DISCUSSION
FUTURE PERSPECTIVES
In the last 10 years the urological community has been reporting new knowledge and developing standards for renal cell carcinoma (RCC). Most were based on large but retrospective series and extend for a very long chronological frame. Selection biases could not be ruled out as selection was made on the availability of the data. Furthermore some of the assertions and conclusion are based on administrative registries data while the limitations of such registries, e.g. absence of clinical data and potential misclassification are well known.

While awaiting for audited prospective data collection sets and well-designed diagnostic and therapeutic studies, there is no reason to neglect the body of knowledge reached as the best available evidence. This considered, together with the change in daily practice it is undeniable that the presentation and treatment of the RCC has completely changed in the last 15 years.

Furthermore most of the prognostic systems and classifications described in the last years deserve a careful validation and testing in other settings than the developmental ones. In this thesis the current knowledge in epidemiology, treatment and prognosis for RCC is evaluated in different and contemporary settings and ways to improve current prognostic systems are assessed.

Part I

In Chapter 2, the gender differences in RCC at presentation in the Netherlands are described. Most data on this subject is derived from studies using USA databases. The main limitation of the two largest, the SEER and NCDB, is their administrative nature and lack of national coverage (26% and 75% respectively).

According to their data women have tumours with smaller size, lower stage and lower grade and their age at presentation is higher. Some studies claimed better cancer related survival in women, whilst others observed equal survival among men and women. Hormonal causes have been designated, as some studies only observed survival advantages in premenopausal women. Evidence for the role of oestrogen in RCC progression however is still preliminary.

PALGA (Pathological Anatomical National Automated Archive) is an obligatory prospectively collected database, which contains excerpts from all pathology reports generated in the Netherlands from 1995 onwards. This nationwide database allowed our study to overcome the limitations of previously mentioned American studies. Data from all RCCs operated between 1995 and 2005 in the Netherlands (n=11619) was extracted from PALGA. Patients were divided into five cohorts based on age at surgery to assess the impact of age on gender differences.

Our study confirmed that women presented with lower staged tumours, more specifically this difference was seen from 40 years onwards. Similarly women in our study presented with smaller tumours, however statistical differences in tumour size were only found in the range 51-60 years. Gender related age differences were observed either in patients older than 70 years or those younger than 40. Only in the age group over 70 years women presented with a higher age. Better overall survival of women in the general population could be a confounder, leading to an overrepresentation of women operated at an older age. Other factors that could have influenced the performance of surgery such as the presence of co-morbidity were not recorded in our database.

The male to female ratio was almost equal in patients over 70 years, compared to 2:1 between 41 and 60 years. Similar observations were done by other studies. Differences in local treatment policies regarding elderly patients could influence these ratios. The main limitations of our study were the absence of survival data and clinical antecedents that could have been acting as confounders or determinants. Due to the pathological nature of the PALGA only operated patients were included, allowing a small selection bias for non-operated patients.

With contemporary cohorts, knowledge of the epidemiology of RCC can be confirmed and even better specified. Complete (or as complete as possible) data is required for testing epidemiological knowledge and assessment would be appropriate on a national/continental level every decade, due to the risk for selection bias and the changing RCC presentation.

In Chapter 3, two recently developed anatomical based classification systems, the PADUA classification and R.E.N.A.L. nephrometry score were assessed for their ability to predict perioperative complications in partial nephrectomy (PN) and tested for their reproducibility in a population of 134 PNs from three Dutch referral centres. Reproducibility was
evaluated by comparing scores of three physicians with different degrees in urological expertise. Classification systems have to be easy to apply and reproducible before widespread clinical use can be recommended. Our study was the first to evaluate the reproducibility of both PADUA and R.E.N.A.L. score. Good reproducibility for both total scores was observed, which was later confirmed by several other studies. Involvement and proximity to the urinary system were more difficult to reproduce in our series compared to other series, however not all studies included scores given by junior residents. According to our results the total PADUA and R.E.N.A.L. scores were associated with complications, however only the highest scores of PADUA, and not the intermediate scores, could predict overall complication risk. The highest scores of R.E.N.A.L. and tumour size showed similar results. The amount of variance in outcomes (complications) explained by the different scores or tumour size, as indicated by the R² value, was equally low amongst all three. Several other studies assessed the ability of both scores to predict complications in either open, laparoscopic or robot-assisted PN. Independent of the method, results on associations between complications and PADUA or R.E.N.A.L. score are inconsistent. Radiological slide thickness could have influenced our results. While older CT images have a 5mm thickness, newer images have 3mm thickness. The same can be said for differences in radiology protocols. A CT with contrast excretion phase is essential for applying both scores, however this is not standard preoperative practice in the Netherlands due to radiation reduction protocols and economic reasons. Other factors explaining the dissimilarities in validation results of PADUA and R.E.N.A.L. scores can be found in cohort composition, surgeon bound factors and patient bound factors such as co-morbidities and age. Limitations of our study include its retrospective nature and the inclusion of patients over a wide period. Anatomical classification systems for renal tumours did not perform as in their developmental setting in regard to predicting complications. Their value in our clinical setting is limited, however we cannot preclude its usefulness in other settings. Due to their good reproducibility both systems can be used for standardized reporting. Taking into account centre and surgeon dependant factors and lack of uniform methods for reporting complications, complication prediction is best tested per centre. Two to five yearly assessment would deem appropriate, since surgical techniques and surgical experience are ever expanding.

In Chapter 4, assumptions on renal function outcomes after radical nephrectomy (RN) for RCC patients and living kidney donors are assessed in a contemporary cohort. The landmark study by Go et al. in which a reduced glomerular filtration rate (GFR) was shown to be associated with higher morbidity and cardiovascular events (CVE), was embraced by urologists to highlight the importance of nephron sparing surgery. When RN is compared to PN, the effects of PN on preservation of renal function are undeniable, however studies relating to better survival and less CVE in PN patients show inconsistent results. Some authors observed differences in survival or CVE almost immediately postoperative, which diminishes in subsequent years, suggesting that differences found merely reflect inherent selection biases. The only randomized trial on the subject, with limitations regarding low accrual and considerable crossover between treatment arms, did not find better survival in RCC patients undergoing PN compared to RN. Current literature for kidney donors describes very low rates of end stage renal disease (ESRD) and excellent survival in kidney donors, with incidences comparable to the general population. Some even stated “Kidney donors live longer”, though the latter most likely reflected the selection of healthy individuals suitable for donation. More studies are being published in which donors are compared to matched healthy non-donors. In these comparisons an increased risk of ESRD in donors was observed, though absolute magnitude is still small and not nearly in the range of after RCC nephrectomy. Adverse CVE and mortality rates were not found to be increased, however a single study indicated increased mortality in elderly donors. More long term follow up is eagerly awaited. In our study 98 RCC patients undergoing RN were matched to 98 living kidney donors by age and gender. Both groups developed comparable renal function at 1 year postoperatively in spite of a higher co-morbidity load and lower baseline renal function in RCC patients. No association between co-morbidity and 1 year postoperative renal
function was found, contrasting with several other studies that stated comorbidities such as hypertension and diabetes influence renal function after surgery, though most presented longer follow up data and larger sample size than the present study. Few studies make direct comparison between donors and RCC patients after RN. A Korean matched cohort study that additionally matched for comorbidities indicated an equal decline in renal function after 1 year, but subsequent recovery of renal function only in donors after 3 years. A study by Timsit et al indicated no difference in renal function decrease when applying donor criteria to RCC patients after 4 years follow up. Our study could not make assumptions beyond the scope of 1 year. As in the other two studies, preoperative GFR was an important predictor of postoperative renal function outcome. None of the donors in current study progressed to an estimated GFR (eGFR) below 30 ml/min/1.73m², however 4 (4.5%) of the RCC patients did progress to eGFRs below 30 ml/min/1.73m². All had preoperative eGFRs ranging between 42 and 61 ml/min/1.73m². Though our numbers are too limited to make definite statements, our observations seem to coincide with results of several other authors. In these studies patients with preoperative chronic kidney disease (CKD) due to medical causes have worse renal function outcomes after renal mass surgery, compared to those with surgical induced CKD. Similar observations were made for survival.

Limitations of our study include length of follow up, no information on proteinuria in RCC patients and estimation of the GFR.

Concluding, our results indicated equal renal function in the first year after nephrectomy in kidney donors and patients with renal masses at an advanced age. Preoperative GFR is a major determinant in outcome after surgery. Research on renal function after nephrectomy in both donors and RCC patients is best done in large populations with long term follow up available, due to relatively low numbers of end stage renal disease and other adverse outcomes.

Changes in tumour stage derived from implementation of the new 2009 TNM, which is mostly based on retrospective studies spread over long time frames, are assessed in contemporary consecutive patient cohort in the latest version of the TNM. T2 tumours are reclassified in T2a (more than 7 but not more than 10 cm) and T2b (larger than 10 cm). Tumour thrombus extension in the renal vein is staged as T3a and extension into the vena cava below the diaphragm as T3b (formerly both T3b). Presence of vena cava wall ingrowth (formerly T3b) is staged as T3c and contiguous tumour extension in the ipsilateral adrenal gland (formerly T3a) as T4. The stage Mx is no longer used and there are no changes for N stages.

Several validation studies with retrospective data have been done, all indicating more or less comparable c-indexes for TNM 2002 and TNM 2009 for cancer specific survival (CSS). Additionally survival differences between adjacent 2009 T-stages were assessed by all. Two of those studies could not confirm the new T2 subclassification. Two studies described no survival differences between pT3b-pT3c stages and the multicentre study by Novarra could not confirm this difference in a subcohort of NxN0M0 patients. Several authors did not observe survival differences between pT3c and pT4 tumours. The new pT3a-pT3b classification is confirmed by all, however in the study of Novarra the significant difference is lost in a NxN0M0 population. One of the problems encountered by the above mentioned studies is a small number of patients included in some T-stages (pT2b, pT3b, pT3c).

Our study is the first to report prospectively collected data on the matter. In a large multicentre database containing 3989 clinical and 3089 pathological cases, both 2002 and 2009 T-stage were assigned and differences in T-stage were analysed. The application of the 2009 TNM staging system to a contemporary cohort induced changes in 18.5% of the clinical renal masses and in 17.9% of the pathologically staged RCCs. The majority of the clinical changes were driven by the new sub-classification of the T2 stage (13.1%). The majority of pathological changes consists of T2 changes (8.4%) and downstaging or upstaging of T3/T4 tumours (8.0% and 1.6% respectively).

The need for further sub-categorisation of T2 tumours as implemented by the 2009 TNM, is questionable. Several studies which contained large numbers of T2 tumours indicated that multiple size cut-offs can be identified for T2 tumours. The cut-offs presented similar discriminative properties, since size as a continuous value is related to survival.
Conversely Brookman –May et al\textsuperscript{74} did not find influence of tumour size on survival in T2 tumours and instead identified high grade and/or collecting system invasion as predictors of poor outcome. Moreover the quantification of the impact of new T2 categories is relatively small. T2 is a rare tumour, representing around 6-10\% of all pathological tumours\textsuperscript{68,69}. The usefulness in clinical staging can be further doubted due to difficulties for radiologists to correctly estimate tumour size in tumours larger than 7 cm\textsuperscript{75,76}.

An important issue, as seen in our data, is that the 2009 TNM creates a large and seemingly heterogeneous group of new pT3a tumours. In support of the new pT3a category, multiple studies observed similar survival between patients with renal vein invasion (RVI) and perirenal fat invasion\textsuperscript{77-79}. However within this category different studies indicate either worse prognosis among those with concomitant presence of fat invasion and RVI\textsuperscript{77,78-80} or among those with tumours larger than 7 or 8 cm respectively\textsuperscript{77,81}.

Two large multicentre studies were published on the prognostic significance of the tumour thrombus. Better survival in patients with RVI compared to those with vena cava invasion (VCI) was observed by Wagner\textsuperscript{80}. In the same study similar survival was observed in patients with VCI below or above the diaphragm, however the number of patients in the latter group was relatively low\textsuperscript{80}. In a study by Martinez –Salamanca a large group of pT3b and pT3c patients was evaluated and prognostic differences between patients with RVI, inferior VCI and supradiaphragmatic VCI were observed, supporting the changes of the 2009 TNM. In a N0M0 population the significant difference was lost between patients with RVI and inferior VCI\textsuperscript{82}, making it difficult to validated these results even in large populations.

Assessment of prognostic systems faces a dilemma. Our study evaluated contemporary data and indicated small immediate changes. Survival data is not yet available. Retrospective data can be analysed immediately however might not represent current situation due to inherent selection bias. Large numbers are needed to detect existing survival differences, especially when assessing less frequent T-stages. Moreover tumour stages could differ among centres depending on the type of centres (peripheral or referral). Therefore evaluation of prognostic systems should be done at least at an international multicentre basis and ideally reassessment or refinements should take place every 5-10 years. It has to be taken into account though that the community will need time to implement this knowledge in to the clinical practice.

**Part II**

Since the new millennium authors started publishing novel prognostic systems based on clinical-histopathological factors. Nowadays an overwhelming amount of these models exists\textsuperscript{83}. One of the first models was published by Motzer in 1999 and his criteria to predict mortality in metastatic RCC patients have since been widely adapted by the urological and oncological community\textsuperscript{84}. It has since undergone multiple adaptations to maintain accuracy\textsuperscript{85,86}, however its usefulness in the era of targeted therapies is doubtful and attention currently has shifted to the model developed by the International Kidney Cancer Work Group\textsuperscript{87}.

Most frequently mentioned postoperative models are the UISS, SSIGN and Karakiewicz model, which all incorporate similar variables\textsuperscript{86}. The SSIGN uses the presence of necrosis\textsuperscript{89}, where Karakiewicz uses the presence of symptoms both in addition to T,N and M stage, tumour size and grade\textsuperscript{86}. UISS incorporates TNM stadium, grade and ECOG score\textsuperscript{91}. Reported c-indexes range between 0.65 and 0.89 for CSS in populations which include patients with lymphnode and distant metastasis\textsuperscript{83,90-99}. From the three models mentioned, Karakiewicz proves best discrimination in comparative studies\textsuperscript{93,99}.

The difficulty in interpreting these results compared to studies reporting c-indexes for the TNM (0.75-0.86)\textsuperscript{67-72}, is that very few studies mention the incremental value of the new prognostic model to the traditional TNM classification. In fact the similar numbers of the c-indexes, makes one assume that the incremental value of these new models to the TNM is limited.

In attempts to further increase the accuracy of clinical-histopathological prognostic systems, a very large variety of other prognostic biomarkers have been identified, which represent the biological components of the tumour. These biomarkers range from immunohistochemistry markers to novel somatic copy number alternations and gene expression signatures\textsuperscript{100,101}.

A small amount of protein biomarkers have been validated in individual studies\textsuperscript{102,103}, however studies validating multiple protein biomarkers
at once have not yet been performed\textsuperscript{104,105}. In a large validation study for genetic biomarkers, the ccA/ccB subtype and to a lesser extent its derived Clearcode 34, were the only markers that retained prognostic value after correction for stage and grade\textsuperscript{100}. The ccA/ccB subtype is a gene expression profile for clearcell RCCs consisting of 110 genes, which divides patient into good (ccA) and poor (ccB) prognosis\textsuperscript{106}. Clearcode 34, assigns the ccA/ccB subtype with 94\% concordance, using 34 genes\textsuperscript{107}. Genetic signatures are complicated by intratumour heterogeneity; expression of both good and poor prognosis signatures in different regions of the same tumour. Whether there is need for multiregional analyses for prognosis estimation remains to be investigated\textsuperscript{108-110}.

In Chapter 6, an evaluation of the adoptability of different prognostic systems among urologists is assessed by means of a questionnaire, send to urologists with known interest in RCC. The answers of 86 responders were assessed, response rate to the questionnaire was 78.2\%. Our results indicated that of the variety of prognostic models and factors available\textsuperscript{83}, the TNM classification is still the most used prognostic tool by 97.7\% of our responders. Almost 90\% of the urologists also take clinical, laboratory and pathological factors into account. Very few respondents made use of markers. At the time of our survey the literature on the use of somatic copy number alternations and gene expression signatures as markers was limited, which precluded their inclusion in our questionnaire. Models that provide uniform methods to represent clinical, laboratory pathological factors and markers, were less frequently utilized. Preoperative models were scarcely used, 14\% mentioned the Raj model\textsuperscript{111}. The generally well performing preoperative model by Sorbellini was not mentioned in our survey\textsuperscript{112}. No more than 40\% of the urologists indicated to use postoperative models, mostly chosen were Kattan nomogram (17.4\%), SSIGN (12.8\%) and UISS (11.6\%). As expected from the literature, the most frequently used model was the Motzer model for metastatic RCC patients (37.2\%).

The main problem with using models according to our responders was the “lack of additional value”, confirmed by current literature described above. “Lack of familiarity” was also frequently mentioned as reason not to use models. At the time of our survey, European guidelines mentioned the possibility of using prognostic models in a small paragraph without further recommendations\textsuperscript{113}, which perhaps contributed to the lack of familiarity of prognostic models. In the most recent guideline the UISS, SIGN and Karackiewicz models are mentioned as most relevant prognostic systems for localized RCC, however no definite statements are made about their accuracy and usefulness\textsuperscript{114}. Limitations of our study included model selection for the questionnaire. Our results cannot be extrapolated to the entire urological community, since our questionnaire was spread among urologists with known interest in RCC and most of the responders were Europeans. Concluding, TNM still is the main prognostic system used in RCC, other prognostic systems are scarcely utilized, with exception of the Motzer model. Main reasons not to use models is their lack of incremental value to the TNM and their unfamiliarity on validation studies.

Chapter 7 aimed to improve prognosis estimation by using immunohistochemistry biomarkers in addition to other known prognostic factors and assesses their incremental value in a national cohort of localized/locally advanced RCCs. For our study patients with pT1b-pT3 clearcell RCC from three referral hospitals were selected. In total 143 patients were included with a median follow up of 63 months. Tissue microarrays were constructed, stained with 6 immunohistochemistry markers and images were analysed automatically. Combined CAIX under-expression (<30\%) and Vimentin over-expression (>50\%) was an independent predictor for CSS, recurrence free and overall survival and the combined marker displayed additional value to conventional prognosis estimation. The prognostic role of CAIX underexpression has been subject of much debate\textsuperscript{115-119}, a similar notion can be made for vimentin expression\textsuperscript{120-123}. Differences in techniques, cut-offs and tested population are the reasons for controversy. Results of current study indicate that combination of two easy and available immunohistochemistry markers can provide a slight but significant improvement in addition to conventional T-stage/grade. Several other protein marker combinations have been identified which provided incremental value to prognosis estimation by T-stage or clinical-pathological based models\textsuperscript{101,122-126}. In contrast to those studies, our research was done in selected T-stages and a non-metastatic population, possibly explaining our lesser increase in c-index. Validation of both our results and results of previous studies is necessary. Limitations of our research included the retrospective nature, long inclusion period and no availability of a validation cohort. Automatic image analysis was used to limit biases regarding interpretation.

In conclusion, biomarkers CAIX and vimentin can provide slight incremental
value to existing prognostic systems, however validation studies are essential before clinical implementation. The assessment of the excessive amount of prognostic biomarkers should preferably be done by multiple centres using large databases where uniform techniques and interpretation methods can be used.

FUTURE PERSPECTIVES

Epidemiological data will become more accurate as new databases are containing lesser missing data, while the implementation of clinical variables will provide a meaningful insight and interpretation. The still increasing incidental presentation of RCC could level out or further increase gender differences. Mortality in highly developed countries is expected to continue to decrease due to early diagnosis and treatment in curative settings, although a certain degree of overtreatment is currently unavoidable at the expenses of the possibly indolent SRMs.

PN will continue to be the preferred treatment in T1 tumours, when technically feasible and when the patients’ health allows it. Ideally a randomized trial should be conducted in patients with clinical T1b/T2a renal masses, assigning them to either minimal-invasive PN or laparoscopic RN. Proper documentation of comorbidities and nephrological status is essential for interpretation of these results, however as with the previous EORTC trial on PN and RN, poor accrual might be a problem. Prospective research can nonetheless focus on identifying preoperative patients who are at risk for adverse outcomes after renal surgery using amongst others preoperative GFR, presence of proteinuria and comorbidities. Preoperative medical CKD or lower baseline eGFR values are strong predictors of renal function decline and mortality. Patients with medical risk factors have additional increased risk of deterioration of renal function. The presences of proteinuria increases risk of mortality, severe kidney disease and myocardial infarction at any given GFR and is associated with an increased rate of renal function decline regardless of baseline GFR.

For elderly patients in whom long term benefits of PN are doubtful and in those who are prone to develop severe complications based on either tumour complexity or large tumour size and comorbidities, RN and for the very frail ablative therapies, can once again be adopted as acceptable options.

Increased understanding of underlying mechanisms of tumour genesis has led to identification of biomarkers, some of which have shown to have incremental value in prognosis estimation as indicated in the current thesis. The promise of a biomarker panel which in addition to current prognostic systems can identify patients with low and intermediate risk of recurrence, has never been more urgently needed. Not the subject of this thesis but an important objective for the future may be the possibility to predict the need for treatment based on biomarkers determined in biopsy specimen.

Furthermore determination of the true incidence and significance of intratumour heterogeneity remains a pending challenge as well as is proper characterization in biopsy tissue. Once identified, prognostic biomarkers need to be assessed in large-scale validation studies. Currently very few of such studies have been published for RCC.

So far few of the numerous biomarkers identified have been implemented in clinical practice. Reasons include their marginal prognostic value when compared to universal clinicopathological systems or their complexity. Easy determination, simplicity and accessibility as well as identification of the target population will be major points to be considered in addition to their strict prognostic accuracy.

REFERENCE LIST

Chapter 8

An analysis of patients with T2 renal cell carcinoma. BJU Int 2013; 111:E191-195


Assessing the accuracy and generalizability of the preoperative and postoperative Karakiewicz nomograms for renal cell carcinoma: results from a multicentre European and US study. BJU Int 2013; 112:578-584

Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. Nat Genet 2014; 46:229-233


114. B. Ljungberg (Chair), K. Bensalah, A. Bex. 
The EAU guidelines on renal cell carcinoma 2015.

Carbonic anhydrase IX is an independent predictor of survival in advanced renal clear cell carcinoma: implications for prognosis and therapy. Clin Cancer Res 2003; 9:802-811


119. Neto GJ, Cheng L. 

120. Garcia-Rag J, Ortiz N, Lokeshwar V. 


122. Adibi M, Karam JA, Wood CG. 
Reporting geographic and temporal trends in renal cell carcinoma: why is this important? Eur Urol 2015; 67:531-532

123. Welch HG, Schwartz LM, Woloshin S. 
Increasing 5-year survival rates evidence of success against cancer? JAMA 2000; 283:2975-2978

Low CAIX expression and absence of VHL gene mutation are associated with tumor aggressiveness and poor survival of clear cell renal cell carcinoma. Int J Cancer 2008; 123:395-400

Carbonic anhydrase IX (CAIX) is not an independent predictor of outcome in patients with clear cell renal cell carcinoma (cRCC) after long-term follow-up. BJU Int 2013; 111:1046-1053

126. Stillebroer AB, Mulders PF, Boerman OC, et al. 


The role of vimentin expression in predicting the long-term outcome of patients with localized renal cell carcinoma. Br J Urol 1997; 80:864-868

130. Shi ZG, Li SQ, Li ZJ, et al. 


Relation between kidney function, proteinuria, and adverse outcomes. JAMA 2010; 303:423-429


136. Chibon F. 

137. Kalia M. 
Biomarkers for personalized oncology: recent advances and future challenges. Metabolism 2015; 64:516-21

138. Laird A, Harrison DJ, Stewart GD. 
The development of prognostic and predictive biomarkers in renal cell cancer are not one and the same thing. Eur Urol 2015; 67:21-22

139. Yousef GM. 