



Effective Health Care Program

Management of Renal Masses and Localized Renal Cancer

Executive Summary

Background

Renal masses are a biologically heterogeneous group of tumors ranging from benign masses to cancers that can be indolent or aggressive.^{1,2} The true incidence of renal masses (including benign lesions) is unknown, but benign lesions comprise approximately 20 percent of surgically resected tumors.^{1,3}

Kidney cancer affects approximately 65,000 new patients each year, with more than 13,000 deaths annually.⁴ The incidence of kidney cancer has increased significantly by 2-3 percent per year over the past few decades – presumably due to the increased use of cross-sectional imaging such as computed tomography.⁵ Tumors are often discovered incidentally and are asymptomatic at presentation. The greatest increase in incidence has been noted in small (less than 4 cm), clinically localized tumors (within the kidney with no evidence of local spread, lymph node involvement, or distant metastases), which now account for upwards of 40 percent of all kidney cancers.^{6,7}

Renal cell carcinoma is the most common type of cancer affecting the kidneys in the United States accounting for more than

Effective Health Care Program

The Effective Health Care Program was initiated in 2005 to provide valid evidence about the comparative effectiveness of different medical interventions. The object is to help consumers, health care providers, and others in making informed choices among treatment alternatives. Through its Comparative Effectiveness Reviews, the program supports systematic appraisals of existing scientific evidence regarding treatments for high-priority health conditions. It also promotes and generates new scientific evidence by identifying gaps in existing scientific evidence and supporting new research. The program puts special emphasis on translating findings into a variety of useful formats for different stakeholders, including consumers.

The full report and this summary are available at www.effectivehealthcare.ahrq.gov/reports/final.cfm.

94 percent of kidney malignancies.⁴ While renal cell carcinoma only represents two percent of adult cancers, it is amongst the most lethal; approximately 35 percent of patients die within 5 years of diagnosis.⁴

Archived: This report is greater than 3 years old. Findings may be used for research purposes, but should not be considered current.



Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

Effective
Health Care

However, the cancer-specific survival is highly stage dependent, with a greater than 95 percent 5-year disease specific survival for stage T1 tumors, and greater than 85 percent 5-year disease specific survival for stage T2 tumors. The deaths due to renal cell carcinoma are driven by the failure of systemic treatments in metastatic (later stage) patients and up to 40 percent of clinically localized tumors are determined to be locally advanced cancers (stage T3, with invasion of perinephric fat or venous structures) at pathological examination.

Diagnostic Evaluation and Detection of Disease

All solid renal masses and cystic lesions with solid components are suspicious for renal cell carcinoma. 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 renal masses and indolent versus aggressive renal cancers. Demographic, clinical, and imaging characteristics are used to risk-stratify patients, and nomograms exist that combine these characteristics into composite models to predict the malignant potential of tumors preoperatively.⁸⁻¹¹

Percutaneous renal mass sampling may be offered as a diagnostic adjunct to imaging studies such as computed tomography, magnetic resonance imaging, or ultrasonography. Percutaneous renal mass sampling can be performed by fine needle aspiration with a reading of the sample by a cytopathologist or via core biopsy with a reading by a surgical pathologist.

Therapeutic Interventions and Outcomes

Several options exist for the management of clinically localized renal masses suspicious for renal cell carcinoma including active surveillance, thermal ablation, and surgery (partial or radical nephrectomy). Given the increased incidence in early, low-stage tumors without improvement in cancer-related deaths, active surveillance has emerged as an option for patients with small renal masses, a low likelihood of aggressive malignancy, procedure limiting comorbidity, and/or a limited life expectancy. It is important to note a difference between active surveillance with curative intent versus watchful waiting. The latter constitutes a strategy where treatment is never entertained and surveillance imaging is infrequent or does not occur at all. Studies of watchful waiting are not examined in this report. Surgery includes partial nephrectomy or radical nephrectomy, which can be performed through a minimally invasive or open approach. Minimally invasive options include both standard laparoscopy and robot-assisted

laparoscopy. Surgical removal (either radical or partial nephrectomy) is the gold standard for the treatment of renal cell carcinoma. The American Urological Association (AUA) Guideline, which only considers clinical stage 1 renal masses, considers partial nephrectomy and radical nephrectomy as “standard” treatment modalities for clinical stage T1a tumors (≤ 4 cm in diameter) and T1b (4-7cm) tumors. Thermal ablation and active surveillance are considered “options” or “recommendations” for T1a tumors, but are only considered “options” (no longer a “recommendation”) for T1b tumors.¹² Thermal ablation, which may include cryoablation or radiofrequency ablation, can either be performed laparoscopically or percutaneously. While most urologists would consider radical nephrectomy as the standard treatment for clinical stage 2 renal masses, there are no professional organization or guideline standards for the management of clinically localized, stage 2 tumors.

Scope and Key Questions

We conducted a systematic review of the effectiveness and comparative effectiveness of different strategies for treating patients with a renal mass suspicious for renal cell carcinoma. We developed analytic frameworks to illustrate the questions and outcomes we considered (Figures A and B), and we sought to address the following Key Questions (KQ):

KQ 1: In patients that undergo surgery for a renal mass that is suspicious for stage I or II renal cell carcinoma, how does the pathologic diagnosis compare to the likelihood of malignancy predicted by using a preoperative 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 RCC includes all solid renal masses and cystic renal masses with a solid component.

KQ 2a: In patients who 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 (using fine needle aspiration with cytopathology or core biopsy with surgical pathology) in establishing a diagnosis (e.g., malignancy, histology, and grade)?

KQ 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 KQ2a) to estimate the risk of malignancy, including direct

complications (e.g., pain, infection, hemorrhage, and radiation exposure) and harms related to false positives, false negatives, or nondiagnostic results?

KQ 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?

KQ 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?

Methods

With input from key informants, we refined the questions, including eligibility criteria, and developed a protocol (PROSPERO registration CRD42015015878).

We searched MEDLINE®, Embase®, and the Cochrane Central Register of Controlled Trials (CENTRAL) from January 1, 1997, (the year the TNM Classification of Malignant Tumor staging system for renal cell carcinoma was modified and the distinctions of T1a/T1b and T2a/T2b were created) through May 1, 2015. We also requested information from device manufacturers and searched Clinicaltrials.gov.

Citations were screened independently by two reviewers using predefined eligibility criteria (see Table A). One reviewer completed data abstraction and a second reviewer checked abstraction for accuracy. Two reviewers independently assessed risk of bias for individual studies. We used the Cochrane Collaboration's tool for assessing the risk of bias of randomized controlled trials (RCTs).¹³ For nonrandomized studies of treatment interventions, we used the Cochrane Risk of Bias Assessment Tool for Non-Randomized Studies of Interventions (ACROBAT-NRSI).¹⁴ For diagnostic studies, we used the quality assessment tool for diagnostic accuracy studies (QUADAS -2).¹⁵ Differences between reviewers were resolved through consensus.

We conducted meta-analyses for an outcome when there were sufficient data and studies were sufficiently homogenous with respect to key variables (population characteristics, intervention, and outcome measurement) using a random effects model with the DerSimonian and Laird method. We identified substantial statistical heterogeneity as an I-squared statistic with a value greater than 50 percent. All meta-analyses were conducted using STATA 12.1 (College Station, TX).

We graded the strength of evidence using the scheme recommended by the "Methods Guide for Effectiveness and Comparative Effectiveness Reviews."¹⁶

Results

Figure C summarizes the results of our searching for relevant studies. This review focuses on 147 studies, reported in 150 articles that met the inclusion criteria.

KQ 1: Pathologic Diagnosis Compared With Likelihood of Malignancy Based on Preoperative Composite Profile of Patient and Tumor Characteristics

Twenty studies (12,149 patients) evaluated composite models to predict pathologic diagnosis, adjusting for imaging characteristics, demographic characteristics, clinical characteristics, and other diagnostic tests (i.e., blood and urine). This body of evidence included 2 prospective studies and 18 retrospective observational studies that ranged in sample size from 84 to 1,726 patients. Nineteen of 20 studies used imaging characteristics while only one evaluated laboratory testing. The overall risk of bias for these studies was low (Table B).

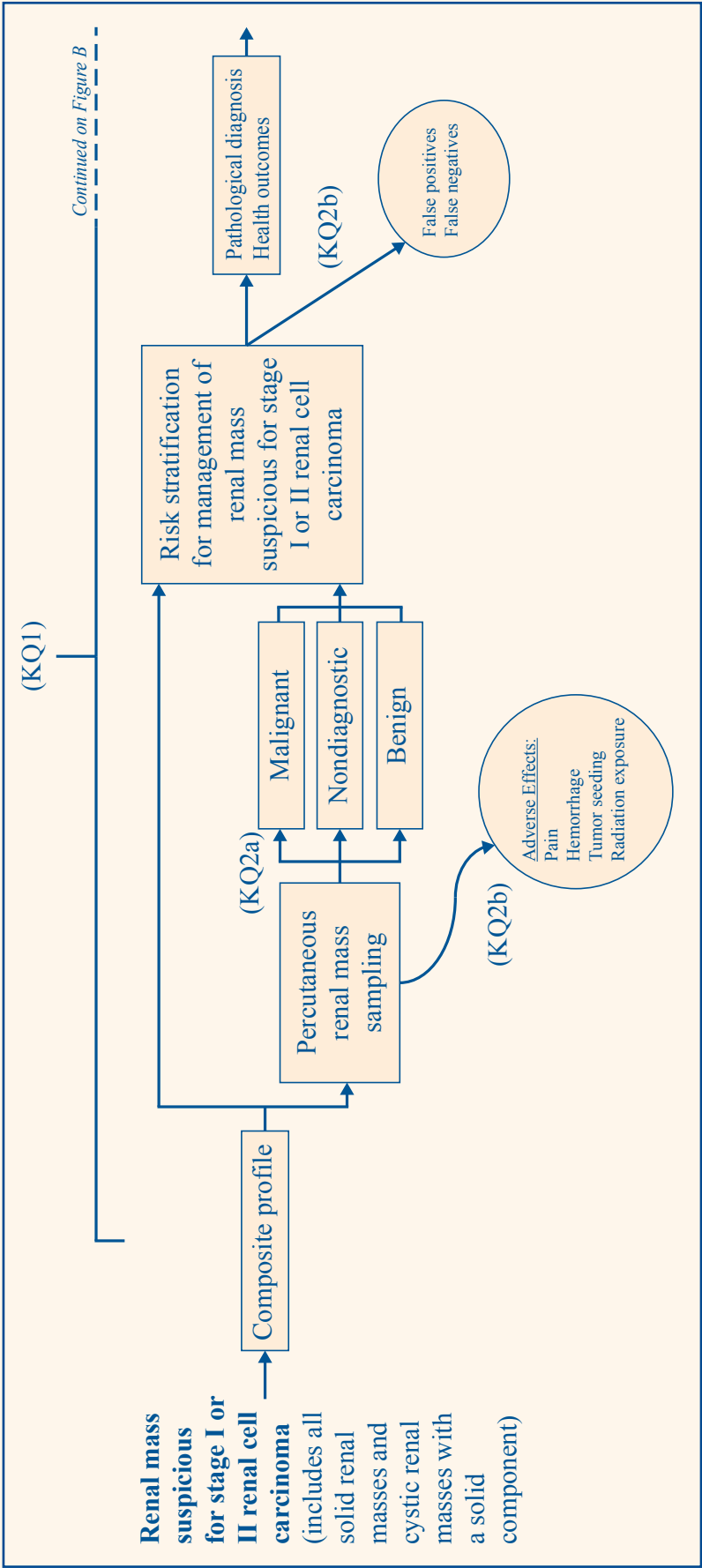
The most common variables included in composite profiles were tumor size, age, sex, body mass index (BMI), and incidental presentation. Increased tumor size was consistently predictive of malignant pathology in studies that evaluated tumor size as a categorical variable and in studies evaluating size as a continuous variable (effect size in meta-analysis: 1.3 times increased risk of malignancy per cm increase in tumor size; 95% confidence interval (CI): 1.22 to 1.43) with moderate strength of evidence. Additionally, 14 of 16 studies and subsequent meta-analysis demonstrated that male sex predicted malignant pathology (effect size: 2.70 times increased risk of malignancy with male sex; 95% CI: 2.39 to 3.02) with moderate strength of evidence. The strength of evidence was moderate that incidental presentation was not predictive of pathology, and the strength of evidence was low that age was not predictive of pathology. The evidence was insufficient on BMI.

KQs 2a and 2b: Accuracy and Harms of Percutaneous Renal Mass Sampling

Twenty studies (2,979 patients) evaluated the performance characteristics of percutaneous renal mass sampling, of which 16 evaluated harms. Only one study evaluated fine needle aspiration with cytopathology; all other studies evaluated core biopsy with surgical pathology. Four

Figure A. Analytic framework for systematic review of the management of renal masses and localized kidney cancer

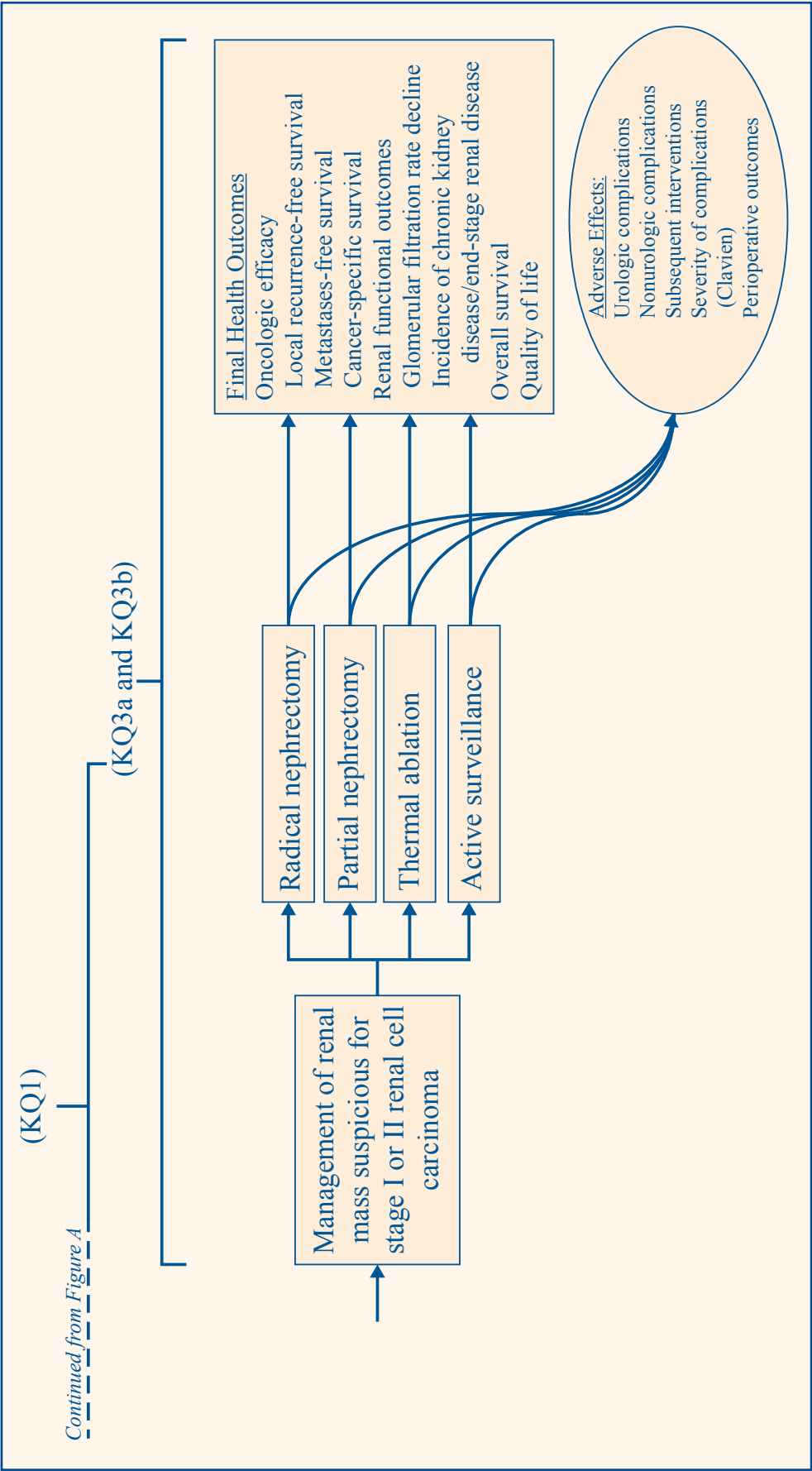
Part I: Diagnostic framework



KQ = Key Question

Figure B. Analytic framework for systematic review of the management of renal masses and localized kidney cancer

PART II: Management strategies



KQ = Key Question

Table A. PICOTS (population, interventions, comparators, outcomes, timing, and setting) for the KQs

PICOTS	Diagnostic (KQs 1, 2a, and 2b)	Management (KQs 3a and 3b)
Population(s)	Newly diagnosed adults (18 years or older) 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 (less than 7 cm and organ confined) or T2 (greater than 7 cm and organ confined) renal masses	
Interventions	<ul style="list-style-type: none"> • Percutaneous renal mass sampling (fine needle aspiration or biopsy) • Composite models (e.g., 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: computed tomography, ultrasonography, magnetic resonance imaging 	<ul style="list-style-type: none"> • Radical nephrectomy (open and minimally invasive) • Partial nephrectomy (open and minimally invasive) • Thermal ablation (e.g., radiofrequency ablation, cryoablation; surgical versus image-guided) • Active surveillance • Minimally invasive surgery may refer to standard laparoscopy or robot-assisted laparoscopy • No microwave ablation
Comparators	Comparisons are between biopsy results, composite models, and pathologic diagnosis after surgical intervention	Comparisons include all of the management options listed above
Outcomes	<p>Diagnostic test-related Outcomes</p> <ul style="list-style-type: none"> • False positives • False negatives • Radiation exposure <p>Adverse effects of percutaneous renal mass sampling</p> <ul style="list-style-type: none"> • Pain • Hemorrhage • Tumor seeding 	<p>Final health outcomes</p> <ul style="list-style-type: none"> • Oncologic efficacy <ul style="list-style-type: none"> – Local recurrence-free survival – Metastasis-free survival – Cancer-specific survival • Renal functional outcomes <ul style="list-style-type: none"> – Glomerular filtration rate decline – Incidence of chronic kidney disease – Incidence of end-stage renal disease • Overall survival • Quality of life

Table A. PICOTS (population, interventions, comparators, outcomes, timing, and setting) for the KQs (continued)

PICOTS	Diagnostic (KQs 1, 2a, and 2b)	Management (KQs 3a and 3b)
Outcomes (continued)		Adverse effects of management strategies <ul style="list-style-type: none"> • Urologic complications <ul style="list-style-type: none"> – Acute kidney Injury – Hemorrhage – Urine leak – Hematuria – Loss of kidney – Ureteral injury (any injury of collecting system and ureter) – Urinary tract infection • Nonurologic complications (by organ system) <ul style="list-style-type: none"> – Hematologic (thromboembolic) – Gastrointestinal – Cardiovascular – Respiratory – Neurologic – Wound complications (e.g. hernia and dehiscence) – Infectious disease – Listed by severity of complications (using the Clavien Grading System if available): – Minor versus major <ul style="list-style-type: none"> o Minor (Clavien 1-2)^a: conservative management or medications only o Major (Clavien 3-4)^b: requiring intervention, resulting in permanent disability or death – Need for subsequent interventions: embolization, drain placement, stent placement, etc. • Perioperative outcomes <ul style="list-style-type: none"> – Blood loss (cc or mL) – Blood transfusion (yes or no) – Conversion to open surgery (%) – Conversion to radical nephrectomy (%) – Length of stay (days)

Table A. PICOTS (population, interventions, comparators, outcomes, timing, and setting) for the KQs (continued)

PICOTS	Diagnostic (KQs 1, 2a, and 2b)	Management (KQs 3a and 3b)
Type of study	Any study design except case report	<p>Controlled studies (randomized controlled trials, nonrandomized controlled trials, and comparative cohort studies): All comparisons between interventions</p> <p>Uncontrolled studies (single cohort studies): Data from uncontrolled studies that addressed active surveillance are described in the report. Every other uncontrolled study that addressed KQ 3 is listed in the appendix with the following data:</p> <ul style="list-style-type: none"> • Author, publication year • Study design • Intervention name • Number of patients • Followup • List of outcomes
Timing and Setting	Any time point and setting	

KQ= Key Question

Clavien-Dindo system currently used for reporting of complications related to urologic surgical interventions:

^aGrade I: Any deviation from the normal postoperative course without the need for pharmacological treatment or surgical, endoscopic and radiological interventions. Allowed therapeutic regimens are: drugs as antiemetics, antipyretics, analgetics, diuretics and electrolytes and physiotherapy. This grade also includes wound infections opened at the bedside.

^aGrade II: Requiring pharmacological treatment with drugs other than such allowed for grade I complications. Blood transfusions and total parenteral nutrition are also included.

^bGrade III: Requiring surgical, endoscopic or radiological intervention.

1.(a) not under general anesthesia

2.(b) under general anesthesia

^bGrade IV: Life-threatening complication (including CNS complications) requiring IC/CU-management.

1.(a) single organ dysfunction (including dialysis)

2.(b) multi-organ dysfunction

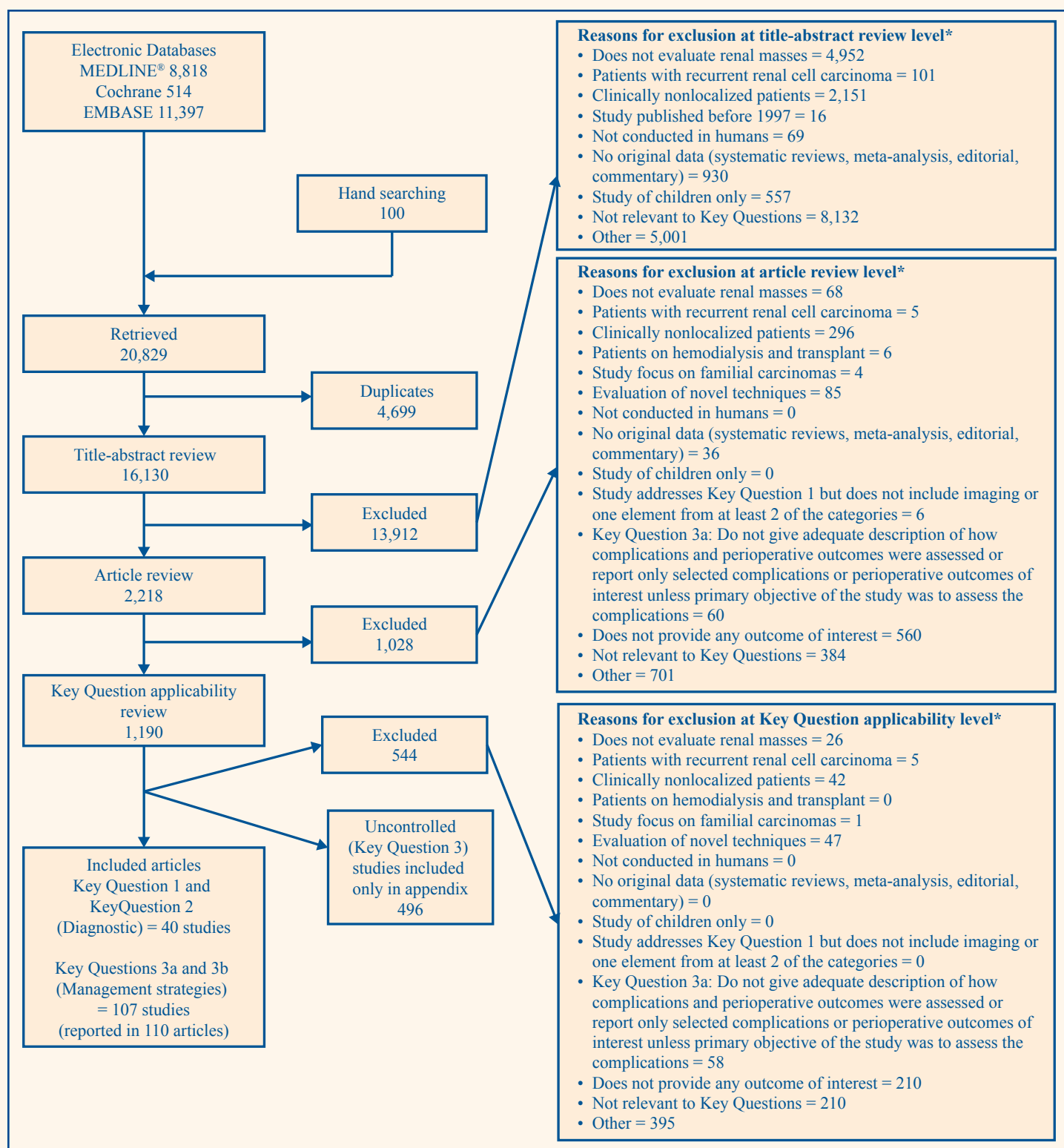
studies were of prospective cohorts while the remainder were retrospective studies; all studies were single center experiences. Risk of bias was low in 5 studies and high in the remaining 15 studies based on the potential risk of bias due to missing reference standard evaluations (surgical pathology) among patients with benign biopsy results (Table C).

Only one study of fine needle aspiration met the inclusion criteria in this review and revealed the following performance characteristics (sensitivity 62.5 percent, specificity not able to be calculated, positive predictive value 100 percent). In comparison, core biopsy revealed better diagnostic abilities: sensitivity of 97.5 percent, specificity of 96.2 percent, positive predictive value of 99.8 percent (0.21 percent of malignant biopsies were false positives), false positive rate 4.0 percent, and negative predictive value of 68.5 percent, but 14 percent of biopsies were non-diagnostic. The majority of nondiagnostic

biopsies were found to correspond with malignant surgical pathology (90.4 percent). Verification bias exists in these studies as benign or nondiagnostic biopsies do not necessarily proceed to surgical extirpation, limiting the analysis and making the exact false negative rate difficult to ascertain. In addition, there is bias in who proceeds to surgery as patient or tumor characteristics (i.e., male sex, larger tumors) influence the decision to proceed to surgery. Therefore, the strength of evidence for diagnostic accuracy of renal mass sampling (core biopsy) was graded as moderate. It is more difficult to make conclusions on final needle aspiration given only one older study met inclusion criteria.

Percutaneous renal mass sampling was associated with infrequent direct complications, including hematoma (4.9 percent), clinically significant pain (1.2 percent), gross hematuria (1.0 percent), pneumothorax (0.6 percent), and hemorrhage (0.4 percent). The strength of evidence was

Figure C. Summary of literature search



* Reviewers were allowed to mark more than one reason for exclusion.

Table B. Summary of the strength of evidence for individual predictors of malignant or benign pathology

Outcomes	No. Studies	Strength of Evidence Finding
Tumor Size	12	Moderate Increasing tumor size consistently is associated with an increased risk of malignancy.
Tumor Characteristics	9	Low Increasing RENAL nephrometry score is consistently associated with malignancy. The data regarding individual components of the RENAL nephrometry score and other tumor characteristics is insufficient to draw conclusions.
Age	15	Low While the relationship between age and malignant pathology varies among studies, the effect size due to age is small in all studies.
Sex	16	Moderate Women are more likely to have benign tumors in all studies. The effect size varied by inclusion criteria and other variables (i.e. age, tumor size).
Body Mass Index	5	Insufficient Conflicting and non-significant results in studies make it difficult to form meaningful conclusions. In addition, geographic and population-based differences in body mass index make interpretation of the association of body mass index with malignant disease difficult.
Incidental Presentation	5	Moderate All studies demonstrate no relationship between an incidental finding and malignant pathology.
Harms	12	Low A small, but notable, proportion of patients experience harms due to renal mass biopsy, with hematoma (5%) being the most common direct complication. Studies were inconsistent in which harms, if any, were reported.

Table C. Summary of the strength of evidence for renal mass biopsy outcomes

Outcomes	No. Studies	Strength of Evidence Finding
Diagnostic Accuracy	18	Moderate Renal mass biopsy has a high positive predictive value (99.8%) for the diagnosis of renal malignancy but also a notable non-diagnostic (~14%) rate and low negative predictive value (<70%). The primary limitation is the absence of surgical pathology for benign biopsies, but sensitivity and specificity of a diagnostic biopsy result appear to be over 90%.
Fuhrman Grade	12	Low Fuhrman upgrading on final pathology occurred in 20.5% of biopsies, but many studies did not provide data on grade concordance.
Harms	16	Low A small, but notable, proportion of patients experience harms due to renal mass biopsy, with hematoma (5%) being the most common direct complication. Studies were inconsistent in which harms, if any, were reported.

low on the harms associated with percutaneous renal mass sampling.

KQs 3a and 3b: Comparative Effectiveness and Harms of the Management Strategies for Clinically Localized Renal Masses

One hundred seven studies (reported in 110 articles) addressed KQs 3a and 3b. Ninety-nine comparative studies (reported in 102 articles, with 179,740 patients) addressed the effectiveness of management strategies for localized renal masses concerning for renal cell carcinoma. Only one study was an RCT (reported in 3 articles). Eight studies, evaluating active surveillance, were uncontrolled studies. The remainder were comparative cohort studies (Table D).

Overall Survival and Oncological Outcomes

Sixty studies (reported in 61 articles) evaluated oncological outcomes such as cancer-specific survival, metastasis-free survival, and local recurrence-free survival. This included one RCT, 48 institutional cohort studies, and 11 studies (reported in 12 articles) of the Surveillance, Epidemiology, and End Results (SEER) dataset. The risk of bias associated with the cohort studies was moderate in 34 (58 percent) studies and serious in 25 (42 percent) studies. Forty-eight studies (reported in 49 articles) evaluated overall survival, including one RCT, 38 institutional cohorts, and 9 studies (reported in 10 articles) of the SEER dataset. The risk of bias associated with cohort studies was moderate in 30 (63.8 percent) studies and serious in 17 (36.2 percent) studies. The single randomized study was determined to have an unclear risk of bias for both overall survival and oncological outcomes. Of note, few comparative studies evaluated active surveillance, necessitating evaluation of seven uncontrolled studies of active surveillance.

The available literature suggested that overall survival and oncological outcomes were similar among all management strategies. In fact, cancer-specific survival was excellent among all modalities, and median 5-year survival approached 95 percent for all included studies. Importantly, cancer-specific survival was associated with tumor size/stage, but not partial or radical nephrectomy (these were the only management strategies to offer stage-specific outcomes): for patients with clinical stage T1a (≤ 4 cm), T1b (> 4 -7 cm) and T2 (> 7 cm) tumors, resulting cancer-specific survival was 97-99 percent, 90-91 percent, and 83-87 percent, respectively. The strength of evidence was *moderate* for the finding of equivalent cancer-specific survival for radical versus partial nephrectomy based on data from one RCT,

23 institutional cohort studies, and 10 SEER analyses. The strength of evidence was *moderate* for the finding of equivalent cancer-specific survival for thermal ablation versus radical nephrectomy, and *low* for thermal ablation versus partial nephrectomy.

Overall 5-year survival was similar for patients undergoing partial nephrectomy when compared to radical nephrectomy (*low* strength of evidence). Thermal ablation was generally associated with similar or poorer overall 5-year survival compared with partial nephrectomy (*low* strength of evidence) due to the selection of older patients with greater comorbidity to undergo the procedure. Uncontrolled active surveillance studies reported a range of overall survival from 69 to 94 percent, but had shorter followup (median 12-35 months) than studies of the other treatment modalities.

Metastasis-free survival did not differ between any treatment modalities with *low* strength of evidence on pairwise comparisons except for partial nephrectomy versus thermal ablation, where there was *moderate* strength of evidence for equivalent metastasis-free survival.

Thermal ablation was associated with worse local recurrence-free survival compared with radical nephrectomy (*low* strength of evidence) and partial nephrectomy (*moderate* strength of evidence). After a repeat treatment, secondary efficacy of thermal ablation appeared to more closely approximate the local cancer control rates of radical nephrectomy and partial nephrectomy (Figure D).

Renal Functional Outcomes Early and Late

Fifty-three studies (reported in 54 articles, 17,784 patients) evaluated renal functional outcomes, including changes in creatinine and/or estimated glomerular filtration rate, incidence of chronic kidney disease stage 3, 3b, and/or 4 [or greater], and incidence of end-stage renal disease. Earlier stages of chronic kidney disease were not evaluated or synthesized, since these typically depend on the presence of albuminuria, a factor not evaluated in these studies. One study was an RCT (reported in two articles) and the remainder were retrospective observational studies. Thirty-eight (38) studies compared radical nephrectomy and partial nephrectomy, eight (8) studies compared radical nephrectomy and thermal ablation, 21 studies compared partial nephrectomy and thermal ablation, and 2 studies compared active surveillance with the other management strategies. Studies varied in the reporting of both continuous (estimated glomerular filtration rate, and serum creatinine) and categorical renal functional outcomes

Table D. Summary of the strength of evidence for health outcomes of the effectiveness and comparative effectiveness of the management strategies						
Outcome	Partial Nephrectomy Versus Active Surveillance	Partial Nephrectomy Versus Thermal Ablation	Radical Nephrectomy Versus Active Surveillance	Radical Nephrectomy Versus Partial Nephrectomy	Radical Nephrectomy Versus Thermal Ablation	Thermal Ablation Versus Active Surveillance
Cancer-specific survival	Insufficient	Low 9 studies Cancer-specific survival was comparable between partial nephrectomy and thermal ablation. One study, at high risk of bias, suggested partial nephrectomy may be associated with better long-term cancer-specific survival.	Low 1 study A single study demonstrated a similar cancer-specific survival despite greater oncologic potential of tumors undergoing radical nephrectomy.	Moderate 37 studies Cancer-specific survival was comparable for radical nephrectomy and partial nephrectomy across the SEER and institutional studies. The one RCT reported few cancer deaths.	Moderate 2 studies Both studies reported comparable cancer-specific survival for radical nephrectomy compared to thermal ablation.	Insufficient
Metastasis-free survival	Insufficient	Moderate 8 studies Metastasis-free survival was comparable between partial nephrectomy and thermal ablation.	Low 1 study A single study showed similar metastasis-free survival for radical nephrectomy vs. active surveillance.	Low 13 studies Metastasis-free survival for radical nephrectomy compared to partial nephrectomy was similar across all 13 studies.	Low 2 studies Both studies reported comparable metastasis-free survival for radical nephrectomy compared to thermal ablation but included few patients and events.	Insufficient
Local recurrence-free survival	Insufficient	Moderate 14 studies Partial nephrectomy was associated with better local recurrence-free survival compared to thermal ablation across studies. Allowing for multiple retreatments led to a more comparable secondary efficacy rate for thermal ablation.	Insufficient 1 study No local recurrences were reported in this single study.	Moderate 21 studies Local-recurrence free survival for radical nephrectomy compared to partial nephrectomy was similar across studies. No study reported a statistically significant difference.	Low 2 studies Both studies reported better local recurrence-free survival for radical nephrectomy compared to thermal ablation but included small sample sizes.	Insufficient

Table D. Summary of the strength of evidence for health outcomes of the effectiveness and comparative effectiveness of the management strategies (continued)

Outcome	Partial Nephrectomy Versus Active Surveillance	Partial Nephrectomy Versus Thermal Ablation	Radical Nephrectomy Versus Active Surveillance	Radical Nephrectomy Versus Partial Nephrectomy	Radical Nephrectomy Versus Thermal Ablation	Thermal Ablation Versus Active Surveillance
Overall survival	Insufficient	Low 13 studies All 13 studies demonstrated worse overall survival for thermal ablation compared to partial nephrectomy, likely due to age and comorbidity.	Low 1 study Single study demonstrated comparable overall survival with radical nephrectomy and active surveillance with a wide confidence interval [hazard ratio 0.75 (95% CI, 0.45 to 1.26)].	Low 36 studies Overall survival was similar for radical nephrectomy and partial nephrectomy, but studies were inconsistent. SEER analyses showed a survival advantage for partial nephrectomy while institutional cohorts and the 1 RCT did not demonstrate this.	Insufficient 3 studies The results of the three studies were too inconsistent to support a conclusion, especially given the limitations of the studies.	Insufficient
Continuous renal functional outcomes ^a	Insufficient 2 studies One study combined both partial nephrectomy and cryoablation without the ability to separate the groups. The other study found no difference in eGFR change between groups. The evidence was insufficient to determine effectiveness of partial nephrectomy alone.	Low 19 studies eGFR fell more with partial nephrectomy than with thermal ablation, by an average of 1.0 ml/min/1.73 m ² (95% CI -0.2-2.1 ml/min/1.73 m ²), but the result was not statistically significant and there was significant heterogeneity.	Low 2 studies While results are limited by having only two studies, decline in eGFR was 14 ml/min/1.73 m ² less in those assigned active surveillance.	Moderate 34 studies eGFR fell more with radical than partial nephrectomy, by an average of 3.6 ml/min/1.73 m ² (95% CI 3.2-4.1 ml/min/1.73 m ²), with significant heterogeneity in the magnitude of the difference.	Moderate 7 studies eGFR fell more with radical nephrectomy than with thermal ablation, by an average of 9.9 ml/min/1.73 m ² (95% CI 7.6-12.3 ml/min/1.72 m ²).	Insufficient 2 studies One study combined both partial nephrectomy and cryoablation without the ability to separate the groups. The other study found no difference in eGFR change between groups. The evidence was insufficient to determine effectiveness of thermal ablation alone.

Table D. Summary of the strength of evidence for health outcomes of the effectiveness and comparative effectiveness of the management strategies (continued)

Outcome	Partial Nephrectomy Versus Active Surveillance	Partial Nephrectomy Versus Thermal Ablation	Radical Nephrectomy Versus Active Surveillance	Radical Nephrectomy Versus Partial Nephrectomy	Radical Nephrectomy Versus Thermal Ablation	Thermal Ablation Versus Active Surveillance
Categorical renal functional outcomes ^a	<p>Insufficient 2 studies</p> <p>One study combined both partial nephrectomy and cryoablation without the ability to separate the groups. The other study found no difference in rates of CKD between groups. The evidence was insufficient to determine effectiveness of partial nephrectomy alone.</p>	<p>Low 11 studies</p> <p>No statistically significant differences were seen in rates of CKD stage ≥ 3, $\geq 3b$, ≥ 4, or ESRD.</p>	<p>Low 2 studies</p> <p>While results are limited by having only two studies, rates of new onset CKD Stage ≥ 3 were 3-6% with active surveillance and 40-76% with radical nephrectomy.</p>	<p>Moderate 24 studies</p> <p>Incidence of all stages of CKD were lower in those undergoing partial nephrectomy compared to radical nephrectomy, with risk 0.39 times lower for CKD stage 3, 0.37 times lower for CKD stage 3b, 0.76 times lower for CKD stage 4, and 0.47 times lower for ESRD. Heterogeneity did exist in the magnitude of the findings.</p>	<p>Moderate 4 studies</p> <p>Rate of CKD Stage >3 was 3.5 fold higher (95% CI 1.1-12.7) for those receiving radical nephrectomy. Rates of CKD stage 3b and ESRD were limited to two studies.</p>	<p>Insufficient 2 studies</p> <p>One study combined both partial nephrectomy and cryoablation without the ability to separate the groups. The other study found no difference in rates of CKD between groups. The evidence was insufficient to determine effectiveness of thermal ablation alone.</p>

Table D. Summary of the strength of evidence for health outcomes of the effectiveness and comparative effectiveness of the management strategies (continued)

Outcome	Partial Nephrectomy Versus Active Surveillance	Partial Nephrectomy Versus Thermal Ablation	Radical Nephrectomy Versus Active Surveillance	Radical Nephrectomy Versus Partial Nephrectomy	Radical Nephrectomy Versus Thermal Ablation	Thermal Ablation Versus Active Surveillance
QOL	Insufficient	Insufficient	Insufficient	Insufficient 4 studies Conclusions cannot be drawn based on limited number of studies, heterogeneity of outcome measures, and inconsistency of results.	Insufficient	Insufficient
Perioperative Outcomes	Insufficient	Moderate 15 studies Estimated blood loss, transfusion rate, rate of conversions, and length of hospital stay favored thermal ablation consistently.	Insufficient	Moderate 23 studies Partial nephrectomy demonstrated consistently higher estimated blood loss and transfusion rate with similar conversion to open rate and length of hospital stay.	Low 3 studies No study evaluated estimated blood loss. Blood transfusion rate was similar, and length of hospital stay favored thermal ablation. However, no more than two studies reported each outcome.	Insufficient

Table D. Summary of the strength of evidence for health outcomes of the effectiveness and comparative effectiveness of the management strategies (continued)

Outcome	Partial Nephrectomy Versus Active Surveillance	Partial Nephrectomy Versus Thermal Ablation	Radical Nephrectomy Versus Active Surveillance	Radical Nephrectomy Versus Partial Nephrectomy	Radical Nephrectomy Versus Thermal Ablation	Thermal Ablation Versus Active Surveillance
Harms	Insufficient	Low 21 studies Rates of harms (specifically urologic, non-urologic, minor and major) varied significantly among studies. Some urologic and non-urologic complications occurred less often after partial nephrectomy and other urologic and non-urologic complications occurred less often after thermal ablation, but the rate of acute kidney injury and the rate of minor or major Clavien complications did not differ between partial nephrectomy and thermal ablation.	Insufficient	Low 32 studies The only RCT in this literature demonstrates higher rates of urologic complications in patients undergoing partial nephrectomy. This is corroborated by the retrospective data. However, rates of harms were modest among studies. The rate of acute kidney injury did not differ between radical and partial nephrectomy, but the rate of major Clavien complications was higher with partial nephrectomy than with radical nephrectomy. Non-urologic complications did not differ between radical and partial nephrectomy.	Low 7 studies Harms were inconsistently reported among the four studies, making it difficult to draw conclusions about the differences that were observed in specific urologic or non-urologic complications. The rate of acute kidney injury did not differ significantly between radical nephrectomy and thermal ablation, but the data were insufficient to rule out a clinically important increased risk with radical nephrectomy. Minor and major Clavien complications were only reported in one study.	Insufficient

CI = confidence interval, CKD = Chronic kidney disease; eGFR = estimated glomerular filtration rate; ESRD = end-stage renal disease, RCT = randomized controlled trial;

SEER = Surveillance, Epidemiology, and End Results

^aContinuous renal functional outcomes included change in serum creatinine and/or change in eGFR; categorical renal functional outcomes included incidence of CKD stage 3, 3b, or 4 or incidence of ESRD.

(incidence of chronic kidney disease and end-stage renal disease). In addition, the time point of reported outcomes varied from 1 month to 10 years. The overall risk of bias across outcomes associated with the observational studies was moderate to serious.

All interventions experienced decreased renal function in the early postoperative period (within the first month) with a subsequent improvement and plateau in renal function within 1 to 6 months. The strength of evidence was moderate that radical nephrectomy is associated with worsened renal functional outcomes, including increasing incidence of chronic kidney disease stage 3, chronic kidney disease stage 3b, end-stage renal disease, as well as change in creatinine and estimated glomerular filtration rate, when compared with partial nephrectomy and thermal ablation, but a low rate of end-stage renal disease among all management strategies (0.4-3 percent by 1 year). Importantly, the only RCT of partial nephrectomy and radical nephrectomy demonstrated a greater decrease in estimated glomerular filtration rate within the first year for those patients undergoing radical nephrectomy and higher initial incidence of chronic kidney disease stage 3 and stage 3b, but similar estimated glomerular filtration rate from 13 to 15 years of followup (moderate strength of evidence). Renal functional outcomes were similar among partial nephrectomy and thermal ablation. The strength of evidence was *insufficient to low* in the remainder of comparisons based on a low number of studies and inconsistencies in reporting of renal functional outcomes.

Health-Related Quality of Life

Four studies (440 patients) compared health-related quality of life outcomes in patients undergoing radical and partial nephrectomy. Three studies were cross-sectional studies and one evaluated the outcome at predetermined time points in a prospective manner. The risk of bias associated with the observational studies was moderate to serious based on selection bias and bias due to confounding.

The strength of evidence was insufficient to support a conclusion about the studies that reported a trend that radical nephrectomy may provide better quality of life regarding perception of cancer control, and partial nephrectomy may be associated with decreased anxiety and depression.

Perioperative Outcomes and Harms

Thirty-eight studies (37 comparative studies and 1 RCT, with a total of 11,802 patients) evaluated perioperative outcomes including estimated blood loss, blood

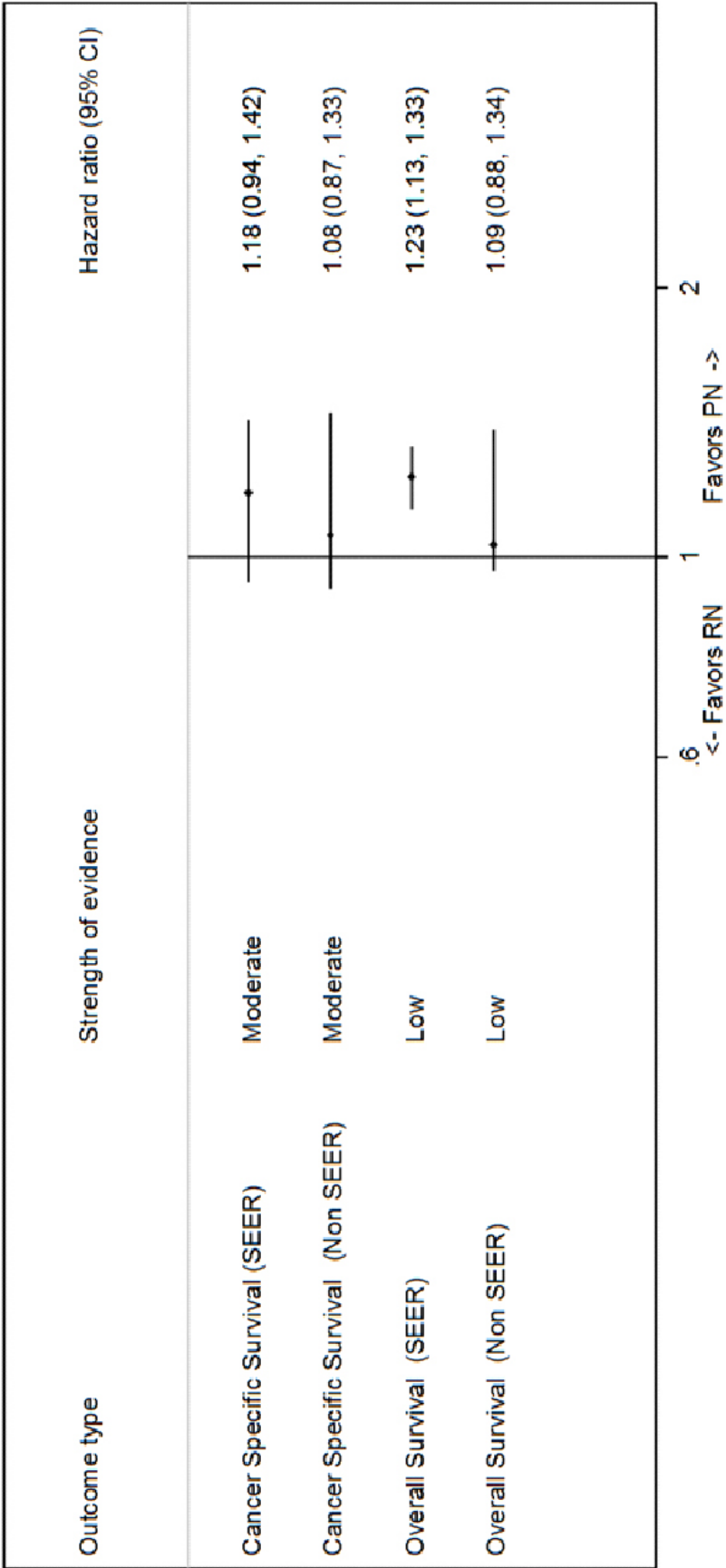
transfusion rate, conversion to open surgery, and length of stay. Twenty-four studies compared radical nephrectomy and partial nephrectomy, 3 studies compared radical nephrectomy versus thermal ablation, and 16 studies compared partial nephrectomy and thermal ablation. Three studies reported multiple comparisons.

Harms were evaluated as urologic and nonurologic complications. Forty-seven studies (46 comparative studies and one RCT (reported in 2 articles), with a total of 180,009 patients) evaluated harms, including 32 studies of radical nephrectomy and partial nephrectomy, 7 studies of radical nephrectomy and thermal ablation, and 21 studies of partial nephrectomy and thermal ablation. Six studies reported multiple comparisons (i.e., three-armed study). There was one RCT, and the remainder were observational studies. The single RCT had unclear risk of bias and the overall risk of bias associated with the observational studies was moderate to serious. No study evaluated perioperative outcomes or harms associated with active surveillance.

Thermal ablation offered the most favorable perioperative outcomes with fewer conversions to open surgery and shorter length of stay when compared to radical nephrectomy (low strength of evidence); and less estimated blood loss, less blood transfusions, no conversions to open surgery or radical nephrectomy, and shorter length of stay when compared to partial nephrectomy (moderate strength of evidence). The strength of evidence was moderate that partial nephrectomy had the highest blood transfusion rate (4.6 to 16.3 percent), which was significantly greater than both radical nephrectomy and thermal ablation.

In general, rates of harms were low among all treatment modalities, with minor (Clavien I-II) and major (Clavien III-IV) complications occurring in 2.6-24.1 percent and 2.8-8.0 percent of patients respectively.¹⁷ When considering specific harms, partial nephrectomy had higher rates of urologic complications (including renal abscess, ureteral injury, urine leak and subsequent interventions) when compared to radical nephrectomy (low strength of evidence) and thermal ablation (low strength of evidence). However, rates of minor and major complications were similar among all three treatment modalities. Thermal ablation had the lowest reported rates of acute kidney injury and non-urologic complications when compared to both radical and partial nephrectomy. The strength of evidence was insufficient to *low* for all other comparisons based on inconsistencies in the reporting of harms (urologic and non-urologic complications) among studies (Figure E to Figure G).

Figure D. Pooled comparisons of cancer-specific survival and overall survival for radical nephrectomy (RN) versus partial nephrectomy (PN) in patients with clinical stage 1 and 2 renal cancer from studies that presented effect estimates as hazard ratios



CI = confidence interval; SEER = Surveillance, Epidemiology, and End Results

Figure E. Pooled comparisons of perioperative outcomes and harms for radical nephrectomy (RN) versus partial nephrectomy (PN) from studies that presented effect estimates as risk ratios

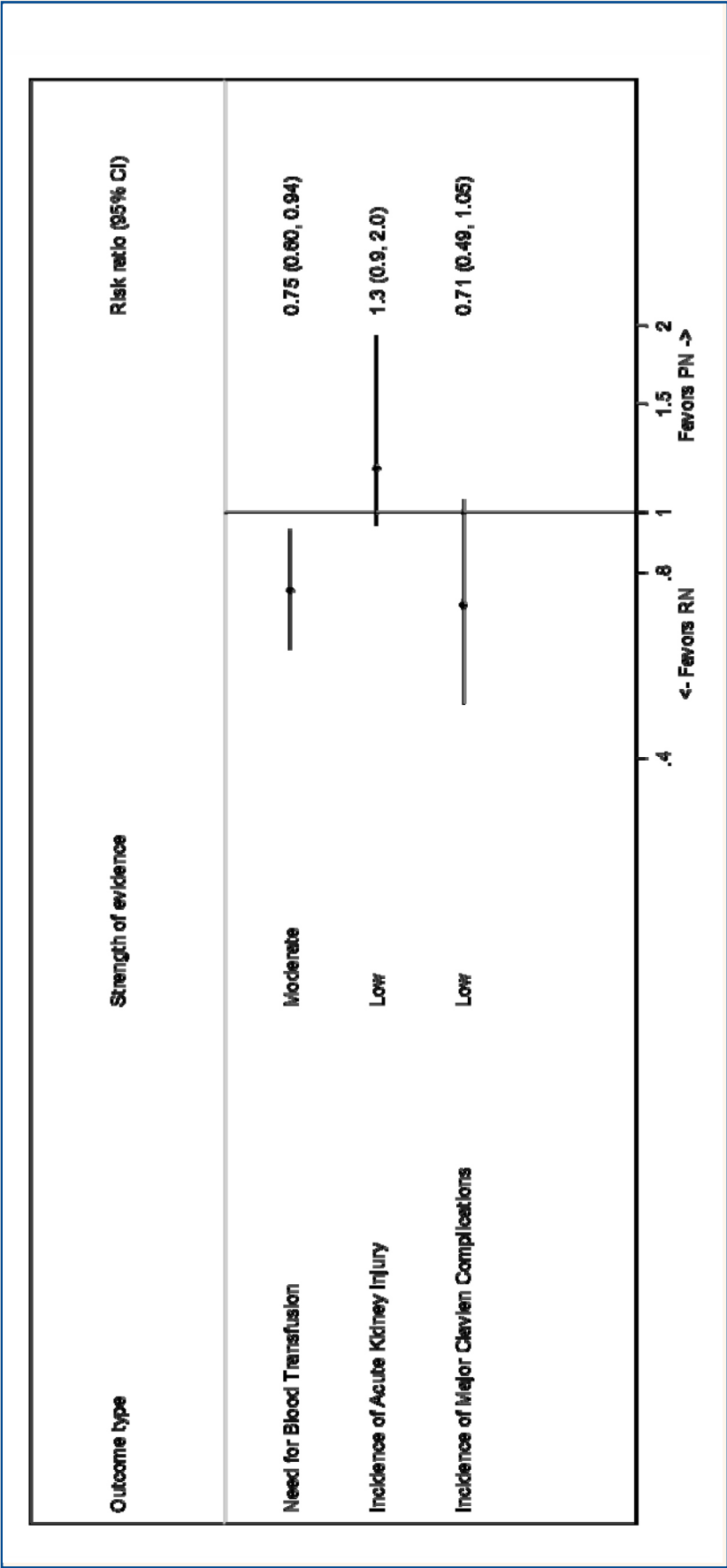
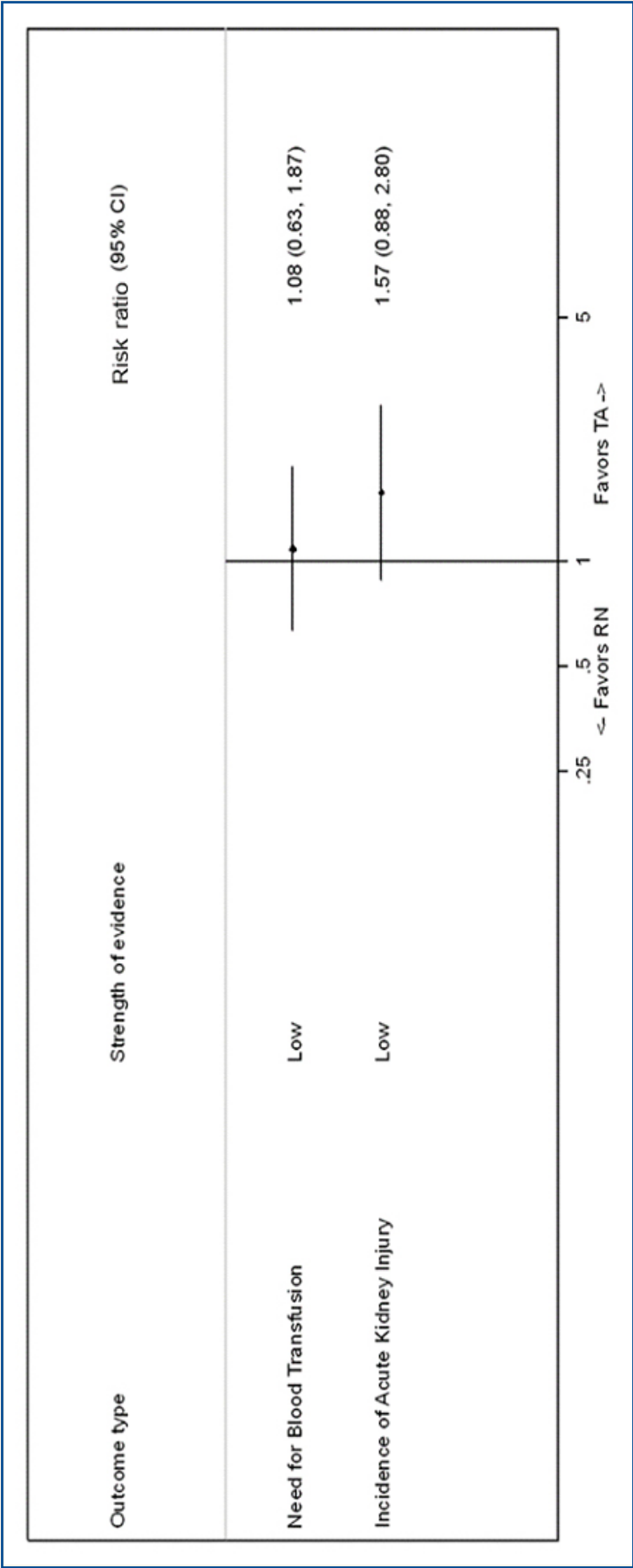
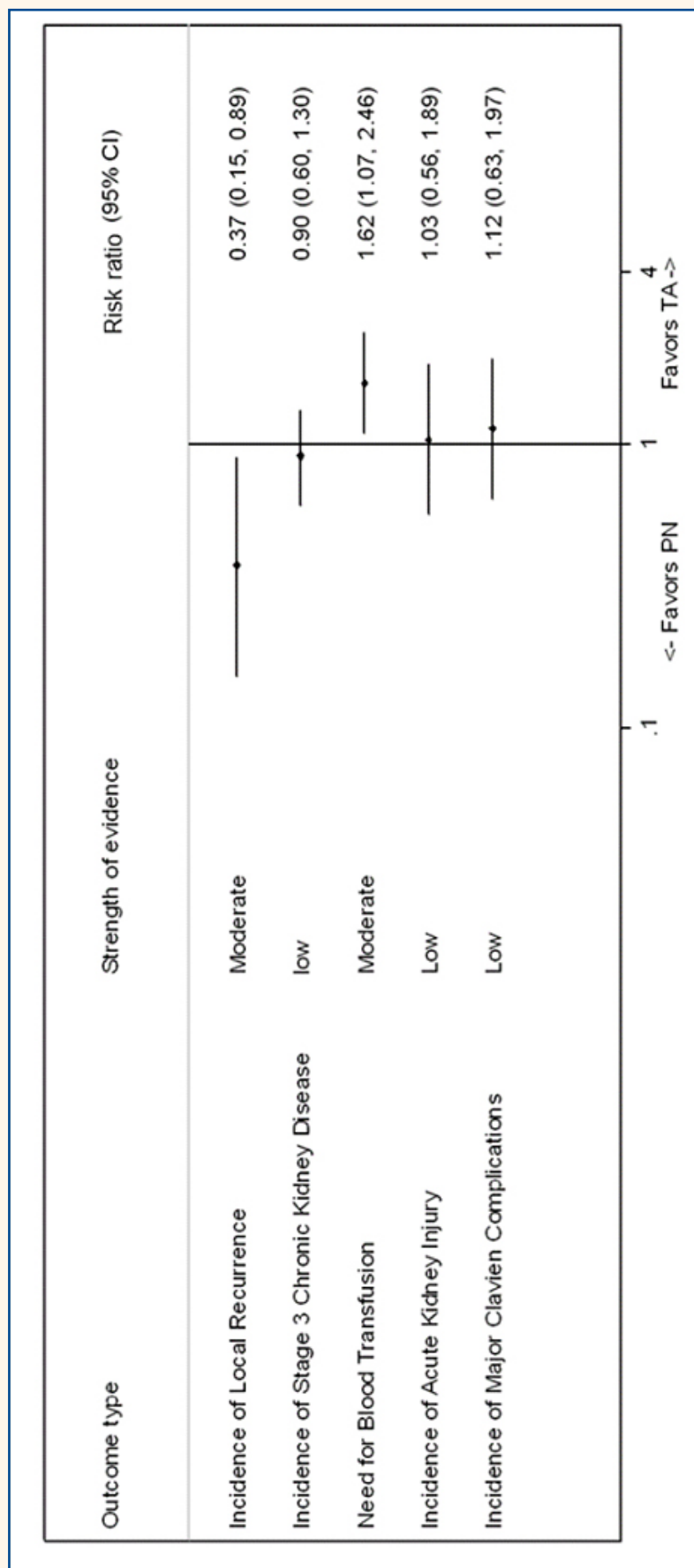


Figure F. Pooled comparisons of perioperative outcomes and harms for radical nephrectomy (RN) versus thermal ablation (TA) from studies reporting risk ratios



CI = confidence interval

Figure G. Pooled comparisons of perioperative outcomes and harms for partial nephrectomy (PN) versus thermal ablation (TA) from studies reporting risk ratios



CI = confidence interval

Predictors of Oncologic Outcomes, Overall Survival, Renal Functional Outcomes, Quality of Life, and Harms

Twenty-one studies evaluated the oncologic outcomes. Seventeen studies (with a total of 101,377 patients) evaluated predictors of cancer-specific survival, one study (475 patients) examined predictors of metastasis-free survival, and 3 studies (360 patients) evaluated predictors of local recurrence-free survival. The evidence was limited regarding the comparative benefits and harms of management strategies based on patient or tumor characteristics. Radical nephrectomy and partial nephrectomy had limited evidence suggesting that age, tumor size, stage, and grade were inversely associated with cancer-specific survival. The strength of evidence was low for cancer-specific survival and insufficient for metastasis-free and local recurrence-free survival (Table E).

Twenty studies (85,939 patients) considered predictors of overall survival. Increasing age and comorbidity predicted overall survival. The strength of evidence was low.

Twenty-five studies (14,272 patients) evaluated predictors of renal functional outcomes. Baseline renal function was associated with long-term renal functional outcomes, regardless of type of surgery. The strength of evidence was low on the predictors of renal functional outcomes in comparative studies, due to inconsistent reporting of variables in prediction models.

Only two studies (247 patients) evaluated predictors of quality of life and three studies (2,168 patients) examined predictors of comparative harms between treatment groups. The strength of evidence from these studies was insufficient to support conclusions about factors predictive of differences between management strategies in quality of life and perioperative outcomes and harms.

Table E. Summary of the strength of evidence for clinical predictors of the comparative benefits and harms of the available management strategies

Outcomes	No. Studies	Strength of Evidence Finding
Cancer-specific survival	19	Low Most data was derived from studies of radical nephrectomy in comparison to partial nephrectomy. Inclusion criteria varied among studies, and the relationship of age, tumor size, stage and grade to oncological outcomes were inconsistent among studies. However, differences in cancer-specific survival among modalities is likely unrelated to age or tumor stage.
Metastases-free survival	1	Insufficient
Local recurrence-free survival	3 (local and metastatic recurrence combined in these studies)	Insufficient Variations in data collection and presentation prevent meaningful conclusions from these studies.
Overall survival	22	Low Based mostly on studies of radical nephrectomy compared to partial nephrectomy, age and comorbidities consistently predicted overall survival.
Renal functional outcomes	27	Low Most data was derived from studies of radical and partial nephrectomy. The effects of baseline renal function and age were consistent among studies, but inconsistencies in other parameters limit the strength of evidence.
Quality of life	2	Insufficient Both studies demonstrated surgical approach (laparoscopic versus open) to predict outcome, but sparse data and inconsistencies among studies prevented determination of whether any factors were predictive of differences in the effects on health-related quality of life.
Perioperative outcomes and harms	3	Insufficient One study evaluated age and two evaluated tumor size. ^{18,19} All studies were inconclusive, preventing meaningful conclusions.

Discussion

This systematic review addresses three key questions evaluating both the diagnostic and therapeutic management of clinically localized renal masses suspicious for malignancy.

Diagnosis of Renal Mass Suspicious for Localized Renal Cell Carcinoma

KQ1: Efficacy of Composite Models in the Diagnosis of a Renal Mass Suspicious for Localized Renal Cell Carcinoma

KQ 2: Accuracy and Efficacy of Renal Mass Biopsy in the Diagnosis of a Renal Mass Suspicious for Localized Renal Cell Carcinoma

The evidence showed that composite models have a predictive utility in differentiating benign and malignant pathology. Imaging characteristics, which included mass size and anatomic location, were the most heavily used variables in the models, but there was not a single variable that was predictive of benign or malignant pathology across all composite models.

In general, increased tumor size and male sex were best correlated with malignant pathology, supporting historical predictors of malignancy in prior guidelines and retrospective studies. The evidence was insufficient to identify any other strong predictors of malignant versus benign pathology in this sample population. Without further prospective studies examining these variables, it is not possible to conclude that any particular composite model variables can be successfully applied as a predictive tool. However, these data can inform clinicians about general variables that have been used to predict benign or malignant pathology, and be used to guide further well-designed clinical trials.

Our review provides support for the current (2009) AUA guidelines regarding the use of tumor size and sex to estimate the risk of malignancy.¹² The findings of this systematic review provide further evidence of the strength of the correlation with tumor size and sex, and may help inform new guideline updates. It is also noteworthy that proposed risk factors from prior research and guidelines, specifically age and BMI, did not have levels of evidence supporting their routine use to predict benign or malignant pathology. Our analysis did not identify any components of a composite model that could be used to definitively distinguish benign from malignant pathology.

The evidence also showed that percutaneous renal mass sampling is associated with a low risk of complications (≤ 5 percent for each evaluated complication) and excellent positive predictive value (97-100 percent). However, the notable nondiagnostic rate (14 percent), low negative predictive value (68 percent), and bias that surgical pathology is not routinely pursued for benign biopsy samples, prevents strong conclusions from being drawn regarding the exact role of renal mass sampling in the clinical practice. The evidence does support the preference of core biopsy over fine needle aspiration, based on the sensitivity (97.5 percent) and negative predictive value (68.5 percent) in an analysis of core biopsy alone, compared with a sensitivity of 62.5 percent and unknown specificity in one study on fine needle aspiration. It is clear renal mass sampling is a safe diagnostic technique as harms from renal biopsy are infrequent and usually do not require additional intervention. Historically, renal mass biopsy was avoided due to concern regarding tumor seeding. In no study included in this systematic review was a case of tumor seeding reported. Based on the available evidence, it is not possible to conclude that renal mass sampling is a universal prerequisite to surgical intervention or active surveillance. More clinical research is needed to better elucidate the utility of renal mass sampling.

Our analysis is consistent with the AUA and European Association of Urology (EAU) guidelines, which recommend using renal mass sampling judiciously, and preferably to use core biopsy over fine needle aspiration in the decision-making algorithm. Our systematic review also demonstrates real limitations to renal mass sampling that may be considered in any recommendation regarding the standard use of renal mass sampling. Given limitations in the data and the performance characteristics of renal mass biopsy, it is difficult to determine the exact clinical scenarios in which renal mass biopsy would influence management. However, there are a number of indications where renal mass biopsy may be considered. For example, in accordance with AUA and EAU guidelines, renal mass biopsy is considered prior to thermal ablation when its results could help determine appropriate followup and treatment efficacy. A young patient determined to have a partial nephrectomy for a small tumor would likely not benefit from biopsy. In contrast, a patient with a solitary kidney in whom surgery will likely lead to an anephric state may benefit from the added information yielded by a biopsy. This decision-making process has to occur thorough discussion of risks and benefits between physician and patient. The implications of the complication

profile on special patient populations such as those on anticoagulant therapy was limited in the studies reviewed.

Management of Renal Mass Suspicious for Localized Renal Cell Carcinoma

KQ 3a: Efficacy and Comparative Efficacy Of Different Interventions for the Management of a Renal Mass Suspicious for Localized Renal Cell Carcinoma

KQ 3b: Comparative Benefits and Harms Of Management Strategies Based on Patient Demographics, Clinical Characteristics, or Disease Severity

The evidence regarding management strategies of renal masses suspicious for localized renal cell carcinoma is based almost entirely on retrospective studies and is susceptible to the inherent limitations of this study design. We included comparative studies regarding radical nephrectomy, partial nephrectomy, and thermal ablation. We included uncontrolled studies on active surveillance because of the lack of comparative studies investigating this treatment modality.

According to the 2009 AUA Guidelines for the Management of the Clinical Stage 1 Renal Mass, physicians should “review with the patient the available treatment options and the attendant benefits and risks, including oncologic considerations, renal functional considerations and potential morbidities.”¹² Our review provides an updated summary of the benefits and risks of the treatment options. Of note, we found that “overall survival rates and cancer-specific survival rates were excellent (95-100 percent) regardless of the surgical management strategy. Interestingly, the AUA, EAU and National Comprehensive Cancer Network (NCCN) guidelines base recommendations for the management of renal masses on the clinical stage of the tumor – recommending nephron-sparing approaches (partial nephrectomy and thermal ablation) for smaller tumors (specifically those with cT1a masses). In our review, we found evidence of improved cancer-specific survival with decreasing tumor stage and size, but we were unable to demonstrate superior cancer-specific survival for any particular management strategy based on tumor size or stage. This may reflect a lack of granularity in these comparative studies or may represent the noninferiority of these management strategies in the treatment of localized renal masses. This could also be further evidence of the generally favorable biology of small tumors, which may supersede the chosen treatment modality.

The 2009 AUA guidelines also recommended thermal ablation as a treatment option for patients at high surgical risk, and active surveillance as an option for patients with decreased life-expectancy or extensive comorbidity. Our review of the evidence showed that thermal ablation and active surveillance were both associated with worse overall survival, reflecting the increased age, comorbidity and competing risks of death in the patients typically selected for less invasive management. Furthermore, thermal ablation was associated with worse local recurrence-free survival compared with radical nephrectomy and partial nephrectomy – as was previously noted in the 2009 AUA Guidelines. Patients should be counseled that an equivalent local control rate with thermal ablation may require more than one treatment. Unfortunately, the evidence remains insufficient to directly compare the outcomes of active surveillance to surgical management options for patients with decreased life-expectancy or extensive comorbidity.

The 2009 AUA guidelines recommend giving consideration to nephron-sparing surgery (partial nephrectomy or thermal ablation) for all patients with a clinical T1 renal mass. To help physicians counsel patients on the potential benefits of nephron-sparing surgery, it is important to have up-to-date information on the comparative effects of the surgical management options on renal functional outcomes. Any analysis of renal functional outcomes in observational studies is inherently biased by the selection of patients into radical versus nephron-sparing management strategies (partial nephrectomy or thermal ablation). Patients with worse baseline function are often selected for nephron-sparing approaches and, as expected, radical nephrectomy was associated with worse renal outcomes when compared with partial nephrectomy or thermal ablation (as measured by estimated glomerular filtration rate, serum creatinine, or incidence of chronic kidney disease). Partial nephrectomy and thermal ablation have similar risks of estimated glomerular filtration rate decline and incidence of chronic kidney disease. Our synthesis of studies suggests that patients with optimal baseline renal function (estimated glomerular filtration rate greater than 90 mL/min/1.73m²) or poor baseline renal function (estimated glomerular filtration rate less than 45 mL/min/m²; chronic kidney disease stage IIIb or worse) may not experience renal functional benefits from nephron-sparing procedures compared with radical nephrectomy. However, this is likely due to decreased numbers of studies reporting these subgroups and outcomes, and the few studies reporting followup beyond 1 year. Further research should strive to identify the patients most likely to benefit from nephron-sparing approaches from a renal functional

standpoint, and in particular long-term development of chronic kidney disease and/or end-stage renal disease. There is also a paucity of data regarding health-related quality of life for patients with clinically localized renal masses suspicious for malignancy. Quality of life in these patients appears to be influenced by a number of factors including cancer control, renal function, physical function, and mental well-being.

In addition to cancer-specific outcomes, overall survival, renal functional outcomes, and quality of life (which all have long-term implications), the choice of management strategy also depends on perioperative outcomes and harms, which may modulate a patient's selection of a given strategy. Based on comparative data, thermal ablation had the most favorable perioperative outcomes (less estimated blood loss, shorter length of stay, and less conversions to open or radical surgery) in comparisons with radical or partial nephrectomy. While the overall rate of postoperative urologic and nonurologic complications was similar among all management strategies, the differential rates of specific postoperative complications varied by strategy. For instance, despite similar overall complication rates, partial nephrectomy had the highest rate of postoperative bleeding while patients undergoing radical nephrectomy had more respiratory harms and acute kidney injury. Since an individual patient's risk factors may play an important role in choosing a management strategy, tailoring management to a specific patient's susceptibility to harms may prove prudent.

While a number of studies evaluated multivariate predictors of oncological efficacy, renal functional outcomes, overall survival, and quality of life, few studies evaluated comparative efficacy of the given management strategies in relation to these predictors. Limited data exists to explain the role of clinical factors in predicting oncologic outcomes, overall survival, renal functional outcomes, quality of life, perioperative outcomes, and harms among the management strategies. Evidence suggests that larger tumors are more likely to be malignant, and uncontrolled studies indicate that large masses may increase the likelihood of complications during partial nephrectomy (comparative data from this review did not demonstrate any increased risk of complications based on tumor size). Therefore, prior guidelines and expert statements may be reasonable in suggesting radical nephrectomy in patients with larger (clinical stage T1b or 2) tumors – despite a lack of evidence in this systematic review. However, studies suggest that baseline renal function is the best predictor of long-term renal functional outcomes regardless of type of surgery – therefore a

patient with a large tumor and chronic kidney disease at baseline (stage 3 or 3b especially), may benefit from a nephron-sparing approach. The choice of management strategy is therefore complex and dependent on patient and tumor characteristics as well as patient and physician preferences regarding the risk of recurrence, survival, renal functional outcomes, and complications. The current data does not provide strong enough evidence to support one management strategy over another for a given patient or clinical scenario. Future research should strive to provide more information to guide the choice of management strategy for different types of patients.

One of the major limitations of the evidence not previously discussed is the imprecise reporting of clinical stage among studies. As nephron-sparing approaches are mostly indicated for clinically localized tumors, these studies were included regardless of the reporting of clinical stage. However, studies of radical nephrectomy were only included if clinical stage was explicitly stated. We urge all studies reporting outcomes on renal masses to consistently report clinical stage.

Applicability

The target population included patients with newly diagnosed, localized renal masses concerning for stage I or II renal cell carcinoma, who were older than age 18, with no family or personal history of renal cell carcinoma.

Regarding diagnostics, we evaluated the accuracy of published composite models (e.g., combination of demographics, clinical characteristics, blood/urine tests, and tumor imaging characteristics) for predicting malignancy. The applicability of our findings was limited by several factors. The patient populations in the reported composite models were relatively old with limited details regarding specific preoperative patient or tumor characteristics. As such, younger patients and those with other comorbidities may have differing risks of malignancy. The literature evaluating renal mass sampling did not routinely report details such as localization and characteristics of the mass that was biopsied. Anterior and hilar tumors may be more difficult to biopsy due to their difficult location, and partially cystic lesions may not yield sufficient biopsy material. Thus, the performance characteristics of renal mass biopsy may not be applicable to these tumors. Furthermore, these findings may not be applicable to patients who had nonmalignant renal mass biopsies as our analysis only included renal mass sampling studies when there was corresponding surgical pathology. Patients on anticoagulant therapy and other special

populations may have different complication profiles than those in the studies analyzed.

The applicability of our findings with respect to management strategies also is limited by several factors. The paucity of prospective comparative data highlights the high risk of bias of the studies reviewed. Selection bias plays a prominent role in treatment selection, thereby limiting the applicability of the findings from retrospective observational studies to specific patient groups. For example, thermal ablation studies were enriched with older patients with multiple comorbid conditions, so their applicability to younger patients may be questioned. The lack of comparative data on active surveillance limits the applicability of our findings related to this management strategy. Specific active surveillance enrollment criteria, followup protocols, and triggers for intervention are not rigorously studied, further limiting our understanding of the applicability of these studies. The emergence of new technologies, and any associated learning curve, could also affect the applicability of studies related to thermal ablation and minimally invasive techniques.

Research Gaps

KQ 1: Efficacy of Composite Models in the Diagnosis of a Renal Mass Suspicious for Localized Renal Cell Carcinoma

The primary gaps in research regarding composite models are the lack of validation of composite models and the limited use of laboratory biomarkers in composite models. The lack of published studies of composite models using biomarkers may be a result of failure to test potential biomarkers within a composite model or tested biomarkers that are nonpredictive. Serum biomarkers include, but are not limited to C-reactive protein, platelet count, and carbonic anhydrase 9. These, along with emerging urine biomarkers such as aquaporin-1 and perilipin-2, should be incorporated into composite models and validated prospectively in well-controlled studies.²⁰ Likewise, future composite models should consider new imaging methods, such as 99m technetium-sestamibi single photon emission computed tomography (SPECT), to better differentiate between malignant and benign pathology.²¹

KQ 2: Accuracy and Efficacy of Renal Mass Biopsy in the Diagnosis of a Renal Mass Suspicious for Localized Renal Cell Carcinoma

Our findings demonstrated a high positive predictive value of renal mass sampling but a significant nondiagnostic

rate as well as a relatively poor negative predictive value. The findings have a high associated risk of bias, as there often was no surgical pathology associated with negative or nondiagnostic biopsies. Further gaps included the lack of a standardized biopsy protocol, lack of correlation with patient characteristics (obesity, anticoagulant therapy, solitary kidney, etc.) or tumor characteristics (size, cystic components, anatomic location within kidney, etc.), and inability to determine how biopsy affects definitive treatment.

To improve analysis of renal mass sampling, future studies should consider standardization and detailed publication of biopsy protocols, including the number of biopsy attempts, number of successful biopsies, and number of patients whose procedures were aborted secondary to technical difficulties. The presence of an on-site pathologist to assess the adequacy of the sample was also not universally reported. Ideally, details on the tumor and its anatomic location should be reported in relationship to the renal mass sampling outcomes. Prospective studies are needed in which all patients undergo biopsy prior to surgery for true assessment of renal mass sampling accuracy. Finally, thorough investigation of renal mass sampling as it affects management strategies and ultimately, oncological outcomes, will be critical to determine its true utility.

KQ 3a: Efficacy and Comparative Efficacy of Different Interventions for the Management of a Renal Mass Suspicious for Localized Renal Cell Carcinoma

Conclusions about the efficacy and comparative efficacy of management strategies are limited by weak study designs, poor reporting of clinical staging, and inconsistent reporting of treatment outcomes. Unreported levels of surgeon/operator expertise allows for confounding of the results.

To address these limitations, greater standardization of treatment data is required. Studies should routinely report both the clinical and pathologic stage of patients, as potentially valuable data was excluded when only pathologic staging was provided. Second, a standardized definition of surgical competence or expertise is needed. This may be achieved either by surgical/procedural case volume or a review of proficiency, success, and complications associated with index cases. Defining surgical or technical proficiency will be an ongoing challenge and standardizing how this is defined is paramount to comparative studies. Third, renal functional and survival outcomes need to be standardized in the

routine reporting of outcomes. Immediate postoperative renal functional data is insufficient and inaccurate for reporting the renal effects of the interventions. We recommend reporting baseline renal function within 1 month of intervention, short-term (1-6 month) and long-term (1 year and longer) outcomes in an attempt to better compare management strategies. Glomerular filtration rate is preferable to serum creatinine, with precise reporting of the data instead of grouping into levels of chronic kidney disease, which are subject to change. In addition, further research should strive to identify the patients most likely to benefit from nephron-sparing approaches from a renal functional standpoint. Survival outcomes (local recurrence, metastasis, cancer-specific, and overall) should be reported at 1, 3, and 5 years, at a minimum. Future research should focus on comparative effects of the management strategies on quality of life to complete the outcome profile associated with each management strategy.

Regarding designing studies that will advance our understanding of the comparative efficacy of each management strategy, it is critical that prospective studies be performed when possible. Retrospective studies may not accurately capture minimally invasive procedures that were converted to open procedures, and may not capture conversions of partial to radical nephrectomies. A trial comparing thermal ablation to partial nephrectomy would be informative. Given the high survival rates of treatment with all modalities studied, quality of life data are lacking and represent an area ripe for discovery. Furthermore, active surveillance should be studied prospectively and in comparison to treatment modalities to better define its place in the management paradigm.

KQ 3b: Comparative Benefits and Harms of Management Strategies Based on Patient Demographics, Clinical Characteristics, or Disease Severity

Patient demographics, clinical characteristics, and disease severity are important in the evaluation of interventions, but were dramatically underreported. To improve understanding of the comparative benefits and harms of the management strategies, studies should be more consistent about reporting clinical stage, tumor characteristics including anatomic location within the kidney, and pre- and postintervention assessments of disease severity and comorbidity.

Conclusions

Diagnosis of Renal Mass Suspicious for Localized Renal Cell Carcinoma

A limited set of studies exists regarding the diagnosis of renal cell carcinoma in our target population. Current composite models do not reliably predict malignancy; however, tumor size and male sex are most highly associated with malignancy. Renal mass sampling is a safe and sensitive procedure, but has a high nondiagnostic rate. The evidence is biased by the failure of nonmalignant biopsies to proceed to intervention. Core biopsy appears to offer improved diagnostic abilities over fine needle aspiration.

Management of Renal Mass Suspicious for Localized Renal Cell Carcinoma

As a result of the paucity of prospective comparative studies on the management of renal masses suspicious for localized renal cell carcinoma, the current literature has a moderate risk of bias. Comparative studies demonstrate comparable cancer-specific survival among all management strategies. However, thermal ablation has a higher local recurrence rate, but favorable perioperative outcome and harms profile. Thermal ablation and partial nephrectomy offer improved renal functional outcomes over radical nephrectomy. Active surveillance may have reasonable survival outcomes in selected populations, but comparative data are lacking. The data are sparse on the quality of life effects of the management options. The evidence also is very limited on how the comparative benefits and harms of management strategies depend on patient characteristics.

References

1. Kutikov A, Fossett LK, Ramchandani P 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):737-40. PMID: 17070344
2. Thompson RH, Hill JR, Babayev Y et al. Metastatic renal cell carcinoma risk according to tumor size. *J Urol* 2009; 182(1):41-5. PMID: 19450840
3. Lane BR, Babineau D, Kattan MW et al. A preoperative prognostic nomogram for solid enhancing renal tumors 7 cm or less amenable to partial nephrectomy. *J Urol* 2007; 178(2):429-34. PMID: 17561141
4. American Cancer Society . *Cancer Facts & Figures* 2014. 2014.

5. Mathew A, Devesa SS, Fraumeni JF Jr, Chow WH. Global increases in kidney cancer incidence, 1973-1992. *Eur J Cancer Prev* 2002; 11(2):171-8. PMID: 11984136
6. Nguyen MM, Gill IS, Ellison LM. 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):2397-400; discussion 2400. PMID: 17085111
7. Kane CJ, Mallin K, Ritchey J, Cooperberg MR, Carroll PR. Renal cell cancer stage migration: analysis of the National Cancer Data Base. *Cancer* 2008; 113(1):78-83. PMID: 18491376
8. Long JA, Arnoux V, Fiard G et al. External validation of the RENAL nephrometry score in renal tumours treated by partial nephrectomy. *BJU Int* 2013; 111(2):233-9. PMID: 22788546
9. Kanao K, Mizuno R, Kikuchi E et al. Preoperative prognostic nomogram (probability table) for renal cell carcinoma based on TNM classification. *J Urol* 2009; 181(2):480-5; discussion 485. PMID: 19100568
10. Haggstrom C, Rapp K, Stocks T et al. Metabolic factors associated with risk of renal cell carcinoma. *PLoS One* 2013; 8(2):e57475. PMID: 23468995
11. Yaycioglu O, Eskicorapci S, Karabulut E et al. A preoperative prognostic model predicting recurrence-free survival for patients with kidney cancer. *Jpn J Clin Oncol* 2013; 43(1):63-8. PMID: 23159766
12. Campbell SC, Novick AC, Belldegrun A et al. Guideline for management of the clinical T1 renal mass. *J Urol* 2009; 182(4):1271-9. PMID: 19683266
13. Higgins JPT, Green S (editors). *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>.
14. 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. <http://www.riskofbias.info>.
15. Whiting PF, Rawlins ME. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011; 155(8):529-36. PMID: 22007046.
16. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality. January 2014. Chapters available at: www.effectivehealthcare.ahrq.gov.
17. Dindo D, DNPCA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg*. 2004; 244:931-7. PMID: 15273542.
18. Roos FC, Brenner W, Thomas C et al. Functional analysis of elective nephron-sparing surgery vs radical nephrectomy for renal tumors larger than 4 cm. *Urology* 2012; 79(3):607-13. PMID: 22386408
19. Tomaszewski JJ, Uzzo RG, Kutikov A et al. Assessing the burden of complications after surgery for clinically localized kidney cancer by age and comorbidity status. *Urology* 2014; 83(4):843-9. PMID: 24680455
20. Rowe SP, Gorin MA, Gordetsky J et al. Initial experience using 99mTc-MIBI SPECT/CT for the differentiation of oncocytoma from renal cell carcinoma. *Clin Nucl Med*. 2015 Apr;40(44):309-13. PMID: 25608174.
21. Morrissey JJ, Mobley J, Figenshau RS et al. Urine aquaporin 1 and perilipin 2 differentiate renal carcinomas from other imaged renal masses and bladder and prostate cancer.: Initial experience using 99mTc-MIBI SPECT/CT for the differentiation of oncocytoma from renal cell carcinoma. *Mayo Clinic Proc*. 2015 Jan;90(1):35-42. PMID: 25572193.

Full Report

This executive summary is part of the following document: Pierorazio PM, Johnson MH, Patel HD, Sozio SM, Sharma R, Iyoha E, Bass EB, Allaf ME. Management of Renal Masses and Localized Renal Cancer. Comparative Effectiveness Review No. 167. (Prepared by the JHU Evidence-based Practice Center under Contract No. 290-2012-00007-I.) AHRQ Publication No. 16-EHC001-EF. Rockville, MD: Agency for Healthcare Research and Quality. February 2016. Available at: www.effectivehealthcare.ahrq.gov/reports/final.cfm.

