Hydronephrosis severity clarifies prognosis and guides management for emergency department patients with acute ureteral colic

Grant D. Innes¹ · Frank X. Scheuermeyer² · Andrew D. McRae¹ · Joel M. H. Teichman³ · Daniel J. Lane⁴

Received: 15 March 2021 / Accepted: 17 June 2021 / Published online: 25 July 2021

© The Author(s), under exclusive licence to Canadian Association of Emergency Physicians (CAEP)/ Association Canadienne de Médecine d'Urgence (ACMU) 2021

Abstract

Objective In emergency department patients with ureteral colic, the prognostic value of hydronephrosis is unclear. Our goal was to determine whether hydronephrosis can differentiate low-risk patients appropriate for trial of spontaneous passage from those with clinically important stones likely to experience passage failure.

Methods We used administrative data and structured chart review to evaluate a consecutive cohort of patients with ureteral stones who had a CT at nine Canadian hospitals in two cities. We used CT, the gold standard for stone imaging, to assess hydronephrosis and stone size. We described classification accuracy of hydronephrosis severity for detecting large (≥ 5 mm) stones. In patients attempting spontaneous passage we used hierarchical Bayesian regression to determine the association of hydronephrosis with passage failure, defined by the need for rescue intervention within 60 days. To illustrate prognostic utility, we reported pre-test probability of passage failure among all eligible patients (without hydronephrosis guidance) to post-test probability of passage failure in each hydronephrosis group.

Results Of 3251 patients, 70% male and mean age 51, 38% had a large stone, including 23%, 29%, 53% and 72% with absent, mild, moderate and severe hydronephrosis. Passage failure rates were 15%, 20%, 28% and 43% in the respective hydronephrosis categories, and 23% overall. “Absent or mild” hydronephrosis identified a large subset of patients (64%) with low passage failure rates. Moderate hydronephrosis predicted slightly higher, and severe hydronephrosis substantially higher passage failure risk.

Conclusions Absent and mild hydronephrosis identify low-risk patients unlikely to experience passage failure, who may be appropriate for trial of spontaneous passage without CT imaging. Moderate hydronephrosis is weakly associated with larger stones but not with significantly greater passage failure. Severe hydronephrosis is an important finding that warrants definitive imaging and referral. Differentiating “moderate-severe” from “absent-mild” hydronephrosis provides risk stratification value. More granular hydronephrosis grading is not prognostically helpful.

Keywords Renal colic · Ureteral colic · Hydronephrosis · Ureteral calculi · Imaging · Diagnosis

Meetings This study has not been previously presented.

✉ Grant D. Innes
grant.innes@ahs.ca

¹ Department of Emergency Medicine and Department of Community Health Sciences, University of Calgary, Calgary, AB, Canada

² Department of Emergency Medicine, St Paul's Hospital, University of British Columbia, Vancouver, BC, Canada

³ Department of Urologic Sciences, University of British Columbia, Vancouver, BC, Canada

⁴ Faculty of Medicine, University of Calgary, Calgary, AB, Canada

Résumé

Objectif Chez les patients des services d'urgence (SU) atteints de colique urétérale, la valeur pronostique de l'hydronéphrose n'est pas claire. Notre objectif était de déterminer si l'hydronéphrose peut différencier les patients à faible risque appropriés pour l'essai de passage spontané de ceux qui ont des calculs cliniquement importants susceptibles de subir un échec de passage.

Méthodes Nous avons utilisé des données administratives et un examen structuré des dossiers pour évaluer une cohorte consécutive de patients atteints de calculs urétraux qui avaient subi une tomodensitométrie dans neuf hôpitaux canadiens de deux villes. Nous avons utilisé la tomodensitométrie, l'étalon-or pour l'imagerie des calculs, pour évaluer l'hydronéphrose et la taille des calculs. Nous avons décrit la précision de la classification de la gravité de l'hydronéphrose pour la détection de gros calculs (> 5 mm). Chez les patients tentant un passage spontané, nous avons utilisé la régression bayésienne hiérarchique pour déterminer l'association de l'hydronéphrose avec l'échec du passage, défini par le besoin d'intervention de sauvetage dans les 60 jours. Pour illustrer l'utilité pronostique, nous avons signalé la probabilité d'échec de passage avant le test chez tous les patients admissibles (sans directives sur l'hydronéphrose) à la probabilité d'échec de passage post-test dans chaque groupe d'hydronéphrose.

Résultats Sur 3251 patients, 70% d'hommes et d'âge moyen 51 ans, 38% avaient un gros calcul, dont 23%, 29%, 53% et 72% avec une hydronéphrose absente, légère, modérée et sévère. Les taux d'échec au passage étaient de 15%, 20%, 28% et 43% dans les catégories d'hydronéphrose respectives et de 23% dans l'ensemble. L'hydronéphrose « absente ou légère » a permis d'identifier un sous-ensemble important de patients (64%) présentant de faibles taux d'échec au passage. Une hydronéphrose modérée prédisait un risque d'échec de passage légèrement plus élevé, et une hydronéphrose sévère un risque sensiblement plus élevé.

Conclusions L'absence d'hydronéphrose et une hydronéphrose légère permettent d'identifier les patients à faible risque, peu susceptibles d'avoir un échec de passage, qui peuvent être appropriés pour un essai de passage spontané sans imagerie CT. Une hydronéphrose modérée est faiblement associée à des calculs plus gros mais pas à un échec de passage significativement plus important. L'hydronéphrose sévère est une constatation importante qui justifie une imagerie définitive et une référence. Différencier l'hydronéphrose « modérée-sévère » de l'« absente-légère » fournit une valeur de stratification du risque. Un classement plus granulaire de l'hydronéphrose n'est pas utile sur le plan pronostique.

Clinician's capsule

What is known about the topic?

The prognostic implications of mild, moderate and severe hydronephrosis are unclear.

What did this study ask?

Can hydronephrosis severity identify patients likely to experience passage failure or have large ureteral stones?

What did this study find?

Absent and mild hydronephrosis reflect favorable prognosis; moderate hydronephrosis has minimal prognostic value, and severe hydronephrosis warrants definitive imaging and referral.

Why does this study matter to clinicians?

Absent-mild hydronephrosis portend low risk (consider trial of passage without CT); more specific hydronephrosis grading is not prognostically helpful.

Background

Renal colic can be diagnosed with relative confidence based on clinical presentation and decision tools like the STONE score [1]. Hydronephrosis on ultrasound increases certainty but seldom confirms stone size and location, which are key predictors of passage success and management approach [2–5]. Stones smaller than 5 mm usually pass spontaneously while larger stones may benefit from early intervention or alpha-blockers [6, 7], but computed tomography (CT) is the only modality that reliably provides this information [8]. Without clues to prognosis, decisions to recommend spontaneous passage, to refer for intervention, or to risk stratify using CT are difficult [2–4]. Despite advances in emergency department point of care ultrasound (ED POCUS), CT utilization has continued rising. Up to 83% of patients now undergo CT imaging during an acute episode [9].

Although hydronephrosis is widely used by emergency physicians to guide decision making, its association with passage failure and stone size are unclear [1, 10–15]. Some researchers suggest that more severe hydronephrosis predicts the need for future intervention [1, 12, 13] while others

report no association with outcomes [14, 15]. Studies suggesting hydronephrosis is prognostically useful are limited by small sample size [13], aggregating patients with and without stones [1], or mixing kidney and ureteral stones, which have differing prognosis and management [12]. Several studies report associations between hydronephrosis and stone size, but these relationships are generally weak [1], and classification accuracy appears insufficient to influence management decisions [10, 11, 14].

If hydronephrosis severity is strongly associated with subsequent passage failure, this could help physicians differentiate low-risk patients appropriate for trial of passage from high-risk patients requiring immediate CT to clarify management approach. This would be particularly helpful for settings with limited imaging capability and stone formers at risk of multiple CTs. If hydronephrosis predicts patient morbidity or the presence of large potentially problematic stones, this identifies a subset of patients who might benefit from alpha-blockers or expedited surgical referral. Our main objective was to determine whether hydronephrosis can differentiate patients at low- versus high-risk of spontaneous passage failure. Secondary objectives were to determine whether hydronephrosis signals the presence of large stones ≥ 5 mm or predicts patient-initiated ED revisits and hospitalizations (morbidity markers).

Methods

Design and setting

This is a secondary analysis of a recently published comparative study of renal colic management in two cities [16]. The methods are previously described but, in brief, we used

administrative data and structured chart review to analyze all ED renal colic visits to nine hospitals in the Calgary Health Zone and Vancouver Coastal Health Region over a 1-year period [16]. To best evaluate the relationship between hydronephrosis and outcomes, we analyzed CT rather than POCUS findings. CT is the preferred diagnostic standard because of reliable image capture and interpretation [5]. Our analyses will therefore advance physician understanding of CT findings but are also relevant to POCUS interpretation because of high agreement between the two modalities for moderate and severe hydronephrosis [20, 21].

Participants

We identified all patients with an ED diagnosis of renal colic based on ICD-10 codes N200, N201, N202, N132, N23 and N209, and included those who had CT confirmed ureteral stones. We excluded patients with missing stone size or hydronephrosis descriptors, and those who had a preceding ED renal colic visit within 30 days, to avoid studying patients already failing conservative management. We also excluded patients with isolated renal stones, patients referred directly to a urologist, and patients with out-of-region postal codes, who could have experienced outcomes outside the study regions.

Data

We gathered patient demographics, arrival mode, index disposition, ED revisits, hospitalizations and procedures from hospital databases. Radiologists at all nine study sites graded CT images using the same standard approach of absent, mild, moderate or severe hydronephrosis. Trained research assistants transcribed radiologist-reported CT findings for

Fig. 1 Study Flowsheet. *ED* emergency department, *CT* computed tomography

stone size, location and hydronephrosis severity. We used stone width, the largest cross-sectional diameter, as the stone size for analyses. To assess data abstraction reliability, *particularly for hydronephrosis severity*, two research assistants blinded to study hypothesis, patient management and outcomes, independently reviewed 100 consecutive imaging reports. We used Cohen's kappa to describe interobserver agreement. Data quality was audited and assured by the Alberta Health Services Strategic Data Analytics unit and the Vancouver Coastal Health Decision Support Unit.

Outcomes

The primary outcome was failure of spontaneous passage, defined by the need for lithotripsy, ureteral stenting or ureteroscopic intervention within 60 days after index discharge. Because intervention is largely physician-driven, we also collated patient-initiated emergency revisits and readmissions as secondary outcomes reflecting morbidity [16, 17]. We analyzed stone size as a dichotomous measure smaller or

larger than 5 mm. A 60-day outcome window was selected based on research showing that almost all relevant outcomes occur within this time [16].

Analysis

We compared patients with absent, mild, moderate and severe hydronephrosis using descriptive statistics with standardized mean differences. To evaluate prognosis, we used hierarchical Bayesian regression and determined the association of hydronephrosis severity with passage failure (rescue intervention). We explicitly modeled baseline differences by city to account for the effect of practice variation on outcomes [18]. We further adjusted for patient sex and age, using a spline function, to isolate the effects of hydronephrosis independent of patient characteristics. To assess the proportion of patients correctly identified as higher risk, we determined area under the curve for predictive discrimination. We used prior predictive simulations to select minimally informative priors, allowing for the data to inform

Table 1 Patient characteristics and main outcomes stratified by hydronephrosis severity

	All patients	Absent	Mild	Moderate	Severe	SMD
1a. Characteristics and index visit events by hydronephrosis category among all eligible patients (unadjusted)						
Hydronephrosis severity: N (%)	3251	417 (13)	1667 (51)	1033 (32)	134 (4)	
EMS arrival	18%	15%	17%	20%	22%	0.10
Male sex	70%	68%	75%	67%	52%	0.25
Mean age	51	47	50	53	56	0.30
Median heart rate	78	77	77	78	81	0.16
Median systolic BP (mmHg)	142	139	143	142	145	0.16
Median creatinine level	92	82	93	95	96	0.26
CTAS Triage Acuity						0.17
1–2 (emergent)	46%	39%	47%	48%	49%	
3–5 (less urgent)	54%	61%	53%	52%	51%	
Median stone width (mm)	4.0	3.0	4.0	5.0	5.8	0.54
Large stone ≥ 5 mm	38%	23%	29%	53%	72%	0.65
Stone location						0.24
Distal ureter	64%	75%	64%	60%	56%	
Middle-ureter	10%	8%	9%	12%	14%	
Proximal ureter	26%	17%	27%	28%	30%	
Index visit outcomes						
Admission	37%	26%	30%	50%	63%	0.47
Stone intervention	34%	23%	27%	47%	58%	0.45
1b. Crude 60-day outcomes by hydronephrosis category in the spontaneous passage (prognosis) cohort: N = 2148						
N (%)	2148	322 (15)	1221 (57)	549 (26)	56 (2.6)	
ED revisit	25%	27%	25%	25%	27%	0.03
Hospitalization	11%	12%	10%	12%	14%	0.07
Intervention	22%	14%	19%	29%	45%	0.39

CTAS is the Canadian Triage and Acuity Scale. SMD is standardized mean difference between groups
Bold values > 0.2 reflect substantial group imbalance

Table 2 Diagnostic performance of hydronephrosis using three decision thresholds

Outcome	Passage failure	Large stone
Perform CT if any hydronephrosis (N = 2834)		
Sensitivity	90%	95%
Specificity	16%	17%
Likelihood ratio (+)	1.07	1.14
Likelihood ratio (−)	0.63	0.29
Perform CT if moderate-severe hydronephrosis (N = 1167)		
Sensitivity	35%	59%
Specificity	74%	76%
Likelihood ratio (+)	1.35	2.46
Likelihood ratio (−)	0.89	0.54
Perform CT if moderate-severe hydronephrosis (N = 134)		
Sensitivity	5%	9%
Specificity	98%	99%
Likelihood ratio (+)	2.5	9.0
Likelihood ratio (−)	0.97	0.92

Fig. 2 Outcomes by hydronephrosis category and using different “high-risk” cutoffs. Top left box shows outcome risk prior to knowledge of hydronephrosis severity (pre-test likelihood). Lower left boxes show post-test likelihood conditional on hydronephrosis cat-

egory. Right boxes show post-test likelihood and change (Δ) in outcome likelihood using different positivity thresholds for hydronephrosis

all estimates. We also calculated sensitivity, specificity and likelihood ratios (LR) for three hydronephrosis decision thresholds in detecting large stones and predicting patients who experienced passage failure.

For stone size analyses we included all patients (Fig. 1). For passage failure analyses we excluded patients who

underwent index stone intervention, because they did not have a trial of passage. We defined *pre-test probability* as the passage failure rate in the overall cohort, prior to knowledge of hydronephrosis severity. We defined *post-test probability* as the passage failure rate within each strata (i.e., failure rate conditional on hydronephrosis severity). We used the same

Fig. 3 Predicted effect of hydronephrosis guidance on CT rates and passage failure. With hydronephrosis-guidance, $\leq 36\%$ would have index CT. 12% would fail passage and require delayed CT despite low-risk hydronephrosis findings. This model does not consider

patients who undergo immediate CT because of ED treatment failure with intractable symptoms, so actual CT rates may be higher and spontaneous passage failure rates lower than suggested here

approach to define pre- and post-test probability of large stone. We reported the probability of a large stone using each of the three possible test thresholds: “any hydronephrosis”, “moderate or severe” hydronephrosis, and “severe hydronephrosis”. After determining that absent and mild hydronephrosis were associated with low rates of passage failure, we used study data to model a risk stratification algorithm in which patients with severe hydronephrosis undergo CT, those with moderate hydronephrosis are considered for possible CT, and those with absent or mild hydronephrosis are discharged for a trial of passage without CT.

We used R for all analyses, including the ‘tableone’ package for descriptive statistics and the ‘brms’ packages with Stan for Bayesian modeling. We adhered to the ROBUST checklist for Bayesian analysis, the STROBE checklist for observational studies, and the STARD checklist for diagnostic tests. This study was approved by Research Ethics Boards at the University of Calgary and University of British Columbia.

Results

Among 3251 consecutive patients with confirmed ureteral stones (Table 1a), mean age was 51 and 70% were male. Hydronephrosis was absent in 13%, mild in 51%, moderate in 32% and severe in 4%. The conditional probability of a large stone was 23%, 29%, 53% and 72% for patients with absent, mild, moderate and severe hydronephrosis respectively. Hydronephrosis was not associated with patient-initiated ED revisits or post-index hospitalization (morbidity outcomes), which were similar across all hydronephrosis

categories (Table 1b). Interobserver agreement on stone characteristics was excellent, with kappa values of 0.97 (length), 0.92 (width), 0.95 (location) and 0.90 (hydronephrosis severity). Table 2 shows that a threshold of “any hydronephrosis” provides poor specificity. “Severe hydronephrosis” is specific but insensitive. A threshold of “moderate-severe” hydronephrosis was 74% specific and 35% sensitive for passage failure.

Among 2148 patients attempting spontaneous passage, adjusted passage failure rates (IQR) were 15% (13–17%), 20% (18–21%), 28% (18–37%) and 43% (36–50%) for patients with absent, mild, moderate or severe hydronephrosis (Fig. 2, left). Figure 2 (right) provides a Bayesian visualization of prognostic utility, illustrating the difference between pre- and post-test likelihood of outcome, using each possible hydronephrosis cutoff. The presence of “any hydronephrosis” reflected a 1% increase in passage failure risk while its absence reflected an 8% decrease. “Moderate-severe” hydronephrosis reflected a 14% higher passage failure risk, from 23 to 37%, while its absence reflected a 5% decrease to 18%. Severe hydronephrosis predicted a 20% higher passage failure risk from 23 to 43%, but only 5% of passage failure patients had severe hydronephrosis.

Figure 3 illustrates the likely effect of using hydronephrosis to guide CT and disposition decisions. In a similar population, 64% would have absent or mild hydronephrosis, qualifying them for trial of passage without CT. Four percent would have severe hydronephrosis and require immediate CT, while 32% would have moderate hydronephrosis and be considered for CT. If all those with moderate hydronephrosis underwent CT (most cautious approach), 48% overall would undergo CT, a substantial reduction from the 100% actually

imaged in this cohort and the ~80% imaged in other studies [9]. Using this approach, 12% would fail spontaneous passage and return for intervention despite low-risk hydronephrosis findings.

Discussion

Interpretation

Many have studied hydronephrosis in renal colic diagnosis. We studied prognosis, asking the question: “Given a diagnosis of ureteral colic, does hydronephrosis severity differentiate low-risk patients appropriate for a trial of passage from higher-risk patients requiring early CT or surgical referral?” We found that absent and mild hydronephrosis carry a favorable prognosis, while severe hydronephrosis predicts a high risk of passage failure justifying CT imaging and expedited urology referral. Moderate hydronephrosis, which increased the likelihood of passage failure from 23% (pretest) to 28%, does not mandate CT imaging but should prompt consideration of other factors including response to ED management, underlying comorbidities and renal function, history of recurrent stones and previous CTs, prior successful stone passages, availability of follow-up, and patient preferences.

Previous studies

Research into the link between hydronephrosis and outcomes is limited. Fields reported that hospitalization was more likely with moderate than mild hydronephrosis (36% vs. 24%), but studied only 11 patients with moderate hydronephrosis [13]. Leo found that no level of hydronephrosis predicted 30-day outcomes, but studied only 37 patients with moderate-severe hydronephrosis [14]. Taylor determined that “moderate-severe” hydronephrosis was 37% sensitive and 88% specific for subsequent intervention, but most patients had parenchymal stones, which differ in management and prognosis [12]. Daniels studied 835 patients with *suspected* ureteric colic and reported that “moderate-severe” hydronephrosis was 36% sensitive and 86% specific for predicting 90-day intervention; however, 47% of patients did not have stones and were not at risk of intervention [1]. In the subset of patients *with* stones, hydronephrosis was not useful in predicting intervention [15].

Most studies addressing hydronephrosis and stone size have small sample sizes. Moak [10] and Goertz [11] found that hydronephrosis is weakly associated with stone size, but studied only 10 and 33 patients with stones ≥ 5 mm. Riddell found that “any hydronephrosis” was 90% sensitive for 60 stones > 6 mm, but did not provide specificity estimates or discriminate hydronephrosis categories [19]. Daniels

determined that 5% of patients without hydronephrosis had stones > 5 mm, but this excellent negative predictive value largely reflected low stone prevalence (53%) in the study population [1]. Daniels also found that 17% of patients with “any hydronephrosis” and 28% with “moderate-severe” hydronephrosis had large stones, but did not differentiate by hydronephrosis category. We stratified outcomes by absent, mild, moderate and severe hydronephrosis to help physicians interpret the significance of individual patient findings.

Strengths and limitations

We evaluated a large consecutive multicentre cohort, including 1223 patients with large stones and 1168 with moderate or severe hydronephrosis. This enabled precise outcome estimates within subgroups. We studied hydronephrosis based on CT rather than POCUS findings because CT is a superior reference standard with more reliable image capture and interpretation. Our data clarify the relationship between hydronephrosis and outcomes but translate to POCUS only if these modalities provide similar interpretations. Prior studies report excellent agreement between POCUS and CT for the presence of hydronephrosis ($\kappa=0.87$) [20]. ED POCUS is 73–87% sensitive and 66–83% specific for hydronephrosis on CT [20–23], but substantially better (~95% specific) for moderate-severe hydronephrosis [20, 21]. Pathan reported that 97% of false-negative POCUS interpretations involved failure to identify mild hydronephrosis [21], which based on our results is not a marker of adverse outcome hence not an important finding. These and other authors have concluded that ED POCUS should prioritize the recognition of moderate-severe hydronephrosis, a conclusion congruent with ours [20, 21].

We studied patients with CT-confirmed ureteral stones, the population that CT-reduction strategies should target. Our findings do not apply to very-low risk patients who did not undergo CT, or to undifferentiated patients with flank pain. Because we studied renal colic—not suspected renal colic—we cannot address the issue of incidental CT findings that lead to other diagnoses. Radiologist interpretations of hydronephrosis severity are subjective; however, our studied approach mitigates interpretive variability because it does not require agreement on each hydronephrosis category. Rather, it requires only that radiologists agree if moderate-severe hydronephrosis is present. Although radiology interpretations are imperfect, they are the source of truth ED physicians depend on in practice.

Differing interventional tendency by city posed a challenge because patients could experience an *intervention* despite differing illness severity; however this was also a strength because it enabled robust statistical modelling across the risk continuum. To address practice variability we performed Bayesian hierarchical regression analyses,

clustering by city to explicitly isolate the independent association of hydronephrosis with passage failure. We cannot account for some important determinants including the use of medical expulsive therapy, adequacy of analgesia, and availability of follow-up care.

Clinical implications

In patients responding favorably to ED management, physicians should consider a risk stratification approach differentiating absent-mild hydronephrosis from moderate-severe hydronephrosis (Fig. 3). Two-thirds of patients had findings of absent-mild hydronephrosis and are likely appropriate for trial of passage without CT. 4% had severe hydronephrosis mandating immediate CT. One third had moderate hydronephrosis, which carried a 28% passage failure rate. Balancing passage failure against radiation risk, moderate hydronephrosis suggests a selective approach incorporating patient preferences, physician risk tolerance and other factors described above. This decision model suggests the need for CT in 4–36% of patients, substantially less than the ~80% reported elsewhere. Using this approach, 12% fell into a false negative group who failed passage despite low risk hydronephrosis findings, but 12% compares favorably with published 17–25% revisit rates for renal colic patients [16, 24]. Our results suggest that the ability to recognize “moderate-severe” hydronephrosis is prognostically important but that more specific categorical grading is not. This binary model for ED POCUS interpretation, requiring only that physicians recognize moderate and severe hydronephrosis, provides a simple and more reliable approach to CT decisions, particularly if imaging access is limited.

Research implications

Hydronephrosis guidance for CT imaging could improve risk stratification and reduce CT utilization, but this strategy requires validation, particularly as relates to POCUS interpretations. Our modeled estimates for passage failure and CT utilization did not consider patients who underwent ED CT because of intractable symptoms; therefore actual CT rates may prove higher and passage failure rates lower than estimated.

Conclusion

Absent and mild hydronephrosis identify patients unlikely to have passage failure or large stones. Moderate hydronephrosis is weakly associated with larger stones but not with significantly greater passage failure. Severe hydronephrosis should trigger definitive imaging and referral. Recognizing “moderate-severe” hydronephrosis provides modest risk

stratification value but more granular hydronephrosis grading is not prognostically helpful.

Author contributions *GDI* conceived and designed the study, acquired research funding, and drafted the manuscript. *FXS* performed the literature search, participated in study design, and assisted with manuscript revisions. *ADM* helped guide development of the design and analysis, and assisted with manuscript revision. *JMHT* was Vancouver Urology study site liaison and contributed to manuscript revisions. *DJL* designed and performed the statistical analyses, and helped revise the draft manuscript.

Funding This study was funded by the *MSI Foundation* (a registered charity and research funding agency under the arms-length oversight of the Alberta College of Physicians and Surgeons).

Declarations

Conflict of interest Dr. Teichman has received grants and personal fees from Boston Scientific, grants from Cook Urologic, personal fees from Urigen and non-financial support from Innova Quartz, although none were related to this research. None of the other investigators have any potential conflicts to report.

References

1. Daniels B, Gross CP, Molinaro A, et al. STONE PLUS: evaluation of emergency department patients with suspected renal colic using a clinical prediction tool combined with point-of-care limited ultrasonography. *Ann Emerg Med*. 2016;67:439–48.
2. Ray AA, Ghiculete D, Pace KT, Honey RJD. Limitations to ultrasound in the detection and measurement of urinary tract calculi. *Urology*. 2010;76:295–300.
3. Yilmaz S, Sindel T, Arslan G, et al. Renal colic: comparison of spiral CT, US and IVU in the detection of ureteral calculi. *Eur Radiol*. 1998;8:212–7.
4. Fowler KA, Locken JA, Duchesne JH, Williamson MR. Ultrasound for detecting renal calculi with nonenhanced CT as a reference standard. *Radiology*. 2002;222:109–13.
5. Wong C, Teitge B, Ross M, et al. The accuracy and prognostic value of point-of-care ultrasound for nephrolithiasis in the emergency department: a systematic review and meta-analysis. *Acad Emerg Med*. 2018;25:684–98.
6. Innes GD, Scheuermeyer FX, McRae AD, et al. Which patients should have early surgical intervention for acute ureteric colic. *J Urol*. 2021;205:152–8.
7. Hollingsworth JM, Canales BK, Rogers M, et al. Alpha blockers for treatment of ureteric stones: systematic review and meta-analysis. *BMJ*. 2016;355:i6112.
8. National Institute for Health and Care Excellence. Renal and ureteric stones: assessment and management. In: *Imaging for diagnosis*. Diagnostic evidence review. London: National Institute for Health and Care Excellence; 2019.
9. Schoenfeld EM, Pekow PS, Shieh MS, et al. The diagnosis and management of patients with renal colic across a sample of US hospitals: high CT utilization despite low rates of admission and inpatient urologic intervention. *PLoS ONE*. 2017;12:e0169160.
10. Moak JH, Lyons MS, Lindsell CJ. Bedside renal ultrasound in the evaluation of suspected ureterolithiasis. *Am J Emerg Med*. 2012;30:218–21.

11. Goertz JK, Lotterman S. Can the degree of hydronephrosis on ultrasound predict kidney stone size? *Am J Emerg Med.* 2010;28:813–6.
12. Taylor M, Woo MY, Pageau P, et al. Ultrasonography for the prediction of urological surgical intervention in patients with renal colic. *Emerg Med J.* 2016;33:118–23.
13. Fields J, Fischer J, Anderson K, et al. The ability of renal ultrasound and ureteral jet evaluation to predict 30-day outcomes in patients with suspected nephrolithiasis. *Am J Emerg Med.* 2015;33:1402–6.
14. Leo M, Langlois BK, Pare J, et al. Ultrasound vs computed tomography for severity of hydronephrosis and its Importance in renal colic. *West J Emerg Med.* 2017;18:559–68.
15. Daniels B, Schoenfeld E, Taylor A. Predictors of hospital admission and urologic intervention in adult emergency department patients with CT-confirmed ureteral stones. *J Urol.* 2017;198:1359–66.
16. Innes G, McRae A, Grafstein E, et al. Variability of renal colic management and outcomes in two Canadian cities. *Can J Emerg Med.* 2018;20:702–12.
17. Ordon M, Urbach D, Mamdani M, et al. The surgical management of kidney stone disease: a population based time series analysis. *J Urol.* 2014;192:1450–6.
18. McGlothlin AE, Viele K. Bayesian hierarchical models. *JAMA.* 2018;320:2365–6.
19. Riddell J, Case A, Wopat R, et al. Sensitivity of emergency department ultrasound to detect hydronephrosis in patients with CTR-proven stones. *West J Emerg Med.* 2014;15:96–100.
20. Herbst MK, Rosenberg G, Daniels B, et al. Effect of provider experience on clinician-performed ultrasonography for hydronephrosis in patients with suspected renal colic. *Ann Emerg Med.* 2014;64:269–76.
21. Pathan SA, Mitra B, Mirza S, et al. Emergency physician interpretation of point-of-care ultrasound for identifying and grading of hydronephrosis in renal colic compared with consensus interpretation by emergency radiologists. *Acad Emerg Med.* 2018;25:1129–37.
22. Watkins S, Bowra J, Sharma P, et al. Validation of emergency physician ultrasound in diagnosing hydronephrosis in ureteric colic. *Emerg Med Australas.* 2007;19(3):188–95.
23. Gaspari RJ, Horst K. Emergency ultrasound and urinalysis in the evaluation of flank pain. *Acad Emerg Med.* 2005;12:1180–4.
24. Schoenfeld EM, Shieh MS, Pekow P, et al. Association of patient and visit characteristics with rate and timing of urologic procedures for patients discharged from the emergency department with renal colic. *JAMA Netw Open.* 2019;2(12):e1916454.