APPEDNDIX A: Supplementary tables

Supplementary Table 1. Characteristics of original studies included in meta-analysis

References	Country	Treatment type		Population characteristics		Mean tumor size, cm (SD)	Mean outcome (SD)	
			Sample size	% Male	Mean age (SD)		Hospitalization	Procedure time
Jeschke, 2001[31]	Austria	Laparoscopic partial nephrectomy	51	27.5	59.8 (11.1)	2.0 (0.9)	5.8 (2.0)	132.0 (50.9)
Matin, 2002[32]	U.S.	Open partial nephrectomy	82	61.0	56.2 (11.5)	2.6 (NR)	5.7 (1.5)	216.2 (91.7)
Gill, 2003[33]	U.S.	Open partial nephrectomy	98	67.0	58.8 (11.6)	3.4 (0.8)	5.7 (1.6)	231.2 (49.7)
		Laparoscopic partial nephrectomy	97	58.2	65.1 (11.4)	2.9 (1.5)	2.3 (0.8)	180.1 (45.2)
Simon, 2003[34]	U.S.	Laparoscopic partial nephrectomy	19	78.9	65.5 (11.7)	2.2 (1.3)	2.4 (1.3)	129.5 (44.1)
Beasley, 2004[35]	Canada	Open partial nephrectomy	22	63.6	51.1 (16.4)	2.9 (1.1)	6.4 (1.8)	144.0 (44.0)

		Laparoscopic partial nephrectomy	27	59.3	53.5 (17.7)	2.4 (1.2)	2.9 (1.5)	210.0 (76.0)
Yoshikawa, 2004[11]	Japan	Laparoscopic partial nephrectomy	17	76.5	55.2 (16.7)	2.5 (1.1)	NR	263.3 (83.4)
Desai, 2005[36]	U.S.	Laparoscopic partial nephrectomy	153	58.2	60.6 (13.2)	2.3 (0.7)	2.3 (2.8)	190.1 (51.7)
Fogarty, 2005[37]	U.S.	Laparoscopic partial nephrectomy	20	75.0	65.8 (11.3)	2.6 (0.9)	3.4 (1.1)	163.0 (30.8)
Aron, 2008[22]	U.S.	Laparoscopic partial nephrectomy	12	66.7	61.0 (13.8)	2.9 (0.7)	4.4 (1.1)	256.0 (70.6)
		Robot-assisted partial nephrectomy	12	66.7	64.0 (13.8)	2.4 (0.7)	4.7 (2.4)	242.0 (69.2)
Bensalah, 2008[38]	U.S.	Laparoscopic partial nephrectomy	50	62.0	56.5 (11.7)	2.6 (0.9)	2.9 (1.6)	282.0 (55.0)
Derweesh, 2008[39]	Canada-U.S.	Percutaneous thermal ablation	26	73.0	69.7 (12.3)	3.1 (1.3)	1.8 (0.7)	106.6 (44.3)

Finley, 2008[40]	U.S.	Percutaneous thermal ablation	18	NR	NR	2.7 (0.8)	1.3 (0.8)	150.0 (32.9)
Gong, 2008[41]	U.S.	Open partial nephrectomy	77	54.5	59.7 (13.6)	2.4 (0.9)	5.6 (3.0)	193.0 (62.9)
		Laparoscopic partial nephrectomy	76	46.1	60.1 (12.5)	2.9 (0.8)	2.5 (2.1)	225.1 (63.8)
DeVoe, 2009[42]	U.S.	Open partial nephrectomy	60	56.5	60.0 (9.8)	2.6 (1.1)	6.9 (4.0)	191.0 (55.0)
		Laparoscopic partial nephrectomy	40	67.3	59.2 (12.4)	2.6 (1.2)	4.9 (1.4)	161.0 (43.0)
Ho, 2009[43]	Austria	Robot-assisted partial nephrectomy	20	65.0	58.2 (7.9)	3.5 (0.5)	4.8 (1.3)	NR
Kural, 2009[18]	Turkey	Laparoscopic partial nephrectomy	20	70.0	58.9 (15.4)	3.1 (1.5)	4.3 (1.1)	226.0 (80.3)
		Robot-assisted partial nephrectomy	11	72.7	50.8 (13.2)	3.2 (0.7)	3.9 (0.7)	185.0 (47.1)
Park, 2010[44]	Korea	Open partial nephrectomy	279	74.2	53.1 (13.2)	2.3 (0.9)	NR	184.0 (68.0)

		Laparoscopic partial nephrectomy	273	70.0	54.6 (13.2)	2.1 (0.8)	NR	221.0 (84.0)
Scoll, 2010[45]	U.S.	Robot-assisted partial nephrectomy	100*	68.0	55.0 (12.0)	3.5 (1.4)	3.2 (1.2)	206 (60.8)
Lavery, 2011[23]	U.S.	Laparoscopic partial nephrectomy	18	77.8	53.6 (11.1)	2.3 (1.2)	2.9 (1.1)	179.7 (25.0)
		Robot-assisted partial nephrectomy	20	55.0	55.4 (11.1)	2.5 (0.9)	2.6 (1.1)	189.2 (49.0)
Lee, 2011[46]	Korea	Robot-assisted partial nephrectomy	69	72.5	53.5 (11.8)	2.4 (1.3)	6.2 (2.0)	192.4 (78.1)
Seo, 2011[19]	Korea	Laparoscopic partial nephrectomy	14	57.1	53.9 (11.6)	2.0 (1.2)	5.3 (0.6)	117.5 (32.0)
		Robot-assisted partial nephrectomy	13	76.9	54.2 (12.4)	2.7 (1.2)	6.2 (1.8)	153.2 (22.3)
Goyal, 2012[47]	U.S.	Percutaneous thermal ablation	141	65.2	69.5 (10.9)	2.4 (0.7)	0.7 (2.7)	NR

Guillotreau, 2012[48]	U.S.	Robot-assisted partial nephrectomy	210	58.6	57.8 (11.8)	2.4 (0.8)	3.3 (0.7)	180.0 (44.8)
Lucas, 2012[49]	U.S.	Open partial nephrectomy	54	70.4	58.0 (13.4)	2.3 (0.8)	NR	147.0 (48.7)
		Laparoscopic partial nephrectomy	15	41.2	49.4 (20.3)	2.2 (1.6)	NR	195.0 (73.6)
		Robot-assisted partial nephrectomy	27	70.4	62.1 (12.0)	2.4 (0.5)	NR	190.0 (47.0)
Petros, 2012[50]	U.S.	Robot-assisted partial nephrectomy	362	67.7	60.0 (11.0)	2.3 (0.6)	2.5 (0.7)	182.0 (54.3)
Simone, 2012[51]	Italy	Laparoscopic partial nephrectomy	101	62.4	59.0 (5.6)	2.6 (0.5)	3.3 (6.0)	81.3 (23.0)
Ceccarelli, 2013[52]	Italy	Robot-assisted partial nephrectomy	32	68.7	60.8 (14.3)	3.6 (1.2)	4.4 (1.9)	149.2 (61.9)
Choi, 2013[17]	Korea	Laparoscopic partial nephrectomy	52	63.5	51.1 (11.3)	2.2 (1.1)	8.2 (1.8)	263.8 (73.3)

		Robot-assisted partial nephrectomy	48	70.8	50.9 (11.4)	2.5 (1.0)	7.8 (0.7)	258.6 (73.3)
Kim, 2013[53]	U.S.	Percutaneous thermal ablation	124	NR	72.6 (10.2)	2.7 (1.1)	1.3 (0.7)	186.0 (40.0)
Masson-Lecomte, 2013[54]	France	Open partial nephrectomy	58	69.0	60.8 (11.2)	3.1 (1.2)	3.5 (1.0)	128.4 (50.5)
		Robot-assisted partial nephrectomy	42	52.4	61.7 (10.9)	2.8 (1.4)	1.1 (3.0)	134.8 (35.3)
Schips, 2013[55]	Italy	Laparoscopic partial nephrectomy	21	66.7	58.4 (9.0)	2.0 (0.3)	4.4 (1.9)	111.0 (41.0)
Tanagho, 2013[56]	U.S.	Robot-assisted partial nephrectomy	267	54.5	57.4 (11.9)	2.9 (1.5)	NR	140.6 (41.6)
		Percutaneous thermal ablation	267	61.0	69.3 (11.0)	2.5 (1.0)	NR	164.8 (60.2)
Williams, 2011[24]	U.S.	Laparoscopic partial nephrectomy	59	69.5	54.6 (11.7)	3.1 (2.2)	2.7 (0.8)	221.4 (54.6)

		Robot-assisted partial nephrectomy	27	63.0	55.7 (11.2)	2.5 (1.2)	2.5 (1.1)	233.0 (43.6)
Youn, 2013[57]	Korea	Open partial nephrectomy	14	57.1	53.9 (16.1)	2.4 (0.8)	12.2 (3.3)	148.6 (40.9)
Emara, 2014[58]	U.K.	Robot-assisted partial nephrectomy	47	46.3	60.5 (9.5)	2.6 (1.0)	1.4 (0.1)	159.0 (50.7)
Ficarra, 2014[59]	Italy	Open partial nephrectomy	200	65.5	62.4 (11.8)	2.8 (1.1)	7.0 (1.5)	130.0 (52.3)
Harris, 2015[60]	U.S.	Robot-assisted partial nephrectomy	321	NR	59.3 (11.7)	2.7 (1.3)	NR	171.7 (46.9)
Kim, 2015[61]	Korea	Laparoscopic partial nephrectomy	195	66.7	54.7 (12.7)	2.3 (1.1)	NR	211.0 (13.6)
		Robot-assisted partial nephrectomy	195	63.6	54.4 (13.0)	2.4 ± 1.2	NR	138.0 (44.8)
Okhunov, 2015[62]	U.S.	Percutaneous thermal ablation	236	67.4	68.2 (10.6)	2.4 (1.0)	1.4 (0.6)	112.9 (40.6)

Huang, 2016[20]	China	Laparoscopic partial nephrectomy	45	62.2	50.0 (10.4)	2.9 (0.6)	6.3 (1.1)	124.8 (34.8)
Li, 2016[63]	Taiwan	Laparoscopic partial nephrectomy	31	61.3	53.0 (7.5)	2.9 (1.1)	6.2 (2.4)	188.0 (41.4)
Oh, 2016[64]	Germany	Open partial nephrectomy	385	69.6	54.9 (13.1)	2.3 (0.8)	NR	140.2 (46.8)
		Robot-assisted partial nephrectomy	317	72.6	52.1 (12.2)	2.2 (0.8)	NR	138.8 (72.4)
Pantelidou, 2016[65]	U.K.	Robot-assisted partial nephrectomy	63	NR	54.0 (7.0)	2.9 (0.1)	2.7 (0.8)	NR
		Percutaneous thermal ablation	63	NR	61.0 (21.0)	2.1 (0.2)	1.0 (0.0)	NR
Robert, 2016[66]	Australia	Laparoscopic partial nephrectomy	50	54.0	56.9 (10.5)	2.5 (1.3)	NR	224.0 (37.9)
Han, 2017[67]	Korea	Open partial nephrectomy	354	76.3	55.3 (12.4)	2.8 (1.4)	7.3 (2.1)	187.2 (43.8)

		Laparoscopic partial nephrectomy	89	69.7	53.6 (9.7)	2.6 (1.1)	5.9 (1.7)	172.9 (42.3)
		Robot-assisted partial nephrectomy	147	73.5	52.5 (11.9)	2.4 (0.9)	5.3 (1.4)	162.3 (32.2)
Luciani, 2017[68]	Italy	Open partial nephrectomy	73	69.9	63.0 (13.0)	3.6 (2.3)	5.0 (1.0)	145.0 (59.0)
		Laparoscopic partial nephrectomy	70	60.0	62.0 (11.0)	3.5 (1.4)	3.0 (2.0)	186.0 (49.0)
		Robot-assisted partial nephrectomy	110	60.9	61.0 (12.0)	3.6 (1.5)	3.0 (3.0)	209.0 (53.0)
Reynolds, 2017[69]	U.S.	Robot-assisted partial nephrectomy	1307	56.2	58.1 (11.8)	2.5 (0.7)	1.3 (0.7)	178.4 (44.3)
Simsek, 2017[70]	Turkey	Laparoscopic partial nephrectomy	20	75.0	50.2 (11.3)	NR	4.4 (1.9)	227.5 (56.3)
		Robot-assisted partial nephrectomy	22	54.5	54.8 (9.6)	NR	6.1 (2.4)	176.0 (23.6)

Azevedo, 2018[71]	Brazil	Percutaneous thermal ablation	60	80.0	63.9 (12.54)	1.6 (0.82)	1.6 (1.2)	98.4 (28.3)
Borghesi, 2018[72]	Italy	Open partial nephrectomy	52	57.7	62.7 (16.0)	3.0 (1.5)	5.7 (1.5)	120.0 (30.5)
		Robot-assisted partial nephrectomy	52	51.1	61.3 (16.0)	3.0 (1.5)	3.3 (0.8)	171.0 (61.0)
Lourenco, 2018[73]	Canada	Percutaneous thermal ablation	121	64.5	71.0 (10.8)	2.6 (0.9)	0.4 (0.3)	NR
Motoyama, 2019[27]	Japan	Open partial nephrectomy	37	62.2	59.8 (10.6)	3.3 (1.5)	22.0 (11.8)	227.5 (54.6)
		Robot-assisted partial nephrectomy	37	59.5	62.0 (12.7)	2.5 (0.8)	12.5 (4.7)	186.3 (36.9)
Park, 2019[74]	Korea	Open partial nephrectomy	53	75.5	53.0 (13.2)	2.5 (0.6)	13.0 (3.9)	173.0 (48.6)
Furukawa, 2020[26]	Japan	Robot-assisted partial nephrectomy	804	72.6	63.0 (11.1)	2.6 (1.0)	9.0 (3.0)	241.0 (69.1)
Watanabe, 2021[75]	Japan	Robot-assisted partial nephrectomy	100	64.0	62.6 (13.6)	2.5 (1.1)	NR	183.2 (38.9)

Benamran, 2022[76]	France	Robot-assisted partial nephrectomy	20	55.0	61.3 (5.6)	2.5 (0.7)	NR	63.3 (23.9)
Bianchi, 2022[77]	Italy	Percutaneous thermal ablation	137	65.7	72.0 (10.5)	2.3 (0.8)	1.7 (0.7)	NR
Furukawa, 2022[28]	Japan	Robot-assisted partial nephrectomy	103	74.8	61.0 (11.6)	2.7 (1.0)	NR	234.0 (66.0)
Junker, 2022[78]	Denmark	Percutaneous thermal ablation	101	71.3	69.2 (10.5)	3.1 (0.9)	2.0 (1.5)	NR
Sri, 2023[79]	U.K.	Robot-assisted partial nephrectomy	784	68.0	54.8 (10.0)	3.1 (1.5)	NR	158.0 (48.6)

*Duration of hospitalization is reported for 98 patients. NR: not reported; SD: standard deviation.

Supplementary Table 2. Risk of bias in included studies						
Study	**Domain**				**Total ROB assessment**	**Comment**
	Study participation	**Outcome measurement**	**Study attrition**	**Statistical analysis and reporting**		
Jeschke, 2001[31]	High	Low	Low	Low	High	Inadequate description of the sampling frame, recruitment, inclusion and exclusion criteria
Matin, 2002[32]	Low	High	Low	Low	High	A clear definition of the outcomes of interest is not provided
Gill, 2003[33]	Low	Low	Low	Low	Low	NA
Simon, 2003[34]	High	High	Low	Low	Very high	Inadequate description of the sampling frame, recruitment, inclusion and exclusion criteria; a clear definition of the outcome (s)

						of interest is not provided
Beasley, 2004[35]	Low	Low	Low	Low	Low	NA
Yoshikawa, 2004[11]	High	High	Low	Low	Very high	Inadequate description of the sampling frame, recruitment, inclusion and exclusion criteria; a clear definition of the outcome (s) of interest is not provided
Desai, 2005[36]	Low	Low	Low	Low	Low	NA
Fogarty, 2005[37]	High	Low	Low	Low	High	Inadequate description of the sampling frame
Aron, 2008[22]	Low	Low	Low	Low	Low	NA
Bensalah, 2008[38]	High	Low	Low	Low	High	Inadequate description of inclusion and exclusion criteria
Derweesh, 2008[39]	Low	Low	Low	Low	Low	NA

Finley, 2008[40]	Low	Low	Low	Low	Low	NA
Gong, 2008[41]	Low	Low	Low	Low	Low	NA
Ho, 2009[43]	Low	Low	Low	Low	Not serious	NA
DeVoe, 2009[42]	High	High	Low	Low	Very high	Inadequate description of inclusion and exclusion criteria; a clear definition of the outcomes of interest is not provided
Kural, 2009[18]	High	Low	Low	Low	High	Inadequate description of inclusion and exclusion criteria
Park, 2010[44]	Low	Low	Low	Low	Low	NA
Park, 2010[44]	Low	Low	Low	Low	Low	NA
Lavery, 2011[23]	Low	Low	Low	Low	Low	NA
Lee, 2011[46]	Low	Low	Low	Low	Low	NA
Seo, 2011[19]	High	Low	Low	Low	High	Inadequate description of

						inclusion and exclusion criteria
Goyal, 2012[47]	High	Low	Low	Low	High	Inadequate description of the sampling frame, recruitment, inclusion and exclusion criteria
Guillotreau, 2012[48]	High	Low	Low	Low	High	Inadequate description of the sampling frame, recruitment, inclusion and exclusion criteria
Lucas, 2012[49]	Low	Low	Low	Low	Low	NA
Petros, 2012[50]	Low	Low	Low	Low	Low	NA
Simone, 2012[51]	Low	Low	Low	Low	Low	NA
Ceccarelli, 2013[52]	Low	Low	Low	Low	Low	NA
Choi, 2013[17]	Low	Low	Low	Low	Low	NA
Kim, 2013[53]	High	Low	Low	Low	High	Inadequate description of exclusion criteria

Masson-Lecomte, 2013[54]	High	Low	Low	Low	High	Inadequate description of the sampling frame, recruitment, inclusion and exclusion criteria
Schips, 2013[55]	High	High	Low	Low	Very high	Inadequate description of the sampling frame and exclusion criteria; inadequate method of outcome measurement - may not be validated or reliable
Tanagho, 2013[56]	High	Low	Low	Low	High	Inadequate description of the inclusion and exclusion criteria
Williams, 2011[24]	High	Low	Low	Low	High	Inadequate description of the sampling frame, recruitment,

						inclusion and exclusion criteria
Youn, 2013[57]	High	Low	Low	Low	High	Inadequate description of the sampling frame, Recruitment, inclusion and exclusion criteria, and place of recruitment
Emara, 2014[58]	Low	Low	Low	Low	Low	NA
Ficarra, 2014[59]	High	High	Low	Low	Very high	Sampling frames were two different databases resulting in heterogeneity in surgical techniques; outcomes of interest were not measured in a similar way for all participants
Harris, 2015[60]	Low	High	Low	Low	High	A clear definition of the outcome (s)

						of interest is not provided
Kim, 2015[61]	High	Low	Low	Low	High	Inadequate description of the sampling frame, recruitment, inclusion and exclusion criteria
Okhunov, 2015[62]	High	Low	Low	Low	High	Inadequate description of the sampling frame, recruitment, inclusion and exclusion criteria
Huang, 2016[20]	High	Low	Low	Low	High	Inadequate participation in the study by eligible persons
Li, 2016[63]	High	Low	Low	Low	High	Inadequate participation in the study by eligible persons
Oh, 2016[64]	Low	Low	Low	Low	Low	NA

Pantelidou, 2016[65]	Low	High	Low	Low	High	Outcome of interest was not measured in a similar way for all participants, setting of outcome measurement is not the same for all study participants
Robert, 2016[66]	High	Low	Low	Low	High	Inadequate description of the sampling frame, recruitment, inclusion and exclusion criteria
Han, 2017[67]	Low	Low	Low	Low	Low	NA
Luciani, 2017[68]	Low	Low	Low	Low	Low	NA
Reynolds, 2017[69]	Low	Low	Low	Low	Low	NA
Simsek, 2017[70]	High	Low	Low	Low	High	Inadequate description of the source population

Azevedo, 2018[71]	Low	High	Low	Low	High	A clear definition of the outcome (s) of interest is not provided
Borghesi, 2018[72]	High	Low	Low	Low	High	Inadequate description of the sampling frame, recruitment, inclusion and exclusion criteria
Lourenco, 2018[73]	High	Low	Low	Low	High	Inadequate description of the sampling frame and recruitment
Motoyama, 2019[27]	Low	Low	Low	Low	Low	NA
Park, 2019[74]	Low	Low	Low	Low	Low	NA
Furukawa, 2020[26]	High	Low	Low	Low	High	Inadequate description of the sampling frame, recruitment, inclusion and exclusion criteria

Watanabe, 2021[75]	High	Low	Low	Low	High	Inadequate description of the sampling frame, recruitment, inclusion and exclusion criteria
Benamran, 2022[76]	Low	Low	Low	Low	Low	NA
Bianchi, 2022[77]	Low	Low	Low	Low	Low	NA
Furukawa, 2022[28]	Low	Low	Low	Low	Low	NA
Junker, 2022[78]	Low	Low	Low	Low	Low	NA
Sri, 2023[79]	Low	Low	Low	Low	Low	NA

Supplementary Table 3. GRADE, quality assessment of included studies, and summary of findings

Outcomes	Interventions	Subgroups	No. of studies	No. of patients	Estimates (mean, 95%CI)	ROB	Inconsistency	Imprecision	Indirectness	Other considerations	Certainty
Hospital stays (day)	**OPN**	North America	5	339	5.9 (5.5–6.2)	Not serious	Not serious	Not serious	Serious[a]	None	⊕⊕⊕◯ Moderate
		Europe	4	383	7.1 (5.7–8.4)	Serious[b]	Serious[c]	Serious[d]	Not serious	None	⊕⊕◯◯ Low
		Asia	4	458	13.4 (7.6–19.3)	Not serious	Not serious	Not serious	Not serious	None	⊕⊕⊕⊕ High
	LPN	North America	11	571	3.0 (2.5–3.6)	Not serious	Serious[c]	Not serious	Serious[e]	None	⊕⊕◯◯ Low
		Europe	4	243	5.4 (3.4–7.4)	Not serious	Serious[c]	Serious[d]	Not serious	None	⊕⊕◯◯ Low
		Asia	7	271	5.8 (4.8–6.8)	Serious[f]	Not serious	Not serious	Not serious	None	⊕⊕⊕◯ Moderate
	RAPN	North America	7	2,036	2.8 (2.1–3.5)	Not serious	Serious[c]	Serious[d]	Not serious	None	⊕⊕⊕◯ Moderate
		Europe	7	366	3.7 (2.6–4.9)	Not serious	Serious[c]	Serious[d]	Not serious	None	⊕⊕⊕◯ Moderate

		Asia	8	1,151	7.1 (5.3–8.9)	Not serious	Serious[c]	Not serious	Not serious	None	⊕⊕⊕◯ Moderate
	PTA	North America	6	666	1.2 (0.7–1.6)	Serious[g]	Not serious	Not serious	Not serious	None	⊕⊕⊕◯ Moderate
		Europe	3	301	1.6 (1.0–2.1)	Not serious	Not serious	Not serious	Not serious	None	⊕⊕⊕⊕ High
		Brazil	1	60	1.6 (1.3–1.9)	Serious[h]	Not applicable	Not serious	Not serious	None	⊕⊕⊕◯ Moderate
Procedure time (minute)	**OPN**	North America	6	393	187 (159–216)	Not serious	Not serious	Not serious	Serious [a]	None	⊕⊕⊕◯ Moderate
		Europe	5	768	132 (124–141)	Serious[b]	Not serious	Not serious	Not serious	None	⊕⊕⊕◯ Moderate
		Asia	5	737	184 (161–208)	Not serious	Not serious	Not serious	Not serious	None	⊕⊕⊕⊕ High
	LPN	North America	12	586	199 (175–223)	Not serious	Serious[i]	Not serious	Serious[j]	None	⊕⊕⊕◯ Moderate
		Europe	4	243	127 (84–171)	Serious[k]	Serious [l]	Serious[m]	Not serious	None	⊕⊕◯◯ Low
		Asia	10	756	200 (169–232)	Serious[n]	Serious[i]	Not serious	Not serious	None	⊕⊕⊕◯ Moderate

		Australia	1	50	224 (213–234)	Serious[h]	Not applicable	Not serious	Not serious	None	⊕⊕⊕◯ Moderate
	RAPN	North America	10	2,651	189 (172–206)	Not serious	Not serious	Not serious	Not serious	None	⊕⊕⊕⊕ High
		Europe	7	620	150 (121–178)	Not serious	Serious[l]	Not serious	Not serious	None	⊕⊕⊕◯ Moderate
		Asia	11	1,549	192 (169-214)	Serious[o]	Serious[l]	Not serious	Not serious	None	⊕⊕⊕◯ Moderate
	PTA	North America	5	671	144 (115–174)	Serious[p]	Not serious	Not serious	Not serious	None	⊕⊕⊕◯ Moderate
		Brazil	1	60	98 (91–106)	Serious[h]	Not applicable	Not serious	Not serious	None	⊕⊕⊕◯ Moderate

[a]Five out of five studies are published before 2010, methods of diagnosis and treatments have changed during last decade. [b]Two studies have high risk of bias in the study participation domain and one study have a high risk of bias in both study participation and outcome measurement domains. [c]The point estimates show a range of normal to long mean days of hospitalization. [d]The lower boundary of confidence interval does not show a long hospital stay while the upper boundary shows a prolonged hospitalization. [e]Nine out of 11 studies are published before 2010, methods of diagnosis and treatments have changed during last decade. [f]Five studies have high risk of bias in the study participation domain. [g]Four studies have high risk of bias in the study participation domain. [h]This study has high risk of bias in the outcome measurement domain. [i]There is about two and a half hours of difference between points estimates of mean procedure time in this subgroup of studies that is a considerable difference. [j]Nine out of twelve studies are published before 2010, methods of diagnosis and treatments have changed during last decade. [k]One study has high risk of bias in the study participation domain and one study has a serious risk of bias in both study population and outcome measurement domains. [l]There are about two hours of difference between points estimates of mean procedure time in this subgroup of

studies that is a considerable difference. [m]The lower boundary of confidence interval does not show a long procedure time while the upper boundary shows a prolonged hospitalization. [n]Six studies have high risk of bias in the study participation domain and one study has a serious risk of bias in both study population and outcome measurement domains. [o]Six studies have high risk of bias in the study participation domain. [p]Three studies have high risk of bias in the study participation domain. LPN: laparoscopic partial nephrectomy; OPN: open partial nephrectomy; PTA: percutaneous thermal ablation; RAPN: robot-assisted partial nephrectomy.

Appendix B: ICEMAN tool

1. Completed ICEMAN tool & Completed ICEMAN tool for assessing the credibility of geographic area analysis in studies on the duration of procedure for patients with small renal masses undergoing open partial nephrectomy

Instrument for assessing the Credibility of Effect Modification Analyses (ICEMAN) in meta-analyses of prognostic studies (*Version 1.0)*

Quick instructions

- Synonyms for effect modification include prognosis *modification*, prognostic effect, and interaction.
- The instrument applies to a single proposed prognostic factor at a time; complete one form per each outcome, time-point, effect measure, and prognostic factor.
- Response options on the left indicate definitely or probably reduced, response options on the right probably or definitely increased credibility.
- Completely unclear goes under probably reduced credibility.
- It is helpful to provide a supporting comment or quotation under each question.
- The manual provides more detailed instructions and examples.

Preliminary considerations

Study reference(s): Hospital stays and procedure time after partial nephrectomy or percutaneous thermal ablation – a systematic review and meta-analysis

If available, protocol reference(s): Registration ID on PROSPERO: CRD42022308375

State a single outcome and, if applicable, time-point of interest (e.g., mortality at 1 year follow-up): Length of hospital stays

State the population of interest of interest (e.g., multiple myeloma patients): Small renal masses patients who underwent laparoscopic partial nephrectomy

State a single effect measure of interest (e.g., risk ratio, proportion or mean): Mean

State a single potential prognosis modifier/prognostic factor of interest (e.g., age or comorbidity): Geographic area of conducting studies

Credibility assessment

1: Is the analysis of prognostic factor based on comparison within rather than between studies?

[ ] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Prognostic factor analysis or meta-regression comparing overall effects of each individual study. This is typical for aggregate data meta-analysis.	*Prognostic factor analysis or meta-regression with most information coming from overall effects, but some studies providing within-study subgroup information*	*Most studies providing within-study prognostic factor information; or individual participant data analysis that combines within and between study information*	*All studies providing within-study prognostic factor information or individual participant data; and the analysis separates within from between study information, e.g., meta-analysis of interactions*

Comment: The analysis of prognostic factor is completely based on comparison between studies. It may not be a very serious threat for credibility of subgroup analysis because there is no variation within studies for area of conducting the study. However, we rated the credibility a little bit down.

2: For within-study comparisons, is the prognosis *modification* similar from study to study ? [ ] Not applicable: no or one within-study comparison

[ ] Definitely not similar	[ ] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Prognosis *modification reported for two or more studies and clearly different directions*	Prognosis *modification not reported for individual studies or too imprecise to tell*	Prognosis *modification reported for two or more studies, mostly similar in direction, but considerable differences in magnitude*	Prognosis *modification reported for two or more studies, similar in direction, only some differences in magnitude*

Comment:

3: For between- study comparisons, is the number of studies large? [ ] Not applicable: no between study comparison

[ ] Very small	[ ] Rather small or unclear	[ ] Rather large	[ ] Large
1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 13 studies are included in total and 4 studies are included in the smallest subgroup.

4: Was the direction of prognosis *modification* correctly hypothesized a priori?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of* prognosis *modification*	*Prior protocol available and includes correct specification of direction of* prognosis *modification, e.g., based on a biologic rationale*

Comment: A prior protocol has been published on PROSPERO. The analysis of prognostic factors in different geographic areas is suggested based on experts' opinions. The hypothesis is that there is a significant difference in the length of hospitalization between various geographic areas worldwide. Previous studies conducted in Asian countries have shown longer hospital stays compared to other studies.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent prognosis modification? (consider irrespective of number of prognosis modifiers)

[ ] Chance a very likely explanation	[ ] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The results of test for subgroup differences showed a p-value of 0.0096.

6: Did the authors test only a small number of prognosis modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Explicitly exploratory analysis or large number of prognosis modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 prognosis modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer prognosis modifiers tested*	*Protocol available and 3 or fewer prognosis modifiers tested or number considered in analysis*

Comment: A protocol is available and prognostic factor analysis is tested only for ROB, sample size and area.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: Meta-analysis for all outcomes in all patient's population groups are based on random effects model.

8: If the prognostic factor is a continuous variable, were arbitrary cut points avoided? not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior study*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment: Area is a nominal variable.

---

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [ ] not applicable

[ ] Yes, probably [ ] Yes, probably increase
decrease

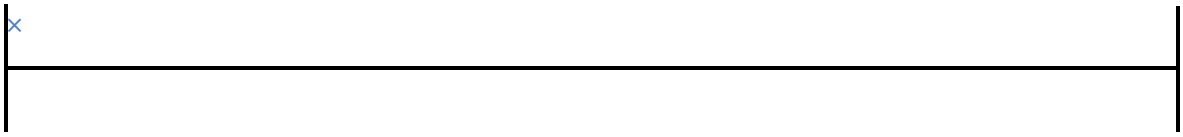
Comment:

10: How would you rate the overall credibility of the proposed prognosis modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Very likely no prognosis modification Use overall effect for each subgroup	Likely no prognosis modification Use overall effect for each subgroup but note remaining uncertainty	Likely prognosis modification Use separate effects for each subgroup but note remaining uncertainty	Very likely prognosis modification Use separate effects for each subgroup

Comment: The credibility of prognostic factor analysis is considered moderate based on the responses.

2. Completed ICEMAN tool for assessing the credibility of geographic area analysis in studies on the length of hospital stays for patients with small renal masses undergoing laparoscopic partial nephrectomy

Instrument for assessing the Credibility of Effect Modification Analyses (ICEMAN) in meta-analyses of prognostic studies (*Version 1.0)*

Quick instructions

- Synonyms for effect modification include prognosis *modification*, prognostic effect, and interaction.
- The instrument applies to a single proposed prognostic factor at a time; complete one form per each outcome, time-point, effect measure, and prognostic factor.
- Response options on the left indicate definitely or probably reduced, response options on the right probably or definitely increased credibility.
- Completely unclear goes under probably reduced credibility.
- It is helpful to provide a supporting comment or quotation under each question.
- The manual provides more detailed instructions and examples.

Preliminary considerations

Study reference(s): Hospital stays and procedure time after partial nephrectomy or percutaneous thermal ablation – a systematic review and meta-analysis

If available, protocol reference(s): Registration ID on PROSPERO: CRD42022308375

State a single outcome and, if applicable, time-point of interest (e.g., mortality at 1 year follow-up): Length of hospital stays

State the population of interest of interest (e.g., multiple myeloma patients): Small renal masses patients who underwent robot assisted partial nephrectomy

State a single effect measure of interest (e.g., risk ratio, proportion or mean): Mean

State a single potential prognosis modifier/prognostic factor of interest (e.g., age or comorbidity): Geographic area of conducting studies

Credibility assessment

1: Is the analysis of prognostic factor based on comparison within rather than between studies?

[ ] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Prognostic factor analysis or meta-regression comparing overall effects of each individual study. This is typical for aggregate data meta-analysis.	*Prognostic factor analysis or meta-regression with most information coming from overall effects, but some studies providing within-study subgroup information*	*Most studies providing within-study prognostic factor information; or individual participant data analysis that combines within and between study information*	*All studies providing within-study prognostic factor information or individual participant data; and the analysis separates within from between study information, e.g., meta-analysis of interactions*

Comment: The analysis of prognostic factor is completely based on comparison between studies. It may not be a very serious threat for credibility of subgroup analysis because there is no variation within studies for area of conducting the study. However, we rated the credibility a little bit down.

2: For within-study comparisons, is the prognosis *modification* similar from study to study ? [ ] Not applicable: no or one within-study comparison

[ ] Definitely not similar	[ ] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Prognosis *modification reported for two or more studies and clearly different directions*	Prognosis *modification not reported for individual studies or too imprecise to tell*	Prognosis *modification reported for two or more studies, mostly similar in direction, but considerable differences in magnitude*	Prognosis *modification reported for two or more studies, similar in direction, only some differences in magnitude*

Comment:

3: For between- study comparisons, is the number of studies large? [ ] Not applicable: no between study comparison

[ ] Very small	[ ] Rather small or unclear	[ ] Rather large	[ ] Large

1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 22 studies are included in total and 7 studies are included in the smallest subgroup.

4: Was the direction of prognosis *modification* correctly hypothesized a priori?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of* prognosis *modification*	*Prior protocol available and includes correct specification of direction of* prognosis *modification, e.g., based on a biologic rationale*

Comment: A prior protocol has been published on PROSPERO. The analysis of prognostic factors in different geographic areas is suggested based on experts' opinions. The hypothesis is that there is a significant difference in the length of hospitalization between various geographic areas worldwide. Previous studies conducted in Asian countries have shown longer hospital stays compared to other studies.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent prognosis modification? (consider irrespective of number of prognosis modifiers)

[ ] Chance a very likely explanation	[ ] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The results of test for subgroup differences showed a p-value of <0.001.

6: Did the authors test only a small number of prognosis modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Explicitly exploratory analysis or large number of prognosis modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 prognosis modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer prognosis modifiers tested*	*Protocol available and 3 or fewer prognosis modifiers tested or number considered in analysis*

Comment: A protocol is available and prognostic factor analysis is tested only for ROB, sample size and area.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: Meta-analysis for all outcomes in all patient's population groups are based on random effects model.

8: If the prognostic factor is a continuous variable, were arbitrary cut points avoided? [ ] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior study*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment: Area is a nominal variable.

---

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [ ] not applicable

[ ] Yes, probably [ ] Yes, probably increase
decrease

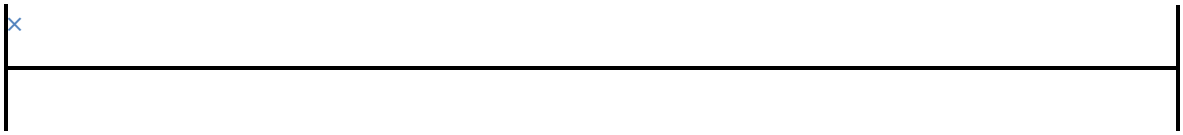
Comment:

10: How would you rate the overall credibility of the proposed prognosis modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Very likely no prognosis modification Use overall effect for each subgroup	Likely no prognosis modification Use overall effect for each subgroup but note remaining uncertainty	Likely prognosis modification Use separate effects for each subgroup but note remaining uncertainty	Very likely prognosis modification Use separate effects for each subgroup

Comment: The credibility of prognostic factor analysis is considered high based on the responses.

Kandi et al. Length of hospital stay and procedure time after partial nephrectomy or percutaneous thermal ablation: A systematic review and meta-analysis

3. Completed ICEMAN tool for assessing the credibility of geographic area analysis in studies on the length of hospital stays for patients with small renal masses undergoing robot-assisted partial nephrectomy.

Instrument for assessing the Credibility of Effect Modification Analyses (ICEMAN) in meta-analyses of prognostic studies (*Version 1.0)*

Quick instructions

- Synonyms for effect modification include prognosis *modification*, prognostic effect, and interaction.
- The instrument applies to a single proposed prognostic factor at a time; complete one form per each outcome, time-point, effect measure, and prognostic factor.
- Response options on the left indicate definitely or probably reduced, response options on the right probably or definitely increased credibility.
- Completely unclear goes under probably reduced credibility.
- It is helpful to provide a supporting comment or quotation under each question.
- The manual provides more detailed instructions and examples.

Preliminary considerations

Study reference(s): Hospital stays and procedure time after partial nephrectomy or percutaneous thermal ablation – a systematic review and meta-analysis

If available, protocol reference(s): Registration ID on PROSPERO: CRD42022308375

State a single outcome and, if applicable, time-point of interest (e.g., mortality at 1 year follow-up): Length of hospital stays

State the population of interest of interest (e.g., multiple myeloma patients): Small renal masses patients who underwent robot assisted partial nephrectomy

State a single effect measure of interest (e.g., risk ratio, proportion or mean): Mean

State a single potential prognosis modifier/prognostic factor of interest (e.g., age or comorbidity): Geographic area of conducting studies

Credibility assessment

1: Is the analysis of prognostic factor based on comparison within rather than between studies?

[ ] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Prognostic factor analysis or meta-regression comparing overall effects of each individual study. This is typical for aggregate data meta-analysis.	*Prognostic factor analysis or meta-regression with most information coming from overall effects, but some studies providing within-study subgroup information*	*Most studies providing within-study prognostic factor information; or individual participant data analysis that combines within and between study information*	*All studies providing within-study prognostic factor information or individual participant data; and the analysis separates within from between study information, e.g., meta-analysis of interactions*

Comment: The analysis of prognostic factor is completely based on comparison between studies. It may not be a very serious threat for credibility of subgroup analysis because there is no variation within studies for area of conducting the study. However, we rated the credibility a little bit down.

2: For within-study comparisons, is the prognosis *modification* similar from study to study ? [ ] Not applicable: no or one within-study comparison

[ ] Definitely not similar	[ ] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Prognosis *modification reported for two or more studies and clearly different directions*	Prognosis *modification not reported for individual studies or too imprecise to tell*	Prognosis *modification reported for two or more studies, mostly similar in direction, but considerable differences in magnitude*	Prognosis *modification reported for two or more studies, similar in direction, only some differences in magnitude*

Comment:

3: For between- study comparisons, is the number of studies large? [ ] Not applicable: no between study comparison

[ ] Very small	[ ] Rather small or unclear	[ ] Rather large	[ ] Large

1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 10 studies are included in total and 3 studies are included in the smallest subgroup.

4: Was the direction of prognosis *modification* correctly hypothesized a priori?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of* prognosis *modification*	*Prior protocol available and includes correct specification of direction of* prognosis *modification, e.g., based on a biologic rationale*

Comment: A prior protocol has been published on PROSPERO. The analysis of prognostic factors in different geographic areas is suggested based on experts' opinions. The hypothesis is that there is a significant difference in the length of hospitalization between various geographic areas worldwide. Previous studies conducted in Asian countries have shown longer hospital stays compared to other studies.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent prognosis modification? (consider irrespective of number of prognosis modifiers)

[ ] Chance a very likely explanation	[ ] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The results of test for subgroup differences showed a p-value of 0.27.

6: Did the authors test only a small number of prognosis modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Explicitly exploratory analysis or large number of prognosis modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 prognosis modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer prognosis modifiers tested*	*Protocol available and 3 or fewer prognosis modifiers tested or number considered in analysis*

Comment: A protocol is available and prognostic factor analysis is tested only for ROB, sample size and area.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: Meta-analysis for all outcomes in all patient's population groups are based on random effects model.

8: If the prognostic factor is a continuous variable, were arbitrary cut points avoided? [ ] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior study*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment: Area is a nominal variable.

---

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [ ] not applicable

[ ] Yes, probably [ ] Yes, probably increase
decrease

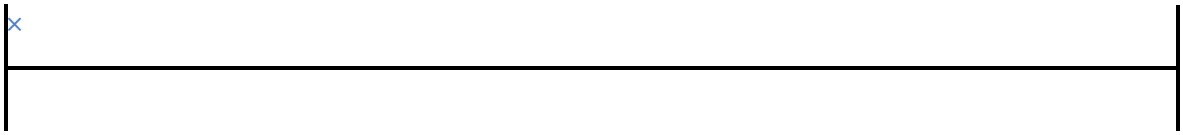
Comment:

10: How would you rate the overall credibility of the proposed prognosis modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Very likely no prognosis modification Use overall effect for each subgroup	Likely no prognosis modification Use overall effect for each subgroup but note remaining uncertainty	Likely prognosis modification Use separate effects for each subgroup but note remaining uncertainty	Very likely prognosis modification Use separate effects for each subgroup

Comment: The credibility of prognostic factor analysis is considered high based on the responses.

Kandi et al. Length of hospital stay and procedure time after partial nephrectomy or percutaneous thermal ablation: A systematic review and meta-analysis

4. Completed ICEMAN tool for assessing the credibility of geographic area analysis in studies on the length of hospital stays for patients with small renal masses undergoing percutaneous thermal ablation.

Instrument for assessing the Credibility of Effect Modification Analyses (ICEMAN) in meta-analyses of prognostic studies (*Version 1.0)*

Quick instructions

- Synonyms for effect modification include prognosis *modification*, prognostic effect, and interaction.
- The instrument applies to a single proposed prognostic factor at a time; complete one form per each outcome, time-point, effect measure, and prognostic factor.
- Response options on the left indicate definitely or probably reduced, response options on the right probably or definitely increased credibility.
- Completely unclear goes under probably reduced credibility.
- It is helpful to provide a supporting comment or quotation under each question.
- The manual provides more detailed instructions and examples.

Preliminary considerations

Study reference(s): Hospital stays and procedure time after partial nephrectomy or percutaneous thermal ablation – a systematic review and meta-analysis

If available, protocol reference(s): Registration ID on PROSPERO: CRD42022308375

State a single outcome and, if applicable, time-point of interest (e.g., mortality at 1 year follow-up): Duration of procedure

State the population of interest of interest (e.g., multiple myeloma patients): Small renal masses patients who underwent percutaneous thermal ablation

State a single effect measure of interest (e.g., risk ratio, proportion or mean): Mean

State a single potential prognosis modifier/prognostic factor of interest (e.g., age or comorbidity): Geographic area of conducting studies

Credibility assessment

1: Is the analysis of prognostic factor based on comparison within rather than between studies?

[ ] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Prognostic factor analysis or meta-regression comparing overall effects of each individual study. This is typical for aggregate data meta-analysis.	*Prognostic factor analysis or meta-regression with most information coming from overall effects, but some studies providing within-study subgroup information*	*Most studies providing within-study prognostic factor information; or individual participant data analysis that combines within and between study information*	*All studies providing within-study prognostic factor information or individual participant data; and the analysis separates within from between study information, e.g., meta-analysis of interactions*

Comment: The analysis of prognostic factor is completely based on comparison between studies. It may not be a very serious threat for credibility of subgroup analysis because there is no variation within studies for area of conducting the study. However, we rated the credibility a little bit down.

2: For within-study comparisons, is the prognosis *modification* similar from study to study ? [ ] Not applicable: no or one within-study comparison

[ ] Definitely not similar	[ ] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Prognosis *modification reported for two or more studies and clearly different directions*	Prognosis *modification not reported for individual studies or too imprecise to tell*	Prognosis *modification reported for two or more studies, mostly similar in direction, but considerable differences in magnitude*	Prognosis *modification reported for two or more studies, similar in direction, only some differences in magnitude*

Comment:

3: For between- study comparisons, is the number of studies large? [ ] Not applicable: no between study comparison

[ ] Very small	[ ] Rather small or unclear	[ ] Rather large	[ ] Large

1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 10 studies are included in total and 3 studies are included in the smallest subgroup.

4: Was the direction of prognosis *modification* correctly hypothesized a priori?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of* prognosis *modification*	*Prior protocol available and includes correct specification of direction of* prognosis *modification, e.g., based on a biologic rationale*

Comment: A prior protocol has been published on PROSPERO. The analysis of prognostic factors in different geographic areas is suggested based on experts' opinions. The hypothesis is that there is a significant difference in the length of hospitalization between various geographic areas worldwide. Previous studies conducted in Asian countries have shown longer hospital stays compared to other studies.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent prognosis modification? (consider irrespective of number of prognosis modifiers)

[ ] Chance a very likely explanation	[ ] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The results of test for subgroup differences showed a p-value of 0.27.

6: Did the authors test only a small number of prognosis modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Explicitly exploratory analysis or large number of prognosis modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 prognosis modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer prognosis modifiers tested*	*Protocol available and 3 or fewer prognosis modifiers tested or number considered in analysis*

Comment: A protocol is available and prognostic factor analysis is tested only for ROB, sample size and area.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: Meta-analysis for all outcomes in all patient's population groups are based on random effects model.

8: If the prognostic factor is a continuous variable, were arbitrary cut points avoided? [ ] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior study*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment: Area is a nominal variable.

---

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [ ] not applicable

[ ] Yes, probably [ ] Yes, probably increase
decrease

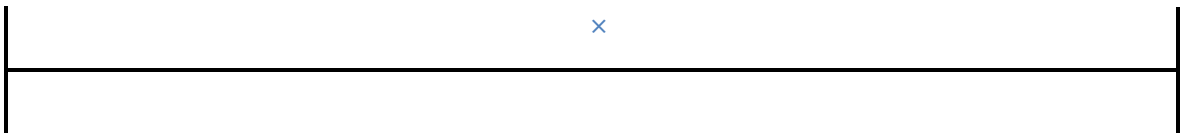
Comment:

10: How would you rate the overall credibility of the proposed prognosis modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Very likely no prognosis modification Use overall effect for each subgroup	Likely no prognosis modification Use overall effect for each subgroup but note remaining uncertainty	Likely prognosis modification Use separate effects for each subgroup but note remaining uncertainty	Very likely prognosis modification Use separate effects for each subgroup

Comment: The credibility of prognostic factor analysis is considered low based on the responses.

5. Completed ICEMAN tool for assessing the credibility of geographic area analysis in studies on the duration of procedure for patients with small renal masses undergoing laparoscopic partial nephrectomy

Instrument for assessing the Credibility of Effect Modification Analyses (ICEMAN) in meta-analyses of prognostic studies (*Version 1.0)*

Quick instructions

- Synonyms for effect modification include prognosis *modification*, prognostic effect, and interaction.
- The instrument applies to a single proposed prognostic factor at a time; complete one form per each outcome, time-point, effect measure, and prognostic factor.
- Response options on the left indicate definitely or probably reduced, response options on the right probably or definitely increased credibility.
- Completely unclear goes under probably reduced credibility.
- It is helpful to provide a supporting comment or quotation under each question.
- The manual provides more detailed instructions and examples.

Preliminary considerations

Study reference(s): Hospital stays and procedure time after partial nephrectomy or percutaneous thermal ablation – a systematic review and meta-analysis

If available, protocol reference(s): Registration ID on PROSPERO: CRD42022308375

State a single outcome and, if applicable, time-point of interest (e.g., mortality at 1 year follow-up): Duration of procedure

State the population of interest of interest (e.g., multiple myeloma patients): Small renal masses patients who underwent open partial nephrectomy

State a single effect measure of interest (e.g., risk ratio, proportion or mean): Mean

State a single potential prognosis modifier/prognostic factor of interest (e.g., age or comorbidity): Geographic area of conducting studies

Credibility assessment

1: Is the analysis of prognostic factor based on comparison within rather than between studies?

[ ] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Prognostic factor analysis or meta-regression comparing overall effects of each individual study. This is typical for aggregate data meta-analysis.	*Prognostic factor analysis or meta-regression with most information coming from overall effects, but some studies providing within-study subgroup information*	*Most studies providing within-study prognostic factor information; or individual participant data analysis that combines within and between study information*	*All studies providing within-study prognostic factor information or individual participant data; and the analysis separates within from between study information, e.g., meta-analysis of interactions*

Comment: The analysis of prognostic factor is completely based on comparison between studies. It may not be a very serious threat for credibility of subgroup analysis because there is no variation within studies for area of conducting the study. However, we rated the credibility a little bit down.

2: For within-study comparisons, is the prognosis *modification* similar from study to study ? [ ] Not applicable: no or one within-study comparison

[ ] Definitely not similar	[ ] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Prognosis *modification reported for two or more studies and clearly different directions*	Prognosis *modification not reported for individual studies or too imprecise to tell*	Prognosis *modification reported for two or more studies, mostly similar in direction, but considerable differences in magnitude*	Prognosis *modification reported for two or more studies, similar in direction, only some differences in magnitude*

Comment:

3: For between- study comparisons, is the number of studies large? [ ] Not applicable: no between study comparison

[ ] Very small	[ ] Rather small or unclear	[ ] Rather large	[ ] Large

1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 16 studies are included in total and 5 studies are included in the smallest subgroup.

4: Was the direction of prognosis *modification* correctly hypothesized a priori?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of* prognosis *modification*	*Prior protocol available and includes correct specification of direction of* prognosis *modification, e.g., based on a biologic rationale*

Comment: A prior protocol has been published on PROSPERO. The analysis of prognostic factors in different geographic areas is suggested based on experts' opinions. The hypothesis is that there is a significant difference in the duration of procedure between various geographic areas worldwide. Previous studies conducted in Asian countries have shown longer procedures compared to other studies and Europe has shorter procedures.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent prognosis modification? (consider irrespective of number of prognosis modifiers)

[ ] Chance a very likely explanation	[ ] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The results of test for subgroup differences showed a p-value of < 0.001.

6: Did the authors test only a small number of prognosis modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Explicitly exploratory analysis or large number of prognosis modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 prognosis modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer prognosis modifiers tested*	*Protocol available and 3 or fewer prognosis modifiers tested or number considered in analysis*

Comment: A protocol is available and prognostic factor analysis is tested only for ROB, sample size and area.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: Meta-analysis for all outcomes in all patient's population groups are based on random effects model.

8: If the prognostic factor is a continuous variable, were arbitrary cut points avoided? [ ] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior study*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment: Area is a nominal variable.

---

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [ ] not applicable

[ ] Yes, probably [ ] Yes, probably increase
decrease

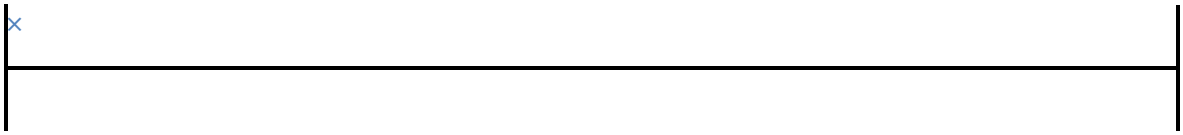
Comment:

10: How would you rate the overall credibility of the proposed prognosis modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Very likely no prognosis modification Use overall effect for each subgroup	Likely no prognosis modification Use overall effect for each subgroup but note remaining uncertainty	Likely prognosis modification Use separate effects for each subgroup but note remaining uncertainty	Very likely prognosis modification Use separate effects for each subgroup

Comment: The credibility of prognostic factor analysis is considered high based on the responses.

Kandi et al. Length of hospital stay and procedure time after partial nephrectomy or percutaneous thermal ablation: A systematic review and meta-analysis

6. Completed ICEMAN tool for assessing the credibility of geographic area analysis in studies on the duration of procedure for patients with small renal masses undergoing open partial nephrectomy

Instrument for assessing the Credibility of Effect Modification Analyses (ICEMAN) in meta-analyses of prognostic studies (*Version 1.0)*

Quick instructions

- Synonyms for effect modification include prognosis *modification*, prognostic effect, and interaction.
- The instrument applies to a single proposed prognostic factor at a time; complete one form per each outcome, time-point, effect measure, and prognostic factor.
- Response options on the left indicate definitely or probably reduced, response options on the right probably or definitely increased credibility.
- Completely unclear goes under probably reduced credibility.
- It is helpful to provide a supporting comment or quotation under each question.
- The manual provides more detailed instructions and examples.

Preliminary considerations

Study reference(s): Hospital stays and procedure time after partial nephrectomy or percutaneous thermal ablation – a systematic review and meta-analysis

If available, protocol reference(s): Registration ID on PROSPERO: CRD42022308375

State a single outcome and, if applicable, time-point of interest (e.g., mortality at 1 year follow-up): Duration of procedure

State the population of interest of interest (e.g., multiple myeloma patients): Small renal masses patients who underwent robot assisted partial nephrectomy

State a single effect measure of interest (e.g., risk ratio, proportion or mean): Mean

State a single potential prognosis modifier/prognostic factor of interest (e.g., age or comorbidity): Geographic area of conducting studies

Credibility assessment

1: Is the analysis of prognostic factor based on comparison within rather than between studies?

[ ] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Prognostic factor analysis or meta-regression comparing overall effects of each individual study. This is typical for aggregate data meta-analysis.	*Prognostic factor analysis or meta-regression with most information coming from overall effects, but some studies providing within-study subgroup information*	*Most studies providing within-study prognostic factor information; or individual participant data analysis that combines within and between study information*	*All studies providing within-study prognostic factor information or individual participant data; and the analysis separates within from between study information, e.g., meta-analysis of interactions*

Comment: The analysis of prognostic factor is completely based on comparison between studies. It may not be a very serious threat for credibility of subgroup analysis because there is no variation within studies for area of conducting the study. However, we rated the credibility a little bit down.

2: For within-study comparisons, is the prognosis *modification* similar from study to study ? [ ] Not applicable: no or one within-study comparison

[ ] Definitely not similar	[ ] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Prognosis *modification reported for two or more studies and clearly different directions*	Prognosis *modification not reported for individual studies or too imprecise to tell*	Prognosis *modification reported for two or more studies, mostly similar in direction, but considerable differences in magnitude*	Prognosis *modification reported for two or more studies, similar in direction, only some differences in magnitude*

Comment:

3: For between- study comparisons, is the number of studies large? [ ] Not applicable: no between study comparison

[ ] Very small	[ ] Rather small or unclear	[ ] Rather large	[ ] Large

1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 27 studies are included in total and 5 studies are included in the smallest subgroup.

4: Was the direction of prognosis *modification* correctly hypothesized a priori?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of* prognosis *modification*	*Prior protocol available and includes correct specification of direction of* prognosis *modification, e.g., based on a biologic rationale*

Comment: A prior protocol has been published on PROSPERO. The analysis of prognostic factors in different geographic areas is suggested based on experts' opinions. The hypothesis is that there is a significant difference in the duration of procedure between various geographic areas worldwide. Previous studies conducted in Asian countries have shown longer procedures compared to other studies and Europe has shorter procedures.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent prognosis modification? (consider irrespective of number of prognosis modifiers)

[ ] Chance a very likely explanation	[ ] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The results of test for subgroup differences showed a p-value of < 0.001.

6: Did the authors test only a small number of prognosis modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Explicitly exploratory analysis or large number of prognosis modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 prognosis modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer prognosis modifiers tested*	*Protocol available and 3 or fewer prognosis modifiers tested or number considered in analysis*

Comment: A protocol is available and prognostic factor analysis is tested only for ROB, sample size and area.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: Meta-analysis for all outcomes in all patient's population groups are based on random effects model.

8: If the prognostic factor is a continuous variable, were arbitrary cut points avoided? [ ] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior study*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment: Area is a nominal variable.

---

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [ ] not applicable

[ ] Yes, probably [ ] Yes, probably increase
decrease

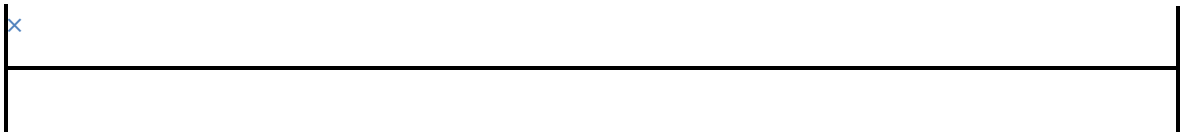
Comment:

10: How would you rate the overall credibility of the proposed prognosis modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Very likely no prognosis modification Use overall effect for each subgroup	Likely no prognosis modification Use overall effect for each subgroup but note remaining uncertainty	Likely prognosis modification Use separate effects for each subgroup but note remaining uncertainty	Very likely prognosis modification Use separate effects for each subgroup

Comment: The credibility of prognostic factor analysis is considered moderate based on the responses.

7. Completed ICEMAN tool for assessing the credibility of geographic area analysis in studies on the duration of procedure for patients with small renal masses undergoing robot-assisted partial nephrectomy

Instrument for assessing the Credibility of Effect Modification Analyses (ICEMAN) in meta-analyses of prognostic studies (*Version 1.0)*

Quick instructions

- Synonyms for effect modification include prognosis *modification*, prognostic effect, and interaction.
- The instrument applies to a single proposed prognostic factor at a time; complete one form per each outcome, time-point, effect measure, and prognostic factor.
- Response options on the left indicate definitely or probably reduced, response options on the right probably or definitely increased credibility.
- Completely unclear goes under probably reduced credibility.
- It is helpful to provide a supporting comment or quotation under each question.
- The manual provides more detailed instructions and examples.

Preliminary considerations

Study reference(s): Hospital stays and procedure time after partial nephrectomy or percutaneous thermal ablation – a systematic review and meta-analysis

If available, protocol reference(s): Registration ID on PROSPERO: CRD42022308375

State a single outcome and, if applicable, time-point of interest (e.g., mortality at 1 year follow-up): Duration of procedure

State the population of interest of interest (e.g., multiple myeloma patients): Small renal masses patients who underwent robot assisted partial nephrectomy

State a single effect measure of interest (e.g., risk ratio, proportion or mean): Mean

State a single potential prognosis modifier/prognostic factor of interest (e.g., age or comorbidity): Geographic area of conducting studies

Credibility assessment

1: Is the analysis of prognostic factor based on comparison within rather than between studies?

[ ] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Prognostic factor analysis or meta-regression comparing overall effects of each individual study. This is typical for aggregate data meta-analysis.	*Prognostic factor analysis or meta-regression with most information coming from overall effects, but some studies providing within-study subgroup information*	*Most studies providing within-study prognostic factor information; or individual participant data analysis that combines within and between study information*	*All studies providing within-study prognostic factor information or individual participant data; and the analysis separates within from between study information, e.g., meta-analysis of interactions*

Comment: The analysis of prognostic factor is completely based on comparison between studies. It may not be a very serious threat for credibility of subgroup analysis because there is no variation within studies for area of conducting the study. However, we rated the credibility a little bit down.

2: For within-study comparisons, is the prognosis *modification* similar from study to study ? [ ] Not applicable: no or one within-study comparison

[ ] Definitely not similar	[ ] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Prognosis *modification reported for two or more studies and clearly different directions*	Prognosis *modification not reported for individual studies or too imprecise to tell*	Prognosis *modification reported for two or more studies, mostly similar in direction, but considerable differences in magnitude*	Prognosis *modification reported for two or more studies, similar in direction, only some differences in magnitude*

Comment:

3: For between- study comparisons, is the number of studies large? [ ] Not applicable: no between study comparison

[ ] Very small	[ ] Rather small or unclear	[ ] Rather large	[ ] Large

1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 27 studies are included in total and 8 studies are included in the smallest subgroup.

4: Was the direction of prognosis *modification* correctly hypothesized a priori?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of* prognosis *modification*	*Prior protocol available and includes correct specification of direction of* prognosis *modification, e.g., based on a biologic rationale*

Comment: A prior protocol has been published on PROSPERO. The analysis of prognostic factors in different geographic areas is suggested based on experts' opinions. The hypothesis is that there is a significant difference in the duration of procedure between various geographic areas worldwide. Previous studies conducted in Asian countries have shown longer procedures compared to other studies and Europe has shorter procedures.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent prognosis modification? (consider irrespective of number of prognosis modifiers)

[ ] Chance a very likely explanation	[ ] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The results of test for subgroup differences showed a p-value of < 0.001.

6: Did the authors test only a small number of prognosis modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Explicitly exploratory analysis or large number of prognosis modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 prognosis modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer prognosis modifiers tested*	*Protocol available and 3 or fewer prognosis modifiers tested or number considered in analysis*

Comment: A protocol is available and prognostic factor analysis is tested only for ROB, sample size and area.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: Meta-analysis for all outcomes in all patient's population groups are based on random effects model.

8: If the prognostic factor is a continuous variable, were arbitrary cut points avoided? [ ] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior study*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment: Area is a nominal variable.

---

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [ ] not applicable

[ ] Yes, probably [ ] Yes, probably increase
decrease

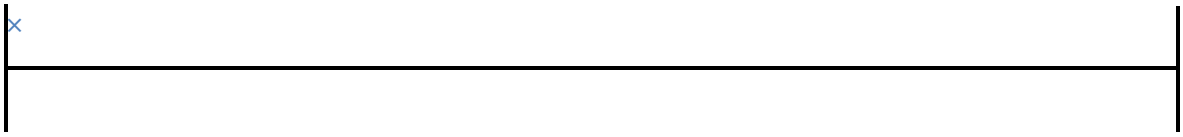
Comment:

10: How would you rate the overall credibility of the proposed prognosis modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)



Very low credibility	Low credibility	Moderate credibility	High credibility
Very likely no prognosis modification Use overall effect for each subgroup	Likely no prognosis modification Use overall effect for each subgroup but note remaining uncertainty	Likely prognosis modification Use separate effects for each subgroup but note remaining uncertainty	Very likely prognosis modification Use separate effects for each subgroup

Comment: The credibility of prognostic factor analysis is considered moderate based on the responses.

8. Completed ICEMAN tool for assessing the credibility of geographic area analysis in studies on the duration of procedure for patients with small renal masses undergoing percutaneous thermal ablation

Instrument for assessing the Credibility of Effect Modification Analyses (ICEMAN) in meta-analyses of prognostic studies (*Version 1.0)*

Quick instructions

- Synonyms for effect modification include prognosis *modification*, prognostic effect, and interaction.
- The instrument applies to a single proposed prognostic factor at a time; complete one form per each outcome, time-point, effect measure, and prognostic factor.
- Response options on the left indicate definitely or probably reduced, response options on the right probably or definitely increased credibility.
- Completely unclear goes under probably reduced credibility.
- It is helpful to provide a supporting comment or quotation under each question.
- The manual provides more detailed instructions and examples.

Preliminary considerations

Study reference(s): Hospital stays and procedure time after partial nephrectomy or percutaneous thermal ablation – a systematic review and meta-analysis

If available, protocol reference(s): Registration ID on PROSPERO: CRD42022308375

State a single outcome and, if applicable, time-point of interest (e.g., mortality at 1 year follow-up): Duration of procedure

State the population of interest of interest (e.g., multiple myeloma patients): Small renal masses patients who underwent percutaneous thermal ablation

State a single effect measure of interest (e.g., risk ratio, proportion or mean): Mean

State a single potential prognosis modifier/prognostic factor of interest (e.g., age or comorbidity): Geographic area of conducting studies

Credibility assessment

1: Is the analysis of prognostic factor based on comparison within rather than between studies?

[ ] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Prognostic factor analysis or meta-regression comparing overall effects of each individual study. This is typical for aggregate data meta-analysis.	*Prognostic factor analysis or meta-regression with most information coming from overall effects, but some studies providing within-study subgroup information*	*Most studies providing within-study prognostic factor information; or individual participant data analysis that combines within and between study information*	*All studies providing within-study prognostic factor information or individual participant data; and the analysis separates within from between study information, e.g., meta-analysis of interactions*

Comment: The analysis of prognostic factor is completely based on comparison between studies. It may not be a very serious threat for credibility of subgroup analysis because there is no variation within studies for area of conducting the study. However, we rated the credibility a little bit down.

2: For within-study comparisons, is the prognosis *modification* similar from study to study ? [ ] Not applicable: no or one within-study comparison

[ ] Definitely not similar	[ ] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Prognosis *modification reported for two or more studies and clearly different directions*	Prognosis *modification not reported for individual studies or too imprecise to tell*	Prognosis *modification reported for two or more studies, mostly similar in direction, but considerable differences in magnitude*	Prognosis *modification reported for two or more studies, similar in direction, only some differences in magnitude*

Comment:

3: For between- study comparisons, is the number of studies large? [ ] Not applicable: no between study comparison

[ ] Very small	[ ] Rather small or unclear	[ ] Rather large	[ ] Large

1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*

Comment: 7 studies are included in total and 1 studies are included in the smallest subgroup.

4: Was the direction of prognosis *modification* correctly hypothesized a priori?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of* prognosis *modification*	*Prior protocol available and includes correct specification of direction of* prognosis *modification, e.g., based on a biologic rationale*

Comment: A prior protocol has been published on PROSPERO. The analysis of prognostic factors in different geographic areas is suggested based on experts' opinions. The hypothesis is that there is a significant difference in the duration of procedure between various geographic areas worldwide. Previous studies conducted in Asian countries have shown longer procedures compared to other studies and Europe has shorter procedures.

5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent prognosis modification? (consider irrespective of number of prognosis modifiers)

[ ] Chance a very likely explanation	[ ] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*

Comment: The results of test for subgroup differences showed a p-value of <0.001.

6: Did the authors test only a small number of prognosis modifiers or consider the number in their statistical analysis?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Explicitly exploratory analysis or large number of prognosis modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 prognosis modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer prognosis modifiers tested*	*Protocol available and 3 or fewer prognosis modifiers tested or number considered in analysis*

Comment: A protocol is available and prognostic factor analysis is tested only for ROB, sample size and area.

7: Did the authors use a random effects model?

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*

Comment: Meta-analysis for all outcomes in all patient's population groups are based on random effects model.

8: If the prognostic factor is a continuous variable, were arbitrary cut points avoided? [ ] not applicable: not continuous

[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior study*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*

Comment: Area is a nominal variable.

---

9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [ ] not applicable

[ ] Yes, probably [ ] Yes, probably increase
decrease

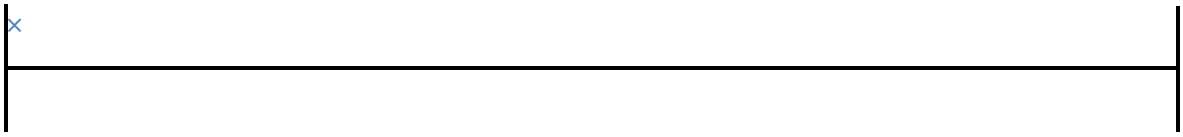
Comment:

10: How would you rate the overall credibility of the proposed prognosis modification?

The overall rating should be driven by the items that decrease credibility. The following provides a sensible strategy:

- All responses definitely or probably decrease credibility or unclear → very low
- Two or more responses definitely decrease credibility → maximum usually low even if all other responses satisfy credibility criteria
- One response definitely decreases credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- Two responses probably decrease credibility → maximum usually moderate even if all other responses satisfy credibility criteria
- No response options definitely or probably decrease credibility → high very likely

Place a mark on the continuous line (or type "x" in editable version)

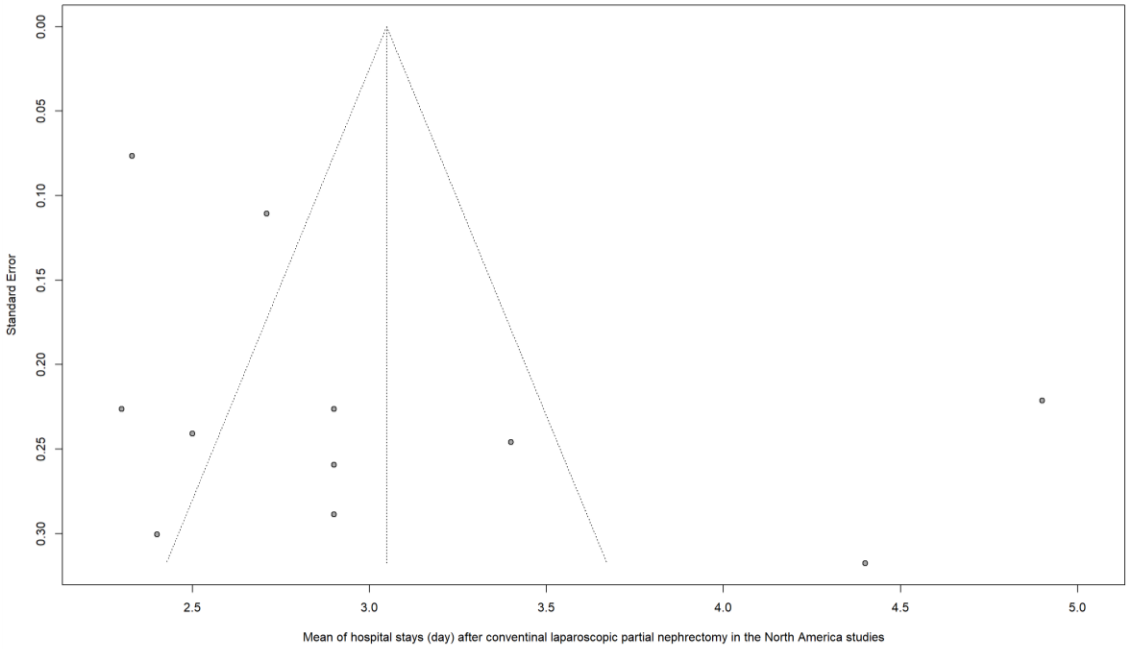


Very low credibility	Low credibility	Moderate credibility	High credibility
Very likely no prognosis modification Use overall effect for each subgroup	Likely no prognosis modification Use overall effect for each subgroup but note remaining uncertainty	Likely prognosis modification Use separate effects for each subgroup but note remaining uncertainty	Very likely prognosis modification Use separate effects for each subgroup

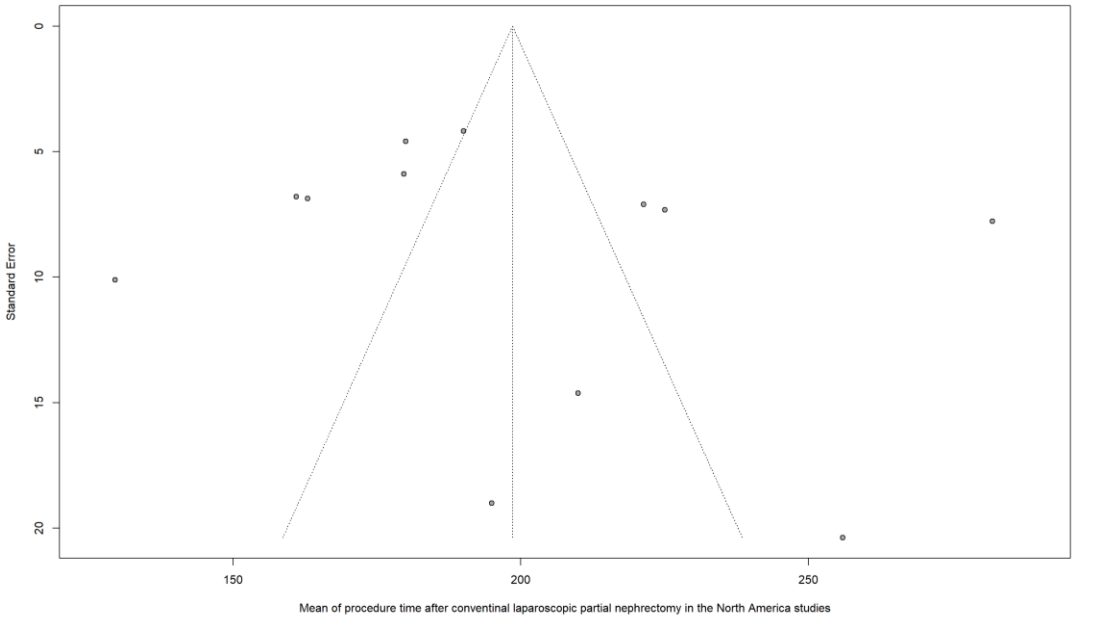
Comment: The credibility of prognostic factor analysis is considered moderate based on the responses.

Appendix C: Supplementary Figures (Publication bias evaluation)

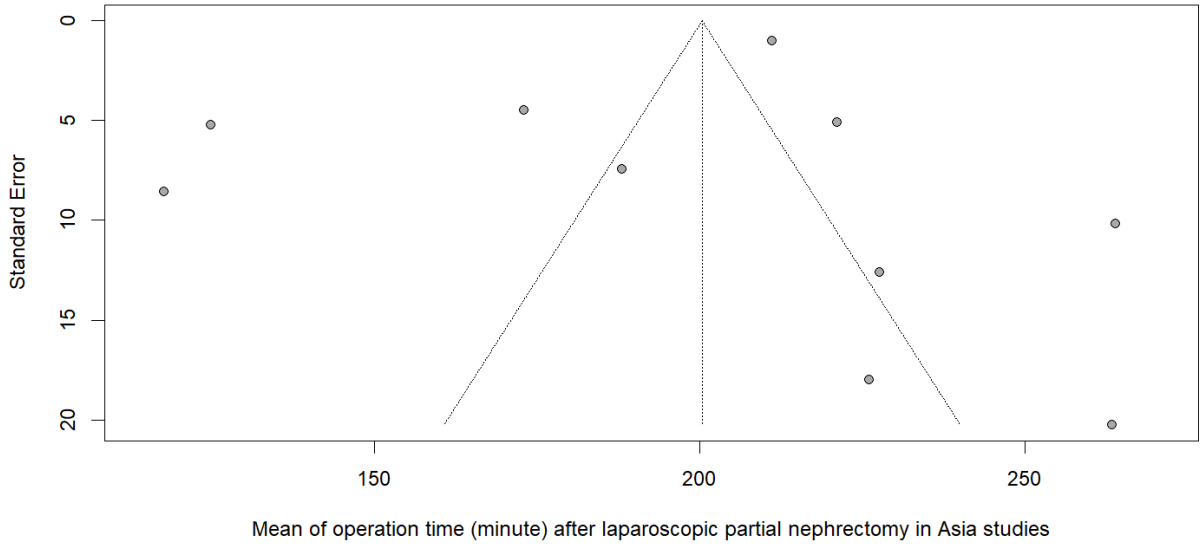
Supplementary Figure 1. Funnel plot of studies on duration of hospital stays after conventional laparoscopic partial nephrectomy in the North America subgroup (Begg test result: Z-value: 0.7, P.V: 0.4835).



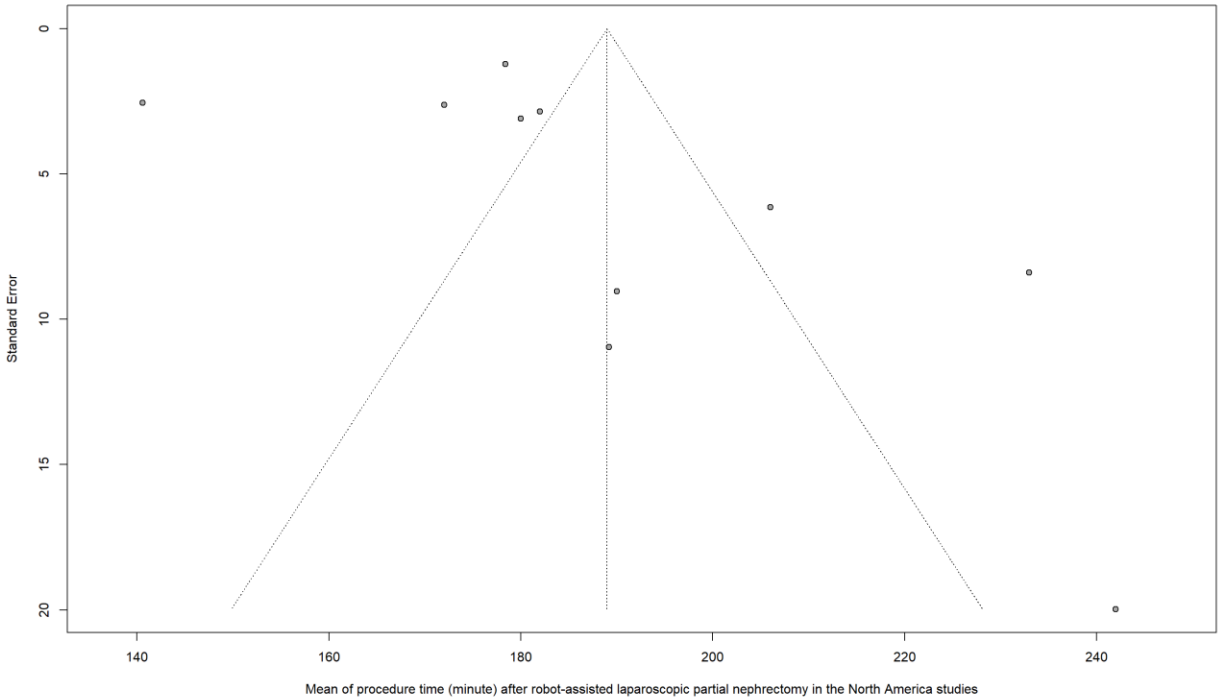
Supplementary Figure 2A. Funnel plot of studies on duration of conventional laparoscopic partial nephrectomy in the North America subgroup (Begg test result: Z-value: 0.96, P.V: 0.337).



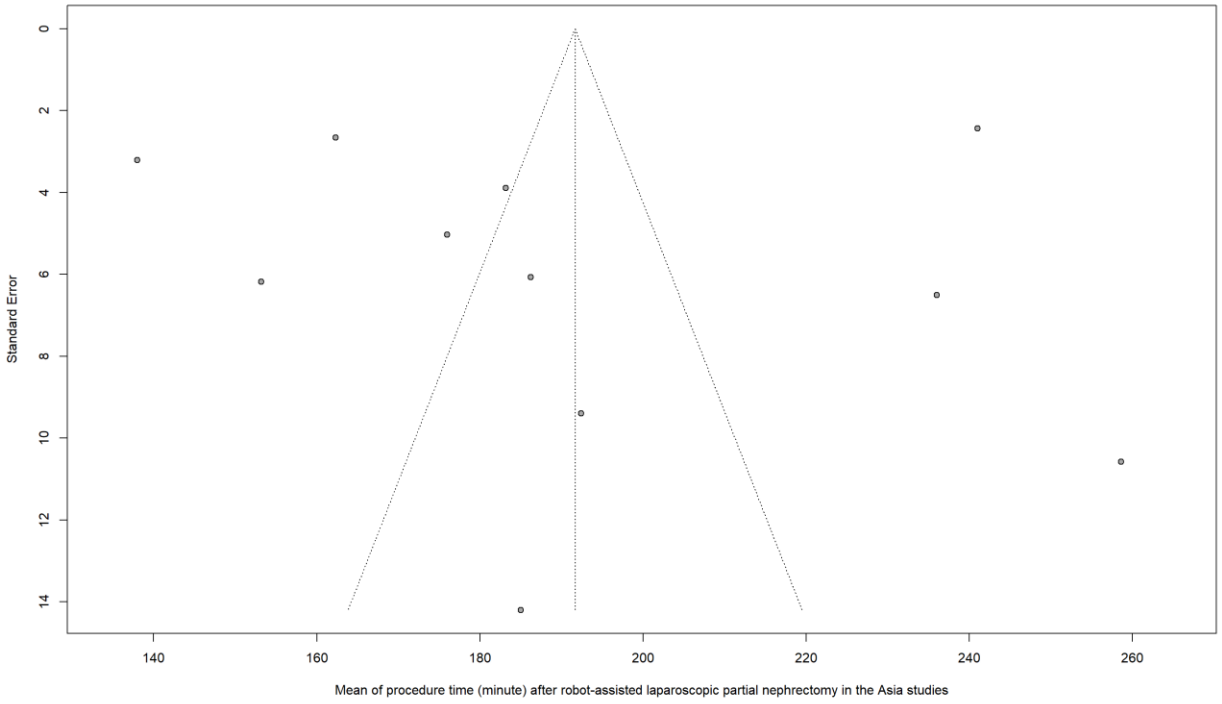
Supplementary Figure 2B. Funnel plot of studies on duration of conventional laparoscopic partial nephrectomy in the Asia subgroup (Begg test result: Z-value: 0.09, P.V: 0.9287).



Supplementary Figure 3A. Funnel plot of studies on duration of robot-assisted partial nephrectomy in the North America subgroup (Begg test result: Z-value: 0.63, P.V: 0.5312).



Supplementary Figure 3B. Funnel plot of studies on duration of robot-assisted partial nephrectomy in the Asia subgroup (Begg test result: Z-value: 1.17, P.V: 0.2429).



Appendix D: Search strategy

Summary of search and strategy Maryam nephrectomy

MEDLINE	1002
EMBASE	1644
Central	22
CINAHL	98
Subtotal	2766
-dupes	636
Total	2130

Oct 15, 2020

MEDLINE

Database: OVID Medline Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present
Search Strategy:
---
1 exp Nephrectomy/ (34694)
2 nephrectom*.mp. (53053)
3 (thermal adj3 ablat*).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (4218)
4 Watchful Waiting/ (3891)
5 (active adj3 surveillance).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (9813)
6 or/1-5 (69709)
7 exp Kidney Neoplasms/ (74958)

8 ((tumor or tumour or mass) adj3 (renal or kidney)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (17338)
9 small.ti,ab. (1449622)
10 7 or 8 (82903)
11 9 and 10 (6389)
12 6 and 11 (2422)
13 Epidemiologic Studies/ (8422)
14 exp Cohort Studies/ (2041265)
15 (cohort adj (study or studies)).tw. (226879)
16 Cohort analy$.tw. (8764)
17 (Follow up adj (study or studies)).tw. (51472)
18 (observational adj (study or studies)).tw. (119947)
19 Longitudinal.tw. (263163)
20 Retrospective.tw. (586595)
21 or/13-20 (2530306)
22 randomized controlled trial.pt. (515040)
23 controlled clinical trial.pt. (93884)
24 randomized.ab. (524659)
25 placebo.ab. (219751)
26 drug therapy.fs. (2242067)
27 randomly.ab. (367032)
28 trial.ab. (553507)
29 groups.ab. (2227708)
30 or/22-29 (4979209)
31 21 or 30 (6699492)
32 exp animals/ not humans.sh. (4743161)
33 31 not 32 (6026868)
34 12 and 33 (1002)

EMBASE (OVID)
Database: Embase <1974 to 2020 October 14>
Search Strategy:
---
1 exp nephrectomy/ (66450)
2 nephrectom*.mp. (70489)
3 (thermal adj3 ablat*).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (6628)
4 watchful waiting/ (4382)
5 (active adj3 surveillance).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (15164)

6 or/1-5 (101332)
7 exp kidney tumor/ (131055)
8 ((tumor or tumour or mass) adj3 (renal or kidney)).mp. (51666)
9 7 or 8 (142713)
10 small.ti,ab. (1750802)
11 9 and 10 (14602)
12 6 and 11 (4940)
13 clinical study/ (156038)
14 case control study/ (162184)
15 family study/ (26115)
16 longitudinal study/ (146053)
17 retrospective study/ (976304)
18 prospective study/ (634391)
19 randomized controlled trials/ (188529)
20 18 not 19 (627517)
21 cohort analysis/ (625505)
22 (Cohort adj (study or studies)).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (320042)
23 (Case control adj (study or studies)).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (220941)
24 (follow up adj (study or studies)).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (67816)
25 (observational adj (study or studies)).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (262275)
26 (epidemiologic$ adj (study or studies)).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (112903)
27 (cross sectional adj (study or studies)).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (418567)
28 or/13-17,20-27 (3003109)
29 randomized controlled trial/ (626125)
30 Controlled clinical study/ (465180)
31 random$.ti,ab. (1588526)
32 randomization/ (88590)
33 intermethod comparison/ (266032)
34 placebo.ti,ab. (312959)
35 (compare or compared or comparison).ti. (521479)
36 ((evaluated or evaluate or evaluating or assessed or assess) and (compare or compared or comparing or comparison)).ab. (2188456)
37 (open adj label).ti,ab. (82301)

38 ((double or single or doubly or singly) adj (blind or blinded or blindly)).ti,ab. (237028)
39 double blind procedure/ (177220)
40 parallel group$1.ti,ab. (26274)
41 (crossover or cross over).ti,ab. (107480)
42 ((assign$ or match or matched or allocation) adj5 (alternate or group$1 or intervention$1 or patient$1 or subject$1 or participant$1)).ti,ab. (339719)
43 (assigned or allocated).ti,ab. (400299)
44 (controlled adj7 (study or design or trial)).ti,ab. (360406)
45 (volunteer or volunteers).ti,ab. (252172)
46 human experiment/ (519532)
47 trial.ti. (311759)
48 or/29-47 (5169444)
49 (random$ adj sampl$ adj7 ("cross section$" or questionnaire$1 or survey$ or database$1)).ti,ab. not (comparative study/ or controlled study/ or randomi?ed controlled.ti,ab. or randomly assigned.ti,ab.) (8259)
50 Cross-sectional study/ not (randomized controlled trial/ or controlled clinical study/ or controlled study/ or randomi?ed controlled.ti,ab. or control group$1.ti,ab.) (249362)
51 (((case adj control$) and random$) not randomi?ed controlled).ti,ab. (17712)
52 (Systematic review not (trial or study)).ti. (155540)
53 (nonrandom$ not random$).ti,ab. (16488)
54 "Random field$".ti,ab. (2423)
55 (random cluster adj3 sampl$).ti,ab. (1313)
56 (review.ab. and review.pt.) not trial.ti. (835852)
57 "we searched".ab. and (review.ti. or review.pt.) (33659)
58 "update review".ab. (110)
59 (databases adj4 searched).ab. (38319)
60 (rat or rats or mouse or mice or swine or porcine or murine or sheep or lambs or pigs or piglets or rabbit or rabbits or cat or cats or dog or dogs or cattle or bovine or monkey or monkeys or trout or marmoset$1).ti. and animal experiment/ (1083488)
61 Animal experiment/ not (human experiment/ or human/) (2281812)
62 or/49-61 (3564463)
63 48 not 62 (4598410)
64 28 not 62 (2638136)
65 63 or 64 (6411648)
66 12 and 65 (1644)

Cochrane Library (Wiley)
Search Name: 2020-10-15 Renal Mass time to event
Date Run: 15/10/2020 15:21:15
Comment:

ID Search Hits
#1 MeSH descriptor: [Nephrectomy] explode all trees 313
#2 nephrectom* 1594

#3	MeSH descriptor: [Watchful Waiting] explode all trees	302
#4	active NEAR/3 surveillance	776
#5	#1 or #2 #3 or #4	1088
#6	MeSH descriptor: [Kidney Neoplasms] explode all trees	1180
#7	((tumor or tumour or mass) NEAR/3 (renal or kidney))	779
#8	#6 or #7	1846
#9	small	90530
#10	#8 and #9	192
#11	#5 and #10 in Trials	22

CINAHL (EBSCO)

#	Query	Results
S46	S33 AND S45	98
S45	S39 AND S44	358
S44	S42 AND S43	951
S43	TI small OR AB small	168,161
S42	S40 OR S41	12,235
S41	TX (tumor or tumour or mass) N3 (renal or kidney)	9,928
S40	(MH "Kidney Neoplasms+")	10,558
S39	S34 OR S35 OR S36 OR S37 OR S38	10,183
S38	TX active N3 surveillance	3,340
S37	"watchful waiting"	855
S36	TX thermal N3 ablat*	913
S35	TX nephrectom*	5,245
S34	(MH "Nephrectomy+")	3,570
S33	S23 OR S32	1,320,265
S32	S31 NOT S21	769,278
S31	S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30	772,997
S30	(MH "Prospective Studies")	449,184
S29	(MH "Case Control Studies+")	83,647

S28	(MH "Correlational Studies")	26,161
S27	(MH "Nonconcurrent Prospective Studies")	248
S26	(MH "Cross Sectional Studies")	194,525
S25	TX (cohort N3 (study or studies))	106,114
S24	TX (observational N3 (study or studies))	56,688
S23	S22 NOT S21	753,422
S22	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15	791,716
S21	S19 NOT S20	186,512
S20	MH (HUMAN)	2,251,790
S19	S16 OR S17 OR S18	214,608
S18	TI (ANIMAL MODEL*)	3,004
S17	MH (ANIMAL STUDIES)	129,641
S16	MH ANIMALS+	93,247
S15	AB (CLUSTER W3 RCT)	364
S14	MH (CROSSOVER DESIGN) OR MH (COMPARATIVE STUDIES)	354,689
S13	AB (CONTROL W5 GROUP)	114,063
S12	PT (randomized controlled trial)	122,692
S11	MH (placebos)	12,581
S10	MH (sample size) AND AB (assigned OR allocated OR control)	4,075
S9	TI (trial)	110,318
S8	AB (random*)	319,865
S7	TI (randomised OR randomized)	108,196
S6	(MH "Cluster Sample")	4,559
S5	(MH "Pretest-Posttest Design")	44,055
S4	(MH "Random Assignment")	63,963

S3	(MH "Single-Blind Studies")	14,377
S2	(MH "Double-Blind Studies")	48,730
S1	(MH "Randomized Controlled Trials")	108,777