Imaging in Suspected Renal Colic: Systematic Review of the Literature and Multispecialty Consensus

Christopher L. Moore,* Christopher R. Carpenter, Marta L. Heilbrun, Kevin Klauer, Amy C. Krambeck, Courtney Moreno,† Erick M. Remer, Charles Scales, Melissa M. Shaw‡ and Kevan M. Sternberg

From the Departments of Emergency Medicine, Yale School of Medicine, New Haven, Connecticut (CLM, MMS), Washington University in Saint Louis, Saint Louis, Missouri (C210) and University of Tennessee, Knoxville, Tennessee (KK), and Department of Radiology, Emory University, Atlanta, Georgia (MLH, CM), Department of Urology, Indiana University School of Medicine, Indianapolis, Indiana (ACK), Imaging Institute and Glickman Urological and Kidney Institute, Cleveland Clinic, Cleveland, Ohio (EMR), Division of Urology, Duke University School of Medicine, Durham, North Carolina (CS) and University of Vermont Medical Center Urology, Burlington, Vermont (KMS)

Purpose: Renal colic is common and CT (computerized tomography) is frequently utilized when the diagnosis of kidney stone is suspected. CT is accurate, but exposes patients to ionizing radiation and has not been shown to alter either interventional approaches or hospital admission rates. This multi-organizational transdisciplinary collaboration sought evidence-based, multispecialty consensus on optimal imaging across different clinical scenarios in patients with suspected renal colic in the acute setting.

Materials and Methods: In conjunction with the ACEP (American College of Emergency Physicians®) E-QUAL (Emergency Quality Network) we formed a nine-member panel with three physician representatives each from the ACEP, the ACR® (American College of Radiology) and the AUA (American Urological Association). A systematic literature review was used as the basis for a 3-step modified Delphi process to seek consensus on optimal imaging in 29 specific clinical scenarios.

Results: From an initial search yielding 6,337 records there were 232 relevant articles of acceptable evidence quality to guide the literature summary. At the completion of the Delphi process consensus, agreement was rated as perfect in 15 (52%), excellent in 8 (28%), good in 3 (10%) and moderate in 3 (10%) of the 29 scenarios. There were no scenarios where at least moderate consensus was not reached. CT was recommended in 7 scenarios (24%) with ultrasound in 9 (31%) and no further imaging needed in 13 (45%).

Conclusions: Evidence and multispecialty consensus support ultrasound or no further imaging in specific clinical scenarios, with reduced-radiation dose CT to be employed when CT is needed in patients with suspected renal colic.

Abbreviations and Acronyms
ACEP = American College of Emergency Physicians®
ACR® = American College of Radiology
ALARA = as low as reasonably achievable
AUA = American Urological Association
BMI = body mass index
CT = computerized tomography
ED = emergency department
E-QUAL = Emergency Quality Network
IV = intravenous
n/a = not applicable
PICO = Population, Intervention, Comparison, Outcome
POCUS = point of care ultrasound
PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analyses
Q = question
RDCT = reduced-radiation dose CT
RPUS = radiology performed ultrasound
SCT = standard CT

Accepted for publication April 5, 2019.

The corresponding author certifies that, when applicable, a statement(s) has been included in the manuscript documenting institutional review board, ethics committee or ethical review board study approval; principles of Helsinki Declaration were followed in lieu of formal ethics committee approval; institutional animal care and use committee approval; all human subjects provided written informed consent with guarantees of confidentiality; IRB approved protocol number; animal approved project number.

Supported by AHRQ (Agency for Healthcare Research and Quality) Grant R18HS023778.

No direct or indirect commercial, personal, academic, political, religious or ethical incentive is associated with publishing this article.

This document is being published without editorial review by the editors of The Journal of Urology®.


* Financial interest and/or other relationship with the AHRQ.
† Financial interest and/or other relationship with GE Healthcare.
‡ Correspondence: Department of Emergency Medicine, Yale University School of Medicine, 484 Congress Ave., New Haven, Connecticut 06519 (e-mail: melissa.m.shaw@yale.edu).

0022-5347/19/2023-0475/0
THE JOURNAL OF UROLOGY®
© 2019 by AMERICAN UROLOGICAL ASSOCIATION EDUCATION AND RESEARCH, INC.
https://doi.org/10.1097/JU.0000000000000342
Vol. 202, 475-483, September 2019
Printed in U.S.A.

www.auajournals.org/jurology 475
BACKGROUND
There are over two million annual ED visits for suspected renal colic in the United States and CT scanning is now performed in more than 90% of patients diagnosed with kidney stone.\textsuperscript{1,2} Despite a dramatic increase in CT use for diagnosis over the last 2 decades, patient-centered outcomes such as admission and intervention do not appear to have been impacted.\textsuperscript{3,4} In 2014, a multicenter prospective trial randomized patients with suspected renal colic to CT, RPUS or POCUS and concluded that initial ultrasonography (either RPUS or POCUS) could lower radiation exposure without adversely affecting patient-centered outcomes.\textsuperscript{5} Despite this evidence, recent data suggest that ultrasound is used in fewer than 7% of patients diagnosed with kidney stone and CT use has continued to increase.\textsuperscript{1}

Similarly, while RDCT is recommended for the evaluation of renal colic, RDCT use has increased only modestly in recent years and is used in fewer than 10% of patients diagnosed with kidney stone.\textsuperscript{6}

Renal colic is a self-limited condition in the majority of patients. However, CT can effectively guide therapy in the subset of patients requiring urological intervention and can detect conditions with signs and symptoms that can mimic renal colic but require intervention (such as appendicitis). The perspective of the urologist considering surgical intervention may differ from that of the clinician initially evaluating and treating the patient (often an emergency physician). Deciding whether patients with suspected renal colic need imaging during the initial evaluation and if so what type is an area with wide practice variation.\textsuperscript{7} While guidelines on “appropriate use” have been developed, they tend to emphasize CT without providing guidance on optimal imaging or scenarios where CT may not be needed.\textsuperscript{8–10}

We convened a multispecialty group with representation from national organizations including emergency medicine, urology and radiology to perform a systematic literature review and seek consensus on imaging approaches in specific clinical scenarios where renal colic was suspected, with an emphasis on situations where CT may not be required.

METHODS
Formation of the Panel, Overview, Definitions and PICO Question
This initiative was formed under the umbrella of the E-QUAL (https://www.acep.org/equal/; a CMS funded SAN (Support and Alignment Network) which is part of a TCPI (Transforming Clinical Practice Initiative) and administered through the ACEP. The E-QUAL network nominated three emergency physician members (CLM, CRC, KK). The ACR and the AUA were contacted and provided 3 representative members from each society, thus forming a nine-member group that conducted this process (ACR: CCM, EMR and MEH; AUA: ACK, KMS and CS). The panel initially defined terms and agreed upon a PICO question. A systematic literature review was conducted followed by a consensus process for specific clinical scenarios.

Consensus definitions are included in supplementary Appendix 1 (https://www.jurology.com). The “optimal diagnostic imaging strategy” was defined as the imaging approach that would maximize patient-centered outcomes: guiding appropriate management while minimizing harms. “Uncomplicated renal colic” was considered to be renal colic without suspicion of infection (based on urinalysis or systemic symptoms such as fever). Cost of imaging was specifically excluded. Improving certainty of diagnosis without an impact on patient care was not considered patient centered. RDCT was not specified as having a specific dose cutoff, but defined as a CT protocol that was specifically tailored to imaging renal colic with lower radiation dose parameters. POCUS was differentiated from RPUS in that it was expected to provide information about the presence or absence of hydronephrosis but would not be uniformly expected to identify stone size or location, while RPUS might (but would not always) be able to determine stone size or location. It is noted that the availability, skill set and use patterns of POCUS may differ between institutions and even between individuals at an institution. While we attempted to suggest when POCUS might be most appropriate, RPUS may be substituted if more appropriate.

The overall PICO question arrived at by the group was: “In patients presenting to the emergency department with pain suspected to be uncomplicated renal colic (P) what imaging (I) should be pursued compared to standard noncontrast computed tomography (CT) scanning (C) to optimize patient-centered outcomes (O)?” From the specific questions submitted by members of the group, three subthemes emerged: imaging in first time renal colic, imaging in recurrent renal colic and special populations (pregnancy, pediatric, obese patients).

Literature Review
A systematic literature review was performed following the PRISMA guidelines. NLM (National Library of Medicine) PubMed\textsuperscript{®} and Embase\textsuperscript{®} databases were queried for English language articles published between January 1995 and May 2018 using a search strategy including MeSH (Medical Subject Headings) as well as key words (supplementary Appendix 2, https://www.jurology.com). To ensure key articles were not missed, each group member provided six articles from their personal libraries (blinded to the search) which were then cross-matched against search results. Articles were reviewed for relevance to the PICO question as level I (definitely relevant), II (probably relevant) or X (not relevant). Interobserver reliability for relevance between nonphysician reviewers and each of the nine physician authors was measured using a random selection of 45 included
and 45 excluded articles that were blindly reviewed (10 articles per author).

Relevant articles were rated for quality of evidence using a hybrid tool that combined the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies) instrument with the ACEP Clinical Policy Committee criteria for diagnostic questions to rate evidence as I, II, III or X (supplementary Appendix 3, https://www.jurology.com). Some articles were not felt to be amenable to rating by this tool for diagnostic accuracy studies (epidemiological and systematic meta-analyses) but were included after review by the group (evidence level n/a). A random selection of thirty-two articles was reviewed by each physician member of the group using this instrument. Discrepancies were discussed on weekly conference calls until the group reached consensus. This was used to help guide the 2 nonphysician reviewers who then assessed the remaining articles independently for evidence level. This evidence was synthesized by the group into a review of principal imaging modalities which formed the basis for the consensus discussion.

Consensus on Clinical Scenarios
After the systematic literature review, discussion and a written summary of the literature for imaging modalities, we attempted to define specific clinical scenarios where consensus could be reached regarding the optimal initial imaging, including scenarios where CT may not be the optimal initial approach. A set of 29 brief clinical vignettes were agreed upon that were felt to represent the best balance between overall number of scenarios and possible permutations (age, gender, pregnancy status, likelihood of stone disease and likelihood of acute alternative diagnosis) (supplementary table, https://www.jurology.com). Consensus was sought using a modified Delphi process that included three rounds of anonymous voting with two group discussions between rounds of voting. All nine members of the group answered the vignettes in a blinded fashion individually for the first round.

For each clinical scenario, the physician was asked to select the “optimal diagnostic imaging strategy” using a priori definitions. Imaging options included no (further) imaging, POCUS, RPUS, RDCT, SCT (noncontrast) and CT with IV contrast. For purposes of defining consensus, imaging modalities were separated into three groups: no further imaging, ultrasound and CT, though subtypes under age 35. BMI, with sensitivities as high as 96% in patients particularly helpful in younger patients with lower BMI, with sensitivities as high as 96% in patients under age 35. While RPUS is imperfectly sensitive for detection of stones when compared to CT, this difference may not be clinically significant. Several studies show that RPUS is unlikely to miss stones requiring intervention.

There were 15 articles relevant to the use of POCUS for renal colic with grades of evidence one (n = 3), two (n = 17), three (n = 20) and n/a (n = 3). Reported sensitivity of RPUS for kidney stone varied widely, ranging from 3% to 98% depending on whether direct stone visualization was required or if indirect evidence of stone presence such as hydronephrosis was sufficient. The preponderance of studies reported sensitivities from 57% to 91%, with improved sensitivity when twinkling artifact is used to help identify stones. RPUS may be particularly helpful in younger patients with lower BMI, with sensitivities as high as 96% in patients under age 35. While RPUS is imperfectly sensitive for detection of stones when compared to CT, this difference may not be clinically significant. Several studies show that RPUS is unlikely to miss stones requiring intervention.

There were 43 relevant articles with grades of evidence one (n = 3), two (n = 17), three (n = 20) and n/a (n = 3). Reported sensitivity of RPUS for kidney stone varied widely, ranging from 3% to 98% depending on whether direct stone visualization was required or if indirect evidence of stone presence such as hydronephrosis was sufficient. The preponderance of studies reported sensitivities from 57% to 91%, with improved sensitivity when twinkling artifact is used to help identify stones. RPUS may be particularly helpful in younger patients with lower BMI, with sensitivities as high as 96% in patients under age 35. While RPUS is imperfectly sensitive for detection of stones when compared to CT, this difference may not be clinically significant. Several studies show that RPUS is unlikely to miss stones requiring intervention.

Ultrasound: RPUS and POCUS
For RPUS there were 43 relevant articles with grades of evidence one (n = 3), two (n = 17), three (n = 20) and n/a (n = 3). Reported sensitivity of RPUS for kidney stone varied widely, ranging from 3% to 98% depending on whether direct stone visualization was required or if indirect evidence of stone presence such as hydronephrosis was sufficient. The preponderance of studies reported sensitivities from 57% to 91%, with improved sensitivity when twinkling artifact is used to help identify stones. RPUS may be particularly helpful in younger patients with lower BMI, with sensitivities as high as 96% in patients under age 35. While RPUS is imperfectly sensitive for detection of stones when compared to CT, this difference may not be clinically significant. Several studies show that RPUS is unlikely to miss stones requiring intervention.

There were 15 articles relevant to the use of POCUS for renal colic with grades of evidence one (n = 3), two (n = 8), three (n = 2) and n/a (n = 1). A systematic review and meta-analysis of POCUS completed in 2018 identified five high quality articles on diagnostic accuracy and six on prognostic value, all of which were identified in our literature review. Diagnostic accuracy was based on presence of hydronephrosis rather than direct visualization of stone and pooled results yielded sensitivity of 70.2% and specificity of 75.4% although specificity increased to 94.4% when moderate or greater hydronephrosis was used as a criterion. Four studies did show a positive association between stone size and larger stones had a higher likelihood of intervention when hydronephrosis was present.

In a multicenter, pragmatic, comparative effectiveness trial published in 2014 investigators randomized over 2,500 patients with suspected kidney stone to POCUS, RPUS or CT and concluded that ultrasound (RPUS or POCUS) is a safe approach for the evaluation of ureterolithiasis, noting that “return emergency department visits, hospitalizations and diagnostic accuracy did not differ significantly among the
RPUS when compared to POCUS was less likely to result in subsequent CT (27% with RPUS vs 41% with POCUS). The study showed a small but significant longer length of emergency department stay for RPUS than either the POCUS or CT groups (7.0 hours vs 6.3 and 6.4 hours, respectively).

**Standard CT (SCT)**

CT is accepted as the reference modality for diagnosis of kidney stones and our group did not seek to review the literature on overall accuracy of CT for kidney stone; rather we sought to understand how often CT identified other pathologies and how CT affected management. There were 36 relevant articles addressing alternative findings on CT with grades of one (n = 3), two (n = 9), three (n = 20) and n/a (n = 4) (table 1). The prevalence of alternative findings ranged from 0% to 33%; however, definitions of acute alternate diagnoses varied widely. Prior to completing the literature review,
our group established definitions (supplementary Appendix 1, https://www.jurology.com), including “acute alternative diagnosis” (requiring a therapeutic intervention, such as appendicitis), “nonacute alternative diagnosis” (causing symptoms but not requiring intervention, such as mesenteric adenitis) and true “incidental findings” (identified as needing followup but not related to presenting symptoms). Incorporating these definitions, the prevalence of alternate findings in suspected renal colic that could be considered acute or clinically important is substantially lower than often reported, typically less than 5%. In the largest prospective study of patients with suspected renal colic (including 2,759 patients randomized to CT or ultrasound as an initial diagnostic test) 4.9% had an “important alternative diagnosis.”\textsuperscript{28} The largest prospective study to randomize patients to CT or ultrasound found no significant difference in high risk diagnoses or adverse events based on initial imaging modality. Notably, this study had a very low rate (0.4%) of “high risk diagnoses.”\textsuperscript{29}

Few studies have separated “incidental findings” from alternate causes of symptoms. A study of over 5,000 patients that used a strict definition of incidental findings (requiring followup but not causative of symptoms) identified incidental findings in 12.7% of CTs performed for renal colic.\textsuperscript{29} While incidental findings may occasionally lead to earlier diagnosis and treatment of pathology (typically malignancy), in many cases these findings may lead to unnecessary workup and even potential morbidity.\textsuperscript{30}

CT may be performed after initial ultrasound imaging if nondefinitive. In the previously cited large randomized pragmatic trial between 27% of patients with RPUS and 41% with POCUS had a subsequent CT.\textsuperscript{5} A single center experience found that while about half of patients evaluated for stone disease underwent CT on initial evaluation, only 10% of those not undergoing CT had it done later in the episode of evaluation and treatment, with 20% of those undergoing initial ultrasound having a CT performed later. They found overall imaging costs

Table 1. Relevant studies reporting an alternate diagnosis on CT for renal colic, listed alphabetically by last name of first author, with year, study type (R: retrospective, P: prospective), ratings of relevance and evidence, overall number of patients, prevalence (%) of renal stone by imaging, and reported prevalence of alternate findings. Note that the definition of “alternate findings” was typically different than the consensus definition of “acute alternative diagnoses”, often including incidental findings and non-acute alternative diagnoses.

<table>
<thead>
<tr>
<th>Year</th>
<th>Author</th>
<th>Study type</th>
<th>Relevance</th>
<th>Evidence</th>
<th>n</th>
<th>Prevalence of stone</th>
<th>Alt findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>Abramson et al</td>
<td>P</td>
<td>2</td>
<td>3</td>
<td>93</td>
<td>60%</td>
<td>18.0%</td>
</tr>
<tr>
<td>2003</td>
<td>Ahmad et al</td>
<td>R</td>
<td>1</td>
<td>3</td>
<td>233</td>
<td>64%</td>
<td>12.0%</td>
</tr>
<tr>
<td>2006</td>
<td>Akay et al</td>
<td>R</td>
<td>1</td>
<td>2</td>
<td>87</td>
<td>54%</td>
<td>16.1%</td>
</tr>
<tr>
<td>2013</td>
<td>Anif et al</td>
<td>P</td>
<td>1</td>
<td>2</td>
<td>124</td>
<td>69%</td>
<td>8.1%</td>
</tr>
<tr>
<td>2009</td>
<td>Ather et al</td>
<td>R</td>
<td>2</td>
<td>3</td>
<td>4,000</td>
<td>78%</td>
<td>9.9%</td>
</tr>
<tr>
<td>2010</td>
<td>Ben Nakh et al</td>
<td>P</td>
<td>2</td>
<td>3</td>
<td>36</td>
<td>31%</td>
<td>11.1%</td>
</tr>
<tr>
<td>2016</td>
<td>Blencer et al</td>
<td>R</td>
<td>1</td>
<td>n/a</td>
<td>626</td>
<td>58%</td>
<td>9.2%</td>
</tr>
<tr>
<td>2007</td>
<td>Broder et al</td>
<td>R</td>
<td>1</td>
<td>2</td>
<td>356</td>
<td>58%</td>
<td>2.0%</td>
</tr>
<tr>
<td>2002</td>
<td>Catalao et al</td>
<td>P</td>
<td>1</td>
<td>2</td>
<td>181</td>
<td>51%</td>
<td>14.0%</td>
</tr>
<tr>
<td>2008</td>
<td>Cullen et al</td>
<td>R</td>
<td>2</td>
<td>3</td>
<td>500</td>
<td>56%</td>
<td>13.0%</td>
</tr>
<tr>
<td>1999</td>
<td>Dorio et al</td>
<td>R</td>
<td>1</td>
<td>3</td>
<td>163</td>
<td>66%</td>
<td>7.0%</td>
</tr>
<tr>
<td>2013</td>
<td>Dyer et al</td>
<td>R</td>
<td>1</td>
<td>3</td>
<td>228</td>
<td>34%</td>
<td>27.0%</td>
</tr>
<tr>
<td>2002</td>
<td>Eshed et al</td>
<td>P</td>
<td>1</td>
<td>3</td>
<td>425</td>
<td>59%</td>
<td>10.0%</td>
</tr>
<tr>
<td>2010</td>
<td>Goldstone et al</td>
<td>R</td>
<td>2</td>
<td>3</td>
<td>231</td>
<td>64%</td>
<td>3.5%</td>
</tr>
<tr>
<td>2004</td>
<td>Ha et al</td>
<td>P</td>
<td>1</td>
<td>3</td>
<td>132</td>
<td>58%</td>
<td>33.0%</td>
</tr>
<tr>
<td>2015</td>
<td>Hall et al</td>
<td>R</td>
<td>2</td>
<td>3</td>
<td>513</td>
<td>45%</td>
<td>14.0%</td>
</tr>
<tr>
<td>2001</td>
<td>Hamm et al</td>
<td>P</td>
<td>1</td>
<td>3</td>
<td>125</td>
<td>73%</td>
<td>3.2%</td>
</tr>
<tr>
<td>2001</td>
<td>Homer et al</td>
<td>P</td>
<td>1</td>
<td>3</td>
<td>228</td>
<td>70%</td>
<td>5.7%</td>
</tr>
<tr>
<td>2006</td>
<td>Hoppe et al</td>
<td>R</td>
<td>2</td>
<td>n/a</td>
<td>1,500</td>
<td>69%</td>
<td>14.0%</td>
</tr>
<tr>
<td>2000</td>
<td>Katz et al</td>
<td>R</td>
<td>2</td>
<td>3</td>
<td>1,000</td>
<td>56%</td>
<td>10.1%</td>
</tr>
<tr>
<td>2012</td>
<td>Khan et al</td>
<td>R</td>
<td>2</td>
<td>3</td>
<td>699</td>
<td>NR</td>
<td>14.0%</td>
</tr>
<tr>
<td>2003</td>
<td>Kobayashi et al</td>
<td>P</td>
<td>2</td>
<td>n/a</td>
<td>560</td>
<td>50%</td>
<td>2.5%</td>
</tr>
<tr>
<td>2001</td>
<td>Messersmith et al</td>
<td>R</td>
<td>2</td>
<td>n/a</td>
<td>21</td>
<td>NR</td>
<td>19.0%</td>
</tr>
<tr>
<td>2014</td>
<td>Moore et al</td>
<td>R</td>
<td>2</td>
<td>1</td>
<td>1,040</td>
<td>50%</td>
<td>2.9%</td>
</tr>
<tr>
<td>2013</td>
<td>Moore et al</td>
<td>R</td>
<td>1</td>
<td>2</td>
<td>5,383</td>
<td>48%</td>
<td>2.8%</td>
</tr>
<tr>
<td>2015</td>
<td>Pernet et al</td>
<td>P</td>
<td>1</td>
<td>3</td>
<td>155</td>
<td>76%</td>
<td>6.0%</td>
</tr>
<tr>
<td>2006</td>
<td>Rafique et al</td>
<td>P</td>
<td>1</td>
<td>3</td>
<td>130</td>
<td>47%</td>
<td>23.0%</td>
</tr>
<tr>
<td>2016</td>
<td>Sarofim et al</td>
<td>R</td>
<td>1</td>
<td>3</td>
<td>215</td>
<td>38%</td>
<td>7.0%</td>
</tr>
<tr>
<td>2015</td>
<td>Schoenfeld et al</td>
<td>R</td>
<td>1</td>
<td>2</td>
<td>291</td>
<td>59%</td>
<td>0.0%</td>
</tr>
<tr>
<td>1999</td>
<td>Sheley et al</td>
<td>R</td>
<td>2</td>
<td>2</td>
<td>180</td>
<td>46%</td>
<td>3.0%</td>
</tr>
<tr>
<td>1996</td>
<td>Smith et al</td>
<td>P</td>
<td>1</td>
<td>3</td>
<td>210</td>
<td>50%</td>
<td>14.8%</td>
</tr>
<tr>
<td>2014</td>
<td>Smith-Bindman et al</td>
<td>P</td>
<td>1</td>
<td>1</td>
<td>926</td>
<td>33%</td>
<td>0.2%</td>
</tr>
<tr>
<td>2002</td>
<td>Strouse et al</td>
<td>R</td>
<td>2</td>
<td>n/a</td>
<td>94</td>
<td>40%</td>
<td>29.0%</td>
</tr>
<tr>
<td>1998</td>
<td>Vieweg et al</td>
<td>P</td>
<td>1</td>
<td>2</td>
<td>105</td>
<td>34%</td>
<td>27.8%</td>
</tr>
<tr>
<td>2008</td>
<td>Xafis et al</td>
<td>R</td>
<td>2</td>
<td>3</td>
<td>638</td>
<td>79%</td>
<td>5.0%</td>
</tr>
</tbody>
</table>
and radiation exposure to be higher in patients undergoing initial CT.\textsuperscript{31} Delayed vs immediate CT does not appear to impact morbidity.\textsuperscript{32}

Despite the potential for CT to predict need for intervention, population based studies have shown little change in rates of admission or intervention with increasing CT use.\textsuperscript{3,4} A 2002 study that prospectively compared CT to the combination of an abdominopelvic radiograph plus ultrasound found a higher sensitivity for CT but that the increased sensitivity “did not result in a change in treatment.” The study concluded that “primary CT causes an over-evaluation of these patients with minor colic… both diagnostic accuracy and patient treatment will be unchanged.”\textsuperscript{33} A secondary analysis of the large prospective multicenter randomized trial comparing ultrasound to CT as an initial imaging modality found that there was no significant difference between timing of urological intervention based on the initial imaging modality.\textsuperscript{27}

Reduced-Radiation Dose CT

While there is controversy around risk of radiation from CT, most practitioners, radiologists and organizations continue to adhere to the ALARA principle for radiation levels in diagnostic imaging. Statements from the ACEP, ACR and AUA recommend using reduced dose CT.\textsuperscript{5,10,34} Despite these recommendations, in 2011 to 2012 only 2% of studies for suspected renal colic were performed with an RDCT technique and as of 2015 to 2016 this percentage had risen to only 8%.\textsuperscript{35} There were 50 relevant articles on RDCT with evidence level one (n = 9), two (n = 27), three (n = 10) and n/a (n = 4). The definition of RDCT in the literature is variable. The consensus definition of RDCT by this group did not involve a specific numerical cutoff; rather it is defined as a protocol that is specific for detection of renal colic and utilizes lower radiation settings than an undifferentiated CT of the abdomen/pelvis at the same institution. It is important to understand that because of the high attenuation of kidney stones, they can be detected using lower radiation levels. Many institutions simply perform the same CT that they would on an undifferentiated abdominal pain but just without contrast. This misses an opportunity to lower dose further, particularly when kidney stone is likely and alternative diagnoses unlikely.\textsuperscript{36} This group specifically defines RDCT as “a CT protocol specific to renal colic that has radiation dose parameters that are lower than an undifferentiated CT abdomen/pelvis at the institution.”

The literature supports the diagnostic accuracy of RDCT compared with SCT. For ureteral stone identification sensitivities of RDCT range from 90% to 95% and specificities from 97% to 99%.\textsuperscript{37} RDCT is particularly sensitive for larger stones which are more likely to require intervention.\textsuperscript{33,38} RDCT is sufficiently sensitive and specific for alternate diagnoses.\textsuperscript{37,38,40} While some guidelines include a BMI cutoff for use of RDCT, the literature suggests that while CT settings may need to be adjusted, diagnostic accuracy can be maintained in obese patients using RDCT techniques.\textsuperscript{10,40,41} Using a definition of RDCT that does not involve a specific cutoff but rather a dedicated protocol using lower radiation (as suggested by our group) should be considered in high BMI patients.

Consensus Results

The level of consensus by round of voting as well as proportions of imaging modalities recommended are shown in table 2. By the final round of voting the group was able to reach perfect consensus in 15 of 29 vignettes (45%), with at least moderate consensus in all scenarios. CT continued to be recommended in 7 of 29 (24%) scenarios. All vignettes along with final level of consensus and voting by modality are shown in the supplementary table (https://www.jurology.com).

DISCUSSION

To our knowledge, this manuscript presents the first systematic, multispecialty, evidence-based consensus regarding imaging in renal colic. The evidence suggests that in many cases ultrasound, either radiology performed or point of care, may provide adequate diagnostic information to guide initial treatment. When CT is needed a reduced radiation approach should be used. Our recommendations are in line with ACR appropriateness criteria which suggest that CT is “usually appropriate” and ultrasound “may be appropriate” for acute onset of flank pain but provide further clarification on when approaches such as ultrasound may be more optimal.\textsuperscript{8}

We were able to reach at least moderate consensus in all scenarios. Question 1 (Q1) is the “base case” scenario: a 35-year-old male with a prior history of kidney stones who presents with typical symptoms and adequate pain relief. No respondents

| Table 2. Consensus level and imaging modality by round in the modified Delphi process. The number and percent of vignettes (out of 29 total) are shown. |
|-----------------|-----------------|-----------------|
| Level of Consensus | Round 1 | Round 2 | Round 3 |
| Perfect | 6 (21%) | 12 (29%) | 15 (52%) |
| Excellent | 3 (10%) | 6 (21%) | 8 (28%) |
| Good | 11 (38%) | 6 (21%) | 3 (10%) |
| Moderate | 7 (24%) | 4 (14%) | 3 (10%) |
| Not reached | 2 (7%) | 1 (3%) | 0 (0%) |
| No imaging | 13 (45%) | 13 (45%) | 13 (45%) |
| Ultrasound | 7 (24%) | 7 (24%) | 9 (31%) |
| CT | 7 (24%) | 8 (28%) | 7 (24%) |
| No consensus | 2 (7%) | 1 (3%) | 0 (0%) |
favored initial CT in this patient. While five recommended POCUS, there was perfect consensus that no additional imaging was needed regardless of the presence or absence of hydronephrosis after POCUS was performed (Q7, 8).

In the same scenario but without prior history of stone (Q4), there was perfect consensus for POCUS as the initial imaging modality, with excellent agreement for no additional imaging regardless of POCUS result (Q10, 11). It could be argued that if the next step regardless of the result of the POCUS (hydronephrosis or no hydronephrosis) is no further imaging, then POCUS is not needed. In group discussion, respondents felt that even if the POCUS result did not influence the decision to obtain a subsequent CT, it could provide a baseline for the patient in case he returned, as well as providing some information regarding prognosis and likelihood of spontaneous stone passage.

One respondent recommended RDCT regardless of POCUS result in the scenario when the patient had no prior history of stone (Q4). This respondent was an emergency physician and this response is highlighted because it is in line with a common teaching in emergency medicine that “every first time stone requires a CT.” However, 8 of the 9 members of the group (including all three urologists) did not feel a CT was required in this young patient with typical symptoms and relief of pain even if it was a “first time stone.”

When the clinical presentation of the case was less typical in a patient with a prior history of stones (Q13) there was still good consensus that ultrasound would be the best initial modality, though more respondents (2/9) favored CT. This suggests that as the presentation becomes less typical CT is favored. Performance of POCUS in a patient with less typical symptoms (Q14, Q15) did alter the recommendation for subsequent CT. While there was still moderate consensus for no further imaging in a young patient with atypical symptoms and no hydronephrosis (5/9, Q15), the absence of hydronephrosis on POCUS led four of nine respondents to recommend RDCT. This is in line with literature to suggest that POCUS may be most helpful in patients who have a moderate likelihood of stone, with presence of hydronephrosis suggesting stone and no need for further imaging, while absence of hydronephrosis suggests that an alternate diagnosis should be considered.

Questions 2-3, 5-6 and 17-18 included identical scenarios to questions 1, 4 and 16 with ages altered to 55 and 75 years old. For age 75, regardless of prior history of stones or presentation, there was good to perfect consensus that CT should be obtained, with RDCT favored (Q3, 6, 18). For age 55, when there was a history of a prior kidney stone and a typical presentation (Q2) there was moderate consensus (5/9) for no imaging with the remainder (4/9) recommending POCUS. For a 55 year old with no prior history of kidney stone there was perfect consensus that RDCT should be obtained (Q4), with perfect consensus that CT be obtained with a less typical presentation (Q17).

When the base case was varied to the female gender in a young patient with no prior history of stones (Q12), there was excellent consensus for ultrasound (8/9), with six respondents recommending POCUS. Notably two respondents chose RPUS (vs none with the male), which would be reasonable particularly if radiology ultrasound was definitive for stone size and location, allowing for avoidance of CT. RPUS may also be diagnostic of pelvic or adnexal pathology that could be present acutely in a female patient.

It should be noted that even in a young patient with a prior history of stones if symptoms cannot be appropriately relieved, RDCT was recommended with perfect consensus (Q9). What might constitute adequate relief of symptoms is a somewhat subjective determination. It is also acknowledged that attempting to achieve analgesia requires at least a short period of observation prior to deciding on imaging but in many cases a single dose of an intravenous nonsteroidal (such as ketorolac) with or without an opioid may achieve adequate analgesia. It is recommended that analgesia be attempted prior to a decision to obtain a CT examination. It is also acknowledged that if adequate analgesia cannot be obtained, RDCT may identify a stone requiring intervention.

There was excellent consensus that in a pregnant patient with typical symptoms RPUS was the preferred test (Q19 and 22). Regardless of presence of hydronephrosis on ultrasound (Q21 and 23), no respondents recommended CT as long as symptoms were relieved. Similarly, in pediatric patients (Q25-27) there was perfect consensus that ultrasound was the favored initial modality, with RPUS favored (7/9, Q25). In pediatric patients with a typical presentation in whom pain was relieved, there was excellent to perfect consensus that CT was not needed regardless of the presence or absence of hydronephrosis if symptoms were relieved.

There was perfect consensus that in a patient with stent placement for kidney stone POCUS was the preferred test (Q29). The presence of hydronephrosis suggests a nonfunctioning stent, while in the absence of hydronephrosis with adequate pain control the patient could be discharged without further imaging. For a patient with pain following lithotripsy there was good consensus that ultrasound should be performed with five respondents favoring RPUS. It was noted by the urologists that
RPUS would be more likely to identify a post-procedural hematoma, favoring RPUS over POCUS.

If CT is not initially performed it is expected that some patients will require subsequent CT if they are in the minority of patients with a stone that does not pass spontaneously. This should be explained to patients prior to discharge (suggested followup instructions, supplementary Appendix 5, https://www.jurology.com).

**Limitations**

There could be nearly unlimited permutations of the clinical presentation described in the vignettes. The decision to include 29 questions was felt to be the best balance of major factors with the number of questions the group felt was within a reasonable scope to address. However, not all clinical scenarios were included. For example, gender was only varied in 1 question (Q12 vs Q4) and there was not a lot of variation in the answers by gender, although it may be more reasonable to forego CT in a female than a male patient (slightly more risk of radiation and somewhat higher likelihood of finding an alternative diagnosis in the pelvis/adnexa) with ultrasound.

**Summary**

Based on systematic literature review with a strength of evidence rating and consensus determination by a modified Delphi method, our recommendations are that CT may be avoided in many common clinical scenarios when uncomplicated renal colic is suspected, summarized in the Take Home Points.

**TAKE HOME POINTS**

- For suspected uncomplicated kidney stone and adequate pain relief, even without prior history of kidney stone, CT can be avoided in younger patients (approximately 35 years old) with a presentation typical for kidney stone.
- In middle-aged patients (approximately 55 years old), CT may be avoided if there is a prior history of kidney stone.
- In older patients (approximately 75 years old), CT should generally be obtained.
- POCUS may help guide clinical suspicion and need for further imaging in patients with less typical signs and symptoms.
- Pregnant and pediatric patients should have RPUS as the initial imaging modality.

**ACKNOWLEDGMENT**

The authors would like to acknowledge the support of the American College of Emergency Physicians, specifically Drs. Jeremiah Schuur and Arjun Venkatesh, PIs, for the ACEP E-QUAL initiative.

**REFERENCES**


