Received: 2 December 2020 | Revised: 26 April 2021 | Accepted: 5 May 2021

DOI: 10.1002/ijc.33701

CANCER EPIDEMIOLOGY

IJC International Journal of Cancer

Differences in risk factors for molecular subtypes of clear cell renal cell carcinoma

Mark P. Purdue[1] | Jongeun Rhee[1] | Lee Moore[1] | Xiaohua Gao[2] | Xuezheng Sun[2] | Erin Kirk[2] | Vladimir Bencko[3] | Vladimir Janout[4] | Dana Mates[5] | David Zaridze[6] | Stacey Petruzella[7] | Abraham Ari Hakimi[7] | William Marston Linehan[8] | Stephen J. Chanock[1] | Paul Brennan[9] | Helena Furberg[7] | Melissa Troester[2] | Nathaniel Rothman[1]

[1]Division of Cancer Epidemiology and Genetics, Department of Health and Human Services, National Cancer Institute, National Institutes of Health, Bethesda, Maryland

[2]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

[3]First Faculty of Medicine, Institute of Hygiene and Epidemiology, Charles University, Prague, Czech Republic

[4]Faculty of Health Sciences, Palacky University, Olomouc, Czech Republic

[5]Department of Environmental Health, Institute of Public Health, Bucharest, Romania

[6]Institute of Carcinogenesis, Cancer Research Centre, Moscow, Russia

[7]Departments of Surgery and Epidemiology, Memorial Sloan Kettering Cancer Center, New York, New York

[8]Department of Health and Human Services, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland

[9]Section of Genetics, International Agency for Research on Cancer, Lyon, France

Correspondence
Mark P. Purdue, Occupational and Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD 20850, USA.
Email: purduem@mail.nih.gov

Funding information
National Cancer Institute, Grant/Award Number: P30 CA008748; National Cancer Institute Intramural Research Program

Abstract

The ccA and ccB molecular subtypes of clear cell renal cell carcinoma (ccRCC) have well-characterized prognostic relevance. However, it is not known whether they possess distinct etiologies. We investigated the relationships between these subtypes and RCC risk factors within a case-control study conducted in Eastern Europe. We analyzed risk factor data for ccA (n = 144) and ccB (n = 106) cases and 1476 controls through case-only and case-control comparisons to assess risk factor differences across subtypes using logistic and polytomous regression models. We also performed a meta-analysis summarizing case-only results from our study and three patient cohorts. Patients with ccB tumors had poorer survival than those with ccA tumors and were more likely to be male (case-only odds ratio [OR] 2.68, 95% confidence interval [CI] 1.43-5.03). In case-control analyses, body mass index was significantly associated with ccA tumors (OR 2.45, 95% CI 1.18-5.10 for ≥35 vs <25 kg/m^2) but not with ccB tumors (1.52, 0.56-4.12), while trichloroethylene was associated with ccB but not ccA (OR 3.09, 95% CI 1.11-8.65 and 1.25, 0.36-4.39 respectively for ≥1.58 ppm-years vs unexposed). A polygenic risk score of genetic variants identified from genome-wide association studies was associated with both ccA and, in particular, ccB (OR 1.82, 1.11-2.99 and 2.87, 95% CI 1.64-5.01 respectively for 90th vs 10th percentile). In a meta-analysis of case-only results including three patient cohorts, we still observed the ccB excess for male sex and the ccA excess for obesity. In conclusion, our findings suggest the existence of etiologic heterogeneity across ccRCC molecular subtypes for several risk factors.

KEYWORDS
case-control study, clear cell renal cell carcinoma, epidemiology, molecular pathology, risk factors

Abbreviations: BMI, body mass index; ccRCC, clear cell renal cell carcinoma; CEERCC, Central and Eastern Europe Renal Cell Cancer Study; CI, confidence interval; GWAS, genome-wide association studies; IARC, International Agency for Research on Cancer; MCC, Moffitt Cancer Center; MSKCC, Memorial Sloan Kettering Cancer Center; OR, odds ratio; PFOA, perfluorooctanoic acid; SNP, single-nucleotide polymorphism; TCE, trichloroethylene; TCGA-KIRC, the Cancer Genome Atlas Kidney Renal Clear Cell Carcinoma.

What's new?

Renal cell carcinoma (RCC) consists of several histologic types, with distinct genetic and clinical characteristics. Clear-cell RCC is classified by two subtypes, defined by distinct gene expression signatures: ccA, and the poorer-prognosis ccB. Do risk factors differ for these subtypes? In this study, the authors found that obesity is associated with increased risk for ccA tumors (but not ccB), while patients with ccB tumors are more likely to be male or have occupational exposure to trichloroethylene. These results suggest that the etiology of RCC may be more complex than previously understood.

1 | INTRODUCTION

Renal cell carcinoma (RCC) is made up of several histologic subtypes with distinct genetic and clinical characteristics. Clear cell RCC is the most common, making up over 75% of cases, and is typically associated with a poorer prognosis than the next most common subtypes, papillary and chromophobe RCC.[1] There has been great interest in characterizing the mutational and transcriptomic landscape of clear cell renal cell carcinoma (ccRCC) and identifying prognostically distinct molecular profiles to better inform the clinical management of patients.[2] One well-validated classification involves two subtypes defined by distinct gene expression signatures: ccA, characterized by overexpression of genes associated with hypoxia, angiogenesis and fatty- and organic acid metabolism, and the poorer-prognosis ccB, overexpressing genes regulating epithelial-to-mesenchymal transition, the cell cycle and wound healing.[3-6] Although the subtypes were originally identified from gene expression array data, a 34-gene panel ("ClearCode34") has been developed enabling the accurate classification of ccA/ccB subtype at a lower cost and using formalin-fixed paraffin-embedded tissue.[6]

Despite being recognized as a collection of distinct diseases with divergent biologic processes,[7] RCC has typically been investigated as a single disease in epidemiologic investigations; identified risk factors include male sex, African American race, obesity, hypertension, smoking, occupational exposure to the chemical trichloroethylene (TCE) and genetic variants identified in genome-wide association studies (GWAS).[1,8] No epidemiologic studies to date have explored whether ccRCC molecular subtypes possess distinct etiologies, although findings from some clinical case series suggest the existence of such heterogeneity, with differences in African American race and obesity observed between patients with ccA and ccB tumors.[9,10] These findings warrant further investigation, ideally in an epidemiologic study including both cases and controls to enable the calculation of subtype-specific relative risk estimates.

To that end, we classified ccRCC tumors within a case-control study of kidney cancer into molecular subtypes and investigated whether risk factor associations differ between ccA and ccB tumors. We also conducted a meta-analysis to combine our findings with those from three patient case series.

2 | MATERIALS AND METHODS

2.1 | Study design

The Central and Eastern European Renal Cell Cancer Study (CEERCC) is a hospital-based case-control study conducted in the Czech Republic, Poland, Romania and Russia between 1999 and 2003.[11] Incident cases of histologically confirmed RCC identified at participating hospitals were eligible for the study. Controls in each center were chosen among subjects admitted as in-patients or out-patients in the same hospital as the cases for noncancer, nongenitourinary or nontobacco-related conditions and were frequency matched with cases by sex, age (±3 years) and study center.

Cases (N = 1097; 90%-99% of contactable eligible individuals across centers) and controls (N = 1476 controls; 90%-96%) participated in in-person interviews conducted by trained interviewers to collect risk factor information. Industrial hygienists assessed potential TCE exposure based on recorded occupational histories and specialized questionnaires regarding workplace exposure to solvents and other chemicals.[11] Genomic DNA was extracted from whole blood buffy coat from 987 of 1097 cases and 1298 of 1476 controls. The survival status of cases in the Czech Republic, Poland and Russia was later captured through medical records abstraction and linkage to vital statistics and cancer registries.

2.2 | Tumor molecular subtype classification

Frozen tumor biopsy specimens were obtained for 524 cases, 477 of which were ccRCC.[12] After manual microdissection to remove non-tumor tissue, frozen sections were placed directly in Trizol reagent (Invitrogen, Carlsbad, CA), homogenized for 2 minutes on ice, and RNA was isolated using the manufacturer's protocol. RNA samples from 355 ccRCC cases were available for this analysis. RNA was quantified based on absorbance at 260 nm and samples with low concentration (n = 8) were excluded. All samples were analyzed to measure gene expression for the Clearcode34 panel (Table S1) and six housekeeping genes using nCounter (NanoString Technologies Inc., Seattle, WA). Samples with poor quality were removed using Nanostring diagnostics; specifically, nSolver tools were used to exclude 26 samples that had poor positive control linearity (positive control R^2 value

<0.95) and nine samples that had low overall expression (>10-fold lower expression of housekeeping genes relative to the average experimental sample) and were outliers in principal component analysis (visual inspection). Data for the remaining 312 cases were normalized using positive control genes and housekeeping genes, log-transformed (base two) and median-centered.

We developed a ccA/ccB classification model using normalized, base-two log-transformed and median-centered RNA sequencing data for the Clearcode34 panel from the Cancer Genome Atlas Kidney Renal Clear Cell Carcinoma (TCGA-KIRC) tumor collection (N = 524) as a training set.[2,6] The classification model, developed through prediction analysis of microarrays, was then applied to the study expression data to determine ccA/ccB subtype. We assessed the reproducibility of subtyping using data from 4 tumors with duplicate samples assayed and 39 tumors previously assayed and subtyped in a pilot project. We observed concordant subtypes across all 4 duplicate pairs and for 38 of the 39 tumors also included in the pilot, with the one discordant tumor having uncertain subtype classifications (ccA probabilities of 46% and 54% in the pilot and main study, respectively). To filter out cases of uncertain subtype, we restricted the statistical analysis to tumors with a predicted subtype probability >70% (n = 250).

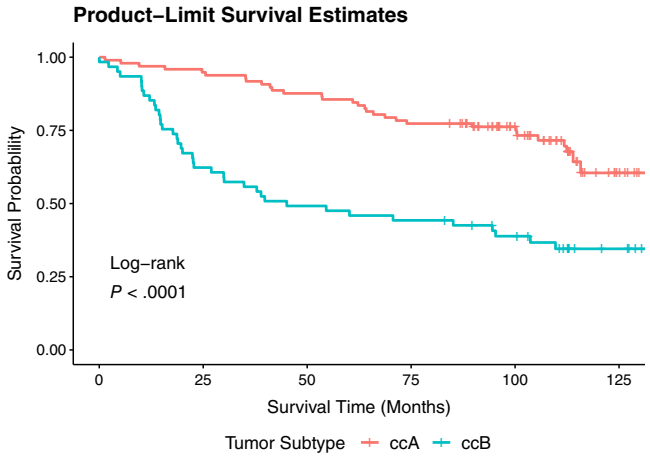

FIGURE 1 Kaplan-Meier overall survival curves by molecular subtype (ccA, ccB) for 158 clear cell renal cell carcinoma (RCC) cases with survival data in the Central and Eastern Europe Renal Cell Cancer (CEERCC) study

2.3 | Statistical analysis

We first assessed the comparability between included cases vs those without known ccA/ccB status and assessed overall survival of patients with ccA and ccB tumors through Kaplan-Meier analysis and Cox regression with adjustment for country, age, sex, education level (primary education [elementary unfinished or finished], secondary and apprenticeships, higher education [high school, university or higher]), body mass index (BMI), stage and grade.

For our investigation, we considered the following RCC risk factors: age, sex, BMI, smoking, hypertension, occupational exposure to TCE, family history of kidney cancer and 13 GWAS-identified SNPs associated with renal cancer. In addition to assessing the SNPs individually, a polygenic risk score including all 13 SNPs was constructed as a potentially informative global assessment of subtype differences

TABLE 1 Case-only analysis comparing selected subject and tumor characteristics between ccA (referent) and ccB tumors among ccRCC cases in the CEERCC study

		N_{ccA}	(%)	N_{ccB}	(%)	OR_{ccB}	(95% CI)[a]	P_{Het}[b]
Country	Czech Republic	103	(72)	62	(58)	1.00		.72
	Poland	10	(7)	11	(10)	1.85	(0.61-5.64)	
	Romania	12	(8)	17	(16)	1.24	(0.46-3.30)	
	Russia	19	(13)	16	(15)	1.02	(0.42-2.48)	
Age	<45	8	(6)	6	(6)	1.00		.45
	45-54	37	(26)	22	(21)	0.97	(0.24-3.92)	
	55-64	48	(33)	32	(30)	0.84	(0.22-3.22)	
	65+	51	(35)	46	(43)	1.47	(0.39-5.61)	
Sex	Female	72	(50)	28	(26)	1.00		.002
	Male	72	(50)	78	(74)	**2.68**	**(1.43-5.03)**	
Stage	I/II	98	(74)	37	(37)	1.00		.0001
	III/IV	34	(26)	64	(63)	**4.00**	**(2.12-7.58)**	
Grade	I	39	(40)	4	(7)	1.00		.004
	II	41	(42)	25	(44)	**5.26**	**(1.57-17.71)**	
	III	18	(18)	28	(49)	**8.49**	**(2.36-30.58)**	

Note: Missing data: stage, n = 17; grade, n = 95. ORs and 95% CI in bold are statistically significant.
Abbreviations: ccRCC; clear cell renal cell carcinoma; CEERCC, Central and Eastern Europe Renal Cell Cancer Study; CI, confidence interval; OR, odds ratio.
[a]ORs adjusted for country, age, sex, stage and grade.
[b]*P*-value from case-only Wald test of OR heterogeneity across subtypes.

across multiple SNPs. Odds ratios (OR) and 95% confidence intervals (CI) relating these variables to ccA or ccB subtype vs controls were calculated using polytomous regression with adjustment for age, sex, country and education level (we adjusted for education level to control for potential confounders correlated with socioeconomic status; this is relevant in particular for analyses of occupational exposures like TCE). Within the polytomous model we conducted a Wald test of OR heterogeneity across subtypes, testing the null hypothesis that the regression coefficient for a given risk factor was the same for each subtype. For age and sex, case-only analyses were conducted using logistic regression (ccB vs ccA) to assess subtype differences, with adjustment for country, stage and grade. This case-only approach was used because age and sex were used as matching factors for control selection and thus could not be investigated in an unbiased manner through case-control analysis.[13] Tests of trend across risk factor categories were performed by modeling the intracategory medians as a continuous variable. To account for multiple testing, we adjusted the *P*-values from tests of OR heterogeneity across the evaluated risk factors by the false discovery rate (FDR) of Benjamini and Hochberg (also known as *q*-values).

To compare our findings with those from recent patient cohorts, we conducted case-only analyses relating subtype to available risk factor information within TCGA-KIRC (205 ccA and 175 ccB tumors; sex evaluated adjusting for age and stage) and a patient cohort at Memorial Sloan Kettering Cancer Center (MSKCC; 109 ccA and 98 ccB tumors; sex and BMI evaluated adjusting for age and stage). For a third case series at Moffitt Cancer Center (MCC; 226 ccA and 56 ccB tumors), we calculated crude case-only ORs relating subtype to sex and BMI from previously reported tabular data.[10] For our meta-analysis we estimated summary ORs using a random-effects model and calculated the Higgin's I^2 statistic to evaluate heterogeneity across studies.

TABLE 2 Case-control analysis investigating associations with ccA and ccB tumors for known RCC risk factors in the CEERCC study

	$N_{Controls}$	(%)	N_{ccA}	(%)	OR_{ccA} [a]	(95% CI)	N_{ccB}	(%)	OR_{ccB}	(95% CI)[a]	P_{Het} [b]
Body mass index (kg/m^2)											.24
<25	538	(36)	33	(23)	1.00		30	(28)	1.00		
25 to <30	620	(42)	64	(44)	1.47	(0.94-2.29)	50	(47)	1.31	(0.82-2.11)	
30 to <35	250	(17)	35	(24)	**2.01**	**(1.20-3.36)**	21	(20)	1.53	(0.85-2.75)	
≥35	68	(5)	12	(8)	**2.45**	**(1.18-5.10)**	5	(5)	1.52	(0.56-4.12)	
					$P_{trend} = .002$				$P_{trend} = .16$		
Per 5 kg/m^2 increase					**1.30**	**(1.08-1.57)**			1.10	(0.87-1.39)	
TCE exposure (ppm-years)											.20
Unexposed	1146	(97)	109	(96)	1.00		81	(91)	1.00		
<1.58	21	(2)	2	(2)	0.69	(0.16-3.04)	3	(3)	1.67	(0.48-5.82)	
≥1.58	19	(2)	3	(3)	1.25	(0.36-4.39)	5	(6)	**3.09**	**(1.11-8.65)**	
					$P_{trend} = .94$				$P_{trend} = .049$		
GWAS polygenic risk score											.21
90th vs 10th percentile					**1.82**	**(1.11-2.99)**			**2.87**	**(1.64-5.01)**	
					$P = .02$[c]				$P = .0002$[c]		
Smoking status											.37
Never smoked	599	(41)	73	(51)	1.00		48	(45)	1.00		
Former smoker	353	(24)	29	(20)	0.84	(0.51-1.38)	20	(19)	0.53	(0.30-0.95)	
Current smoker	521	(35)	42	(29)	0.93	(0.59-1.46)	38	(36)	0.94	(0.57-1.55)	
High blood pressure											.45
No	906	(61)	82	(57)	1.00		54	(51)	1.00		
Yes	569	(39)	62	(43)	1.21	(0.84-1.75)	52	(49)	1.49	(0.98-2.25)	
Family history of kidney cancer											.74
No	1462	(99)	139	(97)	1.00		104	(98)	1.00		
Yes	14	(1)	5	(3)	2.61	(0.90-7.57)	2	(2)	1.97	(0.43-8.92)	

Note: Missing data: smoking status, n = 3; TCE, n = 337; GWAS polygenic risk score, n = 440; high blood pressure, n = 1. ORs and 95% CI in bold are statistically significant.

Abbreviations: CEERCC, Central and Eastern Europe Renal Cell Cancer Study; CI, confidence interval; GWAS, genome-wide association study; OR, odds ratio; TCE, trichloroethylene.

[a]ORs adjusted for country, age, sex and education level.

[b]*P*-value from test of OR heterogeneity across subtypes.

[c]*P*-value from Wald test of GWAS polygenic risk score, modeled as a continuous variable.

3 | RESULTS

Selected characteristics of the 250 ccRCC cases included in the analysis are summarized in Table S2. The cases' average age at diagnosis was 60.2 years, 60% were male and 66% were from the Czech Republic. The characteristics of the cases with subtype data were generally comparable to those of other ccRCC cases enrolled in the CEERCC case-control study with the exception of country of residence (reflecting differences across countries in success procuring tumor tissue) and a smaller proportion of subjects missing data on TCE exposure (Table S2). Kaplan-Meier survival curves indicated that, as expected, patients with ccB tumors had poorer overall survival than those with ccA tumors (Figure 1; log-rank test $P < .0001$). In Cox modeling, ccB status was associated with increased mortality after adjustment for age, sex, country, education level, BMI, stage and grade (hazard ratio 2.34, 95% CI 1.32-4.13; Table S3).

In case-only analyses exploring potential subtype heterogeneity, ccA/ccB status was observed to vary by sex, stage and grade (Table 1). Compared to cases with ccA tumors, ccB cases were more frequently male (OR 2.68, 95% CI 1.43-5.03) and had higher-stage (III/IV vs I/II: OR 4.00, 95% CI 2.12-7.56) and higher-grade (III vs I: OR 8.49, 95% CI 2.36-30.58) tumors. We did not observe statistical differences across subtypes for country or age at diagnosis.

Case-control analyses suggested differences in associations by subtype for BMI, TCE exposure and a GWAS polygenic risk score, although tests of OR heterogeneity did not reach statistical significance (Table 2). Moderate and severe obesity were significantly associated with ccA tumors (OR 2.01, 95% CI 1.20-3.36 and 2.45, 1.18-5.10 for BMI 30.0-34.9 and ≥35.0 vs <25.0 kg/m^2; $P_{trend} = .002$) but not with ccB tumors (1.53, 0.85-2.75 and 1.52, 0.56-4.12; $P_{trend} = .16$). In contrast, TCE exposure was significantly associated with ccB (OR 3.09, 1.11-8.65 for ≥1.58 ppm-years vs unexposed; $P_{trend} = .049$) but not with ccA tumors (1.25, 0.36-4.39; $P_{trend} = .94$). The GWAS polygenic risk score was associated with both ccA and, in particular, ccB tumors (90th vs 10th percentile: OR 1.82, 95% CI 1.11-2.99, $P_{trend} = .02$ and 2.87, 95% CI 1.64-5.01, $P_{trend} = .0002$ respectively). In individual analyses of the 13 SNPs (Table S4), statistically significant heterogeneity across subtypes ($P_{heterogeneity} = .049$) was observed for rs11894252, which was associated with ccB (OR 1.44, 95% CI 1.05-1.96, $P = .02$) but not ccA (OR 0.96, 95% CI 0.73-1.28, $P = .79$). When we excluded rs11894252 and another SNP with suggestive evidence of heterogeneity (rs718314; $P_{heterogeneity} = .09$) from the polygenic risk score, its associations with ccA and ccB were comparable (ORs 2.10 and 2.02 respectively).

We observed no statistical differences across subtypes for smoking, hypertension or family history of kidney cancer.

We adjusted our *P*-values from tests of OR heterogeneity to account for the number of comparisons among risk factors. After FDR adjustment, the test of heterogeneity across subtypes remained statistically significant for sex (q-value = 0.04) but not for rs11894252 (q-value = 0.51).

Figure 2 summarizes forest plots from a meta-analysis comparing case-only findings for sex and BMI from our study to those from TCGA-KIRC and patient cohorts at MSKCC and MCC. A male excess of ccB tumors was observed in three of the four studies and overall (summary OR 1.77, 95% CI 1.06-2.93). Cases with ccB tumors were less likely than ccA cases to have a BMI ≥30 kg/m^2 in each study with BMI data and overall (summary OR 0.65, 95% CI 0.45-0.93).

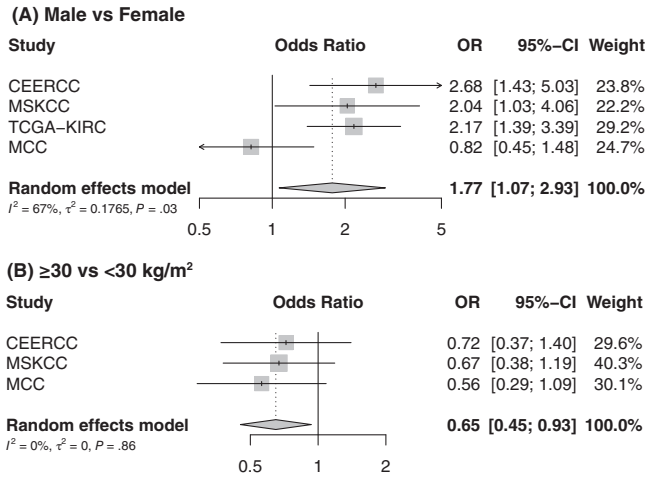

FIGURE 2 Forest plots from meta-analysis comparing case-only findings relating molecular subtype (ccB vs ccA) to (A) sex (male vs female) and (B) body mass index (BMI, ≥30 vs <30 kg/m^2) from Central and Eastern Europe Renal Cell Cancer Study (CEERCC) and three clinical case series (two with results for sex and BMI, one with results for sex only)

4 | DISCUSSION

In this case-control investigation of ccA and ccB subtypes of ccRCC, we found suggestive evidence of etiologic heterogeneity for several risk factors. In case-only analyses, male cases were significantly more likely to be diagnosed with ccB tumors. In case-control analyses, obesity was significantly associated with the risk of developing ccA tumors only, while TCE exposure was significantly associated with ccB tumors. A GWAS polygenic risk score was associated with both subtypes, ccB in particular. When we tested for OR heterogeneity, only the subtype differences by sex reached statistical significance. However, when we combined our case-only findings for sex and BMI with those of three case series in a meta-analysis, the summary results supported a ccB excess for men and a ccA excess for obese cases.

Our report of a ccB excess among male cases is to our knowledge the first of its kind, although male-female differences in ccRCC tumor biology have been previously noted. Clear cell RCC has been reported to have among the most extensive sex-biased molecular signatures across tumors types within TCGA and the highest number of genes differentially expressed by sex.[14] An earlier study of 480 ccRCC tumors also demonstrated substantial differences in gene expression by sex, with many immune and inflammation gene sets overexpressed in tumors from men, and gene sets related to catabolic process overexpressed among tumors from women.[4] Sex differences in mutation

patterns in key RCC genes have also been reported, with a significantly increased mutation frequency of *PBRM1* and the X chromosome encoded *KDM5C* in tumors from male patients and more frequent mutations of *BAP1* in tumors from female patients.[15] Our finding raises the question of whether ccA and ccB are differentially affected by sex hormone pathways and, epidemiologically, whether heterogeneity exists between these subtypes in their relationships with reproductive health factors that have been tentatively associated with RCC risk, including oral contraceptive use, parity and hysterectomy.[16,17]

In our meta-analysis the ccB excess was apparent overall and among the CEERCC, MSKCC and TCGA cases, but not within the MCC study. The reason for this heterogeneity is unclear, although it is notable that for MCC we were only able to report crude ORs calculated from tabular data reported in an earlier publication, as opposed to working with individual-level data.[10] It is possible that our results for this study may have changed if we had the opportunity to analyze the individual-level data and adjust for age and tumor characteristics. It is also notable that the subtype distribution in the MCC case series is different from that of the other studies; the proportion of tumors classified as ccA in MCC (80%) was considerably higher than in the other studies (MSK, 52%; TCGA, 54%; CEERCC, 55%). As the tumors in the MCC study were selected on the basis of tissue quantity, it is possible that the case series is not representative of all cases and could have introduced bias.

Our findings further support the existence of heterogeneity in the relationship with obesity between subtypes, as previously reported by the MCC case series.[10] Through case-control analyses, we found BMI to be more strongly associated with the risk of developing ccA than ccB tumors. These findings suggest that the underlying biological mechanisms through which obesity affects RCC development may operate through subtype-specific pathways. The stronger BMI association with ccA is notable given that this subtype is characterized by overexpression of genes involved in hypoxia and angiogenesis. Obesity-induced kidney injury can lead to chronic renal hypoxia, potentially inducing persistent upregulation of the von Hippel-Lindau (VHL)/hypoxia-inducible factor pathway and downstream oncogenic factors known to stimulate ccRCC development.[18] Given these observations, we speculate that hypoxia-driven effects of obesity may be particularly potent for driving progression of ccA tumors. Although our finding of a stronger BMI association with the better-prognosis ccA subtype is seemingly consistent with the postulated RCC "obesity paradox," whereby obese patients in some studies have had better clinical outcomes than those of normal weight,[19,20] we did not observe an association between BMI and survival in our cases (Table S3).

Our findings suggest that TCE exposure is particularly important in the development of ccB tumors. TCE, an industrial solvent used to clean metal parts, has been classified as a human kidney carcinogen by IARC.[21] We note that a smaller proportion of cases with ccA/ccB classification had missing data on TCE than the ccRCC cases not included in the analysis. This is attributable to the smaller proportion of subtyped cases from Russia, where tumor procurement was less successful and missing data for TCE was greater. As our analysis of TCE is adjusted for country, this missingness is unlikely to have biased our findings. The mechanisms of TCE carcinogenicity remain to be fully elucidated, although there is strong toxicologic evidence that metabolites formed from glutathione conjugation pathways in the kidney are genotoxic.[21] Our observation that TCE exposure is more strongly associated with the more-lethal ccB subtype is an important new finding, but requires confirmation in other studies employing similarly high-quality exposure assessment before meaningful conclusions can be drawn. There is also a need to investigate molecular subtype differences in the association with other suspected renal carcinogens, such as the dry cleaning solvent perchloroethylene, the surfactant perfluorooctanoic acid (PFOA) and the mold toxin aristolochic acid.[22-24]

We also found ccB tumors to be more strongly associated with a polygenic risk score of GWAS-identified RCC susceptibility variants, although tests of OR heterogeneity were not statistically significant for the risk score and most individual SNPs. The one exception was a ccB-specific effect observed for rs11894252, mapping to an intron of *EPAS1*. *EPAS1*, which encodes the VHL-regulated transcription factor HIF-2α, is strongly expressed in ccA tumors, and part of the Clearcode34 gene panel; it is thus unexpected to find this variant to be more strongly associated with ccB. We cannot rule out chance as an explanation for this finding, although it is notable that the rs11894252 association with RCC has been observed to be stronger for men, among whom ccB tumors are more frequent, than women.[25]

Strengths of our case-control investigation of etiologic heterogeneity across ccA/ccB subtypes, to our knowledge the first of its kind, include the inclusion of data from controls to directly calculate estimates of subtype-specific relative risks (as opposed to case-only results), the availability of information on important RCC risk factors, and the inclusion of a meta-analysis to assess the consistency of subtype differences by sex and BMI across clinical case series. Our study was limited by the sample size for ccRCC molecular subtypes, which weakened our statistical power for tests of OR heterogeneity, and our inability to assess potential intratumor heterogeneity in ccA/ccB phenotype with the available specimens.[26] However, the confirmation through meta-analysis of our observed subtype differences by sex and BMI supports the validity of these findings. Additional evidence is needed to confirm the novel subtype differences we observed for TCE exposure and GWAS susceptibility variants.

In conclusion, our investigation of etiologic heterogeneity provides new evidence to suggest that the etiology of ccRCC is more complex than previously recognized, with differences by ccA/ccB subtype observed for sex and BMI in multiple case groups. We also observed suggestive evidence in CEERCC of subtype differences in associations with TCE exposure and genetic susceptibility; however, as statistical tests of OR heterogeneity did not reach significance and these associations have not yet been investigated in other studies, our findings should be interpreted with caution. Additional epidemiologic research is needed to confirm our findings and extend investigations of etiologic heterogeneity to other suspected RCC risk factors. More broadly, our findings underscore the importance of collecting tumor tissue in epidemiologic studies of renal cancer where feasible.

ACKNOWLEDGMENTS

This research was supported by the Intramural Research Program of the National Cancer Institute, NIH. The MSK project was supported by a National Cancer Institute MSK Cancer Center Support Grant/Core Grant (P30 CA008748).

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT

The study protocol was approved by institutional review boards of all participating centers, the International Agency for Research on Cancer (IARC) and the National Cancer Institute. All study subjects and their physicians provided written informed consent.

ORCID

Mark P. Purdue https://orcid.org/0000-0003-1177-3108
Jongeun Rhee https://orcid.org/0000-0003-1873-9359
David Zaridze https://orcid.org/0000-0002-2824-3704
Paul Brennan https://orcid.org/0000-0002-0518-8714

REFERENCES

1. Chow WH, Scelo G and Tarone RE. Renal cancer. In: MJ Thun, MS Linet, JR Cerhan, CA Haiman and D Schottenfeld. *Cancer Epidemiology and Prevention*. 4th ed. New York: Oxford University Press, 2018:961–76.
2. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature*. 2013;499:43-49.
3. Brannon AR, Reddy A, Seiler M, et al. Molecular stratification of clear cell renal cell carcinoma by consensus clustering reveals distinct subtypes and survival patterns. *Genes Cancer*. 2010;1:152-163.
4. Brannon AR, Haake SM, Hacker KE, et al. Meta-analysis of clear cell renal cell carcinoma gene expression defines a variant subgroup and identifies gender influences on tumor biology. *Eur Urol*. 2012;61:258-268.
5. Ghatalia P, Rathmell WK. Systematic review: ClearCode 34 – a validated prognostic signature in clear cell renal cell carcinoma (ccRCC). *Kidney Cancer*. 2018;2:23-29.
6. Brooks SA, Brannon AR, Parker JS, et al. ClearCode34: a prognostic risk predictor for localized clear cell renal cell carcinoma. *Eur Urol*. 2014;66:77-84.
7. Haake SM, Rathmell WK. Renal cancer subtypes: should we be lumping or splitting for therapeutic decision making? *Cancer*. 2017;123:200-209.
8. Scelo G, Purdue MP, Brown KM, et al. Genome-wide association study identifies multiple risk loci for renal cell carcinoma. *Nat Commun*. 2017;8:15724.
9. Krishnan B, Rose TL, Kardos J, Milowsky MI, Kim WY. Intrinsic genomic differences between African American and White patients with clear cell renal cell carcinoma. *JAMA Oncol*. 2016;2:664-667.
10. Haake SM, Brooks SA, Welsh E, et al. Patients with ClearCode34-identified molecular subtypes of clear cell renal cell carcinoma represent unique populations with distinct comorbidities. *Urol Oncol*. 2016;34:122.e1-122.e7.
11. Moore LE, Boffetta P, Karami S, et al. Occupational trichloroethylene exposure and renal carcinoma risk: evidence of genetic susceptibility by reductive metabolism gene variants. *Cancer Res*. 2010;70:6527-6536.
12. Moore LE, Nickerson ML, Brennan P, et al. Von Hippel-Lindau (VHL) inactivation in sporadic clear cell renal cancer: associations with germline VHL polymorphisms and etiologic risk factors. *PLoS Genet*. 2011;7:e1002312.
13. Begg CB, Zhang ZF. Statistical analysis of molecular epidemiology studies employing case-series. *Cancer Epidemiol Biomarkers Prev*. 1994;3:173-175.
14. Yuan Y, Liu L, Chen H, et al. Comprehensive characterization of molecular differences in cancer between male and female patients. *Cancer Cell*. 2016;29:711-722.
15. Ricketts CJ, Linehan WM. Gender specific mutation incidence and survival associations in clear cell renal cell carcinoma (CCRCC). *PLoS One*. 2015;10:e0140257.
16. Karami S, Daugherty SE, Schonfeld SJ, et al. Reproductive factors and kidney cancer risk in 2 US cohort studies, 1993-2010. *Am J Epidemiol*. 2013;177:1368-1377.
17. Michels KA, Brinton LA, Pfeiffer RM, Trabert B. Oral contraceptive use and risks of cancer in the NIH-AARP Diet and Health Study. *Am J Epidemiol*. 2018;187:1630-1641.
18. Klinghoffer Z, Yang B, Kapoor A, Pinthus JH. Obesity and renal cell carcinoma: epidemiology, underlying mechanisms and management considerations. *Expert Rev Anticancer Ther*. 2009;9:975-987.
19. Donin NM, Pantuck A, Klopfer P, et al. Body mass index and survival in a prospective randomized trial of localized high-risk renal cell carcinoma. *Cancer Epidemiol Biomarkers Prev*. 2016;25:1326-1332.
20. Hakimi AA, Furberg H, Zabor EC, et al. An epidemiologic and genomic investigation into the obesity paradox in renal cell carcinoma. *J Natl Cancer Inst*. 2013;105:1862-1870.
21. IARC. *Trichloroethylene, Tetrachloroethylene, and Some Other Chlorinated Solvents*. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Lyon, France: International Agency for Research on Cancer; 2012:35-217.
22. Purdue MP, Stewart PA, Friesen MC, et al. Occupational exposure to chlorinated solvents and kidney cancer: a case-control study. *Occup Environ Med*. 2017;74:268-274.
23. Benbrahim-Tallaa L, Lauby-Secretan B, Loomis D, et al. Carcinogenicity of perfluorooctanoic acid, tetrafluoroethylene, dichloromethane, 1,2-dichloropropane, and 1,3-propane sultone. *Lancet Oncol*. 2014;15:924-925.
24. Turesky RJ, Yun BH, Brennan P, et al. Aristolochic acid exposure in Romania and implications for renal cell carcinoma. *Br J Cancer*. 2016;114:76-80.
25. Laskar RS, Muller DC, Li P, et al. Sex specific associations in genome wide association analysis of renal cell carcinoma. *Eur J Hum Genet*. 2019;27:1589-1598.
26. Gulati S, Martinez P, Joshi T, et al. Systematic evaluation of the prognostic impact and intratumour heterogeneity of clear cell renal cell carcinoma biomarkers. *Eur Urol*. 2014;66:936-948.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Purdue MP, Rhee J, Moore L, et al. Differences in risk factors for molecular subtypes of clear cell renal cell carcinoma. *Int. J. Cancer*. 2021;1–7. https://doi.org/10.1002/ijc.33701