

Supplementary Material, Appendix 1. Sample size justification.

Sample size justification

Since cCLS is an aggregate score derived from the subjective algorithmic assessment of up to six MR imaging features (22), we inferred that a minimum of 10 ccRCC and 10 non-ccRCC SRMs were required per feature to minimize the risk of overfitting cCLS accuracy. This required 60 ccRCC SRMs and 60 non-ccRCC SRMs. Using the lower prevalence of non-ccRCC in SRMs (i.e. 35.6%)(4), we divided the number of non-ccRCC SRMs required to fit the cCLS algorithm (i.e. 60) by the prevalence of non-ccRCC SRMs, which yielded the number of consecutive SRMs required to obtain 60 non-ccRCC SRMs. Therefore, 169 consecutive cT1a renal masses were needed to obtain 60 non-ccRCC tumors ($60/0.365$), which would include an estimated 109 ccRCC SRMs. Our sample of 250 SRMs (50 SRMs per institution) therefore had a higher number of masses than the minimum number required. Moreover, relatively wide 95% CIs for cCLS3 SRMs have been previously reported (i.e., 2%,23% (24) and 9%,30% (23)). We estimated a sample of 250 masses (50 SRMs per site) to most efficiently reduce the 95% CI for cCLS3 to 12%, 24%.

Supplementary Material, Appendix 2. Histological Diagnosis.

All renal masses in the study had a histopathological diagnosis established by either nephrectomy (71% [177/250]) or image-guided percutaneous 18-gauge or larger core needle biopsy (29% [73/250]). All specimens were processed and diagnoses rendered by urological pathologists using standards specified by the available version of the World Health Organization (WHO) classification of renal tumors at the time of processing (44). Histological diagnosis was further grouped into malignant and benign neoplasms (Supplemental Table 1) (42).

Supplementary Table 1. Dichotomization of histology results in malignant and benign categories.

Malignant	Benign
Clear cell RCC	Oncocytoma
Papillary RCC	Fat-poor AML
Chromophobe RCC	Oncocytic neoplasm ¹
RCC unclassified	
Clear cell-papillary tumors	
Metastasis	
Mucinous tubular and spindle cell tumors	
Low grade epithelial neoplasm	
Multilocular cystic neoplasm of low malignant potential	
Lymphoma	
Epithelioid AML	
Sclerosing PEComa (low grade epithelioid neoplasm)	

1-Oncocytic neoplasms are considered a benign/indolent diagnosis (23)

Supplementary Table 2a. Multiparametric MRI techniques used for renal masses at institution 1, including imaging at 1.5 and 3 Tesla^a.

Pulse Sequence	Dual-echo T1W GRE			T2W TSE/FSE	Volume Interpolated T1W 3D GRE ^b		Diffusion-weighted imaging ^c
	2D GRE	3D GRE		Single-shot TSE/FSE	3T	1.5T	Single-shot echo-planar imaging
		3T	1.5T				
Physiology	Breath Hold	Breath Hold	Breath hold	Respiratory Triggered Breath hold	Breath Hold		Breath Hold
Duration	21 sec.	16 sec.	20 sec.	3-4 min. 22 sec.	20 sec.		21 sec
Fat Suppression	N/A	N/A	N/A	N/A	Chemical or Spectral Inversion Recovery		Spectral Inversion Recovery
TE (IP/OP) ^c ; TR (msec)	(4.6/2.3);160-180	(2.5/1.3);5.5 and (2.2/1.1);4.0	(4.6/2.3);7.6	83-88;1030	1.7-2.5; 4.0-4.5	1.4;4.3	60.8-74;2075-4600
Flip angle (degrees)	70	10-12	10	180	10-12	10-12	90
Bandwidth (Hz)	260	700	313	450	325-460	488	250-1446
Number of excitations	1	0.7-1	1	Half-Fourier	1	1	2
Acceleration factor	2	2	1	1 2	2	2	2
Matrix Size	256/320 x 134/152	294 x 224	192 x 320	170 x 256	256 x 320	132 x 320	130-38;96-75
Field of view (cm)	25 x 35	25 x 35	25 x 35	25 x 35	25 x 35	25 x 35	40-380;28-75
Section thickness (mm)	5-6	3-4	3-5	5	2.5-4	2.5-4	6

- a- Imaging performed on clinical 1.5 Tesla (Symphony or Aera, Siemens Healthcare) or 3 Tesla (TRIO, Siemens Healthcare; Discovery 750 W, General Electric Healthcare) systems.
- b- VIBE (Siemens Healthcare), LAVA (General Electric Healthcare).
- c- Diffusion-weighted imaging was performed with two b values (0 and 600 mm²/sec) automatically derived from the ADC map.
- d- IP=in-phase, OP=opposed phase.

Supplementary Table 2b. Multiparametric MRI techniques used for renal masses at institution 2, including imaging at 1.5 and 3 Tesla^a.

a- Imaging performed on clinical 1.5 Tesla (Ingenia, Philips Medical Systems or Espree, Siemens Healthcare)

Pulse Sequence	Dual-echo T1W GRE			T2W		Volume Interpolated T1W			Diffusion-weighted imaging	
	2D GRE	3D GRE		Single-shot TSE/FSE		3T	1.5T ^b	1.5T ^c	Single-shot echo-planar imaging ^d	
Physiology	Breath hold	Breath hold ^b	Breath hold ^c	Respiratory triggered	Breath hold	Breath hold			Respiratory triggered	Free breathing
Duration	62 sec.	15 sec.	20 sec.	1.42 min.	22 sec.	15 sec.	17 sec.	23 sec.	1.51 min.	3 min. 44 sec.
Fat Suppression	N/A	mDixon	N/A	N/A	N/A	N/A		Q fat sat	Spectral Inversion Recovery	SPAIR
TE (IP/OP) ^e ; TR (msec)	(2.38/4.87); 122	(1.35/2.5);3.8	(1.8/4.0);5.8	20;900	120;800	(1.32/2.3);3.7	1.8;4.0	239;5.07	62;3400	84;5700
Flip angle (degrees)	70	10	15	90	150	10	15	10	90	N/A
Bandwidth (Hz)	410	1214	585	569	488	1337	615	390	2200	1302
Number of excitations	1	1	1	1	1	1	1	1	2	4
Acceleration factor	N/A	4.5	4 (C-sense)	2	197	6	4.5 (C-sense)	2 GRAPPA	2	2 GRAPPA
Matrix Size	168 x 256	264 x 264	240 x 223	244 x 213	197 x 256	228 x 200	188 x 189	160 x 320	112 x 170	140 x 192
Field of view (cm)	360	340 x 340	340	340	360	320 x 280	340	360	340	360
Section thickness (mm)	4 x 1	5/2.5	5	5	6 x 1.2	2.8/1.4	4	3.0	7	6 x 1.2

or 3 Tesla (Ingenia, Philips Medical Systems) systems.

b- Imaging parameters for 1.5 Tesla (Ingenia, Philips Medical Systems).

c- Imaging parameters for 1.5 Tesla (Espree, Siemens Healthcare) system.

d- Diffusion-weighted imaging was performed with two b values (20 and 750 mm²/sec) with the ADC map automatically derived (for Philips Medical Systems) or three b values (50, 400, and 800 mm²/sec) with the ADC map automatically derived (for Siemens Healthcare).

f- IP=in-phase, OP=opposed phase

Supplementary Table 2c. Multiparametric MRI techniques used for renal masses at institution 3, including imaging at 1.5 and 3 Tesla^a.

Pulse Sequence	Dual-echo T1W GRE			T2W TSE/FSE		Volume Interpolated T1W 3D GRE ^b		Diffusion-weighted imaging ^c	
	2D GRE	3D GRE		Single-shot TSE/FSE		3T	1.5T	Single-shot echo-planar imaging	
		3T	1.5T	3T	1.5T			3T	1.5T
Physiology	Breath Hold	Breath Hold	Breath hold	Breath hold		Breath Hold		Free breathing	
Duration	15-22 sec	15-22 sec	15-22 sec	45-70 sec 45-70 sec		15-22 sec		15-22 sec	
Fat Suppression	N/A	N/A	N/A	N/A		SPAIR		SPAIR	
TE (IP/OP)^d; TR (msec)	(2.5/3.7);206	(2.5/1.2);3.9	(4.8/2.4); 190	182;1200 97;4680	179;1200 87;4500	1.3;3.2	1.2;3.1	56;7500	57;8500
Flip angle (degrees)	65	9	70	159 67	160 81	9	9	90	90
Bandwidth (Hz)	930/1502 (IP/OP)	1085	445/490 (IP/OP)	505 650	505 650	505	505	2790	2440
Number of excitations	1	1 Ask	1	1	1	1	1	3	3
Acceleration	1-3	1-2	1-3	2-3	2-3	1-3	1-3	2	2

factor									
Matrix Size	256 x 134	280 x 187	320 x 174	320 x 189 256 x 256	320 x 189 256 x 256	320 x 224	320 x 176	128 x 104	128 x 104
Field of view (cm)	36 x 27	46 x 37	36 x 33	46 x 39 38 x 38	36 x 30 38 x 38	36 x 36	36 x 33	50 x 41	38 x 31
Section thickness (mm)	5	4	4	5 5	4 5	3	3	5	5

- a- Imaging performed on clinical 1.5 Tesla (Aera, Siemens Healthcare; AvantoFit, Siemens Healthcare) or 3 Tesla (Skyra, Siemens Healthcare; TrioTim, Siemens Healthcare) systems.
- b- VIBE (Siemens Healthcare), LAVA (General Electric Healthcare).
- c- Diffusion-weighted imaging was performed with b values of 50, 500 and 1000 sec/mm² for Siemens 3T Skyra and b values of 50 and 800 sec/mm² for Siemens 1.5 T Aera, with the ADC map automatically derived.
- d- IP=in-phase, OP=opposed phase.

Supplementary Table 2d. Multiparametric MRI techniques used for renal masses at institution

4, including imaging at 1.5 and 3 Tesla^a.

Pulse Sequence	Dual-echo T1W GRE		T2W TSE/FSE		Volume Interpolated T1W 3D GRE ^b		Diffusion-weighted imaging ^c	
	2D GRE		Single-shot TSE/FSE				Single-shot echo-planar imaging	
	3T	1.5T	3T	1.5T	3T	1.5T	3T	1.5T
Physiology	Breath Hold		Breath hold or Respiratory Triggered		Breath Hold		Free breathing or Respiratory Triggered	
Duration	32-80 sec		44-235 sec		18-24 sec		171-304 sec	
Fat Suppression	N/A		N/A		SPAIR/SPIR		SPAIR/SPIR	
TE (IP/OP) ^d ;	(2.3-2.5/1.2-	(4.6-	70-78;2000-	71-80;2355-	1.3;3	1.8-2.4;3.9-	44-69;1705-	58-64;1747-
TR (msec)	1.3);170-201	4.9/2.3);150-172	2725	3000		4.7	7500	8000

Flip angle (degrees)	55-70	70-75	90-116	90-140	9-10	9	90	90
Bandwidth (Hz)	1090-2828	450-541	977-1261	403-781	450-719	350-434	2488	2332
Number of excitations	1		1	1	1	1	3-7	3-4
Acceleration factor	1.6-2	1.8-2	2-3	2-3	1.2-4	3-4	2-3	2
Matrix Size	180-256 x 153-208	220-320 x 182- 185	256-268 x 256-268	256-276 x 256-276	212-320 x 189-195	212-320 x 192-218	132-134 x 108-115	124-134 x 108- 128
Field of view (cm)	30-32x34-38		30-38x35-38		30-38x38-42		30-37x37-40	
Section thickness (mm)	5-6		6	6	3.3	4	5-6	6

- a- Imaging performed on clinical 1.5 Tesla (Ingenia, Philips Healthcare) or 3 Tesla (Skyra, Siemens Healthcare; or Achieva, Philips Healthcare) systems.
- b- VIBE (Siemens Healthcare) and THRIVE (Philips Healthcare).
- c- Diffusion-weighted imaging was performed with two b values (0 - 800 mm²/sec) automatically derived from the ADC map.
- d- IP=in-phase, OP=opposed phase.

Supplementary Table 2e. Multiparametric MRI techniques used for renal masses at institution 5, including imaging at 1.5 and 3 Tesla^a.

Pulse Sequence	Dual-echo T1W GRE	T2W TSE/FSE	Volume Interpolated T1W 3D GRE^b	Diffusion-weighted
-----------------------	--------------------------	--------------------	---	---------------------------

			Single-shot TSE/FSE			imaging ^c
	3T	1.5T		3T	1.5T	Single-shot echo-planar imaging
Physiology	Breath Hold	Breath hold	Respiratory Triggered Breath hold	Breath Hold		Breath Hold
Duration	18-20 sec.	11-22 sec.	2-3 min. 22-30 sec.	16-22 sec.		16-24 sec
Fat Suppression	N/A	N/A	N/A	Chemical or Spectral Inversion Recovery		Spectral Adiabatic Inversion Recovery
TE (IP/OP)^c; TR (msec)	(2.46/1.23);110-144	(4.4/2.2);120-190	92; 900-1010	1.45; 3.12	1.7; 3.51	59-97;3000-7300
Flip angle (degrees)	52-70	70	160 (1.5T) 90-110 (3T)	10-12	10-12	90
Bandwidth (Hz)	600-1085	500	679-780	750	500	2035 - 2085
Number of excitations	1	1	Half-Fourier	1	1	2-4
Acceleration factor	2	2	1 2	2	2	2
Matrix Size	224-294 x 220-272	192 x 256	256-320 x 203-320	256 x 166	256 x 166-187	160-192 x 96-186
Field of view (cm)	25-37 x 35-39	30-35 x30- 35	27-37 x 30-37	35 x 35	30-35 x 30-35	35-41x 23-36
Section thickness (mm)	4-8	4-8	4-5	2.2-2.5	2-2.2	4-8

- a- Imaging performed on clinical 1.5 Tesla (Symphony, Aera, Avanto or Espree, Siemens Healthcare) or 3 Tesla (Skyra, Vida, Prisma Fit or Biograph mMR Siemens Healthcare) systems.
- b- VIBE (Siemens Healthcare).
- c- Diffusion-weighted imaging was performed with two b values (0, 50 and 800 mm²/sec) automatically derived from the ADC map.
- d- IP=in-phase, OP=opposed phase.

Supplementary Table 3. Clear Cell Renal Cell Carcinoma Likelihood Score ‘Tip Sheet’.

MAJOR CRITERIA	EVALUATION (ALL MASSES)	DATA ENTRY
T2W Signal Intensity (SI) relative to renal cortex	-Evaluate only the SOLID (enhancing) components (determined on the delayed phase) -Assign the T2W SI based on the PREDOMINANT pattern (not the brightest area)	___ HYPERINTENSE ___ ISOINTENSE ___ HYPOINTENSE
CORTICOMEDULLARY (CM) PHASE ENHANCEMENT	QUANTITATIVE: ROI MEASUREMENT -Place ROI (~100 mm ²) in MOST enhancing area -Place ROI (~100 mm ²) in adjacent renal cortex -Calculate the ratio: Tumor SI/Cortex SI * 100%	SI Ratio = _____ % ___ INTENSE (>75%) ___ MODERATE (40-75%) ___ LOW (<40%)
MICROSCOPIC FAT	UNEQUIVOCAL drop of signal intensity comparing OP to IP images Subjective, quantitative for equivocal presence of microscopic fat: (SI.tumor.IP - SI.tumor.OP) > (SD.IP + SD.OP)	___ MICROSCOPIC FAT ___ NO MICROSCOPIC FAT
MINOR CRITERIA (ANCILLARY FEATURES)	EVALUATION (ONLY WHEN ALGORITHM INDICATES)	DATA ENTRY
SEGMENTAL ENHANCEMENT	'FLIP FLOP' enhancement – one portion of the tumor	

INVERSION	'washes out' and another 'washes in' comparing the CM and delayed phase. -Not necessarily on the same section.	____ SEI PRESENT ____ SEI ABSENT
RESTRICTED DIFFUSION	-Evaluate the solid (enhancing component) -Evaluate the PREDOMINANT pattern relative to renal cortex -Restricted diffusion=High signal on high b DWI and low signal on ADC	____ RESTRICTED DIFFUSION ____ NO RESTRICTED DIFFUSION
Arterial to Delayed Enhancement Ratio (ADER)	ADER = $\frac{SI.tumor.CM - SI.tumor.PRE}{SI.tumor.NG - SI.tumor.PRE}$ SUBJECTIVE = Does the tumor washout on the NG Phase compared to the CM phase	ADER = _____ ____ ≥ 1.5 ____ < 1.5 ____ TUMOR WASHES OUT ____ TUMOR DOES NOT WASHOUT

Supplementary Table 4. Individual radiologist diagnostic performance (sensitivity, specificity, positive predictive and negative predictive values) and overall results¹ for the diagnosis of clear cell RCC using a multiparametric MRI clear cell likelihood score of 4 (likely) or 5 (very likely) as a positive test result.

	Sensitivity (95% CI)	Specificity (95% CI)	Positive Predictive Value (95% CI)	Negative Predictive Value (95% CI)	Overall Accuracy (95% CI)
Institution 1					
Radiologist 1	88% (69%, 99%)	60% (39%, 79%)	69% (50%, 84%)	83% (59%, 96%)	74% (62%, 86%)
Radiologist 2	64% (43%, 82%)	72% (51%, 88%)	70% (47%, 87%)	67% (46%, 84%)	68% (55%, 81%)
Institution 2					
Radiologist 1	90% (68%, 99%)	70% (51, 85%)	67% (46%, 84%)	91% (72%, 99%)	80% (69%, 91%)
Radiologist 2	80% (56, 94%)	87% (69, 96%)	80% (56%, 94%)	87% (69%, 96%)	83% (72%, 94%)
Institution 3					
Radiologist 1	72% (51%, 89%)	88% (69%, 98%)	86% (64%, 94%)	76% (57%, 90%)	80% (69%, 91%)
Radiologist 2	72% (51, 88%)	92% (74%, 99%)	90% (68%, 99%)	76% (58%, 90%)	82% (72%, 93%)
Institution 4					
Radiologist 1	77% (54, 92%)	79% (59%, 92%)	74% (52%, 90%)	82% (62%, 94%)	78% (66%, 90%)
Radiologist 2	59% (36%, 79%)	82% (63%, 94%)	72% (47%, 90%)	72% (53%, 86%)	71% (58%, 83%)
Institution 5					
Radiologist 1	74% (54, 89%)	78% (56%, 93%)	80% (59%, 93%)	72% (51%, 88%)	76% (64%, 88%)

Radiologist 2	74% (54%, 89%)	74% (52, 90%)	77% (56%, 91%)	71% (49%, 87%)	74% (62%, 86%)
Overall	75% (68%-81%)	78% (72%, 84%)	76% (69%, 81%)	77% (72%, 82%)	77% (66%-88%)

1-Overall results are derived from a random-effects logistic regression model.

Supplementary Table 5. Individual radiologists' negative predictive value and pooled results¹ for the exclusion of clear cell RCC using a multiparametric MRI clear cell likelihood score of 2 (unlikely) or 1 (very unlikely) as a negative test result.

	Sensitivity (95% CI)	Specificity (95% CI)	Positive Predictive Value (95% CI)	Negative Predictive Value (95% CI)	Overall Accuracy (95% CI)
Institution 1					
Radiologist 1	100% (54%, 100%)	62% (54%, 71%)	57% (41%, 72%)	100% (54.1%, 100%)	62% (54%, 71%)
Radiologist 2	100%	70%	63%	100%	70%

	(70%, 100%)	(60%, 80%)	(46%, 77%)	(69.2%, 100%)	(60%, 80%)
Institution 2					
Radiologist 1	100%	40%	53%	100%	70%
	(83%, 100%)	(23%, 59%)	(36%, 69%)	(73.5%, 100%)	(61%, 79%)
Radiologist 2	95%	43%	53%	93%	69%
	(75%, 100%)	(26%, 63%)	(36%, 70%)	(66%, 99%)	(59%, 79%)
Institution 3					
Radiologist 1	96%	48%	65%	92%	72%
	(80%, 100%)	(28%, 69%)	(48%, 90%)	(64%, 99%)	(61%, 83%)
Radiologist 2	88%	48%	63%	80%	68%
	(69%, 98%)	(28%, 69%)	(45%, 79%)	(52%, 96%)	(56%, 80%)
Institution 4					
Radiologist 1	91%	61%	65%	90%	76%
	(71%, 99%)	(41%, 79%)	(45%, 81%)	(67%, 99%)	(65%, 87%)
Radiologist 2	82%	39%	51%	73%	61%
	(60%, 95%)	(22%, 59%)	(34%, 69%)	(45%, 92%)	(48%, 73%)
Institution 5					
Radiologist 1	93%	30%	61%	78%	62%
	(76%, 99%)	(13%, 53%)	(45%, 76%)	(40%, 97%)	(51%, 72%)
Radiologist 2	96%	26%	86%	86%	61%
	(81%, 100%)	(10%, 48%)	(42%, 100%)	(42%, 100%)	(51%, 71%)
Overall	94%	40%	59%	88%	66%
	(91%, 97%)	(35%, 46%)	(54%, 64%)	(81%, 93%)	(57%, 77%)

1-Pooled results were derived from a random-effects logistic regression model.