

## Evidence-based Practice Center Systematic Review Protocol

### Project Title: Management of Renal Masses and Localized Renal Cancer

Amendment Date applicable: March 2, 2015

(Amendments Details—see Section VII)

#### I. Background and Objectives for the Systematic Review

##### Background

##### Epidemiology and Population of Interest

Renal masses are a biologically heterogeneous group of tumors, ranging from benign masses, to indolent cancers that behave in a benign fashion, and finally, extremely aggressive and deadly cancers.[1, 2] The true incidence of renal masses (including benign lesions) is unknown, but benign lesions comprise approximately 20% of surgically resected tumors.[1, 3] It is known that kidney cancer affects approximately 65,000 new patients each year with more than 13,000 deaths on an annual basis.[4] The incidence of kidney cancer has increased dramatically over the past few decades, believed due to the increased use of axial imaging leading to earlier detection of cancer prior to symptoms.[5] The greatest increase in incidence was noted in small, clinically-localized tumors (i.e. tumors within the kidney with no evidence of local spread, lymph node involvement or distant metastases), now upwards of 40% of all kidney cancers.[6, 7] Interestingly, despite this increase in early-stage cancers, the death rate from kidney cancer has not changed significantly over the same time period. [4] This may be reflective of the stable rates of patients presenting with advanced and metastatic cancer, a changing biology of kidney cancer, or the overtreatment of indolent lesions with resultant complications of that treatment.

Renal cell carcinoma (RCC) and urothelial cell carcinoma (UCC) are the two most common types of kidney cancer. Renal cell carcinoma is the predominant kidney cancer (>94%) in the United States and is clinically distinct from urothelial cancers (<6%) that comprise the remainder of kidney cancers.[4] While renal cell carcinoma only represents 2% of adult cancers, it is among the most lethal, with approximately 35% of patients dying within 5 years of diagnosis.[8] However, the deaths due to renal cell carcinoma are driven by the failure of systemic treatments in metastatic (later stage) patients. Localized renal cell carcinoma (stage T1N0M0, ≤ 7 cm in diameter; stage T2N0M0, > 7 cm in diameter) have excellent outcomes with 5-year survival rates better than 85%.[9, 10] Importantly, approximately 40% of clinically localized tumors are determined to be locally-advanced cancers (stage T3, with invasion of perinephric fat or venous structures) at pathological examination.[11] Locally advanced cancers have a significantly worse prognosis than localized tumors.

##### Diagnostic Evaluation and Detection of Disease

All solid renal masses and cystic lesions with solid components are suspicious for renal cell carcinoma. No test is effective at screening for renal cell carcinoma, and most tumors are detected incidentally during an evaluation for unrelated or non-specific complaints. Preoperative patient and tumor (imaging) characteristics are used to stratify the risk of benign versus malignant disease, and indolent versus aggressive cancers. Demographic, clinical and imaging characteristics are used to risk-stratify patients, and nomograms exist that combine these

characteristics into composite models to classify tumors preoperatively.[12-15] In small studies, such models have modest concordance indices for malignancy in the range of 0.55 to 0.65.[3, 16-18] The best predictors of malignancy are male sex and tumor size, although computed tomography (CT) enhancement patterns have also been able to predict histology in up to 85% of cases.[1, 19-21] In addition, all current standard imaging modalities (CT, magnetic resonance imaging (MRI) and ultrasound) are able to provide insight into whether renal masses are localized or locally advanced, thus suggesting pathologic aggressiveness. Larger or central tumors are more likely to be invasive, but small peripheral tumors may also be invasive.[16, 22]

Percutaneous renal mass sampling is offered as a diagnostic adjunct to traditional axial imaging with contrast. Renal mass sampling can be done by fine needle aspiration with a reading of the sample by a cytopathologist, or by core biopsy with a reading of the sample by a surgical pathologist. In addition, percutaneous sampling can be performed by diagnostic radiologists, interventional radiologists or urologists. In one retrospective study, percutaneous renal core biopsy was diagnostic in only 81% of cases, with 79% of diagnostic biopsies showing evidence of malignancy.[23] However, percutaneous renal sampling carries a small but significant risk of bleeding and tumor seeding – limiting its widespread use and a well-defined role in the management of localized renal masses. Despite the shortcomings in diagnostic testing, management decisions are made based on estimations of malignant potential driven by surgeon decision and patient preference. To help guide such decisions, we plan to perform a systematic review of the evidence on how well malignancy can be predicted using percutaneous renal sampling or composite profiles of demographic, clinical, and imaging characteristics. We anticipate that evidence on the predictive value of renal mass sampling and composite profiles most likely will be limited to studies of patients that have undergone surgery for a renal mass suspicious for carcinoma, without much, if any, evidence from studies of patients that did not undergo surgery for a renal mass. The review of evidence may help to identify sub-populations for whom particular interventions are more impactful. For instance, young women with small (<2 cm) tumors have the highest likelihood of having benign tumors (upwards of 40% in some studies). This sub-population, enriched for benign masses, may have a different risk-benefit ratio for percutaneous renal sampling or nephron-sparing surgeries than the population of men with large, centrally-located tumors – who have the highest risk of having aggressive malignancies.

### **Therapeutic Interventions and Outcomes**

Several options exist for the management of clinically-localized renal masses suspicious for renal cell carcinoma. These include active surveillance, and minimally invasive and open surgical options which include partial nephrectomy, radical nephrectomy and thermal ablation. Given the increased incidence in early, low-stage tumors without improvement in cancer-related deaths (see above), active surveillance has emerged as an option for patients with small renal masses, a low likelihood of aggressive malignancy, and/or a limited life expectancy. Minimally invasive options include both standard laparoscopy and robot-assisted laparoscopic approaches. Extirpative surgery (radical or partial nephrectomy) is the gold-standard for the treatment of renal cell carcinoma. Partial nephrectomy is considered the standard-of-care by the American Urological Association (AUA) for stage T1a ( $\leq 4$ cm in diameter) tumors (when technically feasible); radical nephrectomy is an alternative standard, while active surveillance and thermal ablation are also options.[24] For larger clinically-localized (stage T1b and T2) tumors, partial nephrectomy and radical nephrectomy are standards, while active surveillance and thermal ablation are options. Typically, radical nephrectomy is performed via a minimally invasive

approach (when technically feasible), while it is standard for partial nephrectomy to be performed either through an open incision or minimally invasive approach. Thermal ablation, which may include either cryoablation or radiofrequency ablation, can either be performed laparoscopically or percutaneously. However, professional organizations refrain from defining strict selection criteria (patient or tumor) for particular management strategies, and the existing selection criteria vary by organizational guideline.[24-26] Controversies exist regarding the ideal management for renal masses of different stages. For example, partial nephrectomy has emerged as the recommended treatment for T1 renal masses, yet the single randomized, prospective study demonstrated improved overall survival with radical nephrectomy.[27] The role of age in selecting patients for surgery and type of surgery is not well-established. Older patients may have very different tumors than young patients; older patients are more likely to need nephron-sparing approaches, but may not live long enough to see the detriment of end-stage renal disease.[28] Moreover, cryoablation has comparable oncologic outcomes to partial nephrectomy, although it is not clear which patients are best served with each treatment modality.[29]

The main outcomes of interest in this population are cancer-specific and recurrence-free survival, renal functional outcomes, and the complications associated with each procedure. All extirpative options are associated with an excellent oncologic cure rate (> 95% 5-year disease specific survival for stage T1, and > 85% for stage T2 tumors). Based on the current literature, it is generally believed that nephron-sparing approaches (partial nephrectomy or thermal ablation) are associated with improved renal functional outcomes, but may not have an overall survival benefit.[27, 30, 31] It is also believed that surgical options (partial nephrectomy, radical nephrectomy) may have better oncologic outcomes than active surveillance or thermal ablation. However, surgical interventions (partial nephrectomy, radical nephrectomy) are associated with significantly higher complication rates than thermal ablation or active surveillance; in general, partial nephrectomy has a higher complication rate than radical nephrectomy.[24] All technologies are approved in the U.S. and are established treatment options for renal tumors.

### **Current Guidelines and Shortcomings**

Multiple expert organizations have put forth clinical guidelines on the management of renal masses, including the American Urological Association (AUA), European Association of Urology, and National Comprehensive Cancer Network.[24-26] In addition, the American College of Radiology published a guideline in 2010 to evaluate the appropriateness of radiologic examinations for patients with an indeterminate renal mass – a mass unable to be confidently diagnosed as benign or malignant at the time of discovery.[32] The most widely used guideline within the U.S. urological community was published in 2009 by the AUA. This was largely based on expert opinion and best-available studies that were observational and retrospective in design.[24] Since its publication, multiple significant advances in renal mass detection, diagnosis, risk stratification, and treatment have been made, making a systematic, evidenced-based update necessary. For example, an important recent contribution to the literature is a randomized trial of partial nephrectomy and radical nephrectomy that failed to demonstrate a clinically significant benefit for partial nephrectomy with respect to oncologic or renal functional outcomes.[21] Additionally, a relatively new concept of chronic kidney disease related to surgical and medical disease has emerged that could change how clinicians view existing evidence on the benefits and harms of different strategies for managing a renal mass.[33] Patients with medical chronic kidney disease have a progressive loss of renal function and poor

prognosis with regard to global renal function and overall survival. Patients with surgical chronic kidney disease (defined as chronic kidney disease as a result of surgical nephron loss or injury) is dramatically different, with stable long-term renal function and improved overall survival when compared with chronic kidney disease resulting from medical renal disease.

Determining the best approach to management of clinically-localized renal masses is a complex task. Creating a patient-centered treatment strategy that incorporates factors related to the renal mass (oncologic outcomes, renal functional outcomes, complications) as well as competing health risks was not feasible in the most recent AUA guideline. Treatments such as robotic partial nephrectomy were still in their infancy, and the large clinical data sets lacked the granularity of modern data.

### **Rationale and Relevance for an Evidence Review**

This review will provide a comprehensive review of current data that can help the AUA and other organizations prepare updated guidance on how to evaluate and treat patients with a renal mass suspicious for renal cell carcinoma. First, this review will synthesize current evidence on key issues in the evaluation of patients having a renal mass suspicious for renal cell carcinoma. Second, the review will synthesize evidence on the effectiveness and comparative effectiveness of different strategies for treating patients having a renal mass suspicious for renal cell carcinoma. The review also will highlight areas of controversy and identify needs for future research on the management of renal masses and localized renal cell carcinoma.

## II. The Key Questions

**Key Question 1:** In patients that undergo surgery for a renal mass suspicious for stage I or II renal cell carcinoma, how does the pathologic diagnosis compare to the likelihood of malignancy predicted by using a pre-operative composite profile of patient characteristics including demographics, clinical characteristics, blood/urine markers, and/or imaging?

***For the purpose of this question and further key questions, a renal mass suspicious for stage I or II renal cell carcinoma includes all solid renal masses and cystic renal masses with a solid component.***

**Key Question 2a:** In patients that undergo surgery for a renal mass suspicious for stage I or II renal cell carcinoma, what is the accuracy (i.e., sensitivity, specificity, positive and negative predictive value) of percutaneous renal mass sampling (fine needle aspiration or core biopsy, with cytopathology or surgical pathology) in establishing a diagnosis (malignancy, histology and grade)?

**Key Question 2b:** In patients with a renal mass suspicious for stage I or II renal cell carcinoma, what are the adverse effects associated with using renal mass sampling (see KQ2) to estimate the risk of malignancy, including direct complications (e.g., pain, infection, hemorrhage, radiation exposure) and harms related to false positives, false negatives, or non-diagnostic results?

**Key Question 3a:** In patients with a renal mass suspicious for stage I or II renal cell carcinoma, what is the effectiveness and comparative effectiveness of the available management strategies on health outcomes?

***Available management strategies include: radical nephrectomy (open and minimally-invasive), partial nephrectomy (open and minimally-invasive), thermal ablation (radiofrequency ablation or cryoablation; surgical or image-guided), and active surveillance. The health outcomes of interest include all of the potential benefits and harms listed under outcomes in the PICOTS framework below.***

**Key Question 3b:** Do the comparative benefits and harms of the available management strategies differ according to a patient's demographic or clinical characteristics, or disease severity defined in terms of clinical presentation, tumor characteristics (imaging), renal mass sampling results, or laboratory evaluations?

**Preliminary PICOTS (patients, interventions, comparators, outcomes, timing, setting)**

**Population(s):** Newly diagnosed adults (age greater than or equal to 18 years) with solid renal masses (or cystic renal masses with a solid component) suspicious for stage I and II renal cell carcinoma, which corresponds to clinical stage T1 (< 7 cm, organ confined) or T2 (> 7 cm, organ confined) renal masses

**Subgroups:**

Male and female

Older (65 and over) and younger patients

Patients with comorbidities

**Interventions:**

**Diagnostic interventions:**

Percutaneous renal mass sampling (fine needle aspiration or biopsy)

Composite models (combination of demographics, clinical characteristics, blood/urine tests, and tumor imaging characteristics) for predicting malignancy

Demographic characteristics: Age, sex, smoking, race, marital status, education

Clinical characteristics: obesity, and comorbidities, specifically cardiovascular disease and chronic kidney disease

Blood/urine tests: measures of kidney function, markers of paraneoplastic syndromes and predictors of advanced/metastatic disease (e.g. complete metabolic panel, complete blood count, coagulation parameters, erythrocyte sedimentation rate)

Imaging characteristics: CT scan, Ultrasound, MRI

**Management options:**

Radical nephrectomy (open and minimally invasive\*)

Partial nephrectomy (open and minimally invasive\*)

Thermal ablation (e.g., radiofrequency ablation, cryoablation; surgical vs. image-guided)

Active surveillance

\*Minimally-invasive surgery may refer to standard laparoscopy or robot-assisted laparoscopy

**Comparator(s):** For diagnosis-related questions, the main comparison is between biopsy results, composite models, and the pathologic diagnosis after surgical intervention. For treatment-related questions, the comparisons of interest include all of the management options listed above.

**Outcome(s):**

- **Diagnostic test-related outcomes**
  - False positives
  - False negatives
  - Radiation exposure
- **Adverse effects of percutaneous renal mass sampling:**
  - Pain
  - Hemorrhage
  - Tumor Seeding
- **Adverse effects of management strategies:**
- **Urologic complications:**
  - Acute kidney injury
  - Hemorrhage
  - Urine leak

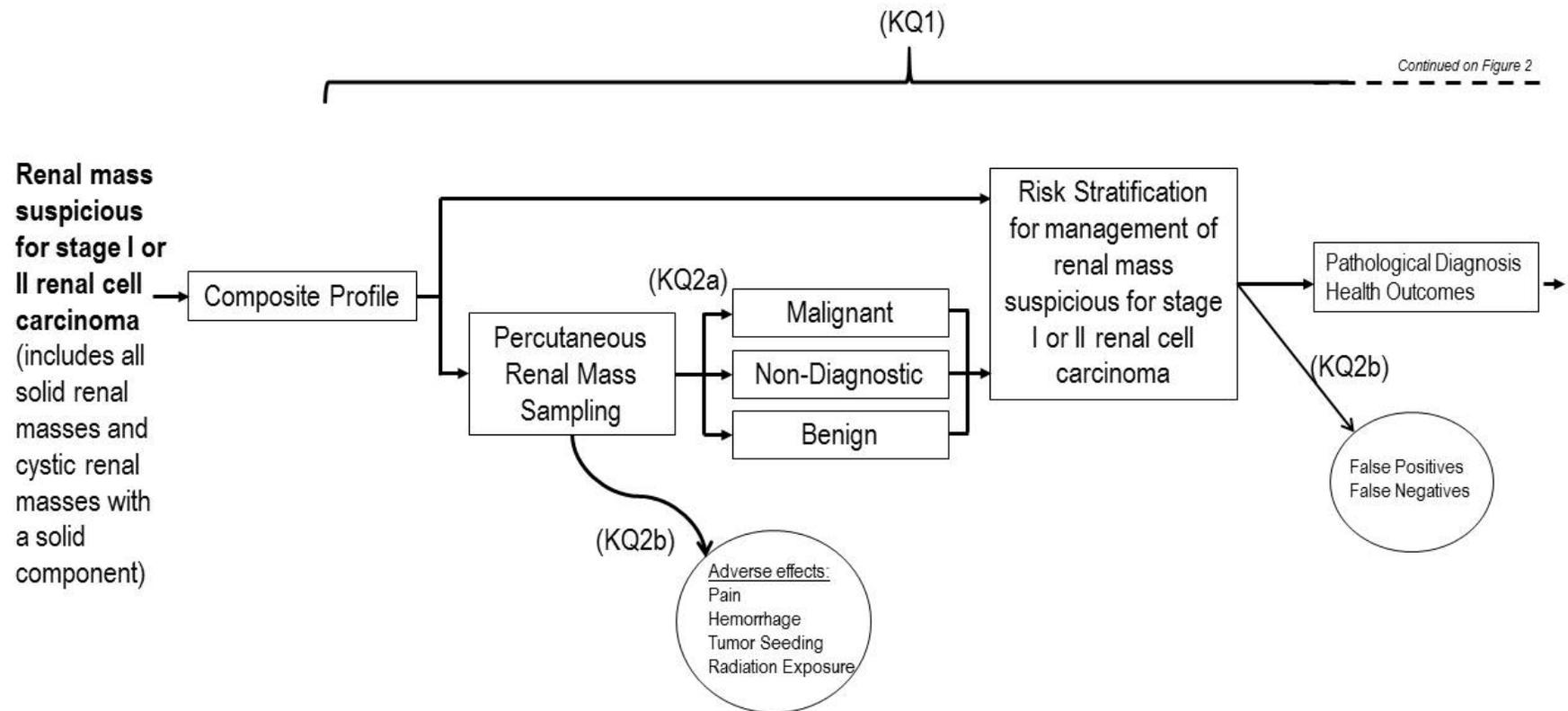
- Hematuria
- Loss of kidney
- Ureteral injury (any injury of collecting system and ureter)
- Urinary tract infection
  
- **Non-urologic complications (by organ-system)**
  - Hematologic (thromboembolic)
  - Gastrointestinal
  - Cardiovascular
  - Respiratory
  - Neurologic
  - Wound complications (hernia, dehiscence)
  - Infectious disease
  - Listed by severity of complications (using the Clavien Grading System if available)
    - Minor vs. major
      - Minor: conservative management or medications only
      - Major: requiring intervention, resulting in permanent disability or death
  - Need for subsequent interventions: embolization, drain placement, stent placement, etc.
- **Peri-operative outcomes**
  - Blood loss (cc or mL)
  - Blood transfusion (yes or no)
  - Length of stay (days)
  - Conversion to:
    - Radical nephrectomy (if initial nephron-sparing approach)
    - Open surgery (if initial minimally-invasive approach)
- **Final health outcomes:**
  - **Oncologic efficacy**
    - Local recurrence-free survival
    - Metastasis-free survival
    - Cancer-specific survival
  - **Renal functional outcomes**
    - Glomerular filtration rate decline
    - Incidence of chronic kidney disease
    - Incidence of end-stage renal disease
  - **Overall survival**
  - **Quality of Life**

**Timing:** All time points

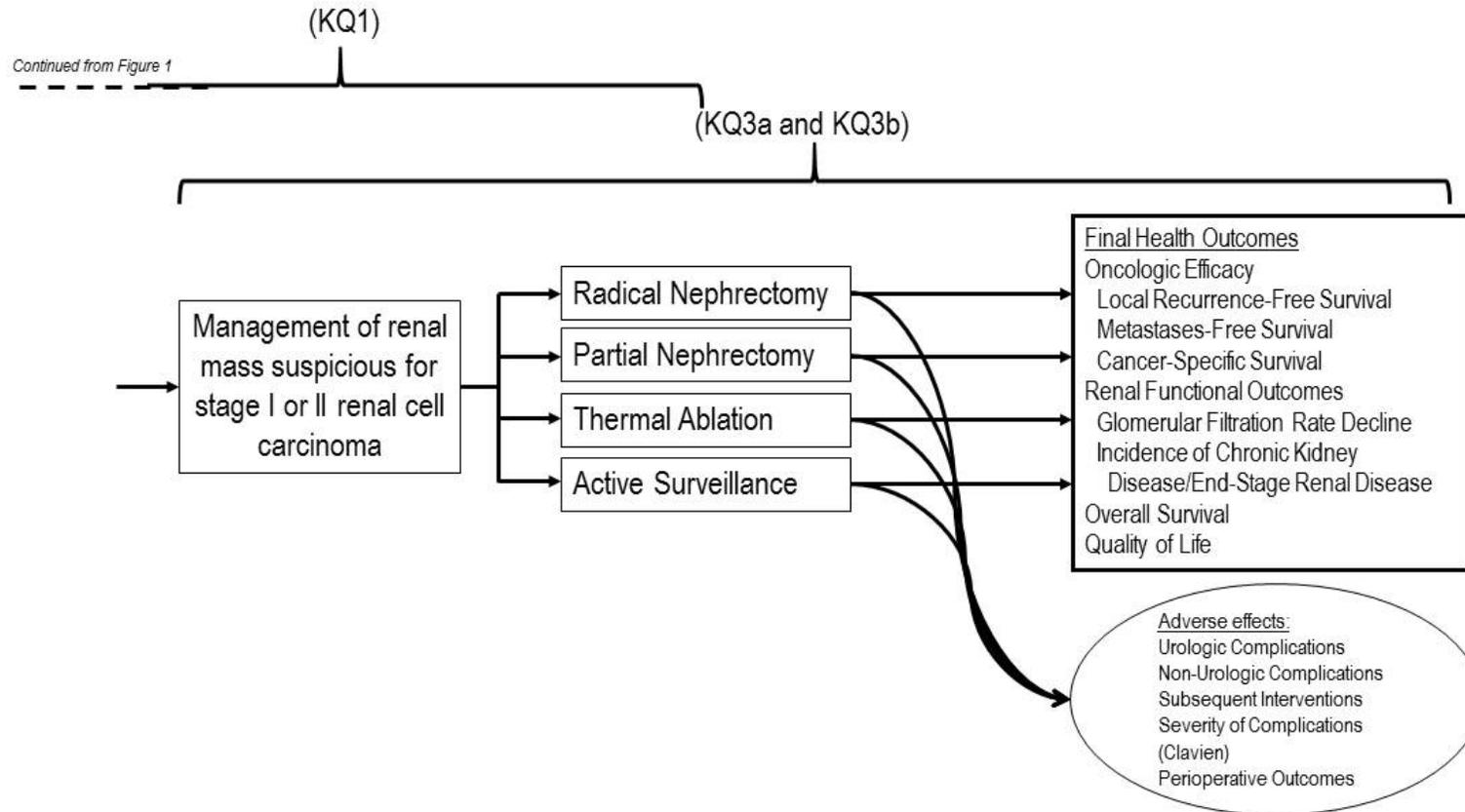
**Settings:** All settings

### III. Analytic Framework

**Figure 1.** Preliminary analytic framework for systematic review of the management of renal masses and localized kidney cancer.  
*PART I: Diagnostic framework.*



**Figure 2.** Preliminary analytic framework for systematic review of the management of renal masses and localized kidney cancer.  
*PART II: Management Strategies.*



## IV. Methods

### A. Criteria for Inclusion/Exclusion of Studies in the Review

Inclusion and exclusion criteria are provided in Table A, based on the PICOTS framework described above.

**Table 1: List of Inclusion/Exclusion Criteria**

	<b>Inclusion</b>	<b>Exclusion</b>
<b>Population</b>	<ul style="list-style-type: none"> <li>Age <math>\geq</math> 18</li> <li>Newly diagnosed adults with solid renal masses (or cystic renal masses with a solid component) suspicious for stage I and II renal cell carcinoma, which corresponds to clinical stage T1 (&lt; 7 cm, organ confined) or T2 (&gt; 7 cm, organ confined) renal masses</li> </ul>	<ul style="list-style-type: none"> <li>Animal studies</li> <li>Age &lt;18</li> <li>Patients with recurrent renal cell carcinoma</li> </ul>
<b>Interventions</b>	<p><b>Diagnostic interventions (KQ1, KQ2)</b>            Percutaneous renal mass sampling (fine needle aspiration or biopsy)</p> <ul style="list-style-type: none"> <li>Composite models (combination of demographics, clinical characteristics, blood/urine tests, and tumor imaging characteristics) for predicting malignancy*               <ul style="list-style-type: none"> <li>➤ <b>Demographic characteristics:</b> Age, sex, smoking, race, marital status, education</li> <li>➤ <b>Clinical characteristics:</b> obesity, and comorbidities, specifically cardiovascular disease and chronic kidney disease</li> <li>➤ <b>Blood/urine tests:</b> measures of kidney function, markers of paraneoplastic syndromes and predictors of advanced/metastatic disease (e.g. complete metabolic panel, complete blood count, coagulation parameters, erythrocyte sedimentation rate)</li> <li>➤ <b>Imaging characteristics:</b> Computerized tomography scan, Ultrasound, magnetic resonance imaging</li> </ul> </li> </ul> <p><b>*Composite models need to adjust for imaging characteristics (i.e. tumor size) or at least one element from 2 of the categories</b></p>	
	<p><b>Management options (KQ3a and 3b):</b></p> <ul style="list-style-type: none"> <li>Radical nephrectomy (open and minimally invasive*)</li> <li>Partial nephrectomy (open and minimally invasive*)</li> <li>Thermal ablation (e.g., radiofrequency ablation, cryoablation; surgical vs. image-guided)</li> <li>Active surveillance (minimum six months)</li> </ul> <p>*Minimally-invasive surgery may refer to standard laparoscopy or robot-assisted laparoscopy</p>	We will exclude studies that do not specify the type of surgery.
<b>Comparisons</b>	<p><b>For diagnosis-related questions:</b>            The main comparison is between biopsy results, composite models, and the pathologic diagnosis after surgical intervention.</p> <p><b>For treatment-related questions:</b></p>	

	All of the management options listed above.	
<b>Outcomes</b>	<p>We will include studies that evaluate at least one of the following outcomes:</p> <p><b>Diagnostic Test-Related Outcomes:</b></p> <ul style="list-style-type: none"> <li>• False positives</li> <li>• False negatives</li> <li>• Radiation exposure</li> </ul> <p><b>Adverse effects of percutaneous renal mass sampling:</b></p> <ul style="list-style-type: none"> <li>• Pain</li> <li>• Hemorrhage</li> <li>• Tumor seeding</li> </ul> <p><b>Final health outcomes:</b></p> <ul style="list-style-type: none"> <li>• Oncologic efficacy <ul style="list-style-type: none"> <li>○ Local recurrence-free survival</li> <li>○ Metastasis-free survival</li> <li>○ Cancer-specific survival</li> </ul> </li> <li>• Renal functional outcomes <ul style="list-style-type: none"> <li>○ Glomerular filtration rate decline</li> <li>○ Incidence of chronic kidney disease</li> <li>○ Incidence of end-stage renal disease</li> </ul> </li> <li>• Overall survival</li> <li>• Quality of Life</li> </ul> <p><b>Adverse effects of management strategies :</b></p> <ul style="list-style-type: none"> <li>• <b>Urologic complications:</b> <ul style="list-style-type: none"> <li>• Acute kidney Injury</li> <li>• Hemorrhage</li> <li>• Urine leak</li> <li>• Hematuria</li> <li>• Loss of kidney</li> <li>• Ureteral injury (any injury of collecting system and ureter)</li> <li>• Urinary tract infection</li> </ul> </li> <li>• <b>Non-urologic complications (by organ-system)</b> <ul style="list-style-type: none"> <li>• Hematologic (thromboembolic)</li> <li>• Gastrointestinal</li> <li>• Cardiovascular</li> <li>• Respiratory</li> <li>• Neurologic</li> <li>• Wound complications (hernia, dehiscence)</li> <li>• Infectious disease</li> <li>• Listed by severity of complications (using the Clavien Grading System if available):</li> <li>• Minor vs. major <ul style="list-style-type: none"> <li>○ Minor: conservative management or medications only</li> <li>○ Major: requiring intervention, resulting in permanent disability or death</li> </ul> </li> </ul> </li> <li>• Need for subsequent interventions: embolization, drain placement, stent placement, etc.</li> </ul> <p>• <b>Peri-operative outcomes</b></p>	<p>We will exclude studies that:</p> <p>do not report the health outcomes of interest</p> <p>OR</p> <p>do not give an adequate description (clavien grading ) of how complications and peri-operative outcomes were assessed</p> <p>OR</p> <p>report only selected complications or peri-operative outcomes of interest unless primary objective of the study was to assess the complications</p>

	<ul style="list-style-type: none"> <li>○ Blood loss (cc or mL)</li> <li>○ Blood transfusion (yes or no)</li> <li>○ Length of stay (days)</li> <li>○ Conversion to: <ul style="list-style-type: none"> <li>○ Radical nephrectomy (if initial nephron-sparing approach)</li> <li>○ Open surgery (if initial minimally-invasive approach)</li> </ul> </li> </ul>	
<b>Type of Study</b>	<p>We will include randomized controlled trials, non-randomized controlled trials, and cohort studies.</p> <p>We will only include case-control studies if they address one of the diagnosis-related key questions.</p>	<p>We will exclude:</p> <ol style="list-style-type: none"> <li>1) publications with no original data (e.g., editorials, letters, comments, reviews;</li> <li>2) case reports</li> <li>3) case series of less than 10 patients for diagnostic studies;</li> <li>4) case series of less than 25 patients for therapeutic studies.</li> </ol> <p>We will limit the systematic review to publications from 1997 or later for two specific reasons. (1) In 1997, the TNM staging system for renal cell carcinoma was modified and the distinctions of T1a/T1b and T2a/T2b were created. T1a tumors, specifically, have the best prognosis and define a category of tumors in which nephron-sparing surgery is considered to be the standard and active surveillance is considered acceptable. (2) In addition, the first laparoscopic radical nephrectomy was performed in 1992. By 1997, minimally-invasive surgery was pervasive in urologic surgery and most surgeons were beyond the learning curve associated with these operations. Therefore, the time period from 1997 to the present accurately reflects the “contemporary” management of clinically localized kidney cancer. Prior to 1997, the staging system and management options do not reflect the current state of practice.</p>
<b>Timing and Setting</b>	<p>All time points All settings</p>	

- B. Searching for the Evidence: Literature Search Strategies for Identification of Relevant Studies to Answer the Key Questions:** We will search the following databases for primary studies: MEDLINE®, Embase®, and the Cochrane Central Register of Controlled Trials from January 1, 1997, through present. We will develop a search strategy for MEDLINE, accessed via PubMed, based on an analysis of the medical subject headings (MeSH) terms for all potential relevant publications and text words of key articles identified a priori. We will also review the reference lists of each included article, relevant review articles and related systematic reviews. The search will be updated during the peer review process. Our preliminary search strategy for MEDLINE is shown in Appendix A.

Additionally, we will search [clinicaltrials.gov](http://clinicaltrials.gov) to identify any relevant ongoing trials. We will review the Scientific Information Packets provided by the device manufacturers (Appendix B). We will review data from the medical devices section of the U.S. Food and Drug Administration.

We will use DistillerSR (Evidence Partners, 2010) to manage the screening process. DistillerSR is a web-based database management program that manages all levels of the review process. All applicable citations identified by the search strategies are uploaded to the system and reviewed in the following manner:

**i. Abstract screening:** Two reviewers will independently review abstracts, which will be excluded if both reviewers agree that the article meets one or more of the exclusion criteria listed in Table 1. Differences between reviewers regarding abstract eligibility will be tracked and resolved through consensus adjudication. Relevant reviews, including systematic reviews and meta-analyses, will be tagged for a references list search.

**ii. Full-text screening:** Citations promoted on the basis of abstract review will undergo another independent parallel review using full-text of the articles to determine if they should be included in the final qualitative and quantitative systematic review and meta-analysis. The differences regarding article inclusion will again be tracked and resolved through consensus adjudication.

- C. Data Abstraction and Data Management:** We will create and pilot test forms for data extraction. Each article will undergo double review for data abstraction. The second reviewer will confirm the first reviewer’s data abstraction for completeness and accuracy. A third reviewer will audit a random sample of articles by the first two reviewers to ensure consistency in the data abstraction of the articles.

Articles referring to the same study will be abstracted on a single review form if reporting the same data or on separate forms if necessary with clear information that the results should be interpreted as from the same study.

For all articles, reviewers will extract information on general study characteristics (e.g., study design, study period, and follow-up), eligibility criteria, study participants (e.g., age, gender, race/ethnicity, body mass index, comorbidities, etc.), renal mass characteristics, interventions, outcome measures and the method of ascertainment, and the results of each outcome, including measures of variability. Data will also be abstracted when available by subgroups such as gender, older (65 and over) and younger patients, or patients with comorbidities.

We will complete the data abstraction process using the Systematic Review Data Repository™ (SRDR). Data will be exported from SRDR into a project-specific Access database (Microsoft, Redmond, WA) to serve as archived or back-up copies and to create detailed evidence tables and summary tables.

- D. Assessment of Methodological Risk of Bias of Individual Studies:** The assessment of risk of bias of included trials of treatment interventions will be conducted independently and in duplicate using the Cochrane Collaboration’s Risk of Bias Tool (Table 2).[35] For non-randomized studies of treatment interventions, we will use the Cochrane Risk Of Bias Assessment Tool for Non-Randomized Studies of Interventions (ACROBAT-NRSI) (Table 3).[36] For diagnostic studies, we will use the quality assessment tool for diagnostic accuracy studies tool (QUADAS -2) (Table 4).[37] Differences between reviewers will be resolved through consensus adjudication.

<b>Table 2: The Cochrane Collaboration’s risk of bias tool for trials</b>	
<b>Domain</b>	<b>Review authors’ judgement</b>
<i>Selection bias</i>	
<b>Random sequence generation</b>	Selection bias (biased allocation to interventions) due to inadequate generation of a randomised sequence.
<b>Allocation concealment</b>	Selection bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment.
<i>Performance bias</i>	
<b>Blinding of participants and personnel</b> <i>Assessments should be made for each main outcome (or class of outcomes).</i>	Performance bias due to knowledge of the allocated interventions by participants and personnel during the study.
<i>Detection bias</i>	
<b>Blinding of outcome assessment</b> <i>Assessments should be made for each main outcome (or class of outcomes).</i>	Detection bias due to knowledge of the allocated interventions by outcome assessors.
<i>Attrition bias</i>	
<b>Incomplete outcome data</b> <i>Assessments should be made for each main outcome (or class of outcomes).</i>	Attrition bias due to amount, nature or handling of incomplete outcome data.
<i>Reporting bias</i>	
<b>Selective reporting</b>	Reporting bias due to selective outcome reporting.
<i>Other bias</i>	
<b>Other sources of bias</b>	Bias due to problems not covered elsewhere in the table.

<b>Table 3: Risk of bias assessment (cohort-type studies)</b>	
Bias due to confounding	1.1 Is confounding of the effect of intervention unlikely in this study?
	<b>If Y or PY to 1.1:</b> the study can be considered to be at low risk of bias due to confounding and no further signaling questions need be considered
	<b>If N or PN to 1.1:</b>
	1.2. Were participants analyzed according to their initial intervention group throughout follow up?
	<b>If Y or PY to 1.2,</b> answer questions 1.4 to 1.6, which relate to baseline confounding
	1.3. <b>If N or PN to 1.2:</b> Were intervention discontinuations or switches unlikely to be related to factors that are prognostic for the outcome?
	<b>If Y or PY to 1.3,</b> answer questions 1.4 to 1.6, which relate to baseline confounding
	<b>If N or PN to 1.1 and 1.2 and 1.3,</b> answer questions 1.7 and 1.8, which relate to time-varying confounding
	<b>If Y or PY to 1.2, or Y or PY to 1.3:</b>
	1.4. Did the authors use an appropriate analysis method that adjusted for all the critically important confounding domains?
	1.5. <b>If Y or PY to 1.4:</b> Were confounding domains that were adjusted for measured validly and reliably by the variables available in this study?
	1.6. Did the authors avoid adjusting for post-intervention variables?
	<b>If N or PN to 1.2 and 1.3</b>
	1.7. Did the authors use an appropriate analysis method that adjusted for all the critically important confounding domains and for time-varying confounding?
1.8. <b>If Y or PY to 1.7:</b> Were confounding domains that were adjusted for measured validly and reliably by the variables available in this study?	
<b>Risk of bias judgment</b>	
Bias in selection of participants into the study	2.1. Was selection into the study unrelated to intervention or unrelated to outcome?
	2.2. Do start of follow-up and start of intervention coincide for all or most subjects?
	2.3. <b>If N or PN to 2.1 or 2.2:</b> Were adjustment techniques used that are likely to correct for the presence of selection biases?
	<b>Risk of bias judgment</b>
Bias in measurement of interventions	3.1 Is intervention status well defined?
	3.2 Was information on intervention status recorded at the time of intervention?
	3.3 Was information on intervention status unaffected by knowledge of the outcome or risk of the outcome?
	<b>Risk of bias judgment</b>
Bias due to departures from intended interventions	4.1. Were the critical co-interventions balanced across intervention groups?
	4.2. Were numbers of switches to other interventions low?
	4.3. Was implementation failure minor?
	4.4. <b>If N or PN to 4.1, 4.2 or 4.3:</b> Were adjustment techniques used that are likely to correct for these issues?
	<b>Risk of bias judgment</b>
Bias due to missing data	5.1 Are outcome data reasonably complete?
	5.2 Was intervention status reasonably complete for those in whom it was sought?
	5.3 Are data reasonably complete for other variables in the analysis?
	5.4 <b>If N or PN to 5.1, 5.2 or 5.3:</b> Are the proportion of participants and reasons for missing data similar across interventions?

	5.5 If N or PN to 5.1, 5.2 or 5.3: Were appropriate statistical methods used to account for missing data?
	<b>Risk of bias judgment</b>
Bias in measurement of outcomes	6.1 Was the outcome measure objective?
	6.2 Were outcome assessors unaware of the intervention received by study participants?
	6.3 Were the methods of outcome assessment comparable across intervention groups?
	6.4 Were any systematic errors in measurement of the outcome unrelated to intervention received?
	<b>Risk of bias judgment</b>
Bias in selection of the reported result	Is the reported effect estimate unlikely to be selected, on the basis of the results, from...
	7.1 multiple outcome <i>measurements</i> within the outcome domain?
	7.2 multiple <i>analyses</i> of the intervention-outcome relationship?
	7.3 different <i>subgroups</i> ?
	<b>Risk of bias judgment</b>
Overall bias	<b>Risk of bias judgment</b>

<b>Table 4. QUADAS-2 questions for assessing risk of bias in diagnostic accuracy studies</b>	
<b>Domain 1: Patient Selection</b>	
Was a consecutive or random sample of patients enrolled? (Yes/No/Unclear)	
Did the study avoid inappropriate exclusions? (Yes/No/Unclear)	
<b>Could the selection of patients have introduced bias? Risk: Low/High/Unclear</b>	
<b>Domain 2: Index Test(s) (complete for each index test used)</b>	
Were the index test results interpreted without knowledge of the reference standard? (Yes/No/Unclear)	
<b>Could the conduct or interpretation of the index test have introduced bias? Risk: Low/High/Unclear</b>	
<b>Domain 3: Reference Standard</b>	
Is the reference standard likely to correctly classify the target condition? (Yes/No/Unclear)	
Were the reference standard results interpreted without knowledge of the results of the index test? (Yes/No/Unclear)	
<b>Could the reference standard, its conduct, or its interpretation have introduced bias? Risk: Low/High/Unclear</b>	
<b>Domain 4: Flow and Timing</b>	
Was there an appropriate interval (6 months) between index test(s) and reference standard? (Yes/No/Unclear)	
Did all patients receive a reference standard? (Yes/No/Unclear)	
Did all patients receive the same reference standard? (Yes/No/Unclear)	
Were all patients included in the analysis? (Yes/No/Unclear)	
<b>Could the patient flow have introduced bias? Risk: Low/High/Unclear</b>	

- E. Data Synthesis:** We will create a set of detailed evidence tables. We plan to conduct meta-analyses of summary data when there are sufficient data (at least 3 studies of the same design) and studies are sufficiently homogenous with respect to key variables (population characteristics, intervention, and outcome) using a random effects model. Randomized controlled trials and nonrandomized studies will be analyzed separately. Statistical significance (will be set at a two sided alpha of 0.05). All studies (diagnostic and management), including those that are not amenable to pooling, will be summarized qualitatively.

We will evaluate for statistical heterogeneity among studies using an  $I^2$  statistic, and anticipate statistical heterogeneity. A value greater than 50% will be considered to have substantial statistical heterogeneity. If we find substantial heterogeneity, we will attempt to determine potential reasons by conducting meta-regression if covariate information (e.g., age, sex, and duration of therapy) is available.

For sparse data meta-analysis we will employ the Peto Odds ratio method when event rates are less than 1 percent. When between event rates are between 5-10%, substantial differences between the N of two arms, or when effect size is large, dichotomous data will be meta-analyzed using the Mantel-Haenszel method without continuity correction. Dichotomous data with zero values in both arms will not be included in meta-analyses.

For studies reporting on test performance outcomes, statistical analyses will be conducted using methods currently recommended for use in Comparative Effectiveness Reviews of diagnostic tests.[38-39] The results of the diagnostic tests will be tabulated and sensitivity, specificity, positive and negative predictive values, positive and negative likelihood ratios and diagnostic odds ratios (DORs) calculated for the diagnosis of Renal Cell Cancer.

All meta-analyses will be conducted using STATA (College Station, TX).

- F. Grading the Strength of Evidence (SOE) for Major Comparisons and Outcomes:** At the completion of our review, two reviewers will independently grade the strength of evidence on key outcomes, including oncologic efficacy, renal functional outcomes, quality of life, and overall survival by adapting an evidence grading scheme recommended by the Methods Guide for Conducting Comparative Effectiveness Reviews. Conflicts will be resolved through consensus or third-party adjudication

We will consider the five required domains: study limitations, directness, consistency, precision, and reporting bias of the evidence body. Additional domains (plausible confounding, dose-response, and magnitude of effect) will be considered if applicable.

We will grade the strength of evidence addressing management studies using the evidence grading scheme recommended in the Methods Guide.[40]

For the diagnostic studies, we will use the grading method recommended in the AHRQ Methods Guide for Diagnostic Tests to rate each domain (Table 6).[41]

<b>Required domain</b>	<b>Definition and Elements</b>	<b>Application to Evaluation of Diagnostic Test Performance</b>
Risk of Bias	<p>Risk of bias is the degree to which the included studies for a given outcome or comparison have a high likelihood of adequate protection against bias (i.e., good internal validity), assessed through main elements:</p> <ul style="list-style-type: none"> <li>• Study design (e.g., RCTs or observational studies)</li> <li>• Aggregate quality of the studies under consideration from the rating of quality (good/fair/poor) done for individual studies</li> </ul>	<p>Use one of three levels of aggregate risk of bias:</p> <ul style="list-style-type: none"> <li>• Low risk of bias</li> <li>• Medium risk of bias</li> <li>• High risk of bias</li> </ul> <p>Well designed and executed studies of new tests compared against an adequate criterion standard are rated as “Low risk of bias.”</p>
Consistency	<p>Consistency is the degree to which reported study results (e.g., sensitivity, specificity, likelihood ratios) from included studies are similar. Consistency can be assessed through two main elements:</p> <ul style="list-style-type: none"> <li>• The range of study results is narrow.</li> <li>• Variability in study results is explained by differences in study design, patient population or test variability.</li> </ul>	<p>Use one of three levels of consistency:</p> <ul style="list-style-type: none"> <li>• Consistent (i.e., no inconsistency)</li> <li>• Inconsistent</li> <li>• Unknown or not applicable (e.g., single study) Single-study evidence bases should be considered as “consistency unknown (single study).”</li> </ul>
Directness	<p>Directness relates to whether the evidence links the interventions directly to outcomes. For a comparison of two diagnostic tests, directness implies head-to-head comparisons against a common criterion standard. Directness may be contingent on the outcomes of interest.</p>	<p>Score dichotomously as one of two levels of directness:</p> <ul style="list-style-type: none"> <li>• Direct</li> <li>• Indirect</li> </ul> <p>When assessing the directness of the overarching question, if there are no studies linking the test to a clinical outcome, then evidence that only provides diagnostic accuracy outcomes would be considered indirect. If indirect, specific which of the two types of indirectness account for the rating (or both, if this is the case); namely, use of intermediate/surrogate outcomes rather than health outcomes, and use of indirect comparisons. If the decision is made to grade the strength of evidence of an intermediate outcome such as diagnostic accuracy, then the reviewer does not need to automatically “downgrade” this outcome for being indirect.</p>
Precision	<p>Precision is the degree of certainty surrounding an effect estimate with respect to a given outcome (i.e., for each outcome separately). If a meta-analysis was performed, the degree of certainty will be the confidence interval around the summary measure(s) of test performance (e.g., sensitivity, or true positive).</p>	<p>Score dichotomously as one of two levels of precision:</p> <ul style="list-style-type: none"> <li>• Precise</li> <li>• Imprecise</li> </ul> <p>A precise estimate is an estimate that would allow a clinically useful conclusion. An imprecise estimate is one for which the confidence interval is wide enough to include clinically distinct conclusions.</p>
Publications bias	<p>Publication bias indicates that studies may have been published selectively, with the result that the estimate of test performance based on published studies does not reflect the true effect. Methods to detect publication bias for medical test studies are not robust. Evidence from small studies of new tests or asymmetry in funnel plots should raise suspicion for publication bias.</p>	<p>Publication bias can influence ratings of consistency, precision, and magnitude of effect – and, to a lesser degree, risk of bias and directness). Reviewers should comment on publication bias when circumstances suggest that relevant empirical findings, particularly negative or no-difference findings, have not been published or are unavailable.</p>

The evidence pertaining to the KQs classify into four basic categories: 1) “high” grade; 2) “moderate” grade; 3) “low” grade; and 4) “insufficient” grade. Table 7 defines each strength of evidence grade.

**Table 7. Strength of evidence grades and definitions**

Grade	Definition
High	We are very confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has few or no deficiencies. We believe that the findings are stable.
Moderate	We are moderately confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.
Low	We have limited confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
Insufficient	We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome. No evidence is available or the body of evidence has unacceptable deficiencies, precluding judgment.

**G. Assessing Applicability:** We will consider elements of the PICOTS framework when evaluating the applicability of evidence to answer our key questions as recommended in the Methods Guide for Comparative Effectiveness Reviews of Interventions. We will consider how important population characteristics (e.g. gender, race, ethnicity), comorbidities (e.g. cardiovascular disease, chronic kidney cancer), and intervention features (e.g. co-intervention) may cause heterogeneity of treatment effects and affect generalizability of the findings.

## V. References

1. Kutikov, A., et al., Incidence of benign pathologic findings at partial nephrectomy for solitary renal mass presumed to be renal cell carcinoma on preoperative imaging. *Urology*, 2006. **68**(4): p. 737-40.
2. Thompson, R.H., et al., Metastatic renal cell carcinoma risk according to tumor size. *J Urol*, 2009. **182**(1): p. 41-5.
3. Lane, B.R., et al., A preoperative prognostic nomogram for solid enhancing renal tumors 7 cm or less amenable to partial nephrectomy. *J Urol*, 2007. **178**(2): p. 429-34.
4. Cancer Facts & Figures 2014, A.C. Society, Editor 2014, American Cancer Society: Atlanta, GA.
5. Mathew, A., et al., Global increases in kidney cancer incidence, 1973-1992. *Eur J Cancer Prev*, 2002. **11**(2): p. 171-8.
6. Nguyen, M.M., I.S. Gill, and L.M. Ellison, The evolving presentation of renal carcinoma in the United States: trends from the Surveillance, Epidemiology, and End Results program. *J Urol*, 2006. **176**(6 Pt 1): p. 2397-400; discussion 2400.
7. Kane, C.J., et al., Renal cell cancer stage migration: analysis of the National Cancer Data Base. *Cancer*, 2008. **113**(1): p. 78-83.

8. Lipworth, L., R.E. Tarone, and J.K. McLaughlin, The epidemiology of renal cell carcinoma. *J Urol*, 2006. **176**(6 Pt 1): p. 2353-8.
9. Ficarra, V., et al., Multiinstitutional European validation of the 2002 TNM staging system in conventional and papillary localized renal cell carcinoma. *Cancer*, 2005. **104**(5): p. 968-74.
10. Frank, I., et al., Independent validation of the 2002 American Joint Committee on cancer primary tumor classification for renal cell carcinoma using a large, single institution cohort. *J Urol*, 2005. **173**(6): p. 1889-92.
11. Pierorazio, P.M., et al., Laparoscopic radical nephrectomy for large renal masses: critical assessment of perioperative and oncologic outcomes of stage T2a and T2b tumors. *Urology*, 2012. **79**(3): p. 570-5.
12. Haggstrom, C., et al., Metabolic factors associated with risk of renal cell carcinoma. *PLoS One*, 2013. **8**(2): p. e57475.
13. Kanao, K., et al., Preoperative prognostic nomogram (probability table) for renal cell carcinoma based on TNM classification. *J Urol*, 2009. **181**(2): p. 480-5; discussion 485.
14. Long, J.A., et al., External validation of the RENAL nephrometry score in renal tumours treated by partial nephrectomy. *BJU Int*, 2013. **111**(2): p. 233-9.
15. Yacyioglu, O., et al., A preoperative prognostic model predicting recurrence-free survival for patients with kidney cancer. *Jpn J Clin Oncol*, 2013. **43**(1): p. 63-8.
16. Kutikov, A., et al., Anatomic features of enhancing renal masses predict malignant and high-grade pathology: a preoperative nomogram using the RENAL Nephrometry score. *Eur Urol*, 2011. **60**(2): p. 241-8.
17. Cindolo, L., et al., A preoperative clinical prognostic model for non-metastatic renal cell carcinoma. *BJU Int*, 2003. **92**(9): p. 901-5.
18. Karakiewicz, P.I., et al., A preoperative prognostic model for patients treated with nephrectomy for renal cell carcinoma. *Eur Urol*, 2009. **55**(2): p. 287-95.
19. Snyder, M.E., et al., Incidence of benign lesions for clinically localized renal masses smaller than 7 cm in radiological diameter: influence of sex. *J Urol*, 2006. **176**(6 Pt 1): p. 2391-5; discussion 2395-6.
20. Pierorazio, P.M., et al., Multiphasic enhancement patterns of small renal masses ( $\leq 4$  cm) on preoperative computed tomography: utility for distinguishing subtypes of renal cell carcinoma, angiomyolipoma, and oncocytoma. *Urology*, 2013. **81**(6): p. 1265-71.
21. Young, J.R., et al., Clear cell renal cell carcinoma: discrimination from other renal cell carcinoma subtypes and oncocytoma at multiphasic multidetector CT. *Radiology*, 2013. **267**(2): p. 444-53.
22. Mullins, J.K., et al., Tumor complexity predicts malignant disease for small renal masses. *J Urol*, 2012. **188**(6): p. 2072-6.
23. Leveridge, M.J., et al., Outcomes of small renal mass needle core biopsy, nondiagnostic percutaneous biopsy, and the role of repeat biopsy. *Eur Urol*, 2011. **60**(3): p. 578-84.

24. Campbell, S.C., et al., Guideline for management of the clinical T1 renal mass. *J Urol*, 2009. **182**(4): p. 1271-9.
25. NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines): Kidney Cancer, 2014, National Comprehensive Cancer Network (NCCN).
26. Ljungberg, B., et al., EAU guidelines on renal cell carcinoma: the 2010 update. *Eur Urol*, 2010. **58**(3): p. 398-406.
27. Van Poppel, H., et al., A prospective, randomised EORTC intergroup phase 3 study comparing the oncologic outcome of elective nephron-sparing surgery and radical nephrectomy for low-stage renal cell carcinoma. *Eur Urol*, 2011. **59**(4): p. 543-52.
28. Aziz, A., et al., Do young patients with renal cell carcinoma feature a distinct outcome after surgery? A comparative analysis of patient age based on the multinational CORONA database. *J Urol*, 2014. **191**(2): p. 310-5.
29. Kim, E.H., et al., Comparison of laparoscopic and percutaneous cryoablation for treatment of renal masses. *Urology*, 2014. **83**(5): p. 1081-7.
30. Huang, W.C., et al., Chronic kidney disease after nephrectomy in patients with renal cortical tumours: a retrospective cohort study. *Lancet Oncol*, 2006. **7**(9): p. 735-40.
31. Kim, E.H., et al., Percutaneous cryoablation of renal masses: Washington University experience of treating 129 tumours. *BJU Int*, 2013. **111**(6): p. 872-9.
32. National Guideline, C. ACR Appropriateness Criteria®; indeterminate renal mass. 9/15/2014]; Available from: <http://www.guideline.gov/content.aspx?f=rss&id=48291&osrc=12>.
33. Demirjian, S., et al., Chronic Kidney Disease Due to Surgical Removal of Nephrons: Relative Rates of Progression and Survival. *J Urol*, 2014.
34. Clavien, P.A., et al., The Clavien-Dindo classification of surgical complications: five-year experience. *Ann Surg*, 2009. **250**(2): p. 187-96.
35. Higgins JPT, Green S (eds). *Cochrane handbook for systemic reviews of interventions* Version 5.1.0. The Cochrane Collaboration. 2011;Oxford, England. [cited 2013 11/4]; Available from: <http://handbook.cochrane.org>.
36. Sterne JAC, Higgins JPT, Reeves BC on behalf of the development group for ACROBATNRSI. A Cochrane Risk Of Bias Assessment Tool: for Non-Randomized Studies of Interventions (ACROBATNRSI), Version 1.0.0, 24 September 2014. Available from <http://www.riskofbias.info>
37. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011 Oct 18;155(8):529-36.
38. Trikalinos TA, Balion CM. Chapter 9: options for summarizing medical test performance in the absence of a "gold standard". *J Gen Intern Med*. 2012 Jun;27 Suppl 1:S67-75.
39. Trikalinos TA, Balion CM, Coleman CI, et al. Chapter 8: meta-analysis of test performance when there is a "gold standard". *J Gen Intern Med*. 2012 Jun; 27 Suppl 1:S56-66.

40. Agency for Healthcare Research and Quality. Grading the Strength of a Body of Evidence When Assessing Health Care Interventions - AHRQ and the Effective Health Care Program: An Update. Rockville, MD: Agency for Healthcare Research and Quality; 2013; Available from:  
[http://effectivehealthcare.ahrq.gov/ehc/products/457/1163/GradingTheStrengthofEvidence\\_DraftMethodsChapter\\_20120625.pdf](http://effectivehealthcare.ahrq.gov/ehc/products/457/1163/GradingTheStrengthofEvidence_DraftMethodsChapter_20120625.pdf).
41. Singh S, Chang SM, Matchar DB, et al. Grading a Body of Evidence on Diagnostic Tests. In: Chang SM, Matchar DB, Smetana GW, et al, eds. *Methods Guide for Medical Test Reviews*. Rockville (MD): Agency for Healthcare Research and Quality (US); 2012 Jun. Chapter 7.

## VI. Definition of Terms

**Local recurrence** - Local recurrence is defined as any persistent or recurrent disease present in the treated kidney or associated renal fossa after a single, curative-intent initial treatment. For partial nephrectomy and ablation, local recurrence includes persistent enhancement of any treated mass, any new or visually enlarging neoplasm, new nodularity, failure of regression in size of the treated lesion(s), or new satellite or port site lesions.

## VII. Summary of Protocol Amendments

Date	Section	Original Protocol	Revised Protocol	Rationale
March 2 <sup>nd</sup> , 2015	A. Criteria for Inclusion/Exclusion of Studies in the Review –Type of study for management questions (KQ3a & 3b)	We will include randomized controlled trials, non-randomized controlled trials, and cohort studies.	<p>For the management section (KQ 3) of this review, we will include controlled studies and uncontrolled studies using the following criteria:</p> <p>a) Controlled studies (RCT, nRCT and comparative cohort studies), we will include studies with any of the following comparisons:</p> <p><b>-Surgical vs. surgical intervention strategy</b>, that is, either of radical nephrectomy, partial nephrectomy or ablative technique.</p> <p><b>-Surgical intervention vs. conservative management strategy</b>, that is, a comparison between any of above mentioned surgical intervention strategies and a conservative management strategy (i.e. active surveillance without surgical intervention).</p> <p>b) Uncontrolled studies (single cohort studies): we will only include uncontrolled studies that look at active surveillance and meet the following inclusion criteria:            Number of patients: 100 or more            Follow-up: 12 months or more</p> <p>Every other uncontrolled study will get listed in the appendix with the following data:</p> <p>Author, publication year            Study design            Intervention name            Number of patients            Follow up            List of outcomes</p>	<p>After elimination of articles captured by the search at title-abstract level, article level, we were left with over 250 articles eligible for data abstraction. After going through and categorizing these studies by study design, we think we have enough controlled studies (90-100) to address KQ3a and 3b.</p> <p>There may not be any controlled study for active surveillance. It's not a surgical treatment. This approach is most often used to identify patients who might safely avoid or defer surgery. The existing literature on active surveillance consists solely of uncontrolled studies</p>

(NOTE THE FOLLOWING PROTOCOL ELEMENTS ARE STANDARD SECTIONS TO BE ADDED TO ALL PROTOCOLS)

### **VIII. Review of Key Questions**

AHRQ posted the key questions on the Effective Health Care Website for public comment. The EPC refined and finalized the key questions after review of the public comments, and input from Key Informants and the Technical Expert Panel (TEP). This input is intended to ensure that the key questions are specific and relevant.

### **IX. Key Informants**

Key Informants are the end users of research, including patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others with experience in making health care decisions. Within the EPC program, the Key Informant role is to provide input into identifying the Key Questions for research that will inform healthcare decisions. The EPC solicits input from Key Informants when developing questions for systematic review or when identifying high priority research gaps and needed new research. Key Informants are not involved in analyzing the evidence or writing the report and have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Key Informants must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals are invited to serve as Key Informants and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

### **X. Technical Experts**

Technical Experts constitute a multi-disciplinary group of clinical, content, and methodological experts who provide input in defining populations, interventions, comparisons, or outcomes and identify particular studies or databases to search. They are selected to provide broad expertise and perspectives specific to the topic under development. Divergent and conflicting opinions are common and perceived as health scientific discourse that results in a thoughtful, relevant systematic review. Therefore study questions, design, and methodological approaches do not necessarily represent the views of individual technical and content experts. Technical Experts provide information to the EPC to identify literature search strategies and recommend approaches to specific issues as requested by the EPC. Technical Experts do not do analysis of any kind nor do they contribute to the writing of the report. They have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Technical Experts must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals are invited to serve as Technical Experts and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

### **XI. Peer Reviewers**

Peer reviewers are invited to provide written comments on the draft report based on their clinical, content, or methodological expertise. The EPC considers all peer review comments on the draft report in preparation of the final report. Peer reviewers do not participate in writing or editing of the final report or other products. The final report does not necessarily represent the views of individual reviewers. The EPC will complete a disposition of all peer review comments. The disposition of comments for systematic reviews and technical briefs will be published three months after the publication of the evidence report.

Potential Peer Reviewers must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Invited Peer Reviewers may not have any financial conflict of interest greater than \$10,000. Peer reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through the public comment mechanism.

## **XII. EPC Team Disclosures**

EPC core team members must disclose any financial conflicts of interest greater than \$1,000 and any other relevant business or professional conflicts of interest. Related financial conflicts of interest that cumulatively total greater than \$1,000 will usually disqualify EPC core team investigators.

## **XIII. Role of the Funder**

This project was funded under Contract No. XXXX from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The Task Order Officer reviewed contract deliverables for adherence to contract requirements and quality. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

**Contacts for Scientific Information Packets (SIPs)**

<b>Device</b>	<b>Name</b>	<b>Manufacturer</b>
Radio Frequency Ablation	RF 3000 ablation system	Boston Scientific
	MicroThermX®	BSD Medical
	Cool-tip RF Ablation System E Series	Covidien
	Evident™ MWA system	Covidien
	HS Amica Microwave System	HS Amica
	Radiofrequency ablation catheter / bidirectional	Medtronic
	MedWaves Ablation system	MedWaves
	Microsulis's Aculis pMTA (percutaneous Microwave Tissue Ablation System)	Microsulis
	Certus 140	NeuWave Medical
	Talon RFA device	StarBurst
	Semi-Flex RFA Device	StarBurst
	XL RFA Device	StarBurst
	Xli-enhanced RFA Device	StarBurst
	Xli-enhanced Semi-Flex RFA Device	StarBurst
	MRI RFA Device	StarBurst
	SDE RFA Device	StarBurst
	RFA Electrode	UniBlate
VivaWave™ Microwave Ablation system	VivaWave	
Cryoablation	Endocare® Cryocare® Systems	HealthTronics
	Visual-ICE System	Galil Medical
Robotic	Da Vinci Surgical System	Intuitive Surgical