

**Critical Review Form
Therapy**

PGY-4

Ultrasonography versus computed tomography for suspected nephrolithiasis. N Engl J Med. 2014 Sep 18;371(12):1100-10

Objectives: "To assess the effect of diagnostic imaging techniques [in ED patients with suspected ureterolithiasis] on patient outcomes." (p. 1101)

Methods: This multicenter, randomized trial was conducted at 15 academic emergency departments in the United States from October 2011 through February 2013. Patients aged 18 to 76 with flank or abdominal pain, in whom the treating physician planned to order diagnostic imaging to establish or rule out the diagnosis of kidney stones were eligible for enrollment. Patients were randomly assigned, in a 1:1:1 fashion, to one of 3 imaging modalities: 1) bedside ultrasonography performed by the emergency physician (point-of-care ultrasound [POCUS]); 2) ultrasonography performed by a radiologist; or 3) abdominal computed tomography (CT). All further treatment decisions were made by the treating physicians. Exclusion criteria included:

1. Patients considered at high risk for serious alternative diagnosis (e.g. bowel disorders, aortic aneurysm, appendicitis, acute cholecystitis)
2. Pregnancy
3. Severe obesity (> 129 kg in men and > 113 kg in women)
4. Single kidney, renal transplantation, or dialysis dependence.

The primary outcomes were 1) high-risk diagnoses with complications that could be related to delayed or missed diagnosis, 2) cumulative radiation exposure within 6 months of randomization, and 3) total cost. High-risk diagnoses were predefined as any of the following occurring within 30 days of the initial emergency department visit: ruptured abdominal aortic aneurysm, pneumonia with sepsis, appendicitis with rupture, diverticulitis with abscess or sepsis, bowel ischemic or perforation, renal infarction, renal stone with abscess, pyelonephritis with urosepsis or bacteremia, ovarian torsion with necrosis, or aortic dissection with ischemia. Secondary outcomes included serious adverse events, return emergency department visits and hospitalization, self-reported pain scores, and diagnostic accuracy for nephrolithiasis.

Follow-up occurred by structured telephone interview with patients at 3, 7, 30, 90, and 180 days after randomization and review of medical records at participating sites. The site principal investigator, the study principal

investigator, and the chair of the data and safety monitoring board performed adjudication of all serious adverse events (n = 466). The reference standard for diagnostic accuracy was either patient's observation of stone passage or surgical removal of a stone.

Out of 3100 eligible patients, a total of 2776 were randomized. Of these, 17 were excluded, resulting in 2759 total patients included. There 908 patients randomized to point-of-care ultrasonography, 893 to radiology ultrasonography, and 958 to CT. A confirmed stone diagnosis occurred in 34.5% of patients in the POCUS group, 31.2% in the radiology ultrasound group, and 32.7% in the CT group (p = 0.39). 40.7% and 27% of patients in the POCUS group and radiology US group, respectively, underwent CT during the ED visit.

Guide		Comments
I.	Are the results valid?	
A.	Did experimental and control groups begin the study with a similar prognosis (answer the questions posed below)?	
1.	Were patients randomized?	Yes. Patients were randomized in a 1:1:1 fashion to one of 3 initial imaging studies: point-of-care ultrasound, radiology ultrasound, or CT scan.
2.	Was randomization concealed (blinded)?	Yes. "Randomization was performed with the use of the RANUNI function in SAS software at the study website." (p. 1101)
3.	Were patients analyzed in the groups to which they were randomized?	Yes. Patients were analyzed based on the initial imaging study to which they were randomized, regardless of what additional imaging was performed (intention to treat analysis).
4.	Were patients in the treatment and control groups similar with respect to known prognostic factors?	Yes. Patients were similar with respect to age, gender, race, self-reported pain score, prior history of kidney stones, medical comorbidities, physical exam findings, and clinicians initial estimates of the likelihood of alternative diagnoses and kidney stone.
B.	Did experimental and control groups retain a similar prognosis after the study started (answer the questions posed below)?	

1.	Were patients aware of group allocation?	Yes. "Patients and providers were aware of the imaging method to which the patients had been assigned." (p. 1101) It seems unlikely that performance bias would affect outcomes.
2.	Were clinicians aware of group allocation?	Yes. "Patients and providers were aware of the imaging method to which the patients had been assigned." (p. 1101) It seems unlikely that performance bias would affect outcomes.
3.	Were outcome assessors aware of group allocation?	Uncertain. Patient interviews and medical record reviews were conducted to evaluate outcomes, and the authors do not specifically mention if personnel involved in these activities were blinded to group allocation. Three study investigators adjudicated serious adverse events, but again there is no mention that they were blinded to group allocation during the adjudication process. Observer bias could therefore have played a role outcome assessment.
4.	Was follow-up complete?	No, although follow-up was quite good. "A total of 113 patients (4.1%) were lost to follow-up, with no significant variation according to study group." (p. 1103)
II.	What are the results (answer the questions posed below)?	
1.	How large was the treatment effect?	<ul style="list-style-type: none"> • A high-risk diagnosis with complications within 30 days of randomization occurred in 6 patients in the POCUS group (0.7%), 3 patients in the radiology ultrasound group (0.3%), and 2 patients in the CT group (0.2%) with no statistically significant difference ($p = 0.30$). • The mean cumulative radiation dose was lower in patients in the POCUS and radiology ultrasound group than in the CT group (10.1 mSv and 9.3 mSv, respectively, vs. 17.2 mSv; $p < 0.001$). • A total of 316 patients suffered adverse events, with no significant difference in the number of patients in each group: 113 (12.4%) in the POCUS group, 96 (10.8%) in the radiology ultrasound group, and 107 (11.2%) in the CT group ($p = 0.50$). • The number of related adverse events was similar between groups: 3 (0.3%) in the POCUS group, 4 (0.4%) in the radiology ultrasound group, and 5 (0.5%) in the CT group ($p = 0.88$). There were no deaths felt to be related to participation in the study. • The median ED length of stay was longer in the radiology ultrasound than in the POCUS group or CT group (7.0 hrs

		<p>vs. 6.4 hrs and 6.3 hrs, respectively; $p < 0.001$).</p> <ul style="list-style-type: none"> • Rates of return ED visits, hospital admission following discharge from the ED, and self-reported pain scores were similar between groups. • Based on the final diagnosis from the index ED visit, the diagnostic accuracy of the various groups was found to be similar (see Table 1). <p>Table 1. Test characteristics based on final ED diagnosis (95% CI)</p> <table border="1"> <thead> <tr> <th></th> <th>POCUS (n = 908)</th> <th>Radiology US (n = 893)</th> <th>CT (n = 958)</th> <th>p value</th> </tr> </thead> <tbody> <tr> <td>Sensitivity</td> <td>85 (80-89)</td> <td>84 (79-89)</td> <td>86 (82-90)</td> <td>0.74</td> </tr> <tr> <td>Specificity</td> <td>50 (45-54)</td> <td>53 (49-57)</td> <td>53 (49-58)</td> <td>0.38</td> </tr> </tbody> </table> <ul style="list-style-type: none"> • Based on the initial imaging study, the diagnostic accuracy had a lower sensitivity and higher specificity than CT (see Table 2). <p>Table 1. Test characteristics based on initial imaging (95% CI)</p> <table border="1"> <thead> <tr> <th></th> <th>POCUS (n = 908)</th> <th>Radiology US (n = 893)</th> <th>CT (n = 958)</th> <th>p value</th> </tr> </thead> <tbody> <tr> <td>Sensitivity</td> <td>54 (48-60)</td> <td>57 (51-64)</td> <td>88 (84-92)</td> <td>< 0.001</td> </tr> <tr> <td>Specificity</td> <td>71 (67-75)</td> <td>73 (69-77)</td> <td>58 (55-62)</td> <td>< 0.001</td> </tr> </tbody> </table>		POCUS (n = 908)	Radiology US (n = 893)	CT (n = 958)	p value	Sensitivity	85 (80-89)	84 (79-89)	86 (82-90)	0.74	Specificity	50 (45-54)	53 (49-57)	53 (49-58)	0.38		POCUS (n = 908)	Radiology US (n = 893)	CT (n = 958)	p value	Sensitivity	54 (48-60)	57 (51-64)	88 (84-92)	< 0.001	Specificity	71 (67-75)	73 (69-77)	58 (55-62)	< 0.001
	POCUS (n = 908)	Radiology US (n = 893)	CT (n = 958)	p value																												
Sensitivity	85 (80-89)	84 (79-89)	86 (82-90)	0.74																												
Specificity	50 (45-54)	53 (49-57)	53 (49-58)	0.38																												
	POCUS (n = 908)	Radiology US (n = 893)	CT (n = 958)	p value																												
Sensitivity	54 (48-60)	57 (51-64)	88 (84-92)	< 0.001																												
Specificity	71 (67-75)	73 (69-77)	58 (55-62)	< 0.001																												
2.	How precise was the estimate of the treatment effect?	See above. No 95% confidence intervals were reported for the outcomes.																														
III.	How can I apply the results to patient care (answer the questions posed below)?																															
1.	Were the study patients similar to my patient?	Yes. These were patients with suspected ureteral colic who were felt to be at low risk of serious, acute alternate pathology. We have access to point-of-care bedside ultrasound and most of our faculty is qualified to perform such ultrasounds. Additionally, our institution was one of the study sites.																														
2.	Were all clinically important outcomes considered?	Yes. The authors considered radiation exposure, pain scores at intermittent follow-up, ED length of stay, return visits, and serious adverse events, including those specifically related to the initial imaging modality to which the patient was randomized.																														
3.	Are the likely treatment benefits worth the potential harm and costs?	Likely yes. This study demonstrated that use of POCUS or radiology US as the initial imaging modality in patients with suspected renal colic, felt to be at low risk of serious alternate pathology, resulted in decreased radiation exposure without																														

		increasing the risk of a serious adverse event within 180 days of randomization. The study is limited by lack of blinding and failure to employ low-dose CT scans in the evaluation of renal colic.
--	--	---

Limitations:

- 1. The chart review methods used to evaluate outcome were not detailed (see [Gilbert](#) and [Worster](#)).**
- 2. There is no mention of blinding of outcomes assessors, including those performing chart reviews, those performing patient interviews, or those adjudicating serious adverse events ([observer bias](#)).**
- 3. Patients were randomly assigned only during hours when all three imaging techniques were feasible, limiting the ability to universally apply the results.**
- 4. The reference standard used to determine the prevalence of kidney stones and calculate diagnostic test characteristics was the self-reported passage of stone or surgical removal of a stone. This reference standard has not been previously for evaluated in the literature, and likely resulted in an under-reporting of stones, and hence an inflated false-positive rate. As a result, the reported specificity of CT (53%) is much lower than that reported in literature ([Sheafor 2000](#), [Eray 2003](#), [Niemann 2008](#)).**
- 5. The mean dose of radiation during the index ED visit in the CT group was 14.1 mSv, reflecting very low rates of use of low-dose CT, despite the high accuracy reported for [CT scans delivering a dose of < 3 mSv](#) in the diagnosis of ureteral stones.**

Bottom Line:

This randomized three-arm study demonstrated that use of point-of-care US, performed by an emergency physician, or radiology US as the initial imaging modality in patients with suspected renal colic, felt to be at low risk of serious alternate pathology, resulted in decreased radiation exposure compared to CT without increasing the risk of a serious adverse event within 180 days of randomization. The study was limited by lack of blinding and failure to employ [low-dose CT](#) scans in the evaluation of renal colic. Given the results, it seems reasonable to use US as the initial imaging modality in patients with suspected renal colic, assuming the patient is felt to be at low risk of a serious alternative pathology.