantly correlated with prolonged survival (Figure 14B-D), with R2 of 0.155 (p-value= 0.01).

A pro-angiogenic phenotype is associated with longer survival. We have identified externally validated predictors of survival, which correspond to distinct molecular subtypes of clear cell RCC. Our results indicate that gene expression profiling is a promising adjunct for survival prediction in clear cell RCC patients. For clear cell RCC, the uncovered relationship between expression of genes associated with hypoxia response, angiogenesis and survival (Figure 12) is difficult to reconcile with current paradigms

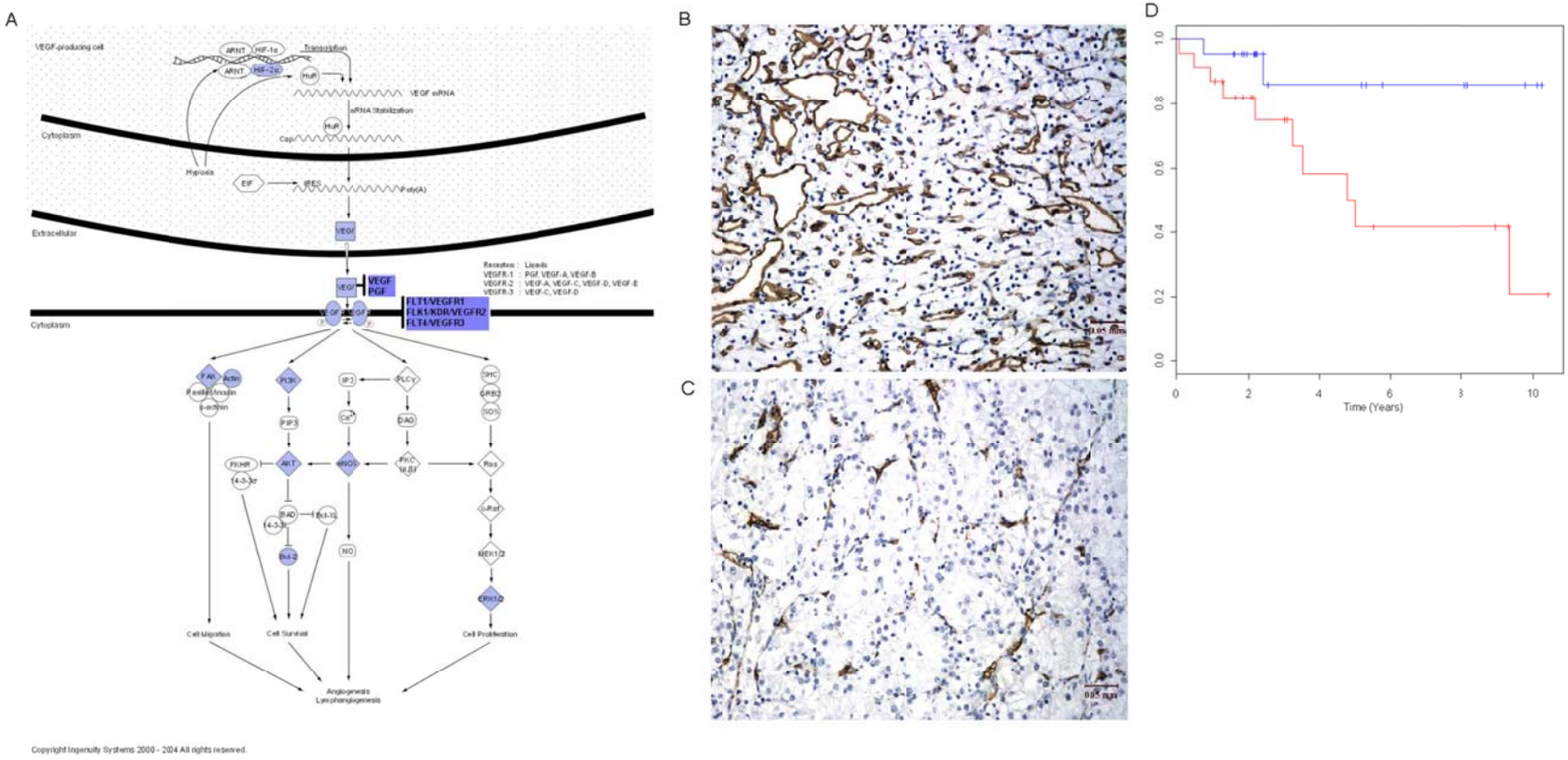

Figure 12 : Angiogenic pathways in RCC
(A) Genes of the canonical VEGF signaling pathway are depicted with blue-shaded genes corresponding to genes that are upregulated in tumours of good-prognosis molecular class. There are no genes in the VEGF pathway upregulated in tumours of poor-prognosis molecular class. It should be noted that *VEGFR* includes *VEGFR1* to *VEGFR3*. (B) Expression of CD31 in good-prognosis and (C) poor-prognosis tumours showing higher vascularity in tumours of the good-prognosis class, and (D) corresponding survival outcomes of highly vascularized (blue line) and poorly vascularized (red line) tumours showing better outcomes in highly vascularized tumours.

of cancer biology. This may represent the most interesting biological insight of this study. This surprising finding has been hinted at previously in an interesting minority of IHC studies (6 of 16 studies) correlating angiogenesis and patient survival in clear cell RCC, as contrasted with most other cancers where a consensus exists that hypoxia and angiogenesis are directly correlated with tumour aggressiveness(*Weidner* 1998; *Weidner* 1999). 6 studies have shown a direct correlation between increasing MVD and prolonged survival in RCC(*Delahunt et al.* 1997; *Herbst et al.* 1998; *Imao et al.* 2004; *Rioux-Leclercq et al.* 2001; *Sabo et al.* 2001; *Sandlund et al.* 2006), 5 studies have shown an inverse relationship(*Dekel et al.* 2002; *Joo et al.* 2004; *Nativ et al.* 1998; *Paradis et al.* 2000; *Yoshino et al.* 1995) and 5 studies have shown no significant correlation(*Gelb et al.* 1997; *Imazano et al.* 1997; *MacLennan and Bostwick* 1995; *Schraml et al.* 2002; *Slaton et al.* 2001). Our results are consistent with previous reports that RCC vascularity is correlated with mRNA expression of *HIF-2* (*Turner et al.* 2002) and *VEGF*(*Takahashi et al.* 1994)*,* both optimal predictors of longer survival in our study. The relationship we observe between tumour aggressiveness and angiogenesis suggests a hitherto unanticipated tumour biology underlying clear cell RCC.

A biological hypothesis explaining the apparently paradoxical relationship between angiogenesis and survival. It is possible that this interesting correlation between hypoxia response, angiogenesis and

prognosis (Figure 3) may be related to an angioblastic origin of clear cell RCC, where different clear cell RCC prognostic groups correspond to various clonal subtypes at different endpoints of misdirected differentiation arrest. The dominant influence on the expression of genes involved in hypoxic response and angiogenesis may thus be similar to the processes driving haematopoetic lineage differentiation. There is circumstantial evidence supporting this hypothesis. A remarkable morphologic resemblance exists between clear cell RCC and haemangioblastoma, both hallmarks of the *VHL* syndrome. Both tumours show immunoreactivity consistent with angioblastic lineage arrest(*Vortmeyer et al.* 2003). Embryonic kidney cells are pluripotent and may function as intrinsic, vasculogenic angioblasts that synthesize microvessels rather than kidney tissue(*Gering et al.* 1998; *Robert et al.* 1996).

Clinical implications of study for therapeutic trials for clear cell RCC. There is considerable current interest in clear cell RCC as a model for targeted therapy interrupting angiogenesis and hypoxia-response pathways(*Rini et al.* 2009). The results of Phase III trials of multi-targeted kinase inhibitors have been announced(*Escudier et al.* 2007; *Motzer et al.* 2007; *Sternberg et al.* 2010), and various agents aimed at blocking *HIF* mediated pathways are also in development. This molecular prognostic predictor is likely to be relevant to Phase II and III trial data interpretation, especially where novel anti-angiogenesis or anti-HIF therapies have been employed. Our hypothesis of an angioblastic origin of RCC, a conceptual

insight paved by expression profiling, opens a new realm of speculation, especially with regards to therapy. Further, in view of the unexpected correlations between markers of hypoxia, angiogenesis and prognosis in clear cell RCC, it will be helpful to determine if good- and poor-prognosis tumours exhibit differential responses to anti-angiogenesis and anti-HIF agents. Our results also indicate that with appropriate stratification, adjuvant therapy should be investigated in individuals with poor-prognosis but apparently localized tumours whom are likely to experience disease recurrence following primary surgery.

The prognostic subtypes are clonally distinct. A previous microarray study has suggested a metastatic expression signature encoded in the majority of cancer cells of a primary tumour(*Ramaswamy et al.* 2003). Given the successful predictions of unsupervised analysis independent of clinico-pathologic variables, an interpretation that fits our data more closely is that biological aggressiveness in clear cell RCC, rather than metastatic ability *per se*, is established by an early clonal event in tumour progression. Results from our test set B, comprising 25 samples of metastatic tumours, clearly highlight that metastatic status and biological aggressiveness are independent, since several metastatic tumours were also classified as good-risk tumours. This phenomenon may have been particularly clear in RCC, which is well known for a proportion of highly indolent metastatic cancers. As such, our results support the existence of at least 2 molecularly distinct clear cell RCC clonal subtypes with distinct cytogenetic aberration patterns, each

subtype growing and metastasizing at different rates. We confirmed a direct prediction of this hypothesis, which was that expression profiles from tumour tissue separated from primary tumour by space (metastasis) and time (relapse) should be expected to exhibit the same correlation between expression signature and survival, as well as similar cytogenetic aberrations. These results account for well-known clinical observations where metastatic clear cell RCC may be startlingly indolent in a small group of untreated patients(*Hughes et al.* 2000) (*Zisman et al.* 2001). Inevitably, the question arises if this conclusion may be extrapolated to other cancers. An *in vitro* study provides support for this concept by demonstrating an independent gene signature for metastasis superimposed upon a poor prognostic gene set in a breast cancer cell line(*Kang et al.* 2003). In establishing that these subtypes are preserved over metastasis and relapse, this result emphasizes the well-known clinical distinction between tumour aggressiveness and metastatic potential *per se*.

Other prognostic microarray studies of clear cell RCC. The gene predictors identified here are externally validated, consistent across array platforms and overlap each other. We were able to demonstrate prognostic utility for overlapping genes in our predictor and the Stanford dataset of Swedish patient(*Zhao et al.* 2006). No overlap was found between the prognostic predictors identified here and a separate microarray study in metastatic RCC(*Vasselli et al.* 2003). This discordance may be a result of the

inclusion of non-clear cell RCC histologies in this particular study as well as specimen quality concerns. However, concordant prognostic predictions among discordant gene predictors are also known(*Fan et al.* 2006). One interesting study (*Jones et al.* 2005) has sought to define a group of 'metastasis-related genes' in kidney cancer by examining a range of histologies of RCC. We have previously speculated that prognostic subtype independent of metastasis forms a dominant contribution to variance in renal cell carcinoma gene expression(*Takahashi et al.* 2001). One group has used unsupervised analysis to study a small group of tumours, linking expression profiles to "tumour aggressiveness" (*Kosari et al.* 2005), defined by the group using a composite index of pathologic scoring, relapse and death.

LIMITATIONS

Although we were able to demonstrate that our predictor remained useful even after adjustment for single variables, including composite variables such as the UISS incorporating stage, grade and functional status, and that it has remained useful in externally generated datasets, ideally a full term model would permit a more measured assessment. Nonetheless, given that the generation of this data was through unsupervised analysis, the result likely represents a biological reality that clinical parameters are a proxy for(*Iliopoulos et al.* 1996). We were unable to use the Karakiewicz nomogram or the Leibovich score due to the absence of data on symptoms and

histologic tumour necrosis; it may well have resulted in a different result on a corresponding regression model.

Secondly, the relatively low number of individuals that received immunotherapy in each group did not permit multivariate analysis by treatment. Nonetheless, given the strength of the predictor in predicting survival adjusted by metastatic status, and the documented modest benefits of immunotherapy in RCC (*Atkins et al.* 2004) it is unlikely that our conclusions will be modified to a major extent. Thirdly, many of the patients in our dataset did not receive targeted therapy, which is the current standard of care. It may be argued that it is difficult to do a study on patients on targeted therapy using overall survival or CSS as an outcome, as the use of these novel agents only became widespread in the last 3 years. There is however good data that the clinical prognostic factors that determine outcomes in metastatic cancer in the pre-targeted therapy era still drive outcomes in the targeted therapy era.

Thirdly, CGMA as an analytic tool remains fundamentally exploratory, focused on virtual karyotyping which requires further definitive validation by cytogenetic approaches, particularly for the copy number aberrations identified. The implications of CGMA results, particularly in interaction with pathway analysis and the generation of expression profile predictors, remain uncertain, and more research is needed.

CONCLUSION

In summary, we have identified clinically useful prognostic gene predictors for clear cell RCC using gene expression profiling. Increased expression of genes classically associated with the VEGF-signaling pathway, angiogenesis and the hypoxic response predicted longer patient survival. Matching genes of our predictor have prognostic value in a public data set that was externally generated. The expression signatures predicted survival in metastatic tumour samples resected following patient relapse, leading to a speculation that the evolution of subtypes with varying biological aggressiveness, propensities for metastasis and cytogenetic aberration patterns is determined in early clonal events during tumour progression. A hypothesis paved by genomic profiling raises the possibility of an angioblastic origin for clear cell RCC.

PAPILLARY RCC

Considering the pathological and clinical issues for papillary RCC discussed in Overall Background and Background (Molecular Models), our goal was to use gene expression profiling as a promising adjunct for cancer diagnosis and prognosis to search for distinct molecular subtypes of papillary RCC that were both biologically and clinically relevant. We also sought to identify novel immunohistochemical markers for each entity.

METHODS

SUBJECTS AND STUDY DESIGN

A total of 43 primary tumour specimens with a diagnosis of papillary RCC after routine pathologic review at 5 medical centers were initially collected following nephrectomy. Participating centres were University of Chicago, Northwestern University, Spectrum Health Grand Rapids, Johns Hopkins University, and University of California Los Angeles. All tumour specimens were collected from participating institutions in the United States, except one case from Japan. 12 non-neoplastic kidney samples were also obtained as controls; these are the same 12 controls for CGMA used for the other microarray studies in other subtypes.

Written informed consent for analysis of clinical samples was obtained from all patients, and all institutional review boards of participating institutions approved the study. Tumour tissue was flash frozen in liquid nitrogen immediately after nephrectomy and stored at –80°C. Portions of the tumours

were fixed in buffered formalin. Each sample was confirmed by pathologic analysis and anonymized prior to the study. A portion of the tumour sample was frozen in liquid nitrogen immediately after surgery and stored at –80˚C. Total RNA was isolated from the frozen tissues using Trizol reagent (Invitrogen, Carlsbad, CA). This was subsequently purified with a RNeasy kit (Qiagen, CA), and quality was assessed on denaturing gel electrophoresis. Representative tissue sections of RCC were cut and stained with hematoxylin-eosin for confirmation of histological diagnosis and confirmation of tumour tissue content (>70%). All samples were examined by a central expert uropathologist. Clinicopathologic information was derived by review of pathologic, radiologic, and case notes by individual clinicians. Tumours with sarcomatoid change were classified as grade 4 tumours. Tumour size was defined as the maximum tumour dimension on direct pathological measurement. Patient follow-up status was assessed by directly contacting patients where possible, or at routine clinical follow-up.

OLIGONUCLEOTIDE ARRAY PROFILING

For oligonucleotide array gene profiling, we extracted and purified total RNA from homogenized samples using Trizol reageant (Invitrogen, CA) followed by RNeasy columns (Qiagen, CA) according to manufacturers' recommendations. Oligonucleotide array gene profiling was performed using the manufacturer's recommended protocol (GeneChip Expression Analysis Technical Manual, April 2003, Affymetrix, CA). Approximately 5-20

micrograms of total RNA was used to prepare antisense biotinylated RNA. A subset of cases were spiked with external poly(A) RNA controls. Synthesis of single-stranged and double-stranded cDNA was done with the use of T7-oligo(dT) primer (Affymetrix). In vitro transcription was done using Enzo Bioarray Transcript Labeling kit (Enzo, Farmingdale, NY). The biotinylated cRNA was subsequently fragmented and 10 micrograms were hybridized to each array at 45°C ovr 16 hours. We used the HGU133 Plus 2.0 GeneChip, containing 54,675 probe sets, representing approximately 47,000 transcripts and variants. Scanning was done in a GeneChip 3000 scanner. Quality indices reviewed for all samples included mean percentage present, mean background, mean scaling factor and a mean GADPH 3'/5' ratio (Median background was 73, median scaling factor was 3.06, and median GADPH 3'/5' ratio was 1.03, indicative of a high overall array and RNA quality). All clinical and microarray data for published data has been uploaded to the Gene Expression Omnibus. The gene expression data can be obtained at the Gene Expression Omnibus (GSE2748). Because of the sample size (n=34), we used supervised analysis, with an internal validation approach. Principal component analysis was used to visualize the 34 expression profiles. Significance analysis of microarrays (SAM) (*Tusher et al.* 2001) based on two-class unpaired analysis, assumption of unequal group variances, and 10,000 permutations was used to derive a list of genes differentially expressed between tumour subclasses and ordered by relative fold change. For derivation of a small gene classifier, we used prediction analysis of

microarrays (PAM) (*Tibshirani et al.* 2002), a *R* implementation of nearest shrunken centroids methodology with 10-fold cross-validation over 30 gene thresholds and an offset percentage of 30%. Gene predictors corresponding to a minimum misclassification error were obtained, with class discriminant scores calculated for class 1 and 2 tumours. Two-tailed Student's *t* test and Fisher's exact testing was used to evaluate correlation between variables and tumour subclassification. For the purpose of this analysis, tumour grade and stage was classified into two categories corresponding to low grade or stage (1 and 2) versus high grade and stage (3 and 4).

PATHWAY ANALYSIS

For pathway analysis in the studies focusing on the molecular aspects of RCC, the Ingenuity Pathways database version 3.0 (Ingenuity Systems, Mountain View, CA) was used. A list of transcripts was generated using significance analysis of microarrays (SAM) based on a 2-class unpaired analysis with 1000 permutations. For the study focusing on chromophobe RCC and oncocytoma, KEGG pathway(*Kanehisa et al.* 2010) and gene ontology (GO) analysis of enriched gene sets was performed using hypergeometric tests available in the GOstats (*Falcon and Gentleman* 2007) package in Bioconductor after having identified unique genes with corresponding annotations. For KEGG pathway analysis, the p-value threshold was 0.01. For GO analysis, conditional testing was performed, and

the threshold for p was 0.001. Molecular function, biologic process, and cellular component analyses were performed.

STATISTICAL ANALYSIS

Statistical analyses were done using R 2.0.1 (*Ihaka and Gentleman* 1996) using packages from the Bioconductor project. The Robust Multichip Average algorithm was used to perform preprocessing of the CEL files, including background adjustment, quartile normalization, and summarization. Principal component analysis was used to visualize the 34 expression profiles. Significance analysis of microarrays (SAM) based on two-class unpaired analysis, assumption of unequal group variances, and 10,000 permutations was used to derive a list of genes differentially expressed between tumour subclasses and ordered by relative fold change. We did pathway analysis on these genes using Ingenuity Pathway Analysis (Ingenuity Systems, Mountain View, CA), and enrichment of canonical pathways was assessed for significance by a hypergeometric

algorithm that did not correct for multiple testing. For derivation of a small gene classifier, we used prediction analysis of microarrays (PAM), a R implementation of nearest shrunken centroids methodology with 10-fold cross-validation over 30 gene thresholds and an offset percentage of 30%(*Tibshirani et al.* 2002). Gene predictors corresponding to a minimum misclassification error

were obtained, with class discriminant scores calculated for class 1 and 2

tumours as described previously. We inferred cytogenetic profiles for the tumours through the use of a refinement of the comparative genomic microarray analysis (CGMA) algorithm(*Crawley and Furge* 2002), which predicts chromosomal

alterations based on regional changes in expression. Relative expression profiles (R) were generated from the single-channel tumour expression profiles (T) and the mean expression values of the 12 single-channel kidney cortical expression profiles (N), such that $R = \log_2(T) - \log_2(N)$. Survival analysis was done by fitting to a Cox proportional hazards model, and significance was determined by the likelihood ratio test. Two-tailed Student's t test and Fisher's exact testing was used to evaluate correlation between variables and tumour subclassification. For the purpose of this analysis, tumour

grade and stage was classified into two categories corresponding to low grade or stage (1 and 2) versus high grade and stage (3 and 4).

IMMUNOHISTOCHEMISTRY

Immunostaining was done on 19 papillary RCC samples that had undergone microarray analysis (10 class 1 tumours and 5 class 2 tumours) as well as an independent set of 15 tumours (10 class 1 tumours and 5 class 2 tumours). We used the biotin-avidin system with mouse monoclonal antibodies specific for cytokeratin 7 (CK7; 1:50 dilution, DAKO, Carpinteria, CA) and DNA topoisomerase II α (TopII α; 1:20 dilution, Vector Laboratories, Burlingame, CA). To verify the differential value of CK7 and TopII α, 21 class

1 tumours were composed of histologic type 1 (*n* = 15), low-grade type 2 tumours (*n* = 3), and mixed type 1/low-grade type 2 tumours (*n* = 8). The 13 class 2 tumours were all high-grade type 2 tumours. The CK7 immunoreactivity was graded as negative (<0.1% positive tumour cells), focally positive (0.1-10% positive tumour cells), or positive (>10% positive tumour cells). The TopII α immunoreactivity was graded as negative (<0.1% positive tumour cells), focally positive (0.1-10% positive tumour cells), or positive (>10% positive tumour cells). The Mann-Whitney test was used to evaluate significance of the differential staining.

RESULTS

Gene expression profiling correspond to morphologic characteristics. Individual patient data is presented in Table 11. Based on histologic features and Fuhrman grading system, we designated four categories of tumours: type 1 (*n* = 14; Fig 13A, page 105) characterized by small tumour cells of low grade ($\leq$2); type 2A (*n* = 4; Fig 13B, page 105) characterized by large eosinophilic tumour cells of low grade ($\leq$2); combined type 1 and 2A (*n* = 5; Figure 13C; and type 2B (*n* = 11; Fig 13D, page 105) characterized by large eosinophilic tumour cells with high grade ($\geq$3).

Table 11 : Individual patient demographic data for papillary RCC dataset

Patient ID	Age (years)	Gender	Histological Classification	Molecular Classification	Size (cm)	TNM Stage	M Stage	Grade	Patient status*	Survival (Months)
P 01	46	F	Type 2B	Class 2	15.2	4	1	3	NA	NA
P 02	59	M	Type 1	Class 1	4.5	1	X	1	NA	NA
P 03	68	M	Type 1	Class 1	4	1	X	2	NA	NA
P 04	32	M	Type 2A	Class 1	7	1	X	2	NA	NA
P 05	71	M	Type 1 and 2A	Class 1	4.7	1	0	2	DOO	30
P 06	70	M	Type 1	Class 1	1.2	3	0	2	NED	26
P 07	72	F	Type 1	Class 1	3	1	0	2	DOO	8
P 08	73	F	Type 1	Class 1	3	1	0	2	NED	25
P 09	84	F	Type 1	Class 1	3.5	1	0	2	NED	27
P 10	56	F	Type 1	Class 1	14	3	0	2	NED	13
P 11	56	M	Type 1	Class 1	3.5	1	0	2	NED	13
P 12	80	F	Type 1	Class 1	2.5	1	0	2	NED	4
P 13	64	M	Type 1	Class 1	5.5	1	0	2	DOD	74
P 14	44	M	Type 1	Class 1	3.5	3	0	2	NED	95
P 15	76	M	Type 1 and 2A	Class 1	4.5	1	0	2	NED	58
P 16	72	M	Type 2B	Class 2	10	3	0	4	DOD	17
P 17	55	M	Type 1	Class 1	5	3	0	2	NED	30
P 18	71	F	Type 2B	Class 2	3	1	0	3	NED	32
P 19	76	M	Type 2B	Class 2	5.7	3	0	3	NED	36
P 20	71	M	Type 1	Class 1	2.7	1	0	2	DOO	63
P 21	80	F	Type 2A	Class 1	3.5	2	0	2	NED	66
P 22	53	M	Type 1 and 2A	Class 1	2.3	1	0	2	NED	61
P 23	54	M	Type 2B	Class 2	8	4	0	4	DOD	12
P 24	50	M	Type 2B	Class 2	11	3	1	3	DOO	37

P 25	44	M	Type 2B	Class 2	14	4	1	4	AWC	12
P 26	75	M	Type 2A	Class 1	9	2	0	2	NA	NA
P 27	74	M	Type 1 and 2A	Class 1	6	1	0	2	NED	24
P 28	37	M	Type 1 and 2A	Class 1	6	1	0	2	DOO	12
P 29	43	M	Type 1	Class 1	5	1	0	2	NA	NA
P 30	63	M	Type 2A	Class 2	6.9	3	0	2	DOO	34
P 31	62	M	Type 2B	Class 2	6.8	4	1	3	DOO	4
P 32	56	M	Type 2B	Class 2	6	4	1	3	DOD	16
P 33	71	F	Type 2B	Class 2	15	4	1	3	DOD	29
P 34	49	F	Type 2B	Class 2	15	4	1	3	DOD	9

We visualized the 34 expression profiles by principal component analysis. We noted overlap between histologic type 1 and 2 tumours, contrary to our expectation of distinct molecular subtypes (Figure 13, page 105). Tumours with mixed type 1 and 2 components (*n* = 5) grouped with type 1 tumours. PAM with 10-fold cross-validation persistently classified three of four low-grade type 2 tumours with type 1 tumours over a wide range of shrinking gene thresholds. The only low-grade type 2 tumour that persistently classified with the high-grade type 2 tumours was P30 (the only tumour we were unable to personally evaluate histologically). These results supported a hypothesis that type 2 tumours were molecularly heterogenous. We analyzed the profiles based on this morphologic subtyping into two classes (class 1 corresponding to type 1, low-grade type 2, and mixed type 1/low-grade type 2 tumours and class 2 corresponding to high-grade type 2 tumours) from a molecular viewpoint. The first two principal components accounted for 95.7% of all variation, allowing for effective clustering. Visualization of the first two principal components now showed distinct differentiation between expression profiles of class 1 and 2 tumours, consistent with distinct tumour subclasses (Fig 15F). Transcripts (*n* = 796) (available online as Supplemental Data http://cancerres.aacrjournals.org/content/vol65/issue13/images/data/5628/DC1/Supplementary_Table_1.xls) differentially expressed between class 1 and 2 tumours were identified using SAM at delta of 1.8, with a false discovery rate (FDR) of 0.01.

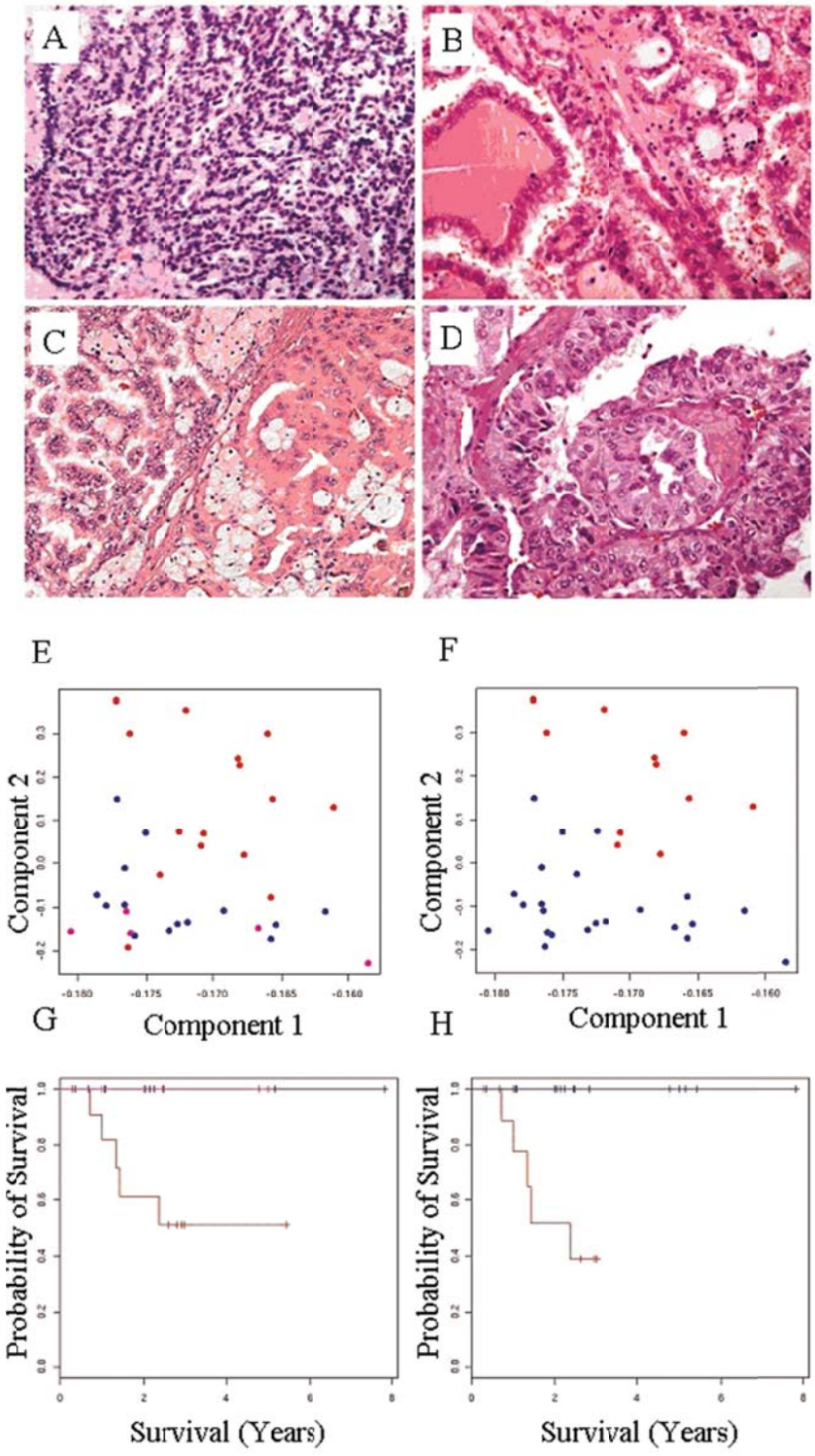

Figure 13 : Histologic and molecular subtypes of papillary RCC
(A) – (D) Respectively, type I, type II low grade, type I/II mixed and type II high grade and (E) – (F) molecular subtyping, with (G) – (H) corresponding survival outcomes

We were able to identify multiple gene classifiers that effectively differentiated class 1 and 2 tumours at 97% accuracy at multiple shrinkage thresholds using PAM (between 7 and 3,881 transcripts) using nearest shrunken centroids methodology. We report here the seven-transcript predictor that achieved this accuracy (Table 12). Only the tumour of P30, initially reported as a type 2 tumour with grade 2, which we were unable to confirm histologically, persistently classified as a class 2 tumour, rather than as a class 1 tumour, throughout these multiple shrinkage thresholds.

Table 12 : The 7 transcript predictor discriminating Class 1 and 2 papillary RCC

Probe ID	Gene Description	Class.1.score	Class.2.score
232151_at	MRNA full length insert cDNA clone EUROIMAGE 2344436	0.1789	-0.3741
1566766_a_at	MRNA full length insert cDNA clone EUROIMAGE 2344436	0.1419	-0.2967
204304_s_at	prominin 1	0.1182	-0.2472
210298_x_at	four and a half LIM domains 1	0.0164	-0.0342
201539_s_at	four and a half LIM domains 1	0.0159	-0.0332
214505_s_at	four and a half LIM domains 1	0.0101	-0.0211
205597_at	chromosome 6 open reading frame 29	0.0044	-0.0092

* These class scores are linear discriminant scores for each class as described in the reference for significance analysis of microarrays in the text.

We list the top 50 transcripts relatively upexpressed in each subclass (Table 13) (this represents an arbitrary cutoff for presentation, and is not based on a p-value threshold) and show a hierarchical clustering of the tumour samples based on these 100 transcripts (Figure 14, page 111).

Table 13 : Top 50 transcripts differentially expressed in Class 1 and 2 papillary RCC

GENES UPEXPRESSED IN CLASS 1 papillary RCC						
Probe ID	Gene Description	D(i)*	Standard deviation	p-value	q-value	Fold-Change**
213456_at	sclerostin domain containing 1	-7.2	0.5	0.0	0.0	33.4
201539_s_at	four and a half LIM domains 1	-10.9	0.3	0.0	0.0	14.8
204304_s_at	prominin 1	-8.0	0.6	0.0	0.0	14.4
210298_x_at	four and a half LIM domains 1	-9.9	0.4	0.0	0.0	14.3
214505_s_at	four and a half LIM domains 1	-11.3	0.3	0.0	0.0	12.1
210299_s_at	four and a half LIM domains 1	-8.1	0.5	0.0	0.0	10.6
209016_s_at	keratin 7	-6.9	0.5	0.0	0.0	8.4
205597_at	chromosome 6 open reading frame 29	-11.4	0.3	0.0	0.0	8.3
232151_at	MRNA full length insert cDNA clone EUROIMAGE 2344436	-9.2	0.4	0.0	0.0	8.0
1566764_at	MRNA full length insert cDNA clone EUROIMAGE 2344436	-8.0	0.4	0.0	0.0	7.5
1553809_a_at	chromosome 9 open reading frame 71	-6.9	0.3	0.0	0.0	6.5
1566766_a_at	MRNA full length insert cDNA clone EUROIMAGE 2344436	-10.0	0.3	0.0	0.0	6.0
224027_at	chemokine (C-C motif) ligand 28	-7.2	0.3	0.0	0.0	5.7
1555203_s_at	chromosome 6 open reading frame 29	-9.7	0.2	0.0	0.0	5.3
238184_at	Transcribed sequences	-9.4	0.2	0.0	0.0	4.8
202820_at	aryl hydrocarbon receptor	-7.2	0.3	0.0	0.0	4.2
202790_at	claudin 7	-9.1	0.2	0.0	0.0	4.0
222764_at	asparaginase like 1	-8.1	0.2	0.0	0.0	3.9
219127_at	hypothetical protein MGC11242	-7.5	0.2	0.0	0.0	3.6
218857_s_at	asparaginase like 1	-9.0	0.2	0.0	0.0	3.5
229084_at	contactin 4	-6.9	0.2	0.0	0.0	3.4
219614_s_at	solute carrier family 6 (neurotransmitter transporter), member 20	-10.1	0.2	0.0	0.0	3.2
210398_x_at	fucosyltransferase 6 (alpha (1,3) fucosyltransferase)	-8.3	0.2	0.0	0.0	3.1
205405_at	Semaphorin 5A	-7.3	0.2	0.0	0.0	3.0
1559361_at	MRNA full length insert cDNA clone EUROIMAGE 2344436	-8.2	0.2	0.0	0.0	2.8
203365_s_at	matrix metalloproteinase 15 (membrane-inserted)	-10.0	0.1	0.0	0.0	2.6
229144_at	KIAA1026 protein	-7.9	0.2	0.0	0.0	2.4

211110_s_at	androgen receptor	-6.9	0.2	0.0	0.0	2.3
223636_at	zinc finger, MYND domain containing 12	-7.5	0.1	0.0	0.0	2.2
231022_at	Transcribed sequences	-7.0	0.2	0.0	0.0	2.2
221665_s_at	EPS8-like 1	-8.0	0.1	0.0	0.0	2.1
235937_at	Occluding	-7.1	0.1	0.0	0.0	2.1
217795_s_at	hypothetical protein MGC3222	-7.5	0.1	0.0	0.0	2.1
211621_at	androgen receptor	-7.3	0.1	0.0	0.0	2.1
210399_x_at	fucosyltransferase 6 (alpha (1,3) fucosyltransferase)	-7.1	0.1	0.0	0.0	2.0
202005_at	suppression of tumourigenicity 14 (colon carcinoma, matriptase, epithin)	-7.1	0.1	0.0	0.0	1.9
243225_at	hypothetical protein LOC283481	-7.2	0.1	0.0	0.0	1.9
91826_at	EPS8-like 1	-8.2	0.1	0.0	0.0	1.7
218779_x_at	EPS8-like 1	-7.5	0.1	0.0	0.0	1.7
205977_s_at	EphA1	-9.1	0.1	0.0	0.0	1.6
235293_at	CDNA FLJ45593 fis, clone BRTHA3014920	-7.5	0.1	0.0	0.0	1.6
238028_at	Similar to hypothetical protein BC006605 (LOC389389), mRNA	-7.9	0.1	0.0	0.0	1.6
221655_x_at	EPS8-like 1	-7.2	0.1	0.0	0.0	1.5
236058_at	hypothetical protein FLJ34633	-8.8	0.1	0.0	0.0	1.5
225778_at	RNA binding motif, single stranded interacting protein 2	-6.9	0.1	0.0	0.0	1.5
223724_s_at	DKFZp434A0131 protein	-6.9	0.1	0.0	0.0	1.5
230111_at	Transcribed sequence with moderate similarity to protein ref:NP_308425.1	-7.4	0.1	0.0	0.0	1.5
226095_s_at	hypothetical protein LOC146517	-7.3	0.1	0.0	0.0	1.4
239812_s_at	hypothetical protein FLJ12476	-8.1	0.1	0.0	0.0	1.4
238318_at	Transcribed sequences	-7.1	0.0	0.0	0.0	1.2

GENES UPEXPRESSED IN CLASS 2 papillary RCC						
Probe ID	Gene Description	D(i)*	Standard deviation	p-value	q-value	Fold-Change**
201292_at	topoisomerase (DNA) II alpha 170kDa	6.0	0.4	0.0	0.0	6.5

202920_at	ankyrin 2, neuronal	5.7	0.4	0.0	0.0	5.9
228776_at	CDNA FLJ40955 fis, clone UTERU2011199	6.3	0.4	0.0	0.0	4.4
201761_at	methylene tetrahydrofolate dehydrogenase (NAD+ dependent)	7.7	0.3	0.0	0.0	4.1
209900_s_at	solute carrier family 16 (monocarboxylic acid transporters), member 1	6.7	0.3	0.0	0.0	3.9
218009_s_at	protein regulator of cytokinesis 1	5.6	0.3	0.0	0.0	3.6
202234_s_at	solute carrier family 16 (monocarboxylic acid transporters), member 1	7.1	0.2	0.0	0.0	3.5
209773_s_at	ribonucleotide reductase M2 polypeptide	6.2	0.3	0.0	0.0	3.2
210052_s_at	TPX2, microtubule-associated protein homolog (Xenopus laevis)	6.5	0.2	0.0	0.0	3.2
218883_s_at	KSHV latent nuclear antigen interacting protein 1	5.5	0.3	0.0	0.0	3.1
225655_at	ubiquitin-like, containing PHD and RING finger domains, 1	5.6	0.3	0.0	0.0	3.1
218039_at	nucleolar and spindle associated protein 1	6.0	0.3	0.0	0.0	3.0
204822_at	TTK protein kinase	5.7	0.2	0.0	0.0	2.9
227607_at	associated molecule with the SH3 domain of STAM (AMSH) like protein	5.8	0.3	0.0	0.0	2.8
201664_at	SMC4 structural maintenance of chromosomes 4-like 1 (yeast)	7.8	0.2	0.0	0.0	2.7
201663_s_at	SMC4 structural maintenance of chromosomes 4-like 1 (yeast)	5.7	0.2	0.0	0.0	2.5
212110_at	solute carrier family 39 (zinc transporter), member 14	5.6	0.2	0.0	0.0	2.5
203554_x_at	pituitary tumour-transforming 1	6.3	0.2	0.0	0.0	2.5
202338_at	thymidine kinase 1, soluble	5.9	0.2	0.0	0.0	2.3
1555758_a_at	cyclin-dependent kinase inhibitor 3	5.8	0.2	0.0	0.0	2.3
202954_at	ubiquitin-conjugating enzyme E2	6.2	0.2	0.0	0.0	2.2
1554408_a_at	thymidine kinase 1, soluble	6.1	0.2	0.0	0.0	2.2
227606_s_at	associated molecule with the SH3 domain of STAM (AMSH) like protein	5.6	0.2	0.0	0.0	2.1
203764_at	discs, large homolog 7 (Drosophila)	5.9	0.2	0.0	0.0	2.1
221923_s_at	nucleophosmin (nucleolar phosphoprotein B23, numatrin)	5.7	0.2	0.0	0.0	2.0
212295_s_at	solute carrier family 7, member 1	5.4	0.2	0.0	0.0	2.0
204092_s_at	serine/threonine kinase 6	5.6	0.2	0.0	0.0	2.0
202705_at	cyclin B2	5.7	0.2	0.0	0.0	1.9
213188_s_at	MYC induced nuclear antigen	5.8	0.2	0.0	0.0	1.9

228245_s_at	ovostatin 2	6.0	0.2	0.0	0.0	1.9
207828_s_at	centromere protein F, 350/400ka (mitosin)	5.6	0.2	0.0	0.0	1.9
205345_at	BRCA1 associated RING domain 1	5.6	0.2	0.0	0.0	1.9
203669_s_at	diacylglycerol O-acyltransferase homolog 1 (mouse)	5.8	0.1	0.0	0.0	1.8
213689_x_at	ribosomal protein L5	5.8	0.1	0.0	0.0	1.7
203562_at	fasciculation and elongation protein zeta 1 (zygin I)	5.6	0.1	0.0	0.0	1.7
214096_s_at	serine hydroxymethyltransferase 2 (mitochondrial)	5.4	0.1	0.0	0.0	1.7
205651_x_at	Rap guanine nucleotide exchange factor (GEF) 4	5.4	0.1	0.0	0.0	1.7
213947_s_at	nucleoporin 210	5.7	0.1	0.0	0.0	1.7
230165_at	shugoshin-like 2 (S. pombe)	5.4	0.1	0.0	0.0	1.6
218950_at	ARF-GAP, RHO-GAP, ankyrin repeat and plekstrin homology domains-containing protein 3	6.5	0.1	0.0	0.0	1.6
203022_at	ribonuclease H2, large subunit	6.4	0.1	0.0	0.0	1.5
203719_at	excision repair cross-complementing rodent repair deficiency	6.3	0.1	0.0	0.0	1.5
218115_at	ASF1 anti-silencing function 1 homolog B (S. cerevisiae)	5.4	0.1	0.0	0.0	1.5
210023_s_at	likely ortholog of mouse nervous system polycomb 1	6.1	0.1	0.0	0.0	1.4
221591_s_at	hypothetical protein FLJ10156	5.6	0.1	0.0	0.0	1.4
204126_s_at	CDC45 cell division cycle 45-like (S. cerevisiae)	6.7	0.1	0.0	0.0	1.3
214426_x_at	chromatin assembly factor 1, subunit A (p150)	5.9	0.1	0.0	0.0	1.3
212313_at	hypothetical protein MGC29816	5.9	0.1	0.0	0.0	1.3
221779_at	molecule interacting with Rab13	5.8	0.1	0.0	0.0	1.3
239680_at	hypothetical protein FLJ12973	6.0	0.1	0.0	0.0	1.3

- * D(i) is a modified t-statistic calculated by SAM.
- ** Fold-change is shown in terms of a relationship between the tumour with higher expression relative to the tumour with lower expression.

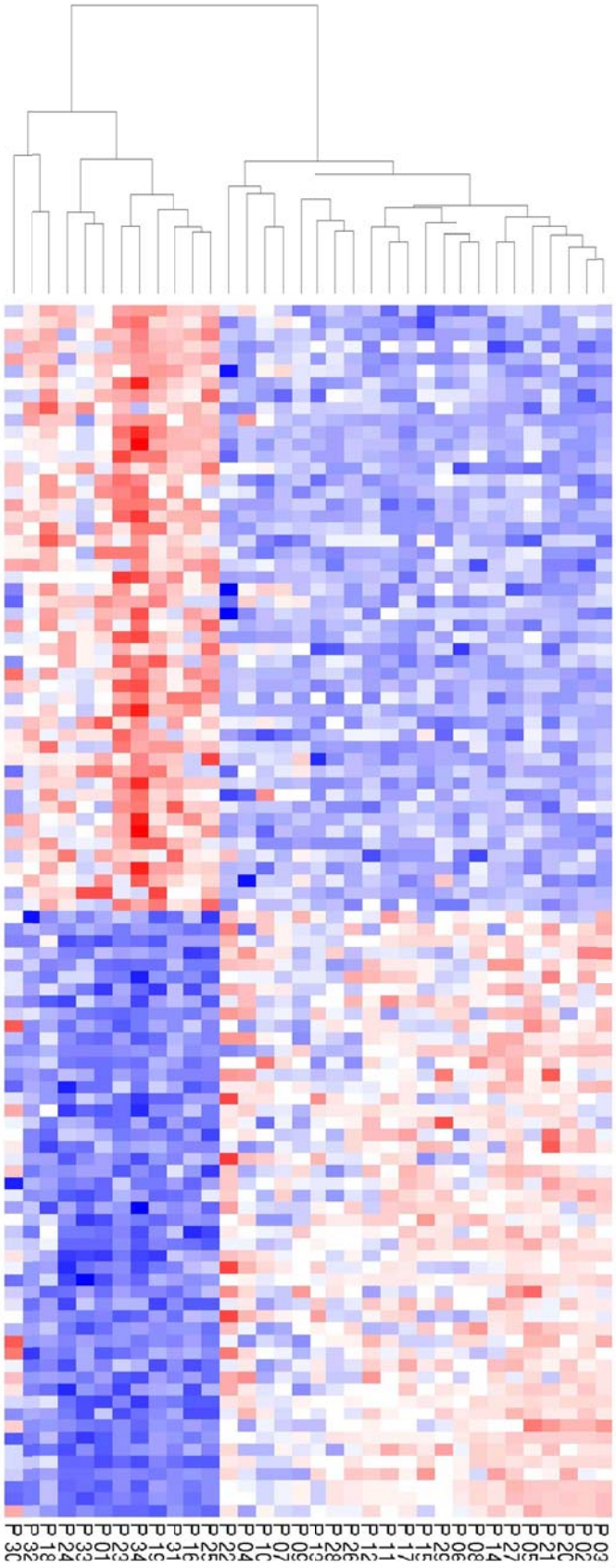

Figure 14 : Hierarchical clustering of papillary RCC expression profiles based on the 100 differentially expressed transcripts.
Red represents relative overexpression and blue represents relative underexpression. The heatmap shows clear partitioning of the dataset into two major molecular classes. Rows represent individual oligonucloetide probes; columns, individual tumour samples. Complete linkage clustering and a Euclidean distance metric was used, and values were scaled by row. On left, class 2 tumours corresponding to all type 2 papillary tumors; on right, class 1 tumours corresponding to all type 1 and type 2A tumours.

Survival analysis showed that this refined morphologic and molecular classification system showed a survival prediction that showed a statistically insignificant edge over the previous morphology-based classification approach (Nagelkerke's R^2 = 0.505 and P = 0.001 versus R^2 = 0.389 and P = 0.005). Class 2 tumours were larger in tumour dimension (P = 0.003), of higher grade (P < 0.001), of higher stage (P < 0.001), and were more likely to exhibit distant metastases at initial surgery (P < 0.001) than class 1 tumours. Indeed, all tumours metastatic at initial surgery were class 2 tumours (n = 7). No significant difference in age (P = 0.37) or gender (P = 0.70) was found between the two classes.

Chromosomal aberrations inferred by CGMA. Distinct cytogenetic profiles for each tumour were generated using high-resolution CGMA (Figure 15, following page) Full-length gains in chromosomes 7, 12, 16, 17, and 20 was found both in class 1 and 2 tumours, consistent with the previously reported trisomies observed by using conventional cytogenetic analysis characteristic of papillary RCC. However, in comparison with class 1 tumours, class 2 tumours exhibited more frequent gains at 1q, 2, and 8q and losses at 3p and 6q and showed fewer gains of chromosome 3, 7, and 16. More frequent losses of 6q and 14q were also evident.

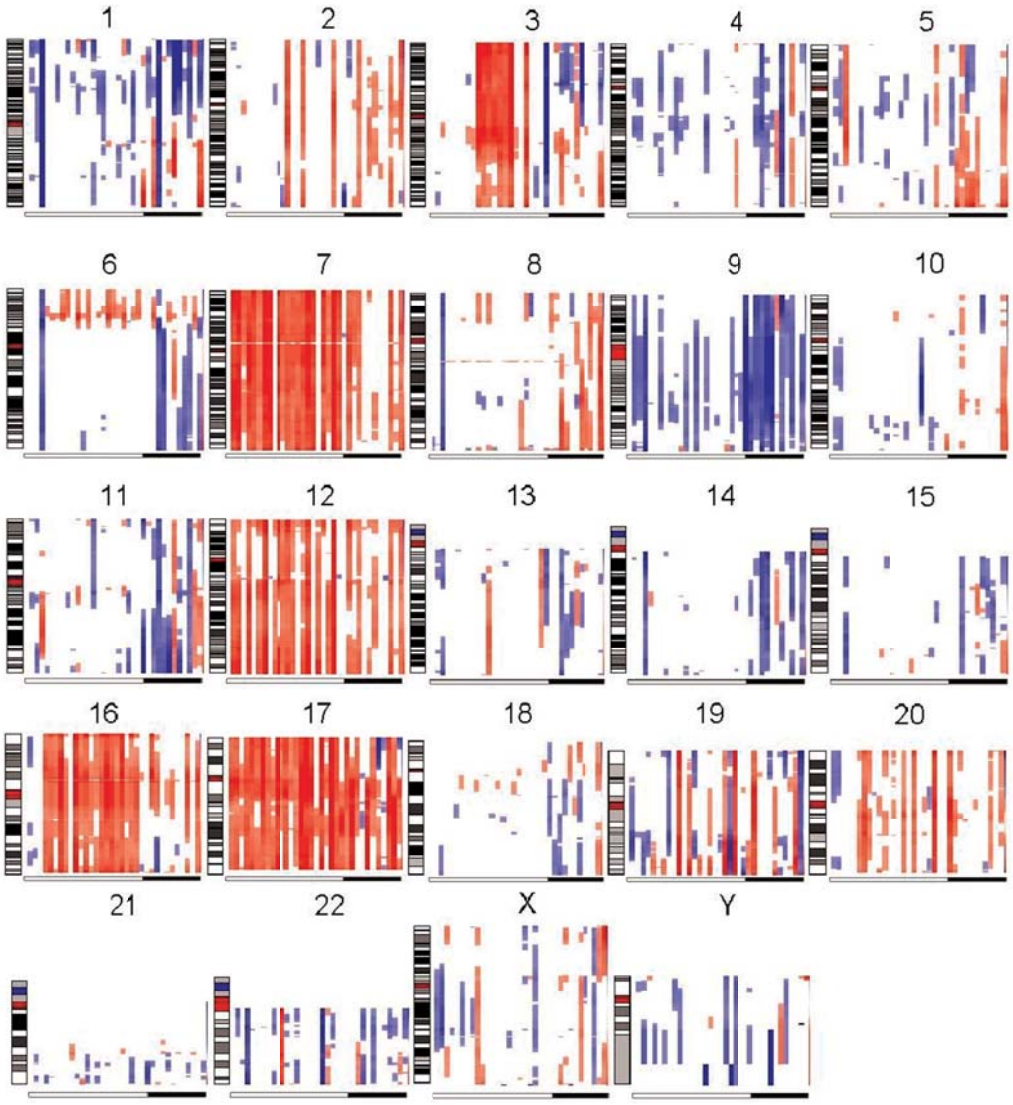

Figure 15 : Chromosomal ideograms depicting regional gene expression biases of papillary RCC.
Red indicates upexpression and blue downexpression in terms of inferred regional expression biases. Each block corresponds to a single chromosome, representing the chromosomal expression profiles of a group of samples, and each sample is represented by a single vertical line in each block. Class 1 papillary RCC are presented in each ideogram as the samples above the white bar and Class 2 papillary RCC as samples above the black bar respectively. As an example, for chr 7 ideogram, Class 1 papillary RCC shows regional chr 7 upexpression, but Class 2 tumours show a lower prevalence. Centromeres are shown in red on the chromosomal map on the left of each block.

Pathway analysis. Genes (*n* = 203) derived from the 796 transcripts were eligible for generation of networks in pathway analysis using hypergeometric analyses as described in Methods. Ranking of canonical pathways yielded three pathways that were significantly enriched within these differentially expressed genes: G_2-M DNA damage checkpoint regulation (*P* = 0.007), arginine and proline metabolism (*P* = 0.011), and G_1-S checkpoint regulation (*P* = 0.018). Genes involved in G_1-S checkpoint regulation (cyclin D2, cyclin-dependent kinase 6, retinoblastoma-like 2, and $p21^{Cip1}$) were relatively upexpressed in class 1 tumours (Figure 16, following page) whereas genes involved in G_2-M checkpoint regulation (cyclin B1, cyclin B2, and TopII α) were relatively upexpressed in class 2 tumours (Figure 16, following page). Multiple oligonucleotide probe sets corresponding to c-met were identified as being upexpressed in class 1 tumours, ranging between 2- and 3-fold upexpression.

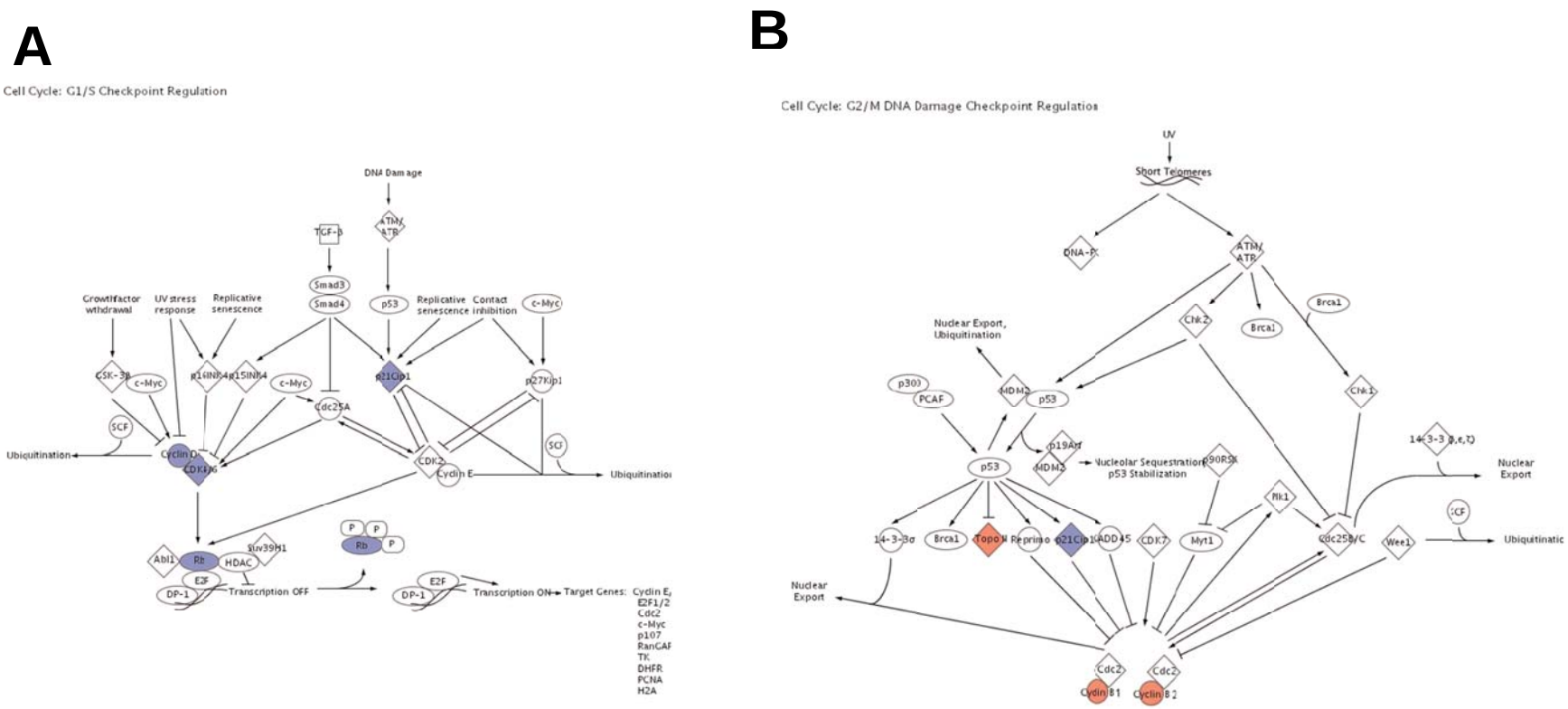

Figure 16 : Pathway analysis for papillary RCC
(A) Analysis in Class 1 papillary RCC showing predominant G1/S checkpoint derangement and (B) Class 2 papillary RCC showing predominant G2/M checkpoint derangement. Blue represents transcripts upexpressed in Class 1 and red represents transcripts upexpressed in class 2 tumours.

Immunohistochemical characteristics.

The IHC findings are reported (Table 14, following page), and are consistent between the sets of profiled and independent tumours. The majority of class 1 tumours (86%), including type 1 (Figure 19A-C) and type 2A (Figure 17, following page) tumours, showed strong CK7 immunoreactivity, whereas the majority of class 2 tumours (Figure 17, following page) showed absent (77%) or reduced (23%) CK7 immunoreactivity in both the set of profiled tumours and the independent set of tumours. In contrast, TopIIα immunoreactivity was focally positive (10%) or negative (90%) in class 1 tumours, including both type 1 tumours and type 2A tumours The majority of class 2 tumours were positive for TopIIα (90% positive and 10% focally positive) in both the set of profiled tumours and the independent set of tumours. No TopIIα immunoreactivity was detected in normal kidney tissue. There was no apparent difference between type 1 and low-grade type 2 (type 2A) tumours in CK7 and TopIIα immunostaining. Summarizing the results, CK7 immunoreactivity was significantly higher in class 1 tumours ($P < 0.001$), and TopIIα immunoreactivity was significantly higher in class 2 tumours ($P < 0.001$).

Table 14 : Immunohistochemical results for Class 1 and Class 2 papillary RCC

		Cytokeratin 7 immunostaining			TopIIα immunostaining		
		Negative	Focally positive	Positive	Negative	Focally positive	Positive
Profiled tumours	Type 1	0	0	10	10	0	0
	Type 2A	0	0	1	1	0	0
	Mixed Type 1 and 2A	0	1	2	3	0	0
	Type 2B	4	1	0	0	1	4
Independent tumours	Type 1	0	1	4	5	0	0
	Type 2A	1	0	1	0	2	0
	Mixed Type 1 and 2A	0	0	0	0	0	0
	Type 2B	6	2	0	0	0	8
TOTAL		11	5	18	19	3	12

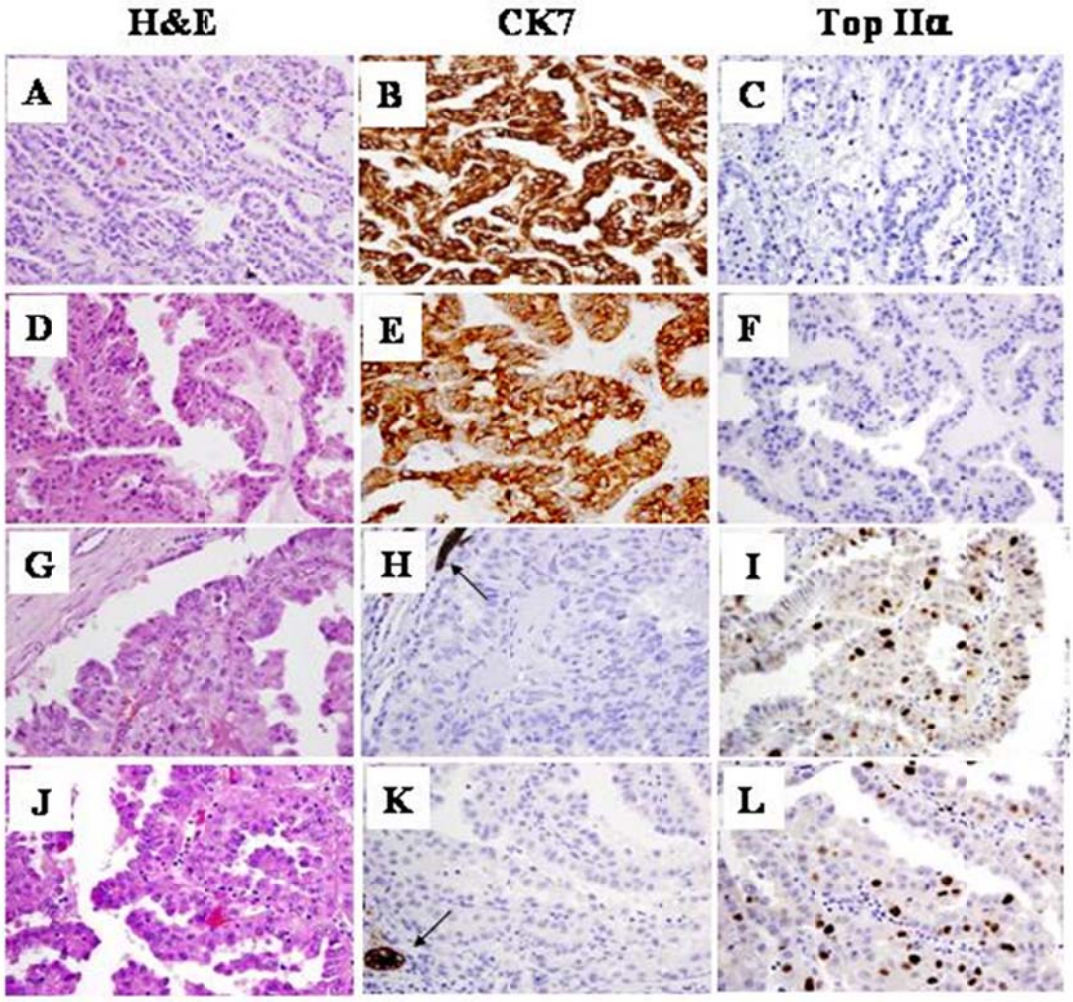

Figure 17: Immunohistochemical staining of papillary RCC
(A)- (C) Type 1 tumour, (D)-(F) Type 2A tumour, (G)-(I) Type 2B tumour which was subjected to microarray analysis and (J)-(L) Type 2B tumour, an independent tumour which was not subjected to microarray analysis. Note that a renal tubule (arrow) stains positive for CK7 as an internal positive control, whereas all tumour cells are otherwise negative This image shows that Type 1 and 2A tumours stain positive for CK7, whereas Type 2B tumours stain positive for topoisomerase II-alpha.

DISCUSSION

Morphologic classification. The morphologic classification of papillary RCC into type 1 and 2 tumours has been supported by several histologic studies, although there is relatively limited molecular evidence to substantiate this subtyping. There remains controversy over the recent proposed morphologic classification system of papillary RCC, preventing its widespread application. For example, there is no agreement whether a tumour with eosinophilic cytoplasm but low nuclear grade should be classified as type 1 or 2. In the initial proposal outlining this morphologic subtyping. 63% of type 2 tumours were assessed as being of low Fuhrman nuclear grade despite pleomorphic nuclei being defined as a characteristic of type 2 tumours (*Delahunt and Eble* 1997). More recently, Allory et al. classified only 1 of 13 (8%) as low-grade type 2 tumours using a modified criteria (*Allory et al.* 2003). The high frequency of tumours with coexisting type 1 and 2 components poses difficulties for such a binary classification, the prevalence of such mixed tumours having been reported as high as 28%. Allory et al. chose to classify these tumours with mixed (type 1 and 2) features as type 1 tumours, an approach in line with our molecular classification.

Molecular classification. Our results provide only partial support for the proposed histologic subtyping of papillary RCC into type 1 and 2 tumours. Type 2 tumours are molecularly heterogenous, with a subset of type 2 (low-grade) tumours and mixed type 1 and 2 tumours demonstrating molecular profiles more consistent with type 1 tumours. These type 2 tumours were all low-grade tumours

and showed excellent clinical outcomes, in contrast with the poor outcomes recorded in high-grade type 2 tumours. Type 2 papillary RCC is composed of at least two genetically distinct subtypes: one subtype (type 2A) resembles type 1 in terms of indolent tumour behavior, excellent survival, low tumour grade, similar expression profiles, immunoreactivity, and inferred cytogenetic profiles; the other subtype (type 2B) is an highly metastatic, aggressive cancer that is molecularly distinct from type 1 or 2A tumours. Our findings support a view that nuclear grade is the key correlate for a molecular classification with both biological and clinical relevance, with features such as cell size or cytoplasmic eosinophilia being more peripheral. Additional distinctive histopathologic features for these subclasses may be defined with a larger series. In this report, the molecular classification showed a statistically insignificant edge in prognostication over the previously proposed histologic classification. However, the molecular approach with correlation to nuclear grade may be more relevant, as it also accurately classifies mixed type 1 and 2 tumours, which are not well accounted for in the histologic classification. This refined classification of papillary RCC based on both morphologic features and molecular studies may be more relevant and is likely to benefit diagnosis, prognostication, clinical follow-up, and experimental selection of therapeutic targets.

We successfully generated an internally validated seven-transcript predictor, which was able to classify class 1 and 2 tumours with 97% accuracy, the only misclassification arising from a tumour (P30) that we were unable to personally evaluate. Consistent with our microarray classification, this tumour

from P30 behaved in an aggressive fashion, the patient relapsing 2 years after surgery. The patient died of a non-cancer-related cause 10 months after relapse. External validation in a second population is required for assessment of true generalizability of these gene predictors, but these results are very encouraging.

Inferred cytogenetic profiles. Aneuploidy is well established as a key driver of global gene expression, and regional DNA copy number correlates well with regional expression in cancer, which we have also shown in RCC classification. Papillary RCC typically shows frequent trisomies of chromosomes 7, 12, 16, 17, and 20 (*Amin et al.* 1997) (*Kovacs et al.* 1991) (*Kattar et al.* 1997). For papillary RCC subclassification, our results are strictly not directly comparable with recent cytogenetic studies that have classified their results by the type 1 and 2 classification (*Jiang et al.* 1998) (*Gunawan et al.* 2003). As expected, our inferred cytogenetic profiles were consistent with previous studies correlating cytogenetic findings with tumour grade; Lager et al. identifying less frequent trisomy of 7 in high-grade tumours (*Lager et al.* 1995) and Renshaw and Corless reporting that trisomy of 3 was found in a defined subset of low-grade papillary RCC tumours (*Renshaw and Corless* 1995). In addition to these findings, in demonstrating that loss of 9q occurred more commonly in class 2 tumours, our results support a report that loss of heterozygosity at 9q is associated with reduced survival.

Immunohistochemical findings. To validate the gene predictor and to derive IHC markers for the pathology laboratory, we used IHC to confirm high

protein expression of CK7 in class 1 tumours and of Topo IIα in class 2 tumours. CK7 immunoreactivity has been reported previously to the vast majority of papillary RCC (*Renshaw and Corless* 1995), but more recent studies suggested that CK may differentiate type 1 and 2 tumours. Our microarray and IHC findings were generally consistent with findings using the morphologic classification that between 87% and 100% of type 1 tumours showed CK7 positivity and ~20% of type 2 tumours showed CK7 positivity (*Delahunt and Eble* 1997). No IHC marker has been reported previously as being specifically upexpressed in type 2 tumours; we showed the usefulness of DNA TopIIα as an IHC marker in class 2 tumours.

Pathway analysis. Our study highlighted an exploratory finding of dysregulation of G_1-S checkpoint genes in class 1 papillary RCC and dysregulation of G_2-M checkpoint genes in class 2 papillary RCC as the most highly ranked pathways identified in the differentially expressed genes. In familial studies, mutations of the *MET* proto-oncogene have been implicated in hereditary type 1 papillary RCC and a small subset (<10%) of sporadic type 1 papillary RCCs(*Schmidt et al.* 1998; *Schmidt et al.* 1999). From a mechanistic point of view, this associative link between *MET* overexpression/mutation and genes associated with G_1-S checkpoint dysregulation is particularly interesting, as hepatocytes in conditional *met*-mutant mice exhibit defective exit from quiescence and diminished entry into the S-phase of the cell cycle(*Borowiak et al.* 2004). Further work is required to delineate the role of met signaling in G_1-S checkpoint dysregulation. Differential expression of the *FH* gene, which is

mutated in a group of families with type 2 papillary RCC (*Tomlinson et al.* 2002), was not observed (data not shown).

The implication of dysregulation of the G_2-M checkpoint regulation in class 2 tumours is particularly interesting from a therapeutic point of view. We took a particular interest in DNA TopII, which we additionally established as a diagnostic marker for class 2 tumours. As there is no effective medical therapy for advanced papillary RCC and this enzyme is associated with the more aggressive papillary RCC subclass, TopII inhibitors are distinct possibilities for a therapeutic trial of papillary RCC. G_2 arrest occurs in response to these agents (*Clifford et al.* 2003) and may therefore be particularly appropriate. Although several kidney cancer trials have reported disappointing results for TopII inhibitors (*Escudier et al.* 2002; *Law et al.* 1994), these trials have predominantly recruited patients with clear cell RCC, a genetically distinct disease. In further support of this suggestion, a previous study reports that this gene is the most overexpressed gene in pediatric Wilms' tumour, for which current therapeutic regimens consisting primarily of TopII inhibitors are very effective(*Takahashi et al.* 2002).

Clonal origin versus progression. It has been hypothesized previously based on cytogenetic findings that type 1 tumours progress to type 2 tumours(*Gunawan et al.* 2003). Prudent evaluation of our results in the context of this hypothesis is required. Although microarrays of gross tumour tissue show a global expression signature presumably reflective of early clonal events (*Ramaswamy et al.* 2003), it is plausible that a competitive growth advantage

may accrue to the transformation of a single cell into a class 2 within a class 1 tumour, resulting in its expansion at the expense of other class 1 tumour cells. Nonetheless, the additional presence of a distinct group of mixed tumours with coexisting type 1 and 2A histology and presenting with molecular profiles resembling other type 1 tumours strongly suggests that type 1 and 2A tumours are clonally more closely related to each other than to type 2B tumours. We did not note the presence of low-grade components in any of our type 2B tumours. Given the divergent survival outcomes following nephrectomy between class 1 (type 1, type 2A, and mixed type 1/2A tumours) and class 2 tumours, we do not favor the idea of progression between class 1 and 2 tumours.

In terms of specific limitations to this study, the relatively small sample size requires subsequent validation in a larger study. Additionally, future pathway analysis is likely to benefit from novel methods to control for multiple testing in pathway analysis. Certainly the findings of dysregulation of G_1-S checkpoint genes in class 1 papillary RCC and dysregulation of G_2-M checkpoint genes in class 2 papillary RCC should be deemed to be exploratory, requiring validation in a second study.

CONCLUSION

In conclusion, using gene expression profiling supported by immunohistochemical and morphologic studies, we have identified two distinct classes of papillary RCC that differ strikingly in their clinical behavior and have dysregulation of genes controlling different parts of the cell cycle. This finding

represents a biologically and clinically relevant refinement to previously proposed morphologic criteria for subclassification of papillary RCC. We summarize our findings that may be practically evaluated in the clinical setting laboratory as follows: class 2 (type 2B) papillary RCC may be distinguished from class 1 (type 1, mixed type 1 and 2A, and type 2A tumours) by the following characteristics: larger gross tumour size, higher nuclear grade (3-4), decreased CK7 staining and increased Top II alpha staining, higher rate
of metastases at surgery, and poorer patient survival. Morphologic findings of less specificity include larger cell size and eosinophilic cytoplasm in class 2 tumours. Our findings may benefit further efforts to elucidate the molecular basis of development and progression of papillary RCC and will be helpful in stratifying patients for additional interventions.

CHROMOPHOBE RCC AND ONCOCYTOMA

Considering the pathological and clinical issues for chromophobe RCC and oncocytoma discussed in Overall Background and Background (Molecular Models), our goal was to perform a comprehensive characterization of both enetities by integrating gene expression and high resolution single-nucleotide polymorphism (SNP) profiling for the identification of a useful and valid molecular predictor. We also sought to identify novel immunohistochemical markers for each entity. Further to identifying novel cytogenetic abnormalities underlying the two entities, we chose to investigate further the specific genes and pathways that may be responsible for dysregulation of gene expression.

METHODS

SUBJECTS AND STUDY DESIGN

A total of 30 frozen primary kidney tumours (15 chromophobe RCC and 15 oncocytomas) were obtained from the French Kidney Tumours Consortium, University of Chicago, Northwestern University, and Spectrum Health Hospital (Grand Rapids, MI). 12 non-neoplastic kidney samples were also obtained as controls for regional expression bias analysis.

Written informed consent for analysis of clinical samples was obtained from all patients, and all IRB boards of participating institutions approved the study. Tumour tissue was flash frozen in liquid nitrogen immediately after nephrectomy and stored at –80°C. Portions of the tumours were fixed in buffered formalin. Each sample was confirmed by pathologic analysis and anonymized

prior to the study. A portion of the tumour sample was frozen in liquid nitrogen immediately after surgery and stored at –80˚C. Total RNA was isolated from the frozen tissues using Trizol reagent (Invitrogen, Carlsbad, CA). This was subsequently purified with a RNeasy kit (Qiagen, CA), and quality was assessed on denaturing gel electrophoresis. Representative tissue sections of RCC were cut and stained with hematoxylin-eosin for confirmation of histological diagnosis and confirmation of tumour tissue content (>70%). All samples were examined by a central expert uropathologist (X.Y.). Clinicopathologic information was derived by review of pathologic, radiologic, and case notes by individual clinicians. Tumours with sarcomatoid change were classified as grade 4 tumours. Tumour size was defined as the maximum tumour dimension on direct pathological measurement. Patient follow-up status was assessed by directly contacting patients where possible, or at routine clinical follow-up. 12 non-neoplastic kidney samples from unrelated patients were also obtained as controls for regional expression bias analysis; these are the same 12 controls for regional expression bias used for the other microarray studies in other subtypes in this thesis.

OLIGONUCLEOTIDE ARRAY PROFILING

For oligonucleotide array gene profiling, we extracted and purified total RNA from homogenized samples using Trizol reageant (Invitrogen, CA) followed by RNeasy columns (Qiagen, CA) according to manufacturers' recommendations. Oligonucleotide array gene profiling was performed using the manufacturer's recommended protocol (GeneChip Expression Analysis Technical Manual, April 2003, Affymetrix, CA). Approximately 5-20 micrograms of total RNA was used to

prepare antisense biotinylated RNA. A subset of cases were spiked with external poly(A) RNA controls. Synthesis of single-stranged and double-stranded cDNA was done with the use of T7-oligo(dT) primer (Affymetrix). In vitro transcription was done using Enzo Bioarray Transcript Labeling kit (Enzo, Farmingdale, NY). The biotinylated cRNA was subsequently fragmented and 10 micrograms were hybridized to each array at 45°C ovr 16 hours. We used the HGU133 Plus 2.0 GeneChip, containing 54,675 probe sets, representing approximately 47,000 transcripts and variants. Scanning was done in a GeneChip 3000 scanner. Quality indices reviewed for all samples included mean percentage present, mean background, mean scaling factor and a mean GADPH 3'/5' ratio). All clinical and microarray data for published data has been uploaded to the Gene Expression Omnibus. The gene expression data can be obtained at the Gene Expression Omnibus (GSE19982). For external validation in a dataset with oncocytoma and chromophobe data, an external GEO data-set of gene expression profiles of oncocytomas and chRCC from Cornell University was obtained for validation (GSE12090) (*Rohan et al.* 2006). Statistical analyses were performed in the statistical environment R 2.6.0, utilizing packages from the Bioconductor project. Data preprocessing was performed using the RMA method as implemented in the *affy* package and using updated probe set mappings such that a single probe set describes each gene (*Dai et al.* 2005) (*Gentleman et al.* 2004). Chromosomal abnormalities were predicted using the comparative genomic microarray analysis (CGMA) method as implemented in the *reb* package (*Furge et al.* 2005). Briefly, for each measured gene, the gene

expression value was normalized such that the average gene expression value in the nondiseased samples was subtracted from the tumour-derived gene expression value. A Welsh's *t*-test was applied to the relative gene expression values that mapped to each chromosome arm. For the smoothing curve, the normalized expression values derived from genes mapping to chromosome 19 were replaced by a summary score that comprised a running two-sided *t*-test statistic using window sizes of 61, 245, and 611 (representing 5%, 20%, and 50% of the length of the chromosome). The results of the three smoothing curves were averaged. For purposes of hierarchical analysis using complete linkage analysis, probe set filtering for coefficient of variation ($\geq$0.05, with at least 2 samples showing $\log_2$ value expression of 8) was performed. Significance analysis of microarrays (SAM) on unfiltered data based on two-class unpaired analysis, assumption of unequal group variances and 10,000 permutations was used to derive a list of probe sets differentially expressed between tumour subclasses, and ordered by relative fold-change. A maximum false discovery rate threshold was defined as 0.05.

PATHWAY ANALYSIS

KEGG pathway(*Kanehisa et al.* 2010) and gene ontology (GO) analysis of enriched gene sets was performed using hypergeometric tests available in the GOstats (*Falcon and Gentleman* 2007) package in Bioconductor after having identified unique genes with corresponding annotations. For KEGG pathway analysis, the p-value threshold was 0.01. For GO analysis, conditional testing

was performed, and the threshold for p was 0.001. Molecular function, biologic process, and cellular component analyses were performed.

SINGLE NUCLEOTIDE POLYMORPHISM ANALYSIS

An independent set of samples were obtained from the Cooperative Human Tissue Network (CHTN) for high throughput single-nucleotide polymorphism (SNP) array analysis (chromophobe RCC n=6 and oncocytoma n=8). A Jetquick DNA extraction kit (Genomed, Lohne, Germany) was used to isolate DNA based on manufacturer's protocol. The SNP assay was performed according to the manufacturer's instructions using the Affymetrix GeneChip mapping 100K array (Affymetrix, Santa Clara, CA). Image quantification was performed using a GeneChip Scanner 3000, and the resulting data was processed using GCOS 1.4 (Affymetrix). Allele calls were generated using Affymetrix GeneChip Genotyping analysis (GTYPE v.4) with a confidence threshold set at 0.25. Raw copy numbers in log-transformed format (non-paired reference and test samples) were exported from the CNAG version 2.0 (Affymetrix) software using normal references downloaded from Affymetrix (http://www.affymetrix.com;ccnt_reference_data). The NCBI human genome reference build 36 was used for analysis. DNA copy number changes were visualized using data smoothing in which raw copy number values were replaced by a summary score that comprised a running 1-sided t test statistic with window size set to 31, where each SNP probe along with 15 5' SNP and 15 3' SNPs were included in the window. DNA copy number data can be obtained at the Gene Expression Omnibus (GSE8271)

IMMUNOHISTOCHEMISTRY

IHC staining was performed on an independent set of chromophobe RCC (n=11) and oncocytomas (n=7). Aquaporin 6 and synaptogrin 3 were selected from the PAM (Table 1). Parafibromin (218578_at) (2-fold mean expression difference) and cytokeratin 7 (209016_s_at) were selected from the SAM analysis of the gene expression profiles for validation. Candidate marker choice was determined by factors including fold-change, specificity, biological and clinical interest. CK7, a previously recognized marker, was selected here for testing to ascertain the additional benefit of routine pathologic practice in the samples. Briefly, following blocking and antigen retrieval, 4-micron sections on coated slides were incubated with the following antibodies: a mouse anti-cytokeratin 7 monoclonal antibody (DakoCytomation, Carpinteria, CA, 1:50, cytoplasmic staining), a polyclonal rabbit anti-human aquaporin 6 (AQP6, Alpha Diagnostic International, San Antonio, TX, 1:100, overnight at 4°C, membranous staining), polyclonal goat anti-synaptogyrin 3 N-18 and C-18 antibodies (SYNGR3, Santa Cruz Biotechnology, Santa Cruz, CA, cytoplasmic and membranous staining), an mouse monoclonal anti-parafibromin antibody (1:250, 1 hour at room temperature, nuclear staining) (*Tan et al.* 2004a), a rabbit monoclonal anti-HER2 antibody (Neomarker RM 9103-S clone SP3, 1:200, membranous staining), a phospho-AKT (Ser473) antibody (Cell Signaling Technology, 1:30, cytoplasmic and nuclear staining). For the latter two antibodies, 22 chromophobe RCC and 8 oncocytoma specimens were available. For p-AKT, staining in the stromal and tumour cell compartments was separately

assessed. Subsequent reactions were performed with biotin-free HRP enzyme labeled polymer of EnVision Plus detection system (DakoCytomation). All slides were examined by a pathologist in a blinded fashion.

FLUORESCENT IN SITU HYBRIZATION AND WHOLE CHROMOSOME PAINTING

For fluorescent in-situ hybridization and whole chromosome painting, bacterial artificial chromosomes (BACs) RP11-157B13 (19p12), RP11-1137G4 (19p13.3), RP11-15A1 (19q13.31) were obtained from the Children's Hospital Oakland Research Institute and BAC CTC-429C10 (19q13.41) was purchased from Invitrogen (Invitrogen Corporation, Carlsbad, CA). These clones were labeled with either SpectrumGreen or SpectrumOrange (Abbott Molecular Inc, Des Plaines, IL) by nick translation and applied to tissue touch preps of oncocytoma samples as described, with the exception that slides were counterstained with VECTASHIELD (Vector Laboratories, Inc. Burlingame, CA) anti-fade 4′,6-diamidino-2-phenylindole (DAPI). Telomere-specific DNA probes, the chr 1,5,19 alpha satellite probe, and the arm-specific paints were purchased from Q-BIOgene (MP Biomedicals, Solon, OH). FISH was performed using these probes according to the manufacturer's supplied protocol. As the alpha satellite probe cross-hybridizes to chromosome 1 and chromosome 5, in all studies chromosome 19 was co-labeled with a probe that maps distal to the centromere, RP11-157B13 (19p12). In addition, analysis of the centromeric probe on the metaphase spreads of control cells revealed that hybridization to chromosome 1

resulted in a significantly brighter signal (data not shown). These hybridization characteristics allowed the discrimination between chr 1 and 5 cross-hybridization. For image quantification, three separate photomicrographs containing five, six, and three cells, respectively, in which the 19q31.31 FISH signals were in the same image plane were obtained. Photomicrographs were processed using the *rtiff* package for the R environment. The fluorescent FISH signals were automatically segmented from background using the method of Ridler and Calvard, individual spots were identified using the connected component algorithm, and the number of pixels per feature were calculated. Twelve doublet FISH signals and eight singlet FISH signals were compared. Differences in size were evaluated using a one-sided Student's *t*-test.

RESULTS

We visualized the 30 expression profiles of chromophobe RCC and oncocytoma by hierarchical clustering upon a filtered data-set of 8,995 transcripts. Clear partitioning of the two entities into separate classes was observed (Figure 18, next page) with PAM yielding excellent cross-validated discrimination over a series of thresholds.

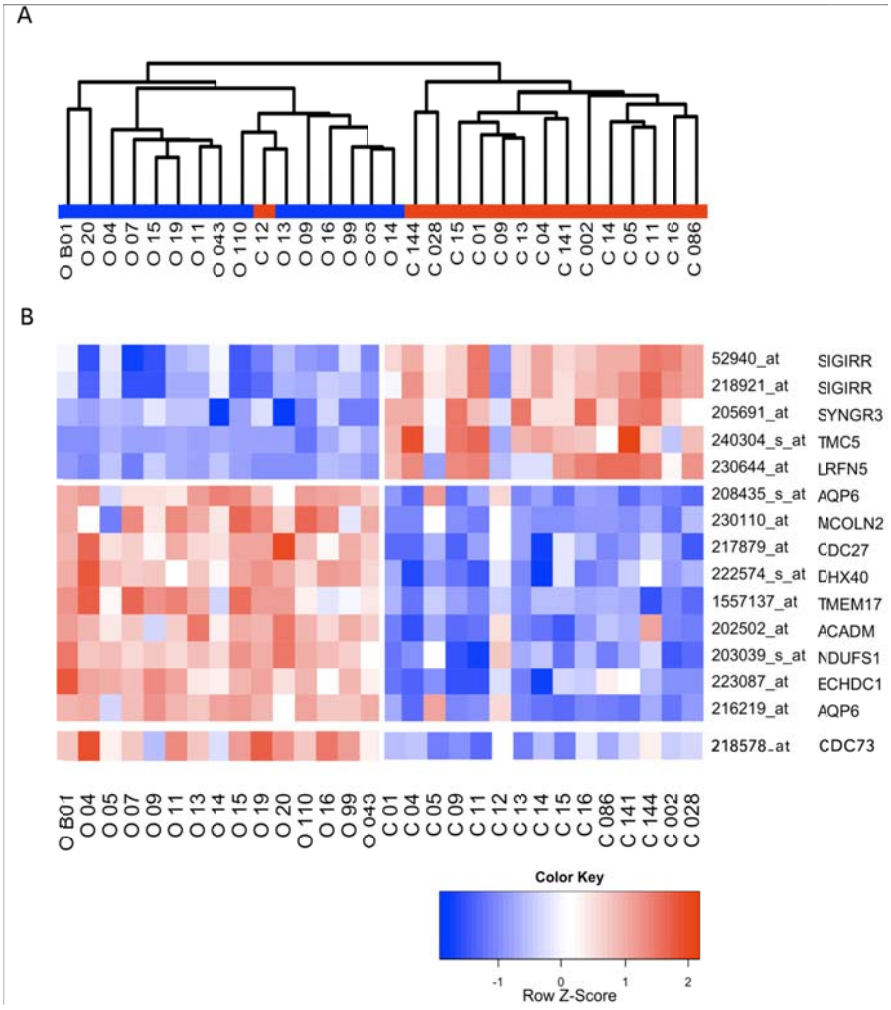

Figure 18 : Distinct clustering of gene expression profiles of chromophobe RCC and oncocytoma.
(A) A dendrogram showing hierarchical clustering of the filtered data showing clustering of oncocytoma and chromophobe RCC. The colour bar here separates oncocytoma [O] from chromophobe RCC [C]. (B) A heatmap of the predictor genes. Red denotes relative overexpression and blue denotes relative under-expression.

A gene predictor comprising 14 probe sets was identified (Table 15), which yielded an overall accuracy of 93% in the internal data-set (28/30) (Table 16, next page). The same predictor successfully classified 17 of 18 samples in the external dataset from Cornell University, corresponding to an overall accuracy of 94% (Table 15). 5,210 probe sets were found to be differentially expressed between the two entities as identified using SAM at a delta of 1.4, with a false discovery rate of 0.03 corresponding to an estimated 222 probe sets. 2,564 number of probe sets were relatively overexpressed in chromophobe RCC, and 2,646 transcripts relatively underexpressed in chromophobe RCC.

Table 15 : Predictor derived by nearest shrunken centroid methodology for sample classification of chromophobe RCC and oncocytoma

Affymetrix Probe ID	Gene description	ChRCC-score*	Oncocytoma-score*	Fold change**
216219_at	aquaporin 6	-0.1972	0.1972	0.20
240304_s_at	transmembrane channel-like 5	0.1247	-0.1247	13.0
208435_s_at	aquaporin 6	-0.1218	0.1218	0.22
205691_at	synaptogyrin 3	0.1108	-0.1108	3.75
230110_at	mucolipin 2	-0.0731	0.0731	0.15
52940_at	single immunoglobulin and toll-interleukin receptor (TIR) domain	0.0577	-0.0577	3.28
217879_at	cell division cycle 27 homolog (S. cerevisiae)	-0.0403	0.0403	0.
222574_s_at	DEAH (Asp-Glu-Ala-His) box polypeptide 40	-0.0228	0.0228	0.48

218921_at	single immunoglobulin and toll-interleukin receptor (TIR) domain	0.0205	-0.0205	3.16
230644_at	leucine rich repeat and fibronectin type III domain containing 5	0.0172	-0.0172	4.90
223087_at	enoyl Coenzyme A hydratase domain containing 1	-0.0167	0.0167	0.42
203039_s_at	NADH dehydrogenase (ubiquinone) Fe-S protein 1	-0.0093	0.0093	0.46
202502_at	acyl-Coenzyme A dehydrogenase	-0.0089	0.0089	0.35

* A class discriminant score derived from nearest shrunken centroids methodology.

** Fold change of gene expression in chromophobe RCC relative to oncocytoma

Table 16 : Predictor performance in sample classification in distinguishing chromophobe RCC and oncocytoma in internal and external datasets

		Gene predictor (14 probe sets)	
		Predicted chromophobe RCC	Predicted oncocytoma
Internal Data-Set	chromophobe RCC	13/15 (87%)	2/15 (13%)
	Oncocytoma	0/15 (0%)	15/15 (100%)
External Data-Set (Cornell)	chromophobe RCC	8/9 (89%)	1/9 (11%)
	Oncocytoma	0/9 (0%)	9/9 (100%)

Immunohistochemical findings The IHC profiling is summarized in Table 17 and Figure 19 (next page), results of which were consistent with the microarray quantitation. The immunoreactivity of chromophobe RCC to cytokeratin 7 was higher than that of oncocytomas and normal kidney. For parafibromin, clear differential staining was noted, with predominantly nuclear expression in oncocytomas, and absent expression in chromophobe RCC. For synaptogyrin-3, both N-18 and C-18 antibodies yielded a similar signal, but the N-18 antibody yielded a crisper result though the maximal signal was distinctly weaker compared to AQP6, for which crisp membranous staining was noted in oncocytoma, but not in chromophobe RCC. For p-AKT, there was an apparent but non-significant higher immunoreactivity in chromophobe RCC than oncocytoma, particularly in the stromal cells relative to the tumour cells. For extracellular HER2, all samples were unreactive.

Table 17 : Results of IHC staining showing sample discrimination between chromophobe RCC and oncocytoma.

	chromophobe RCC		Oncocytoma		
Protein	**Positive**	**Negative**	**Positive**	**Negative**	**P-value**
AQP6	3/11 (28%)	8/11 (72%)	6/7 (86%)	1/7 (14%)	0.05
Parafibromin	1/11 (9%)	10/11 (91%)	5/7 (71%)	2/7 (29%)	0.01
CK7	8/11 (72%)	3/11 (27%)	1/7 (14%)	6/7 (86%)	0.05
SYNGR3	9/11 (82%)	2/11 (18%)	0/7 (0%)	7/7 (100%)	0.002
p-AKT (stromal)	5/22 (28%)	17/22 (72%)	0/8 (0%)	8/8 (100%)	0.29
p-AKT (tumour)	13/22 (59%)	8/22 (41%)	4/8 (50%)	4/8 (50%)	0.68
Extracellular HER2	0/22 (0%)	22/22 (100%)	0/8 (0%)	8/8 (100%)	NA

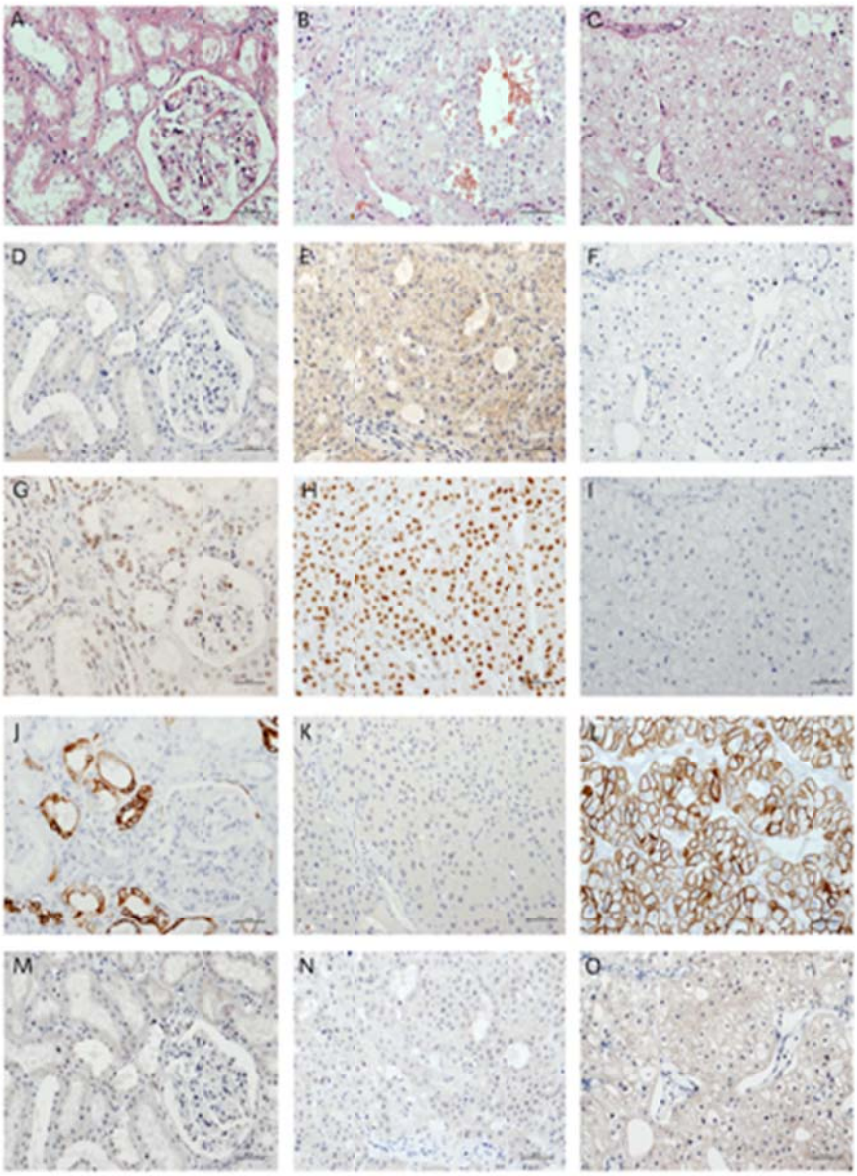

Figure 19: Immunohistochemical profiling of renal oncocytoma and chromophobe RCC. (A) – (C) Hematoxylin and eosin stains of normal cortical kidney tissue, oncocytoma and chromophobe RCC respectively; (D) – (F) Aquaporin 6 immunostaining showing membranous staining in oncocytoma but absent staining in chromophobe RCC; (G) – (I) Parafibromin immunostaining showing strong nuclear expression in oncocytoma and tubular epithelium but absent staining in chromophobe RCC; (J) – (L) Cytokeratin 7 immunostaining showing distinct cytoplasmic staining in chromophobe RCC but absent staining in oncocytoma ; (M) – (O) Synaptogyrin 3 immunostaining showing cytoplasmic staining in chromophobe RCC but absent staining in oncocytoma.

Pathway Analysis. Pathway and GO analysis was performed on the SAM analysis, demonstrating an enrichment of genes involved in metabolic pathways in oncocytomas relative to chromophobe RCC (Table 18, next page). These metabolic pathways include oxidative phosphorylation, amino acid metabolism, and fatty acid metabolism. Conversely, high expression of genes involved in cell adhesion, immune receptor signaling as well as proliferative pathways such as c-erbB2 (Her-2/neu) and mammalian target of rapamycin (mTOR) signaling are detected in chromophobe RCC. Further gene ontology analyses (not presented) performed supported these results highlighting that mitochondrial genes were highly overrepresented among genes relatively overexpressed in oncocytomas, whereas tight junction genes were similarly overrepresented among genes overexpressed in chromophobe RCC.

Table 18 : Molecular pathways discriminating chromophobe RCC and oncocytoma

Pathways relatively upregulated in oncocytoma

KEGGID	Pvalue	OddsRatio	ExpCount	Count	Size	Term
280	0	6.206	4	17	44	Valine, leucine and isoleucine degradation
640	0	6.349	3	13	33	Propanoate metabolism
190	0	3.093	11	27	114	Oxidative phosphorylation
970	0	4.488	4	12	38	Aminoacyl-tRNA biosynthesis
20	0.001	4.833	3	9	27	Citrate cycle (TCA cycle)
330	0.001	4.032	3	10	34	Arginine and proline metabolism
4120	0.003	3.13	4	11	45	Ubiquitin mediated proteolysis

Pathways relatively upregulated in chromophobe RCC

KEGGID	Pvalue	OddsRatio	ExpCount	Count	Size	Term
4660	0	3.263	9	23	93	T cell receptor signaling pathway
4662	0	3.945	6	18	63	B cell receptor signaling pathway
4514	0	2.505	12	26	129	Cell adhesion molecules (CAMs)
4670	0	2.676	10	23	108	Leukocyte transendothelial migration
5220	0	3.05	7	18	76	Chronic myeloid leukemia
5212	0	2.977	7	17	73	Pancreatic cancer
4520	0.001	2.823	7	17	76	Adherens junction
5130	0.001	3.335	5	13	51	Pathogenic Escherichia coli infection - EHEC
5131	0.001	3.335	5	13	51	Pathogenic Escherichia coli infection - EPEC
4530	0.001	2.277	11	22	117	Tight junction
4664	0.001	2.65	7	16	75	Fc epsilon RI signaling pathway
4620	0.003	2.268	10	19	101	Toll-like receptor signaling pathway
4012	0.003	2.372	8	17	87	ErbB signaling pathway
564	0.003	2.621	6	14	66	Glycerophospholipid metabolism
4150	0.004	2.964	4	11	47	mTOR signaling pathway
5120	0.004	2.523	6	14	68	Epithelial cell signaling in Helicobacter pylori infection
4210	0.005	2.293	8	16	84	Apoptosis
4070	0.006	2.317	7	15	78	Phosphatidylinositol signaling system
4540	0.007	2.124	9	17	95	Gap junction
4912	0.009	2.07	9	17	97	GnRH signaling pathway
5221	0.01	2.536	5	11	53	Acute myeloid leukemia

DNA copy number profiling and CGMA

We report copy number gains that were detected in chromosomes 4, 7, 11, 12, 14q, and 18q(Figure 20). Chromosomal losses of chromosome 1, 2, 6, 10, 13, 17, and 21 represent a unique copy number loss profile for chromophobe RCC. For renal oncocytoma, losses of chromosome 1p were noted.

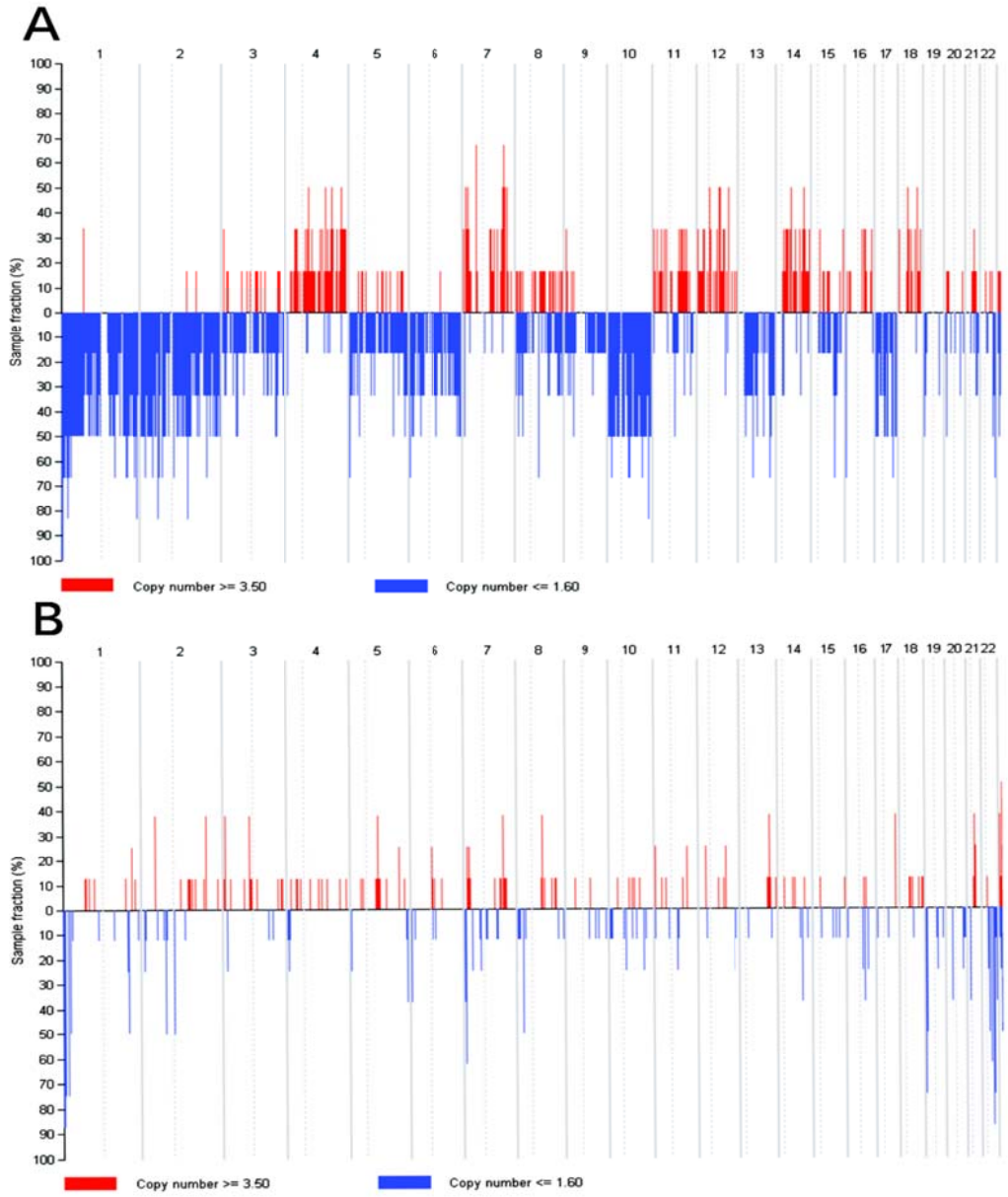

Figure 20: High throughput SNP analysis in chromophobe RCC (above) and oncocytoma (below). The percentage sample fraction with chromosomal alteration is reported. Multiple chromosomal copy number alterations are visualized in chromophobe RCC but not in oncocytoma. Red denotes gain in copy number, and blue denotes loss. The number running along the top represents the chromosome.

A CGMA (Figure 21) derived from the expression data yielded regional expression biases consistent with that reported by copy number profiling.

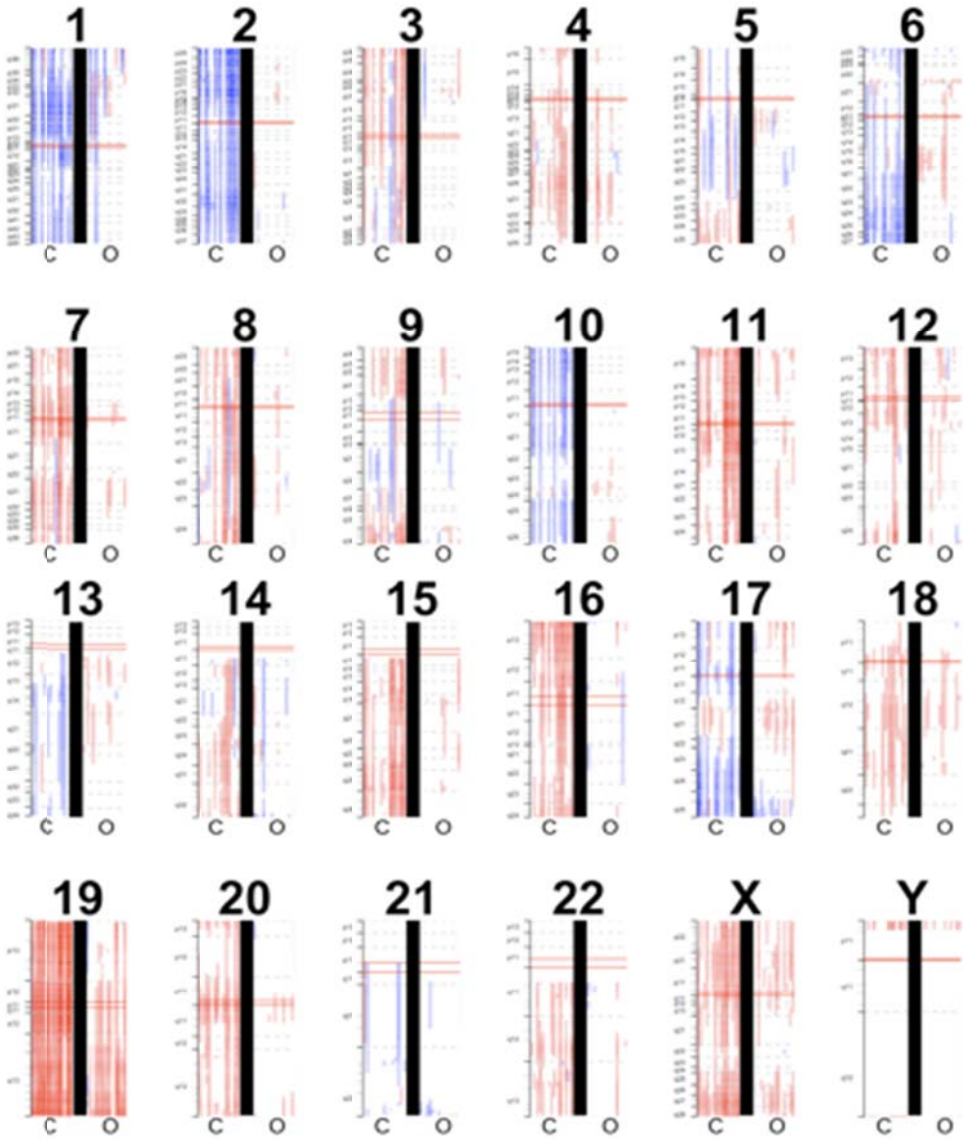

Figure 21 : Chromosomal ideograms showing regional gene expression biases in chromophobe RCC and oncocytoma.
For each ideogram, a tumour is represented by an individual vertical bar with chromophobe RCC (C below the ideogram, left) and oncococytoma (O below the ideogram, right). Red represents regional upexpression, and blue regional downexpression. E.g. chromosome 2 shows regional downexpression throughout for all chromophobe RCC samples, but not oncocytoma.

We report in particular that our high throughput methods demonstrate that there is a common gene alteration to both tumours (loss of chromosome 1p), which may represent an early event common in the histogenesis of both tumours. Otherwise, consistent with previous cytogenetic studies, the renal oncocytoma cells were largely devoid of transcriptional abnormalities that would reflect a DNA amplification or deletion. In contrast, losses of chromosomes 2, 6, 10, and 17 are frequently found in chromophobe RCC. In our chromophobe RCC samples, these well-established chromosomal losses were strongly reflected in the gene expression profiling data. In addition, a transcriptional abnormality involving genes mapping to chromosome 19 was frequently identified in both the renal oncocytomas and the chromophobe RCCs. In renal oncocytomas, the transcriptional abnormality primarily involved the q arm of chromosome 19, while in chromophobe RCC the abnormality involved the entire chromosome (Figure 22, next page). For each gene on chromosome 19, the average $\log_2$-transformed expression ratio comparing oncocytoma or chromophobe RCC to non-diseased kidney was plotted relative to genomic location (Figure 22, next page). A smoothing curve was fit to the $\log_2$-transformed data to highlight regions that contain a disproportionate number of up-regulated genes.

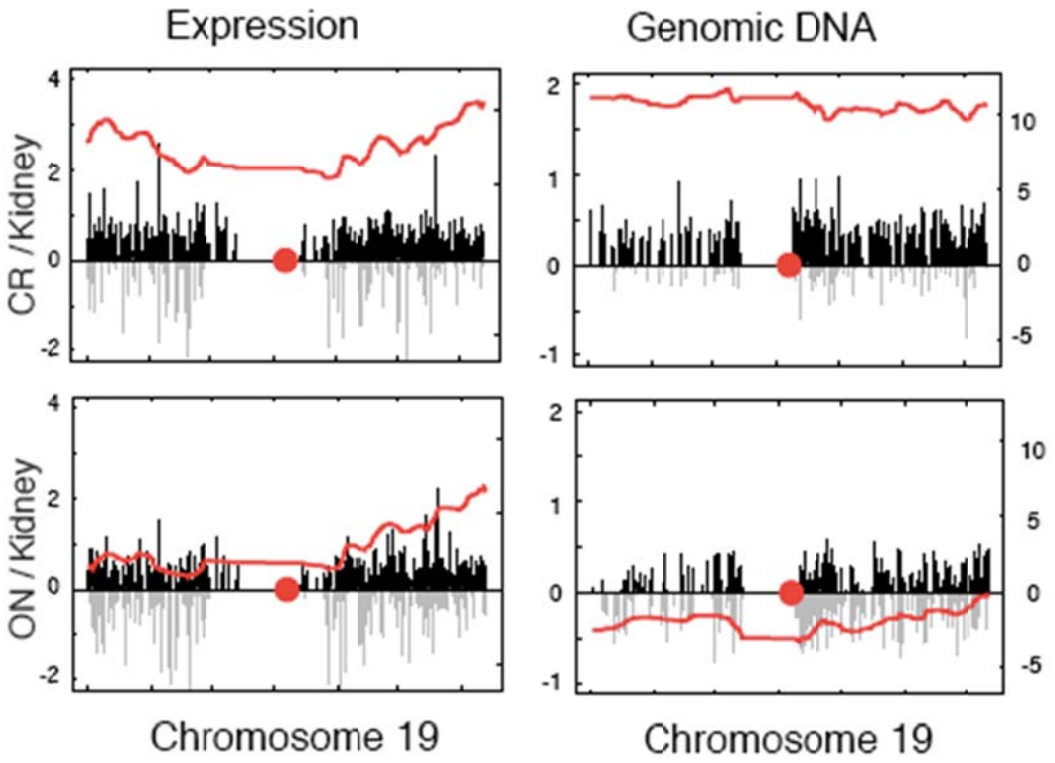

Figure 22 : Depiction of the transcriptional changes along chromosome 19, and corresponding copy number profiles of chromophobe RCC and oncocytoma.
The image shows the presence of chromosomal amplification for chromosome 19q in chromophobe RCC, but not in oncocytoma, but a similar 19q regional increased expression in both pathologic entities. The red circle indicates the location of the centromere; the p arm on the left and q arm on the right. For each SNP on chr 19, the average $\log_2$-transformed DNA copy number ratio comparing oncocytoma to a pooled normal reference is plotted relative to genomic location as described.

From the SNP data, an amplification of the entirety of chromosome 19 was detected in the chromophobe RCC samples (Figure 22). This whole-chromosome amplification was confirmed by fluorescence in-situ hybridization (FISH) using locus-specific probes that mapped to the p and q arms of chromosome 19 (Table 19, next page). In contrast, no change in DNA copy number was detected in the renal oncocytoma samples.

Table 19 : Chromosome 19 FISH patterns in chromophobe RCC

Cell type	Probe	Location	Signal 1*	Signal 1	Signal 2	Signal 3+	Other	Cells Counted
Kidney	RP11-1137G4	19p13.3	5.5%	5.0%	**88.5%**	2.0%	NC	200
Kidney	RP11-15A1	19q13.31	6.5%	2.5%	**90.0%**	1.0%	NC	200
CR 86	RP11-1137G4	19p13.3	0.0%	0.0%	40.0%	**48.0%**	NC	52
CR 86	RP11-15A1	19q13.31	8.0%	0.0%	40.0%	**48.0%**	NC	52
CR 15	PTEL 19P	19p tel	0.0%	0.0%	13.0%	**87.0%**	NC	100
CR 15	PTEL 19Q	19q tel	0.0%	0.0%	13.0%	**87.0%**	NC	100
CR 9	RP11-37G4	19p13.3	2.0%	4.0%	65.0%	**29.0%**	NC	100
CR 9	RP11-15A1	19q13.31	7.0%	3.0%	58.0%	**31.0%**	NC	100

[a]Tumor touch preps were labeled with the indicated probes and the number of signals per cell were counted. NC indicates not counted. CR indicate chromophobe RCC and numbers indicate samples. Results from normal kidney are also shown (Kidney) An asterisks (*) indicates a large signal or two signals in close proximity

FLUORESCENT IN SITU HYBRIDIZATION AND WHOLE CHROMOSOME PAINTING

To determine the status of chromosome 19 in more detail in the renal oncocytoma cells, this chromosome was evaluated further using a panel of FISH probes. Two distinct and well-separated FISH signals, typical of diploid cells in interphase, were frequently observed when probes specific to the chr 19p arm were used (Figure 23, next page). In contrast, a single, large FISH signal (singlet) or two FISH signals that were in close proximity (proximal doublet) were frequently observed when probes specific to the chr 19q arm were used. Approximately 35% of cells examined contained proximal doublets (data not shown).

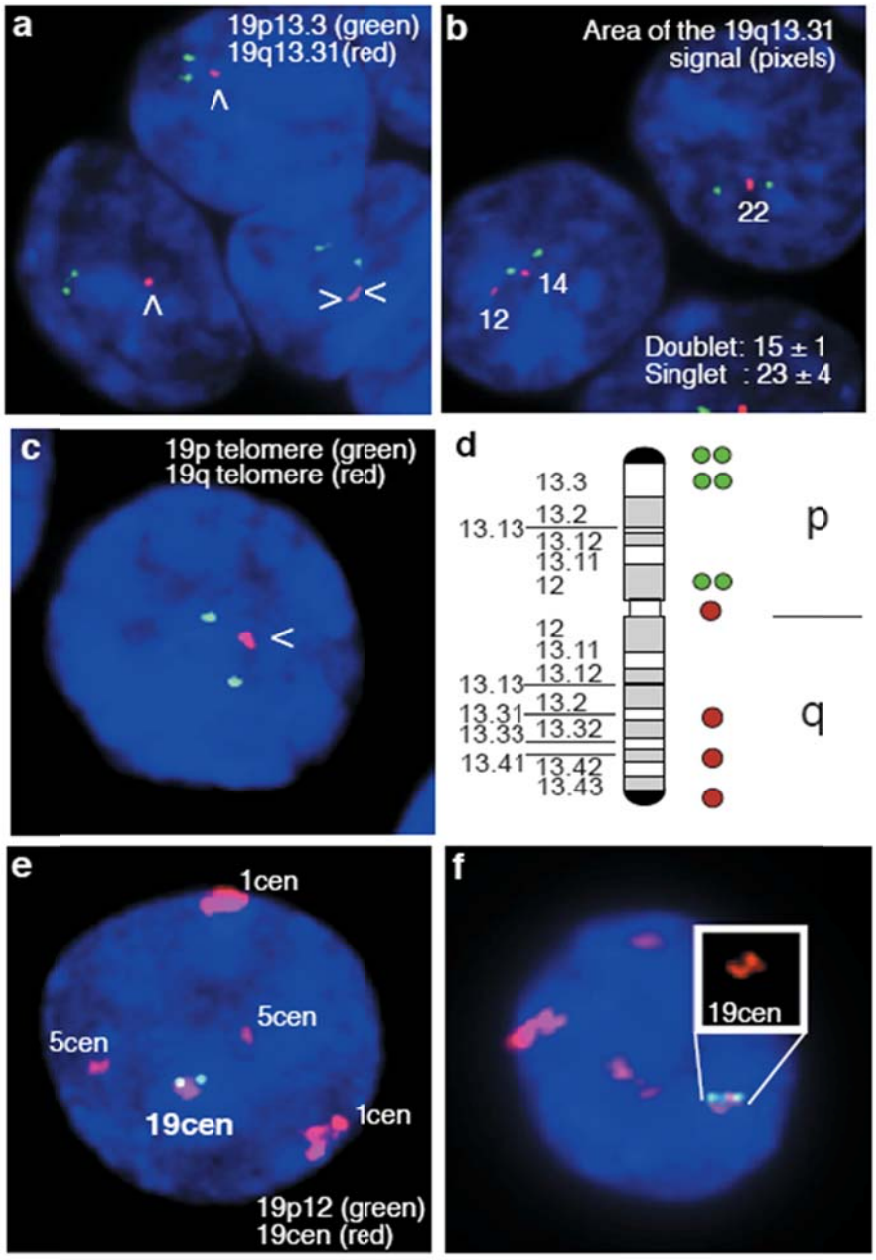

Figure 23 : Somatic pairing in renal oncocytoma.
Representative photomicrographs of tri-colour interphase FISH on renal oncocytoma touch preps. White arrows indicate large singlet or proximal doublet signals. In all images DAPI counterstaining is shown in blue. (A,B) Labeling with 19p13.3 (green) and 19q13.31 (red) probes. Image area of the 19q13.31 signal was quantified across multiple cells (n = 25) in the same image plane. Area mean and standard error are shown. (C,E,F) Labeling with 19p telomere (green) and 19q telomere (red) probes or 19p12 (green) and an alpha satellite probe for chr 19 that also cross-hybridizes to chromosomes 1 and 5 (red). Inset highlights centromeric pattern. (D) Schematic representation of frequently observed FISH patterns.

Semi-quantitative image analysis was used to examine the characteristics of the large FISH singlet (Figure 23, preceding page) This analysis demonstrated that the size of the singlet FISH signal was on average 1.5-fold larger than the size of two well-separated 19q FISH signals (P = 0.02). This large signal was observed using multiple probes directed against the q arm of the chromosome, including centromeric and telomeric probes (Figure 23, preceding page). The large FISH singlet had striking similarities to the FISH signals observed in studies of somatically paired chromosomes. Somatic pairing refers to the close association of homologous chromosomes and is typically associated with chromosomes in meiotic prophase. However, somatic pairing has also been observed in interphase in normal human cells and some tumour cells. The presence of a large FISH singlet reflects the overlapping FISH signals generated from two chromosomal regions in very close proximity. The lack of evidence for a DNA copy number change coupled with the presence of large FISH singlets and proximal doublets using multiple locus-specific probes, suggested that chr 19q was somatically paired. To confirm that the q arms of chr 19 were somatically paired in the renal oncocytoma cells, the p and q arms of chr 19 were visualized simultaneously using whole-arm chromosome painting (WCP) (Figure 24, next page).

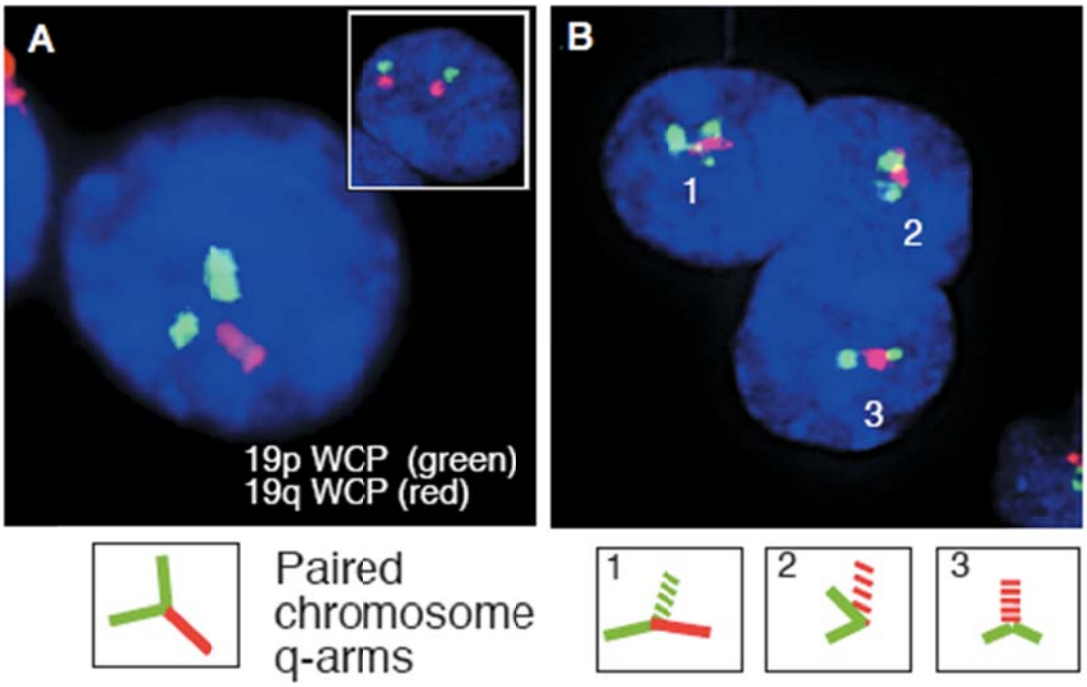

Figure 24 : Whole-arm chromosome paint (WCP) for chromosome 19 in oncocytoma. This image shows the presence of somatic pairing of chromosome 19q. The p arm is in red and the q-arm in green. The inset shows a normal cell. Schematic representations of the paired chromosomes are shown below. Dashed lines represent chromosomal regions perpendicular to the plane of the image.

Using this approach, two distinct p arms, typical of diploid cells in interphase, were frequently observed in renal oncocytoma cells (Figure 24). However, the majority of cells contained a single q-arm signal that was located proximal to the two p-arm signals. While the diffuse nature of the WCP prevented the quantification the fluorescence signal, this pattern is consistent with the locus-specific FISH analysis and further indicates that the q arms of the chromosomes are in close proximity or are paired in these cells.

DISCUSSION

ChRCC and oncocytoma are morphologic and genetically related entities, and distinction between these two tumours is important because of their different biological behaviors. However, these entities can be difficult to distinguish morphologically. We report the derivation of a novel and useful gene predictor validated both on an internal and an independent external data-set, implying its generalizability. Our results suggest that it is possible to classify accurately histopathologically challenging tumours. The degree of accuracy achieved at 93% is reasonable for a genetic classifier. However integration into clinical practice requires a comprehensive evaluation of these classifiers within a clinical setting, comparing clinical outcomes in routine pathologic evaluation relative to that derived from novel classifiers. This may be most practically if not most ideally done in a retrospective fashion on paraffin-embedded tissue in a large multi-institutional collaboration, which we are currently pursuing. This issue may become progressively more important with the increase in incidentally detected small tumours on radiologic surveillance, where the dilemma between observation or intervention is commonly posed.

Novel cytogenetic alterations both common and discriminating. Integrating RNA and DNA genomic data allows us to verify genomic alterations in tumour samples and distinguish the genomic signatures of different tumour subtypes. Frequent losses of chromosome 1, 2, 6, 10, 13, 17, and 21 and gains in chromosome 4, 7, 11, 12, 14q and 18q were observed in chromophobe RCC, consistent with previously reported data(*Bugert and Kovacs* 1996). For renal

oncocytoma, we show a high prevalence of chromosome 1p loss. Both chromophobe RCC and oncocytoma share this chromosomal alteration, consistent with a speculation that this may represent an early event in neoplastic transformation of a common progenitor cell.

Chromosome 1p loss represents a common cytogenetic alteration in both chromophobe RCC and renal oncocytoma identified by high-throughput SNP studies. This may suggest that this is an early event in the histogenesis of both tumours, before additional cellular events lead to malignancy in lesions that progress to chromophobe RCC, similar to chromosome 3p loss in clear cell RCC, which is thought to be an early event in carcinogenesis. Loss of chromosome 1p has been identified recently in renal oncocytoma(*Picken et al.* 2008), but this has not been previously shown to be a common cytogenetic alteration common to both entities, which is the key insight. Our delineation of the nature of chromosome 1p loss in renal oncocytoma provides the opportunity to identify novel tumour suppressor genes in future studies, and in establishing a possible carcinogenesis progression sequence.

Pathway identification in oncocytoma and chromophobe RCC. There has been a recent advent of targeted therapies for a wide variety of cancers. Given the relative rarity of chromophobe RCC, there is no current standard of care and it is unlikely that any specific clinical trial is feasible or will be initiated. Here, we report two clinically relevant pathways—the c-erbB2/HER2 pathway and the mTOR signaling pathway—are dysregulated in chromophobe RCC on exploratory pathway analysis of mRNA expression, but our evaluation of

cxtracellular HER2 and phospho-AKT IHC expression has not provided direct support for this mRNA finding. On a clinical trial level, in a subgroup analysis of a Phase III trial of temsirolimus, an mTOR inhibitor, in poor-prognosis RCC of all subtypes, patients of non-clear cell histology benefited as much as patients with clear cell histology, if not more(*Hudes et al.* 2007). Our findings do not permit a single definitive conclusion about the nature of pathway activation in these two entities. Currently, mTOR inhibitors remain a clinical standard of care for poor-risk metastatic non-clear cell RCC. HER2 expression has been evaluated in chromophobe RCC and oncocytoma, with distinct patterns of peptide expression varying according to epitope(*Seliger et al.* 2000). Interestingly, this study showed that strong intracellular HER2 expression (as defined by a 3+ expression) was strongly expressed in chromophobe RCC (9/19) but not in oncocytoma (1/11), whereas neither chromophobe RCC nor oncocytoma showed strong extracellular HER2 expression. Further evaluation of this is warranted, in conjunction with relevant fluorescent in-situ hybridization studies.

It has been previously reported that oxidative phosphorylation and energy pathway genes are overexpressed in chromophobe RCC and renal oncocytoma relative to the other subtypes of RCC(*Schuetz et al.* 2005). We are able to clarify this issue, demonstrating that even between these two entities, there are major differences in quantitative expression of the same pathways discriminating the two entities. Consistent with these results, it has been recently reported that oncocytomas exhibit mitochondrial DNA mutations with clonal expansion and complex I deficiencies(*Gasparre et al.* 2008; *Mayr et al.* 2008). Oncocytoma

contains a large number of mitochondria, and the overexpression of these genes involved in cellular metabolism may reflect the relative quantitative excess of the mitochondria. A similar profound modification in energy metabolism genes has been observed in thyroid oncocytomas(*Baris et al.* 2004), with high activity of the aerobic respiratory pathway. It may be speculated that potential inhibition of autophagy in the chromophobe RCC may correspond to this difference as well. Rohan et al have previously reported in a smaller data-set that gene expression profiling is able to discriminate oncocytomas and chromophobe RCC(*Rohan et al.* 2006), and has reported that vesicular transport and cell junction proteins are relatively upregulated in chromophobe RCC.

Novel biomarker identification discriminating chromophobe RCC and oncocytoma. In the process of validating our high-throughput expression studies, we report three novel markers discriminating between chromophobe RCC and oncocytoma: parafibromin, aquaporin 6, and synaptogyrin 3. Parafibromin, the protein product of the *HRPT2* tumour suppressor gene, has been reported to be downregulated in a variety of tumours(*Selvarajan et al.* 2008; *Tan et al.* 2004a), and a role has been assigned to it in the Wnt signaling pathway(*Mosimann et al.* 2006). While the mechanism of parafibromin downregulation in parathyroid carcinoma appears to be mediated through gene mutation, this does not seem to be the mechanism in chromophobe RCC, as we have not identified any *HRPT2* mutations after analyzing DNA samples from 5 chromophobe RCC tumours (data not shown). Similarly, other investigators have reported allelic imbalances in the HRPT2 gene in oncocytoma and chromophobe

RCC, but no mutations(*Zhao et al.* 2007). Aquaporin 6 is an intracellular vesicle water channel protein reported to be expressed in the intercalated cells of the collecting duct(*Yasui et al.* 1999), which is hypothesized to be the originating cell for oncocytoma and chromophobe RCC. Little is known about synaptogyrin-3, a tyrosine-phosphorylated protein that is expressed in synaptic vesicles(*Belizaire et al.* 2004). The reasons underlying the reduced expression of aquaporin 6 and increased expression of synaptogyrin-3 in chromophobe RCC, relative to oncocytoma are uncertain.

Based on our transcriptional and genomic DNA studies, we reveal somatic pairing of chr 19q as a recurrent cytogenetic abnormality in renal oncocytoma that results in dramatic changes in transcription within the paired region. Somatic pairing refers to the close association of homologous chromosomes and is typically associated with chromosomes in meiotic prophase. However, somatic pairing has also been observed in interphase in normal human cells and some tumour cells (*Atkin and Jackson* 1996; *Brown et al.* 1994; *Lewis et al.* 1993; *Zhang et al.* 1997). The functional consequence of chromosome joining is formally unknown but it is may disrupt chromatin structure causing the juxtaposition of *cis* and *trans* regulatory regions that modulate the transcription of a large set of genes. The changes in gene expression that accompanied the somatic pairing suggested that deregulation of a gene, or multiple genes, associated with tumour development mapped within the paired chr 19q region. The identification of *EGLN2* as a significantly deregulated gene that maps within the paired chromosome 19q region directly implicates defects in the oxygen-

sensing network to the pathobiology of renal oncocytoma. These results suggest that in addition to numerical and structural chromosomal abnormalities, somatic pairing should be considered as a chromosomal event that associates with tumourigenesis.

Hypoxia-response dysregulation in chromophobe RCC and oncocytoma. A proper oxygen-sensing response is vital to the maintenance of normal cellular functions. Deregulation of *HIF*, the principal driver of the adaptive response to hypoxia, is associated with the pathogenesis of several diseases, including cancer. While the hypoxic tumour microenvironment - by the virtue of the ubiquitous oxygen-sensing pathway - results in modulation of HIF activity, loss-of-function mutations in a growing list of tumour suppressor genes also can affect HIF function. Mutations in *PTEN*, *PML*, *TSC*, and *VHL* have been identified in tumour cells that result in the deregulation of HIF via multiple distinct mechanisms involving Akt/PI3K, mTOR and the ubiquitin pathway. Emerging evidence now implicates cancer-causing mutations that directly impinge on EGLNs. For example, mutations in succinate dehydrogenase (*SDH*) result in the cytosolic accumulation of succinate, which inhibits EGLNs, leading to the stabilization and activation of *HIF-1α*. Inactivating germline mutations in *EGLN1* have been identified to cause erythrocytosis (*Percy et al.* 2006; *Takeda et al.* 2008) and deregulation of *EGLN3* has been linked to the development of pheochromocytoma, a neuroendocrine tumour of the adrenal glands(*Lee et al.* 2005).

The disruption of HIF activity has been associated with kidney cancer related to VHL disease, sporadic clear cell RCC, and hereditary papillary RCC(*Isaacs et al.* 2005; *Pollard et al.* 2005; *Tomlinson et al.* 2002).

CONCLUSION

In conclusion, we have comprehensively characterized the molecular proiles of chromophone RCC and oncocytoma using high throughput expression and SNP profling. We have consequently derived discriminating expression signatures, pathways, cytogenetic profiles and protein markers that are of biologic, clinical and therapeutic interest. The present study reveals deregulation of the oxygen-sensing response in renal oncocytoma, as well as chromophobe RCCs (which display DNA amplification mediated up-regulation of *EGLN2*) and thereby supporting the dysfunction of HIF pathway as a common and perhaps central theme in the pathogenesis of kidney cancer.

OVERALL LIMITATIONS

Given that multiple studies in this thesis rely on the use of high throughput technology such as gene expression profiling and single nucleotide polymorphism analysis, it is important to evaluate novel issues and considerations of this technology in common. Here I present several limitations common to all the studies in the Molecular Models section of this study, from both a classical as well as from a pathological viewpoint.

STUDY DESIGN

I will broadly focus on the issues of study design and validity in this discussion of the epidemiologic issues generated within the work on the molecular models, rather than focus on individual limitations arising within each individual paper. Describing these high-level overall limitations here also aids in reducing unnecessary duplication in the various sections. These issues are primarily related to the high-throughput technologies employed, before continuing with discussion of the individual issues in each paper. In essence, standard epidemiologic issues remain as relevant, if not more so, in this era of high throughput technology.

VALIDITY

In terms of internal validity, many conclusions for each study was generated from an initial sample set of frozen tissue samples from a wide range of centres that were made available to the Van Andel Research Institute for

research. As such, the issue of selection bias needs to be comprehensively discussed. Frozen tissue samples are relatively inaccessible. Legacy samples are often archived in tissue banks overseen by pathologists, and often with no clinical data available. Hence, while modern tissue banks have been addressing these issues, these are necessary and practical limitations to working with retrospective tissue samples of the past (and thus with long follow-up). While we were careful to define the time period for samples, it is necessary to accept that there may be some bias introduced in the assessment of the importance of these genetic markers by a retrospective design. Hence, for all studies, it was critical to introduce a degree of external validation by an independent sample set. For the study focusing on clear cell RCC, in addition to our internal sample set of primary clear cell RCC samples, our results were separately validated in terms of (i) an independent Swedish cDNA microarray data-set, across platforms (*Larkin et al.* 2005) (ii) immunohistochemistry on an independent paraffin-embedded tissue set. This was validated in terms of immunohistochemistry on an independent paraffin-embedded tissue set. For the chromophobe RCC study, this was validated in terms of (i) a separate oligonucleotide array data-set of corresponding samples that is openly accessible and (ii) an independent set of tissue samples. With more sample sets becoming available in the foreseeable future (albeit largely selected hospital-based samples), external validation studies will certainly become easier.

EXPOSURE ASSESSMENT

In terms of analytic validity, our study uses primarily high throughput techniques to elicit genome-level biomarkers against conventional survival based outcomes. Analytic validity is a major concern of high throughput technology – in particular, issues of bias introduced at the various steps of experiment design, sample processing, measurement, quantification, scanning and data analysis(*Zervakis et al.* 2009). Further, on a meta-analytic level, the use of the same samples in different studies also may result in bias. Our studies use a common commercial platform – the Affymetrix U133Plus 2.0 microarray platform, with its own internal controls for sample quality (e.g. GADPH 3'-5' ratios). Multiple replicates, both biologic and technical, are required for reducing the noise generated by analysis of high-dimensional data such as the microarray data we generate here (*Larkin et al.* 2005). It should be noted that multiple probesets for each gene may exist on this platform, and that only a subset of these probesets may be identified in a predictor. It is unclear that summarizing the multiple probeset data into a single gene expression value is superior to interpreting individual probeset expression values.

It is well recognized that tumour tissue is heterogenous(*Liu* 2007). The studies here are based on microarrays of bulk tissue, inclusive of both tumour and stroma. Laser microdissection studies have shown that these two compartments differ in terms of gene expression profiles(*Gregg et al.* 2010). Nonetheless, it is important to note that the epidemiologic outcomes measured

here are primarily survival based, and thus, our primary interest is essentially external validation of a clinical based predictor. Indeed, the use of bulk tissue samples for microarray analysis may even yield insights, best seen in our work on clear cell RCC, where we suggest that the molecular determinants of prognosis may occur early in clonal development of a cancer cell, rather than arise later in individual subclones.

OVERALL CONCLUSIONS AND FUTURE RESEARCH

In this work, we evaluate multiple clinical models for use in predicting outcomes in RCC, and determine that molecular studies may improve or complement these results, improving both clinical predictions and yielding useful biological insights. We describe the contributions in these specific areas as follows.

For clinical models, we are able to determine that in terms of evaluating all the relevant clinical models, the Karakiewicz nomogram is superior to all other tested models in terms of predicting survival outcomes in localized RCC. In comparing models in current use in ongoing pharmaceutical trials, the Leibovich clinical trial criteria is superior to the UISS clinical trial criteria in terms of prediction of relapse free survival, but is equivalent to the UISS trial criteria in the prediction of CSS and OS. Exploratory analysis that we have performed is able to determine a potentially useful survival cutoff (a 5 year estimated cancer-specific survival of 0.9) for the Karakiewicz nomogram for dichotomization. This cut-off should be considered for use in future adjuvant trial design.

We have surveyed multiple pathological subtypes of RCC for our molecular model analysis. For clear cell RCC, we have identified clinically useful prognostic gene predictors for clear cell RCC using gene expression profiling. Increased expression of genes classically associated with the VEGF-signaling pathway, angiogenesis and the hypoxic response predicted longer patient survival.

For papillary RCC, we have identified two distinct molecular classes of papillary RCC that differ strikingly in their clinical behavior and have dysregulation of genes controlling different parts of the cell cycle. This finding represents a biologically and clinically relevant refinement to previously proposed morphologic criteria for subclassification of papillary RCC.

For chromophobe RCC, we have comprehensively characterized the molecular profiles of chromophobe RCC and oncocytoma using high throughput expression and SNP profiling. We have consequently derived discriminating expression signatures, pathways, cytogenetic profiles and protein markers that are of biologic, clinical and therapeutic interest. Additionally, we show that while chromophobe RCC cells contain an extra copy of chromosome 19, the renal oncocytoma cells contain a rarely reported chromosomal abnormality. Both of these chromosomal abnormalities result in transcriptional disruptions of *EGLN2*, a gene that is located on chromosome 19. Defects in oxygen sensing are found in other types of kidney tumours, and the identification of *EGLN2* directly implicates defects in the oxygen-sensing network in these neoplasias as well.

Overall, this thesis thus provides insights into both classic and molecular epidemiology, representing a useful evaluation of existing statistical prognostic models for RCC, with immediate practical value for clinical practice, epidemiologic research and trial design. The use of molecular profiling for all major subtypes of RCC in this thesis has yielded novel subtypes for clear cell RCC and papillary RCC, with attending biologic, clinical and epidemiologic implications. The molecular predictors differentiating chromophobe RCC and

oncocytoma have provided useful insights as to the underlying biology, as well as provided opportunities for practical differentiation between these two highly related entities in the pathology laboratory.

FUTURE RESEARCH

This thesis lays out the molecular epidemiology of RCC as discerned through both a clinical and epidemiologic lens, predominantly with gene expression profiling techniques. These results clearly show the heterogeneity of RCC in terms of gene expression and survival outcomes. As such, it would be important to utilize the insights afforded by this subtyping in future epidemiologic studies of RCC, thereby clarifying the associated factors predisposing to, and influencing outcomes of RCC such as prognosis and drug response.

From a biological viewpoint, evaluation of the molecular basis of these subtypes is crucial. Toward this goal, cancer genetics and epigenetics may be viewed as fundamental to investigating the dysregulated gene expression identified here. In particular, with improvements in next-generation sequencing technology and the identification of new somatic alterations such as mutations of *PBRM1*, such approaches will likely provide more information on the basis of the molecular epidemiology observed here. With better characterization of each sample, it is inevitable that more complex analyses will be possible.

From a clinical viewpoint, these novel subtypes are useful for utilization in the context of clinical research. It is likely that different RCC subtypes even within the same pathological subtype may yield different outcomes when treated with different targeted therapies, particularly if the underlying genetics of these

subtypes differ. Hence, the use of our clinical and molecular predictors would be very relevant in the context of international drug trials. It is already recognized that even within clear cell RCC, indolent tumours respond better to anti-angiogenic therapy and more aggressive tumours respond better to mTOR inhibitors. The extension of clinical trials as currently designed to accommodate novel insights of molecular epidemiology can improve study recruitment and outcomes.

In particular, this thesis focuses primarily on the epidemiology of RCC, in terms of diagnosis and prognosis. With the approval of many novel targeted therapies for cancer in the last five years, the use of these agents for the treatment of RCC has become of major interest. Hence, the issue of how biology affects therapeutic response is a key future research area, and this work provides a clear biological foundation for this challenge.

BIBLIOGRAPHY

Abdullah-Sayani, A, JM Bueno-de-Mesquita, and MJ van de Vijver (2006), 'Technology Insight: tuning into the genetic orchestra using microarrays--limitations of DNA microarrays in clinical practice.', *Nat Clin Pract Oncol*, 3 (9), 501-16.

Al-Radi, O. O., et al. (2007), 'Case complexity scores in congenital heart surgery: a comparative study of the Aristotle Basic Complexity score and the Risk Adjustment in Congenital Heart Surgery (RACHS-1) system', *J Thorac Cardiovasc Surg*, 133 (4), 865-75.

Allory, Y., et al. (2003), 'Papillary renal cell carcinoma. Prognostic value of morphological subtypes in a clinicopathologic study of 43 cases', *Virchows Arch*, 442 (4), 336-42.

Amin, M. B., et al. (1997), 'Papillary (chromophil) renal cell carcinoma: histomorphologic characteristics and evaluation of conventional pathologic prognostic parameters in 62 cases', *Am J Surg Pathol*, 21 (6), 621-35.

Atkin, NB and Z Jackson (1996), 'Evidence for somatic pairing of chromosome 7 and 10 homologs in a follicular lymphoma.', *Cancer Genet Cytogenet*, 89 (2), 129-31.

Atkins, Michael B., Meredith Regan, and David McDermott (2004), 'Update on the Role of Interleukin 2 and Other Cytokines in the Treatment of Patients with Stage IV Renal Carcinoma', *Clin Cancer Res*, 10 (18), 6342S-6346.

Baddi, L. and A. Benson, 3rd (2005), 'Adjuvant therapy in stage II colon cancer: current approaches', *Oncologist*, 10 (5), 325-31.

Baris, O, et al. (2004), 'Transcriptional profiling reveals coordinated up-regulation of oxidative metabolism genes in thyroid oncocytic tumors.', *J Clin Endocrinol Metab*, 89 (2), 994-1005.

Belizaire, R, et al. (2004), 'Characterization of synaptogyrin 3 as a new synaptic vesicle protein.', *J Comp Neurol*, 470 (3), 266-81.

Borowiak, M., et al. (2004), 'Met provides essential signals for liver regeneration', *Proc Natl Acad Sci U S A*, 101 (29), 10608-13.

Bostwick, D. G. and J. N. Eble (1999), 'Diagnosis and classification of renal cell carcinoma', *Urol Clin North Am*, 26 (3), 627-35.

Brown, JA, et al. (1994), 'Chromosomal aneusomies detected by fluorescent in situ hybridization analysis in clinically localized prostate carcinoma.', *J Urol*, 152 (4), 1157-62.

Bugert, P. and G. Kovacs (1996), 'Molecular differential diagnosis of renal cell carcinomas by microsatellite analysis', *Am J Pathol*, 149 (6), 2081-88.

Chen, DY and RG Uzzo (2009), 'Optimal management of localized renal cell carcinoma: surgery, ablation, or active surveillance.', *J Natl Compr Canc Netw*, 7 (6), 635-42; quiz 643.

Chen, F, et al. (1995), 'Suppression of growth of renal carcinoma cells by the von Hippel-Lindau tumor suppressor gene.', *Cancer Res*, 55 (21), 4804-07.

Cheville, J. C., et al. (2003), 'Comparisons of outcome and prognostic features among histologic subtypes of renal cell carcinoma', *Am J Surg Pathol*, 27 (5), 612-24.

Chow, W. H., et al. (1999), 'Rising incidence of renal cell cancer in the United States', *JAMA*, 281 (17), 1628-31.

Chow, W. H., et al. (2000), 'Obesity, hypertension, and the risk of kidney cancer in men', *N Engl J Med*, 343 (18), 1305-11.
Cindolo, L, et al. (2005), 'Comparison of predictive accuracy of four prognostic models for nonmetastatic renal cell carcinoma after nephrectomy: a multicenter European study.', *Cancer*, 104 (7), 1362-71.
Clifford, B., et al. (2003), 'G2 arrest in response to topoisomerase II inhibitors: the role of p53', *Cancer Res*, 63 (14), 4074-81.
Crawley, J. J. and K. A. Furge (2002), 'Identification of frequent cytogenetic aberrations in hepatocellular carcinoma using gene-expression microarray data', *Genome Biol*, 3 (12), RESEARCH0075.
Dai, M, et al. (2005), 'Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data.', *Nucleic Acids Res*, 33 (20), e175.
Dekel, Y., et al. (2002), 'Significance of angiogenesis and microvascular invasion in renal cell carcinoma', *Pathol Oncol Res*, 8 (2), 129-32.
Delahunt, B., P. B. Bethwaite, and A. Thornton (1997), 'Prognostic significance of microscopic vascularity for clear cell renal cell carcinoma', *Br J Urol*, 80 (3), 401-04.
Delahunt, B. and J. N. Eble (1997), 'Papillary renal cell carcinoma: a clinicopathologic and immunohistochemical study of 105 tumors', *Mod Pathol*, 10 (6), 537-44.
Di Blasio, C. J., et al. (2003), 'Predicting clinical end points: treatment nomograms in prostate cancer', *Semin Oncol*, 30 (5), 567-86.
Eastham, JA, et al. (2003), 'Cancer and Leukemia Group B (CALGB) 90203: a randomized phase 3 study of radical prostatectomy alone versus estramustine and docetaxel before radical prostatectomy for patients with high-risk localized disease.', *Urology*, 62 Suppl 1 55-62.
Eisen, Tim (2007), 'Adjuvant Therapy in Renal Cell Carcinoma: Where Are We?', *European Urology Supplements*, 6 (7), 492-98.
Escudier, B, et al. (2007), 'Sorafenib in advanced clear-cell renal-cell carcinoma.', *N Engl J Med*, 356 (2), 125-34.
Escudier, B., et al. (2002), 'Doxorubicin and ifosfamide in patients with metastatic sarcomatoid renal cell carcinoma: a phase II study of the Genitourinary Group of the French Federation of Cancer Centers', *J Urol*, 168 (3), 959-61.
Falcon, S and R Gentleman (2007), 'Using GOstats to test gene lists for GO term association.', *Bioinformatics*, 23 (2), 257-58.
Fan, C, et al. (2006), 'Concordance among gene-expression-based predictors for breast cancer.', *N Engl J Med*, 355 (6), 560-69.
Ficarra, V, et al. (2006), 'External validation of the Mayo Clinic Stage, Size, Grade and Necrosis (SSIGN) score to predict cancer specific survival using a European series of conventional renal cell carcinoma.', *J Urol*, 175 (4), 1235-39.
Ficarra, V, et al. (2009), 'The 'Stage, Size, Grade and Necrosis' score is more accurate than the University of California Los Angeles Integrated Staging System for predicting cancer-specific survival in patients with clear cell renal cell carcinoma.', *BJU Int*, 103 (2), 165-70.
Flanigan, R. C., et al. (2001), 'Nephrectomy followed by interferon alfa-2b compared with interferon alfa-2b alone for metastatic renal-cell cancer', *N Engl J Med*, 345 (23), 1655-59.

Flanigan, Robert C. (2004), 'Debulking Nephrectomy in Metastatic Renal Cancer', *Clin Cancer Res*, 10 (18), 6335S-6341.
Frank, I., et al. (2002), '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*, 168 (6), 2395-400.
Fujii, Y, et al. (2008), 'External validation of the Mayo Clinic cancer specific survival score in a Japanese series of clear cell renal cell carcinoma.', *J Urol*, 180 (4), 1290-5; discussion 1295-6.
Furge, K. A., et al. (2005), 'Comparison of array-based comparative genomic hybridization with gene expression-based regional expression biases to identify genetic abnormalities in hepatocellular carcinoma', *BMC Genomics*, 6 (1), 67.
Galfano, A, et al. (2008), 'Mathematical models for prognostic prediction in patients with renal cell carcinoma.', *Urol Int*, 80 (2), 113-23.
Gasparre, G., et al. (2008), 'Clonal expansion of mutated mitochondrial DNA is associated with tumor formation and complex I deficiency in the benign renal oncocytoma', *Hum Mol Genet*,
Gautier, L., et al. (2004), 'affy--analysis of Affymetrix GeneChip data at the probe level', *Bioinformatics*, 20 (3), 307-15.
Gelb, A. B., et al. (1997), 'Appraisal of intratumoral microvessel density, MIB-1 score, DNA content, and p53 protein expression as prognostic indicators in patients with locally confined renal cell carcinoma', *Cancer*, 80 (9), 1768-75.
Gentleman, R. C., et al. (2004), 'Bioconductor: open software development for computational biology and bioinformatics', *Genome Biol*, 5 (10), R80.
Gering, M., et al. (1998), 'The SCL gene specifies haemangioblast development from early mesoderm', *EMBO J*, 17 (14), 4029-45.
Gettman, M. T., et al. (2001), 'Pathologic staging of renal cell carcinoma: significance of tumor classification with the 1997 TNM staging system', *Cancer*, 91 (2), 354-61.
Gregg, JL, et al. (2010), 'Analysis of gene expression in prostate cancer epithelial and interstitial stromal cells using laser capture microdissection.', *BMC Cancer*, 10 165.
Gunawan, B., et al. (2003), 'Cytogenetic and morphologic typing of 58 papillary renal cell carcinomas: evidence for a cytogenetic evolution of type 2 from type 1 tumors', *Cancer Res*, 63 (19), 6200-05.
Haas, NB and R Uzzo (2008), 'Adjuvant therapy for renal cell carcinoma.', *Curr Oncol Rep*, 10 (3), 245-52.
Han, KR, et al. (2003), 'Validation of an integrated staging system toward improved prognostication of patients with localized renal cell carcinoma in an international population.', *J Urol*, 170 (6 Pt 1), 2221-24.
Harrell, F.E. (2001), *Regression Modelling Strategies: With Applications to Linear Models, Logistic Regression and Survival Analysis*, (New York: Springer-Verlag).
Harrell, FE Jr, KL Lee, and DB Mark (1996), 'Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors.', *Stat Med*, 15 (4), 361-87.
Herbst, C., et al. (1998), 'Evaluation of microvessel density by computerised image analysis in human renal cell carcinoma. Correlation to pT category, nuclear grade, proliferative activity and occurrence of metastasis', *J Cancer Res Clin Oncol*, 124 (3-4), 141-47.

Higgins, J. P., et al. (2003), 'Gene expression patterns in renal cell carcinoma assessed by complementary DNA microarray', *Am J Pathol*, 162 (3), 925-32.
Hochreiter, S, DA Clevert, and K Obermayer (2006), 'A new summarization method for Affymetrix probe level data.', *Bioinformatics*, 22 (8), 943-49.
Hubbell, E, WM Liu, and R Mei (2002), 'Robust estimators for expression analysis.', *Bioinformatics*, 18 (12), 1585-92.
Hudes, G, et al. (2007), 'Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma.', *N Engl J Med*, 356 (22), 2271-81.
Hughes, T. R., et al. (2000), 'Widespread aneuploidy revealed by DNA microarray expression profiling', *Nat Genet*, 25 (3), 333-37.
Hutson, TE, et al. (2008), 'Targeted therapies for metastatic renal cell carcinoma: an overview of toxicity and dosing strategies.', *Oncologist*, 13 (10), 1084-96.
Ihaka, R. and R. Gentleman (1996), 'R: A Language for Data Analysis and Graphics', *J Comput Graph Stat*, 5 (3), 299-314.
Iliopoulos, O., et al. (1995), 'Tumour suppression by the human von Hippel-Lindau gene product', *Nat Med*, 1 (8), 822-26.
Iliopoulos, O., et al. (1996), 'Negative regulation of hypoxia-inducible genes by the von Hippel-Lindau protein', *Proc Natl Acad Sci U S A*, 93 (20), 10595-99.
Imao, T., et al. (2004), 'Inverse correlation of microvessel density with metastasis and prognosis in renal cell carcinoma', *Int J Urol*, 11 (11), 948-53.
Imazano, Y., et al. (1997), 'Correlation between thymidine phosphorylase expression and prognosis in human renal cell carcinoma', *J Clin Oncol*, 15 (7), 2570-78.
Irizarry, RA, et al. (2003), 'Summaries of Affymetrix GeneChip probe level data.', *Nucleic Acids Res*, 31 (4), e15.
Isaacs, JS, et al. (2005), 'HIF overexpression correlates with biallelic loss of fumarate hydratase in renal cancer: novel role of fumarate in regulation of HIF stability.', *Cancer Cell*, 8 (2), 143-53.
Jemal, A, et al. (2009), 'Cancer statistics, 2009.', *CA Cancer J Clin*, 59 (4), 225-49.
Jiang, F., et al. (1998), 'Chromosomal imbalances in papillary renal cell carcinoma: genetic differences between histological subtypes', *Am J Pathol*, 153 (5), 1467-73.
Jones, J, et al. (2005), 'Gene signatures of progression and metastasis in renal cell cancer.', *Clin Cancer Res*, 11 (16), 5730-39.
Joo, H. J., et al. (2004), 'Increased expression of caveolin-1 and microvessel density correlates with metastasis and poor prognosis in clear cell renal cell carcinoma', *BJU Int*, 93 (3), 291-96.
Kaelin, W. G., Jr. (2004), 'The von Hippel-Lindau tumor suppressor gene and kidney cancer', *Clin Cancer Res*, 10 (18 Pt 2), 6290S-5S.
Kanehisa, M, et al. (2010), 'KEGG for representation and analysis of molecular networks involving diseases and drugs.', *Nucleic Acids Res*, 38 (Database issue), D355-60.
Kang, Y., et al. (2003), 'A multigenic program mediating breast cancer metastasis to bone', *Cancer Cell*, 3 (6), 537-49.
Karakiewicz, P. I., et al. (2007), 'Multi-institutional validation of a new renal cancer-specific survival nomogram', *J Clin Oncol*, 25 (11), 1316-22.
Karakiewicz, PI and GC Hutterer (2007), 'Predicting cancer-control outcomes in patients with renal cell carcinoma.', *Curr Opin Urol*, 17 (5), 295-302.

Kattan, M. W., et al. (2001), 'A postoperative prognostic nomogram for renal cell carcinoma', *J Urol*, 166 (1), 63-67.
Kattan, M. W. (2008), 'Do we need more nomograms for predicting outcomes in patients with prostate cancer?', *Nat Clin Pract Urol*, 5 (7), 366-67.
Kattan, M. W. (2010), 'Editorial comment', *Urology*, 75 (6), 1371; author reply 1371-72.
Kattar, M. M., et al. (1997), 'Clinicopathologic and interphase cytogenetic analysis of papillary (chromophilic) renal cell carcinoma', *Mod Pathol*, 10 (11), 1143-50.
Khoo, S. K., et al. (2001), 'Birt-Hogg-Dube syndrome: mapping of a novel hereditary neoplasia gene to chromosome 17p12-q11.2', *Oncogene*, 20 (37), 5239-42.
Kosari, F, et al. (2005), 'Clear cell renal cell carcinoma: gene expression analyses identify a potential signature for tumor aggressiveness.', *Clin Cancer Res*, 11 (14), 5128-39.
Kovacs, G. and P. Brusa (1988), 'Recurrent genomic rearrangements are not at the fragile sites on chromosomes 3 and 5 in human renal cell carcinomas', *Hum Genet*, 80 (1), 99-101.
Kovacs, G., et al. (1991), 'Cytogenetics of papillary renal cell tumors', *Genes Chromosomes Cancer*, 3 (4), 249-55.
Lager, D. J., et al. (1995), 'Papillary renal tumors. Morphologic, cytochemical, and genotypic features', *Cancer*, 76 (4), 669-73.
Larkin, JE, et al. (2005), 'Independence and reproducibility across microarray platforms.', *Nat Methods*, 2 (5), 337-44.
Law, T. M., P. Mencel, and R. J. Motzer (1994), 'Phase II trial of liposomal encapsulated doxorubicin in patients with advanced renal cell carcinoma', *Invest New Drugs*, 12 (4), 323-25.
Lee, S, et al. (2005), 'Neuronal apoptosis linked to EglN3 prolyl hydroxylase and familial pheochromocytoma genes: developmental culling and cancer.', *Cancer Cell*, 8 (2), 155-67.
Leibovich, B. C., et al. (2003), 'Prediction of progression after radical nephrectomy for patients with clear cell renal cell carcinoma: a stratification tool for prospective clinical trials', *Cancer*, 97 (7), 1663-71.
Lewis, JP, et al. (1993), 'Somatic pairing of centromeres and short arms of chromosome 15 in the hematopoietic and lymphoid system.', *Hum Genet*, 92 (6), 577-82.
Li, C and WH Wong (2001), 'Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection.', *Proc Natl Acad Sci U S A*, 98 (1), 31-36.
Liu, ET (2007), 'Stromal effects in breast cancer.', *N Engl J Med*, 357 (25), 2537-38.
Liu, Z, et al. (2009), 'Validation of the current prognostic models for nonmetastatic renal cell carcinoma after nephrectomy in Chinese population: a 15-year single center experience.', *Int J Urol*, 16 (3), 268-73.
MacLennan, G. T. and D. G. Bostwick (1995), 'Microvessel density in renal cell carcinoma: lack of prognostic significance', *Urology*, 46 (1), 27-30.
Mayr, JA, et al. (2008), 'Loss of complex I due to mitochondrial DNA mutations in renal oncocytoma.', *Clin Cancer Res*, 14 (8), 2270-75.
Mejean, A., S. Oudard, and N. Thiounn (2003), 'Prognostic factors of renal cell carcinoma', *J Urol*, 169 (3), 821-27.
Meloni-Ehrig, A. M. (2002), 'Renal cancer: cytogenetic and molecular genetic aspects', *Am J Med Genet*, 115 (3), 164-72.

Morgan, T. M. and R. M. Elashoff (1987), 'Effect of covariate measurement error in randomized clinical trials', *Stat Med*, 6 (1), 31-41.
Mosimann, C, G Hausmann, and K Basler (2006), 'Parafibromin/Hyrax activates Wnt/Wg target gene transcription by direct association with beta-catenin/Armadillo.', *Cell*, 125 (2), 327-41.
Motzer, R. J., et al. (1999), 'Survival and prognostic stratification of 670 patients with advanced renal cell carcinoma', *J Clin Oncol*, 17 (8), 2530-40.
Motzer, R. J., et al. (2004), 'Prognostic factors for survival in previously treated patients with metastatic renal cell carcinoma', *J Clin Oncol*, 22 (3), 454-63.
Motzer, RJ, et al. (2007), 'Sunitinib versus interferon alfa in metastatic renal-cell carcinoma.', *N Engl J Med*, 356 (2), 115-24.
Motzer, RJ, et al. (2008), 'Efficacy of everolimus in advanced renal cell carcinoma: a double-blind, randomised, placebo-controlled phase III trial.', *Lancet*, 372 (9637), 449-56.
Nativ, O., et al. (1998), 'Clinical significance of tumor angiogenesis in patients with localized renal cell carcinoma', *Urology*, 51 (5), 693-96.
Nickerson, M. L., et al. (2002), 'Mutations in a novel gene lead to kidney tumors, lung wall defects, and benign tumors of the hair follicle in patients with the Birt-Hogg-Dube syndrome', *Cancer Cell*, 2 (2), 157-64.
Nickerson, ML, et al. (2008), 'Improved identification of von Hippel-Lindau gene alterations in clear cell renal tumors.', *Clin Cancer Res*, 14 (15), 4726-34.
Oken, M. M., et al. (1982), 'Toxicity and response criteria of the Eastern Cooperative Oncology Group', *Am J Clin Oncol*, 5 (6), 649-55.
Paoletti, X., et al. (2010), 'Benefit of adjuvant chemotherapy for resectable gastric cancer: a meta-analysis', *JAMA*, 303 (17), 1729-37.
Paradis, V., et al. (2000), 'Expression of vascular endothelial growth factor in renal cell carcinomas', *Virchows Arch*, 436 (4), 351-56.
Parker, A. S., et al. (2002), 'High expression levels of insulin-like growth factor-I receptor predict poor survival among women with clear-cell renal cell carcinomas', *Hum Pathol*, 33 (8), 801-05.
Parker, AS, et al. (2009), 'Development and evaluation of BioScore: a biomarker panel to enhance prognostic algorithms for clear cell renal cell carcinoma.', *Cancer*, 115 (10), 2092-103.
Patard, J. J., et al. (2004), 'Use of the University of California Los Angeles integrated staging system to predict survival in renal cell carcinoma: an international multicenter study', *J Clin Oncol*, 22 (16), 3316-22.
Percy, MJ, et al. (2006), 'A family with erythrocytosis establishes a role for prolyl hydroxylase domain protein 2 in oxygen homeostasis.', *Proc Natl Acad Sci U S A*, 103 (3), 654-59.
Picken, MM, et al. (2008), 'Analysis of chromosome 1p abnormalities in renal oncocytomas by loss of heterozygosity studies: correlation with conventional cytogenetics and fluorescence in situ hybridization.', *Am J Clin Pathol*, 129 (3), 377-82.
Pollard, PJ, et al. (2005), 'Accumulation of Krebs cycle intermediates and over-expression of HIF1alpha in tumours which result from germline FH and SDH mutations.', *Hum Mol Genet*, 14 (15), 2231-39.

Quackenbush, J (2006a), 'Computational approaches to analysis of DNA microarray data.', *Yearb Med Inform*, 91-103.
Quackenbush, J (2006b), 'Microarray analysis and tumor classification.', *N Engl J Med*, 354 (23), 2463-72.
Ramaswamy, S., et al. (2003), 'A molecular signature of metastasis in primary solid tumors', *Nat Genet*, 33 (1), 49-54.
Renshaw, A. A. and C. L. Corless (1995), 'Papillary renal cell carcinoma. Histology and immunohistochemistry', *Am J Surg Pathol*, 19 (7), 842-49.
Rini, B. I. (2009), 'Vascular endothelial growth factor-targeted therapy in metastatic renal cell carcinoma', *Cancer*, 115 (10 Suppl), 2306-12.
Rini, BI, SC Campbell, and B Escudier (2009), 'Renal cell carcinoma.', *Lancet*, 373 (9669), 1119-32.
Rioux-Leclercq, N., et al. (2001), 'Clinical significance of cell proliferation, microvessel density, and CD44 adhesion molecule expression in renal cell carcinoma', *Hum Pathol*, 32 (11), 1209-15.
Robert, B., et al. (1996), 'Evidence that embryonic kidney cells expressing flk-1 are intrinsic, vasculogenic angioblasts', *Am J Physiol*, 271 (3 Pt 2), F744-53.
Rohan, S, et al. (2006), 'Gene expression profiling separates chromophobe renal cell carcinoma from oncocytoma and identifies vesicular transport and cell junction proteins as differentially expressed genes.', *Clin Cancer Res*, 12 (23), 6937-45.
Sabo, E., et al. (2001), 'Microscopic analysis and significance of vascular architectural complexity in renal cell carcinoma', *Clin Cancer Res*, 7 (3), 533-37.
Sandlund, J, et al. (2006), 'Endoglin (CD105) expression in human renal cell carcinoma.', *BJU Int*, 97 (4), 706-10.
Sayers, EW, et al. (2010), 'Database resources of the National Center for Biotechnology Information.', *Nucleic Acids Res*, 38 (Database issue), D5-16.
Schmidt, L., et al. (1998), 'Two North American families with hereditary papillary renal carcinoma and identical novel mutations in the MET proto-oncogene', *Cancer Res*, 58 (8), 1719-22.
Schmidt, L., et al. (1999), 'Hereditary papillary renal carcinoma: pathology and pathogenesis', *Contrib Nephrol*, 128 11-27.
Schraml, P., et al. (2002), 'VHL mutations and their correlation with tumour cell proliferation, microvessel density, and patient prognosis in clear cell renal cell carcinoma', *J Pathol*, 196 (2), 186-93.
Schuetz, A. N., et al. (2005), 'Molecular classification of renal tumors by gene expression profiling', *J Mol Diagn*, 7 (2), 206-18.
Seliger, B, et al. (2000), 'HER-2/neu is expressed in human renal cell carcinoma at heterogeneous levels independently of tumor grading and staging and can be recognized by HLA-A2.1-restricted cytotoxic T lymphocytes.', *Int J Cancer*, 87 (3), 349-59.
Selvarajan, S., et al. (2008), 'Parafibromin expression in breast cancer: a novel marker for prognostication?', *J Clin Pathol*, 61 (1), 64-67.
Sengupta, S, et al. (2005), 'Histologic coagulative tumor necrosis as a prognostic indicator of renal cell carcinoma aggressiveness.', *Cancer*, 104 (3), 511-20.
Seruga, B., et al. (2010), 'Absolute benefits of medical therapies in phase III clinical trials for breast and colorectal cancer', *Ann Oncol*, 21 (7), 1411-18.

Shariat, S. F., et al. (2008), 'Comparison of nomograms with other methods for predicting outcomes in prostate cancer: a critical analysis of the literature', *Clin Cancer Res*, 14 (14), 4400-07.
Slaton, Joel W., et al. (2001), 'Expression Levels of Genes that Regulate Metastasis and Angiogenesis Correlate with Advanced Pathological Stage of Renal Cell Carcinoma', *Am J Pathol*, 158 (2), 735-43.
Smyth, GK (2004), 'Linear models and empirical bayes methods for assessing differential expression in microarray experiments.', *Stat Appl Genet Mol Biol*, 3 Article3.
Sorbellini, M, et al. (2005), 'A postoperative prognostic nomogram predicting recurrence for patients with conventional clear cell renal cell carcinoma.', *J Urol*, 173 (1), 48-51.
Sternberg, CN, et al. (2010), 'Pazopanib in Locally Advanced or Metastatic Renal Cell Carcinoma: Results of a Randomized Phase III Trial.', *J Clin Oncol*,
Storkel, S, et al. (1989), 'The human chromophobe cell renal carcinoma: its probable relation to intercalated cells of the collecting duct.', *Virchows Arch B Cell Pathol Incl Mol Pathol*, 56 (4), 237-45.
Takahashi, A., et al. (1994), 'Markedly increased amounts of messenger RNAs for vascular endothelial growth factor and placenta growth factor in renal cell carcinoma associated with angiogenesis', *Cancer Res*, 54 (15), 4233-37.
Takahashi, M., et al. (2001), 'Gene expression profiling of clear cell renal cell carcinoma: gene identification and prognostic classification', *Proc Natl Acad Sci U S A*, 98 (17), 9754-59.
Takahashi, M., et al. (2002), 'Gene expression profiling of favorable histology Wilms tumors and its correlation with clinical features', *Cancer Res*, 62 (22), 6598-605.
Takahashi, M., et al. (2003), 'Molecular subclassification of kidney tumors and the discovery of new diagnostic markers', *Oncogene*, 22 (43), 6810-18.
Takeda, K, et al. (2008), 'Regulation of adult erythropoiesis by prolyl hydroxylase domain proteins.', *Blood*, 111 (6), 3229-35.
Tan, M. H., et al. (2004a), 'Loss of parafibromin immunoreactivity is a distinguishing feature of parathyroid carcinoma', *Clin Cancer Res*, 10 (19), 6629-37.
Tan, M. H., et al. (2004b), 'Gene expression profiling of renal cell carcinoma', *Clin Cancer Res*, 10 (18 Pt 2), 6315S-21S.
Thoenes, W, S Storkel, and HJ Rumpelt (1986), 'Histopathology and classification of renal cell tumors (adenomas, oncocytomas and carcinomas). The basic cytological and histopathological elements and their use for diagnostics.', *Pathol Res Pract*, 181 (2), 125-43.
Tibshirani, R., et al. (2002), 'Diagnosis of multiple cancer types by shrunken centroids of gene expression', *Proc Natl Acad Sci U S A*, 99 (10), 6567-72.
Tomlinson, IP, et al. (2002), 'Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer.', *Nat Genet*, 30 (4), 406-10.
Turner, K. J., et al. (2002), 'Expression of hypoxia-inducible factors in human renal cancer: relationship to angiogenesis and to the von Hippel-Lindau gene mutation', *Cancer Res*, 62 (10), 2957-61.

Tusher, VG, R Tibshirani, and G Chu (2001), 'Significance analysis of microarrays applied to the ionizing radiation response.', *Proc Natl Acad Sci U S A*, 98 (9), 5116-21.
Twine, N. C., et al. (2003), 'Disease-associated expression profiles in peripheral blood mononuclear cells from patients with advanced renal cell carcinoma', *Cancer Res*, 63 (18), 6069-75.
Varela, I., et al. (2011), 'Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma', *Nature*, 469 (7331), 539-42.
Vasselli, J. R., et al. (2003), 'Predicting survival in patients with metastatic kidney cancer by gene-expression profiling in the primary tumor', *Proc Natl Acad Sci U S A*, 100 (12), 6958-63.
Vickers, A. J., et al. (2009), 'Clinical benefits of a multivariate prediction model for bladder cancer: a decision analytic approach', *Cancer*, 115 (23), 5460-69.
Vortmeyer, A. O., et al. (2003), 'Developmental arrest of angioblastic lineage initiates tumorigenesis in von Hippel-Lindau disease', *Cancer Res*, 63 (21), 7051-55.
Waldert, M, et al. (2008), 'Comparison of type I and II papillary renal cell carcinoma (RCC) and clear cell RCC.', *BJU Int*, 102 (10), 1381-84.
Weidner, N. (1998), 'Tumoural vascularity as a prognostic factor in cancer patients: the evidence continues to grow', *J Pathol*, 184 (2), 119-22.
Weidner, N. (1999), 'Tumour vascularity and proliferation: clear evidence of a close relationship', *J Pathol*, 189 (3), 297-99.
Westfall, PH, DV Zaykin, and SS Young (2002), 'Multiple tests for genetic effects in association studies.', *Methods Mol Biol*, 184 143-68.
Yamazaki, K., et al. (2003), 'Overexpression of KIT in chromophobe renal cell carcinoma', *Oncogene*, 22 (6), 847-52.
Yasui, M, et al. (1999), 'Aquaporin-6: An intracellular vesicle water channel protein in renal epithelia.', *Proc Natl Acad Sci U S A*, 96 (10), 5808-13.
Yoshino, S., M. Kato, and K. Okada (1995), 'Prognostic significance of microvessel count in low stage renal cell carcinoma', *Int J Urol*, 2 (3), 156-60.
Young, A. N., et al. (2001), 'Expression profiling of renal epithelial neoplasms: a method for tumor classification and discovery of diagnostic molecular markers', *Am J Pathol*, 158 (5), 1639-51.
Zervakis, M, et al. (2009), 'Outcome prediction based on microarray analysis: a critical perspective on methods.', *BMC Bioinformatics*, 10 53.
Zhang, FF, et al. (1997), 'Toward the validation of aneusomy detection by fluorescence in situ hybridization in bladder cancer: comparative analysis with cytology, cytogenetics, and clinical features predicts recurrence and defines clinical testing limitations.', *Clin Cancer Res*, 3 (12 Pt 1), 2317-28.
Zhao, H., et al. (2006), 'Gene expression profiling predicts survival in conventional renal cell carcinoma', *PLoS Med*, 3 (1), e13.
Zhao, J, et al. (2007), 'Sporadic human renal tumors display frequent allelic imbalances and novel mutations of the HRPT2 gene.', *Oncogene*, 26 (23), 3440-49.
Zisman, A, et al. (2002), 'Risk group assessment and clinical outcome algorithm to predict the natural history of patients with surgically resected renal cell carcinoma.', *J Clin Oncol*, 20 (23), 4559-66.

Zisman, A., et al. (2001), 'Improved prognostication of renal cell carcinoma using an integrated staging system', *J Clin Oncol*, 19 (6), 1649-57.

THE END