National Institute for Health and Care Excellence

Consultation

Renal and ureteric stones: assessment and management

Imaging for diagnosis

NICE guideline

Diagnostic evidence review

July 2018

Consultation

This evidence review was developed by the National Guideline Centre

Contents

1 Imaging for diagnosis of renal and ureteric stones

1.1 Review question: In people with suspected (or under investigation for) renal and ureteric stones, how accurate is ultrasound, plain abdominal radiograph or MRI to identify whether a renal or ureteric stone is present, as indicated by the reference standard, non-contrast CT?

1.2 Introduction

Imaging which provides an accurate and timely diagnosis of a stone in a patient presenting with acute renal colic is essential to manage the patient in the most appropriate way. An accurate diagnosis is essential as the site and size of the stone and anatomical features of the patient are important in defining the most appropriate treatment options. There are a variety of imaging modalities used to assess patients with suspected renal colic including ultra sound, CT scanning with radiation and MRI scanning. CT is more expensive than ultrasound but the extra cost may be outweighed by avoiding additional investigations if the first test misses the diagnosis. There is uncertainty about which imaging modality should be the first line investigation in the acute setting of suspected renal colic. Similarly there are concerns regarding radiation doses in certain groups, children and pregnant women and the question will address the most suitable imaging test in these groups.

1.3 PICO table

For full details see the review protocol in appendix A.

Table 1: PICO characteristics of review question

Population	People (adults, children and young people) with suspected (or under investigation for) renal and ureteric stones
Target condition	Renal and ureteric stones
Index tests	• Plain abdominal radiograph (conventional, KUB) • Ultrasound • Magnetic resonance imaging
Reference standard	• Non contrast computed tomography
Statistical measures	• Specificity • Sensitivity • Positive Predictive Value • Negative Predictive Value
Study design	Cross-sectional studies, cohort studies Case–control and case series studies should be included only if there is no other evidence

1.4 Clinical evidence

1.4.1 Included studies

A search was conducted for prospective and retrospective cohort studies assessing the diagnostic test accuracy of ultrasound, MRI or plain abdominal radiograph to identify whether the condition is present (as indicated by the reference standard) in people under investigation for renal and ureteric stones.

Thirteen studies were included in the review;[16, 18, 20, 42, 52-54, 57, 63, 95, 105, 119, 124] these are summarised in Table 2 below. Evidence from these studies is summarised in the clinical evidence summary below (Table 3).

See also the study selection flow chart in appendix C, sensitivity and specificity forest plots and receiver operating characteristics (ROC) curves in appendix E, and study evidence tables in appendix D.

1.4.2 Excluded studies

See the excluded studies list in appendix H.

1.4.3 Summary of clinical studies included in the evidence review

Table 2: Summary of studies included in the evidence review

Study	Population	Target condition	Index test	Reference standard	Comments
Chan 2008[16]	n=100 People being investigated for suspected urolithiasis Ireland	Urolithiasis	Plain abdominal radiograph (KUB)	Unenhanced CT	Diagnosing number of patients with stones
Cifci 2016[18]	n=159 People who were admitted to the urology department with suspected acute urinary calculi (flank pain, hematuria or patients with a history of urinary calculi) Turkey	Ureteral calculi	MRI (MRU with a B-TFE sequence)	Unenhanced CT	Diagnosing number of patients with stones Calculi located in the kidneys and the bladder were not documented. Only calculi located in ureters were documented. Reports results from two independent observers
de Souza 2007[20]	n=52 People referred for evaluation of acute renal colic Brazil	Ureterolithiasis	US	Non-contrast helical CT	Diagnosing number of patients with stones

Study	Population	Target condition	Index test	Reference standard	Comments
Haroun 2010[42]	n=156 People who underwent UHCT scan and US for suspicion of urolithiasis Jordan	Renal stones Ureteral stones	US (B-mode)	Unenhanced CT	Diagnosing number of patients with stones Study reports sensitivity and specificity for all stones (including urinary bladder) or for renal and ureteric stones separately. Currently extracted all stones as no raw data is reported for separate types of stone.
Kanno 2017[52]	n=822 People with acute flank pain, hematuria, or a history of urinary stones who had KUB, US and NCCT on the same day Japan	Renal stones	US (greyscale) Plain abdominal radiograph (KUB)	Non-contrast CT	Diagnosing number of kidneys with stones
Kanno 2014a[54]	n=428 People with acute flank pain, hematuria, or a history of urinary stones who had US and NCCT on the same day Japan	Renal stones	US	Non-contrast CT	Diagnosing number of kidneys with stones Study reports results for ‘individual stone’ and for ‘specific stone’. Currently extracted ‘individual stone’ data. The study includes 856 kidneys, but in the results there are only 853.

Study	Population	Target condition	Index test	Reference standard	Comments
Kanno 2014b[53]	n=428 People with symptoms such as acute flank pain or hematuria Japan	Ureteral stones	US	Non-contrast CT	Diagnosing number of ureters with stones
Kielar 2012[57]	n=51 People with flank pain Canada	Urolithiasis	US (greyscale)	Unenhanced CT	Diagnosing number of stones
Levine 1997[63]	n=152 People with acute flank pain who had a CT within 4 hours of plain radiography USA	Ureteral stones	Plain abdominal radiography	Unenhanced helical CT	Diagnosing number of patients with stones
Passerotti 2009[95]	n=50 Children who had signs, symptoms or a history suggestive of urolithiasis who had US and CT within 0.5-8 hours of each other	Nephrolithiasis/urinary stones	US	Non-contrast CT	Diagnosing number of patients with stones 24% of participants were asymptomatic at presentation but were evaluated because of a history of urolithiasis

Study	Population	Target condition	Index test	Reference standard	Comments
	USA				
Resorlu 2015[105]	n=500 People with acute flank pain who had CT and US within 10 days Turkey	Urinary stones (kidney and ureter)	US (grey scale)	Non-contrast CT	Diagnosing number of patients with stones
Semins 2013[119]	n=22 People with suspected acute ureteral calculus USA	Obstructing stones	MRI (MRU (non-contrast HASTE [Half-Fourier single shot turbo spin-echo]))	Non-contrast CT	Diagnosing number of patients with stones
Sternberg 2016[124]	n=155 People with suspected renal colic who had US and CT within 1 day Lebanon	Urinary calculi	Renal US	Non-contrast CT	Diagnosing number of patients with stones

See appendix D for full evidence tables.

1.4.4 Quality assessment of clinical studies included in the evidence review

Table 3: Clinical evidence summary: diagnostic test accuracy for Ultrasound, plain abdominal radiograph and MRI in adults

Index Test (Threshold)	Number of studies	n	Quality	Sensitivity % (95% CI)	Specificity % (95% CI)
Ultrasound					
Ultrasound	7	4189	VERY LOW[a,b,d] due to risk of bias, very serious inconsistency and very serious imprecision	Pooled[e]: 0.60 (0.38, 0.79)	Pooled[e]: 0.90 (0.79, 0.97)
Plain abdominal radiograph					
Plain abdominal radiograph	3	1895	VERY LOW[a,d] due to risk of bias, very serious imprecision	Pooled[e]: 0.58 (0.29, 0.58)	Pooled[e]: 0.90 (0.41, 1.00)
MRI					
MRI	3	159	LOW[a,d] due to risk of bias, serious imprecision	0.66 (0.55, 0.76)[f]	0.96 (0.89, 0.99)[f]
		159	MODERATE[a] due to risk of bias	0.72 (0.61, 0.81)[f]	1.00 (0.95, 1.00)[f]
		22	LOW[a,d] due to risk of bias, serious imprecision	0.84 (0.60, 0.97)[f]	1.00 (0.29, 1.00)[f]

The assessment of the evidence quality was conducted with emphasis on sensitivity as this was identified by the committee as the primary measure in guiding decision-making. The committee set the sensitivity threshold at 95% as the acceptable level to recommend a test.

(a) Risk of bias was assessed using the QUADAS-2 checklist. The evidence was downgraded by 1 increment if the majority of studies were rated at high risk of bias, and downgraded by 2 increments if the majority of studies were rated at very high risk of bias.

(b) Inconsistency was assessed by inspection of the sensitivity and specificity plots

(c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. The evidence was downgraded by 1 increment if the majority of studies were seriously indirect, and downgraded by 2 increments if the majority of studies are very seriously indirect

(d) Imprecision was assessed based on inspection of the confidence region of sensitivity in the diagnostic meta-analysis or, where diagnostic meta-analysis has not been conducted, assessed according to the range of confidence intervals in the individual studies. The evidence was downgraded by 1 increment when the range of the confidence interval around the point estimate was 20–40%, and downgraded by 2 increments when there was a range of >40%

(e) Pooled sensitivity/specificity from diagnostic meta-analysis

(f) Could not be pooled due to insufficient data. Individual sensitivity values and their coupled specificity is presented.

Table 4: Clinical evidence summary: diagnostic test accuracy for Ultrasound, plain abdominal radiograph and MRI in children

Index Test (Threshold)	Number of studies	n	Quality	Specificity % (95% CI)	Sensitivity % (95% CI)
Ultrasound					
Ultrasound	1	52	LOW[a,d] due to risk of bias, serious imprecision	1.00 (0.79, 1.00)	0.76 (0.59, 0.89)

The assessment of the evidence quality was conducted with emphasis on sensitivity as this was identified by the committee as the primary measure in guiding decision-making. The committee set the sensitivity threshold at 95% as the acceptable level to recommend a test.

(a) Risk of bias was assessed using the QUADAS-2 checklist. The evidence was downgraded by 1 increment if the majority of studies were rated at high risk of bias, and downgraded by 2 increments if the majority of studies were rated at very high risk of bias.

(b) Inconsistency was assessed by inspection of the sensitivity and specificity plots

(c) Indirectness was assessed using the QUADAS-2 checklist items referring to applicability. The evidence was downgraded by 1 increment if the majority of studies were seriously indirect, and downgraded by 2 increments if the majority of studies are very seriously indirect

(d) Imprecision was assessed based on inspection of the confidence region of sensitivity in the diagnostic meta-analysis or, where diagnostic meta-analysis has not been conducted, assessed according to the range of confidence intervals in the individual studies. The evidence was downgraded by 1 increment when the range of the confidence interval around the point estimate was 20–40%, and downgraded by 2 increments when there was a range of >40%

1.5 Economic evidence

1.5.1 Included studies

No relevant health economic studies were identified.

1.5.2 Excluded studies

No health economic studies that were relevant to this question were excluded due to assessment of limited applicability or methodological limitations.

See also the health economic study selection flow chart in appendix F.

1.5.3 Unit costs

Relevant unit costs are provided below to aid consideration of cost effectiveness.

Table 5: UK costs of diagnostic imaging techniques

Diagnostic imaging	Detail	Unit cost
Plain abdominal radiograph	Direct access plain film Currency code: DAPF	£29.78
Ultrasound	Ultrasound Scan with duration of less than 20 minutes, without Contrast Currency code: RD40Z	£51.59
Computerised Tomography (CT)	Adults: Computerised Tomography Scan of One Area, without Contrast, 19 years and over Currency code: RD20A Children: Computerised Tomography Scan of One Area, without Contrast, between 6 and 18 years Currency code: RD20B Computerised Tomography Scan of One Area, without Contrast, 5 years and under Currency code: RD20C	£85.56 £91.67 £94.72
Magnetic Resonance Imaging (MRI)	Adults: Magnetic Resonance Imaging Scan of One Area, without Contrast, 19 years and over Currency code: RD01A Children: Magnetic Resonance Imaging Scan of One Area, without Contrast, between 6 and 18 years Currency code: RD01B Magnetic Resonance Imaging Scan of One Area, without Contrast, 5 years and under Currency code: RD01C	£138.24 £135.88 £160.59

Source: NHS reference costs 2016/17[87].

1.6 Resource costs

The recommendations made by the committee based on this review (see section 1.8) are not expected to have a substantial impact on resources.

1.7 Evidence statements

1.7.1 Clinical evidence statements

Thirteen studies that evaluated 3 diagnostic tests were included in the review. Of these, the committee noted that all tests demonstrated poor sensitivity for identifying renal and ureteric stones. The evidence was of Moderate to Very Low quality. Evidence was identified for the following diagnostic tests: ultrasound, plain abdominal radiograph, MRI.

- **Ultrasound**: Low quality evidence from 1 study with 52 children with suspected stones showed that ultrasound has a sensitivity of 76% and a specificity of 100%. Very Low quality evidence from 7 studies with 4189 adults with suspected stones showed that ultrasound has a sensitivity of 60% and a specificity of 90%.
- **Plain abdominal radiograph**: Very Low quality from 3 studies with 1895 adults with suspected stones demonstrated that plain abdominal radiograph had a sensitivity of 58% and a specificity of 90.
- **MRI**: Three studies investigated the diagnostic accuracy of MRI, however they were unable to be pooled due to insufficient data. Moderate quality evidence from one study of 159 adults showed that the MRI has a sensitivity and specificity respectively of 72% and 100%. Low quality evidence from 1 study with 159 adults with suspected stones showed that the MRI has a sensitivity and specificity respectively of 66% and 96%. Low quality evidence from 1 study with 22 adults with suspected stones showed that the MRI has a sensitivity and specificity respectively of 84% and 100%.

1.7.2 Health economic evidence statements

- No relevant economic evaluations were identified.

1.8 Recommendations

B1. Offer urgent (within 24 hours of presentation) low-dose non-contrast CT to adults with suspected renal colic. If a woman is pregnant, offer ultrasound instead of CT.

B2. Offer urgent (within 24 hours of presentation) ultrasound as first-line imaging for children and young people with suspected renal colic.

B3. If there is still uncertainty about the diagnosis of renal colic after ultrasound for children and young people, consider low-dose non-contrast CT.

1.8.1 Research recommendations

None.

1.9 Rationale and impact

1.9.1 Why the committee made the recommendations

Limited evidence showed that MRI, ultrasound and plain abdominal radiograph were not as good as non-contrast CT for detecting renal and ureteric stones in adults. CT is more expensive than ultrasound or plain abdominal radiograph but the extra cost is likely to be outweighed by avoiding additional investigations when a first test misses the diagnosis. The committee agreed that CT should be performed as soon as possible because renal function can decline quickly. However, they acknowledged that it could be delayed for up to 24 hours if needed (for example, in some locations and when first presentation is out of hours). The committee agreed that CT should not be offered to everyone with abdominal pain, only those with suspected renal colic. They also noted that CT should not be used for pregnant women due to the radiation exposure, and agreed that ultrasound is the preferred imaging modality in this group.

No evidence was found for the use of MRI or plain abdominal radiograph in diagnosing renal and ureteric stones in children. Limited evidence on the use of ultrasound showed that it was not as good as CT and there is known to be widespread variation among ultrasonographers. The committee acknowledged that although CT is a better test, there is serious concern about radiation exposure in children and young people and they were keen to minimise this. They agreed that ultrasound should be offered as first-line imaging, and that low-dose non-contrast CT should only be considered if there is still uncertainty about the diagnosis of renal colic after ultrasound.

1.9.2 Impact of the recommendations on practice

The recommendation reflects current practice in adults so the committee agreed there should be no change.

Usual practice is to use ultrasound as first line imaging for children and young people because of concerns about radiation dosages. CT is not common practice for this population but it may be used when first-line imaging is negative or unclear, or to confirm the diagnosis. Therefore the recommendations should not change current practice.

1.10 The committee's discussion of the evidence

1.10.1 Interpreting the evidence

1.10.1.1 The diagnostic measures that matter most

Diagnostic accuracy for renal and ureteric stones was the outcome prioritised for this review. Sensitivity was considered the most important measure by the committee for this review question because it was considered that false positives are rare and not hugely problematic. The consequences of missing a patient with a renal or ureteric stone could include being sent for further imaging or investigations and a delay in treatment, potentially resulting in damage to the kidney.

1.10.1.2 The quality of the evidence

The evidence for ultrasound in adults was very low quality, due to very serious imprecision and very serious inconsistency, and in children the evidence was low quality due to imprecision. There was also a risk of bias for both populations, due to uncertainty regarding whether the results of the index or reference standard tests were interpreted without knowledge of the other test, and uncertainty regarding participants excluded from the analysis. The evidence for plain abdominal radiograph was low quality, due to very serious imprecision and a high risk of bias. The evidence for MRI ranged from low to moderate quality, due to risk of bias and imprecision. The committee noted a number of other factors that varied between the studies, such as the amount of time between the index and reference standard test, and the expertise of the clinician interpreting the test results.

1.10.1.3 Benefits and harms

Evidence for adults and children and young people, and for those with symptomatic and asymptomatic suspected stones was searched for, however no evidence was identified where the majority of the population was people with a suspected asymptomatic stone. The committee agreed that the recommendations should only apply to those who were symptomatic of a renal or ureteric stone.

The committee noted that no mode of imaging met the pre-specified threshold for sensitivity, which was set at 95%, and therefore concluded that none of the imaging modalities were as effective as the reference standard; non-contrast CT. The committee considered the consequences of a low sensitivity, and noted that this could include being sent for other imaging or investigations, and potentially a delay to treatment. The committee noted that there are risks associated with CT such as the exposure to radiation, however considered that this did not outweigh the risks associated with not diagnosing a stone.

On the other hand it was discussed how there may be some groups where the radiation risk is a concern; such as women who may be pregnant, in which case ultrasound is current practice in place of CT. The committee discussed the timing of CT imaging for diagnosis, and noted that access to CT is not currently the same across the country. It was also noted that there are harms associated with not carrying out imaging urgently, such as delay to diagnosis, and delay to treatment which may increase the risk of deterioration in renal function. Therefore, the committee agreed that based on clinical expertise and opinion of the committee, CT imaging should be done within 24 hours. The committee highlighted that CT should only be offered for those with suspected renal colic, rather than any abdominal pain without additional indicators or reasons for suspicion of renal colic. Other imaging modalities may be more appropriate where renal colic is not the suspected cause of abdominal pain.

The committee considered that for the paediatric population, there was no evidence for MRI or plain abdominal radiograph, and the evidence for ultrasound suggested that it did not meet the pre-specified threshold for sensitivity. The committee noted that in current UK practice, ultrasound is often used as the first line method of diagnosis. They considered the benefits of ultrasound include no dose of ionising radiation; however it is not as sensitive as CT and is open to wide operator variation. The committee discussed the risks associated with using CT in young children concerning the radiation, such as the increased lifetime malignancy risk, and considered that this was a very serious and potentially severe harm. The committee also discussed that renal and ureteral colic is often lower on the differential diagnosis list in children compared to adults, and presentation is much more commonly atypical. Therefore, they agreed that it was important not to increase the amount of unnecessary CT's being carried out, given that they are associated with harms. Based on this the committee agreed that US should be offered as first line imaging, and that low-dose non-contrast CT should be considered only when there is diagnostic uncertainty following US.

1.10.1.4 Cost effectiveness and resource use

No economic evidence was identified for this question.

The committee were presented with the costs of the different imaging techniques. MRI is the most costly, followed by CT, ultrasound, and plain abdominal radiograph.

A test with a low sensitivity will miss people and create a lot of false negatives. Poor specificity will result in more people being diagnosed as having stones when they do not (false positives). The implications of low sensitivity would be that people's condition could get worse as they have not been diagnosed as having renal stones, which could result in more emergency care or higher risks, and require further investigation. The implication of low specificity would be unnecessary management that the patient doesn't need, and delays identification of the true underlying condition.

The reference standard is a CT. This is one of the more expensive tests. MRI is the most expensive test, and as it is less effective than CT (because CT is the reference standard and therefore assumed to be 100% accurate); this makes MRI a dominated alternative.

Plain abdominal radiograph or ultrasound are also assumed to be by default less accurate than CT. This has to be traded off against their lower cost.

In adults, the sensitivity of ultrasound was 0.6 and specificity was 0.9. Assuming a prevalence of renal stones of 60% in a population being imaged because of pain, then if 1000 people are imaged using ultrasound, this will mean out of 600 that have a stone, only 360 are correctly identified as having a stone, and 240 will be false negatives (i.e. missed people with renal stones). There will also be 40 false positives.

Plain abdominal radiograph had similar sensitivity of 0.58 and specificity of 0.9. Using the same assumptions as above, this would lead to only 348 people identified as correctly having a stone, 252 false negatives and 40 false positives.

The cost of the different types of imaging over a cohort of 1000 people would be around £86,000 for CT, £52,000 for ultrasound, and £30,000 for plain abdominal radiograph. Spreading the cost difference between CT and ultrasound, and CT and plain abdominal radiograph, over the individuals correctly identified as having stones, would lead to a cost per correct stone diagnosis of £143 for both CT and ultrasound, and £86 per correct diagnosis for plain abdominal radiograph. Those who were missed with the less accurate techniques of ultrasound and plain abdominal radiograph will consume further resources, as they will be diagnosed at a later point, which means more imaging/tests, and any GP attendances/hospital admissions because they have been misdiagnosed and still have a stone, or potential adverse events because of the delay in diagnosis. Therefore, it could cost more in the long run to use a lower cost technique initially, because of the lower accuracy. These calculations didn't consider the false positives who would also face delay in achieving their real diagnosis and consume unnecessary resources.

There is also the impact on quality of life to consider from being misdiagnosed.

Additionally, more accurate techniques such as CT can have other benefits unrelated to the condition in question. For example, it was committee opinion that there are around 10% incidental findings such as abdominal aneurysm, which is a life threatening condition if missed.

The committee felt that the evidence had not shown that any other technique was as good as CT, based on their pre-specified threshold for sensitivity of 95%. CT is generally already current practice for diagnosing renal stones in adults, and therefore wasn't considered likely to have a large resource impact. The committee felt that the costs of misdiagnosing people were likely to outweigh the additional cost of undertaking CT over other techniques, and there were also other benefits like incidental findings or other diagnoses with similar presentations to acute renal colic (for example a leaking abdominal aortic aneurysm). Therefore a recommendation was made to offer CT in adults with suspected renal colic. The urgency with which the CT should take place was debated, as although immediately would be ideal, this may not always be feasible in all locations particularly out of hours, and some hospitals would ask someone to come back the following day. So although CT is generally the gold standard for diagnosis of renal stones and this is established practice, how quickly this takes place can be variable. The committee felt that by specifying 'urgent' and outlining that this means within 24 hours, would allow some flexibility in hospitals where this cannot happen within a few hours.

It is important to make clear that the recommendation is specifying a population with suspected renal colic, which means this suspicion (based on history and clinical examination) has to be in place clinically, and this would be a subset of people presenting to hospital with abdominal/flank pain in general.

For children, there was only evidence for ultrasound, which showed a sensitivity of 0.76 compared to CT. There is more caution around imaging with a higher radiation burden in children because of their age and the cumulative effect of imaging over their lifetime. Because of these concerns, a recommendation was made for ultrasound as first line imaging and low dose non-contrast CT, if there is still uncertainty around the diagnosis and there is a high degree of suspicion for renal colic. This is in keeping with current practice where ultrasound is more likely to be first line imaging. The paediatric population suspected with renal colic is likely to be small as renal stones in this population itself is very small.

References

1. Abdel-Gawad M, Kadasne R, Anjikar C, Elsobky E. Value of Color Doppler ultrasound, KUB and urinalysis in diagnosis of renal colic due to ureteral stones. International Brazilian Journal of Urology. 2014; 40(4):513-9

2. Abdel-Gawad M, Kadasne RD, Elsobky E, Ali-El-Dein B, Monga M. A prospective comparative study of color doppler ultrasound with twinkling and noncontrast computerized tomography for the evaluation of acute renal colic. Journal of Urology. 2016; 196(3):757-62

3. Ahmad SK, Abdallah MM. The diagnostic value of the twinkle sign in color Doppler imaging of urinary stones. Egyptian Journal of Radiology and Nuclear Medicine. 2014; 45(2):569-74

4. Ahn SH, Mayo-Smith WW, Murphy BL, Reinert SE, Cronan JJ. Acute nontraumatic abdominal pain in adult patients: abdominal radiography compared with CT evaluation. Radiology. 2002; 225(1):159-64

5. Albani JM, Ciaschini MW, Streem SB, Herts BR, Angermeier KW. The role of computerized tomographic urography in the initial evaluation of hematuria. Journal of Urology. 2007; 177(2):644-8

6. Andresen R, Wegner HE. Intravenous urography revisited in the age of ultrasound and computerized tomography: diagnostic yield in cases of renal colic, suspected pelvic and abdominal malignancies, suspected renal mass, and acute pyelonephritis. Urologia Internationalis. 1997; 58(4):221-6

7. Arif U, Ijaz M, Shah ZA, Khan MA, Khan MAR. Diagnostic accuracy of non-contrast-enhanced helical CT scan in comparison with ultrasonography in patients with acute flank pain. Pakistan Journal of Medical and Health Sciences. 2013; 7(2):462-4

8. Assi Z, Platt JF, Francis IR, Cohan RH, Korobkin M. Sensitivity of CT scout radiography and abdominal for revealing ureteral calculi on helical CT: implications for radiologic follow-up. American Journal of Roentgenology. 2000; 175(2):333-7

9. Ather MH, Jafri AH, Sulaiman MN. Diagnostic accuracy of ultrasonography compared to unenhanced CT for stone and obstruction in patients with renal failure. BMC Medical Imaging. 2004; 4(1):2

10. Ben Nakhi A, Gupta R, Al-Hunayan A, Muttikkal T, Chavan V, Mohammed A et al. Comparative analysis and interobserver variation of unenhanced computed tomography and intravenous urography in the diagnosis of acute flank pain. Medical Principles and Practice. 2010; 19(2):118-21

11. Blandino A, Minutoli F, Scribano E, Vinci S, Magno C, Pergolizzi S et al. Combined magnetic resonance urography and targeted helical CT in patients with renal colic: a new approach to reduce delivered dose. Journal of Magnetic Resonance Imaging. 2004; 20(2):264-71

12. Bozdar HR, Phul AH, Ahmed F, Shaikh NA. Comparison of ultrasound, intravenous urography and plain CT KUB in the diagnosis of ureteric stone. Rawal Medical Journal. 2016; 41(1):36-8

13. Cabrera FJ, Kaplan AG, Youssef RF, Tsivian M, Shin RH, Scales CD et al. Digital tomosynthesis: a viable alternative to noncontrast computed tomography for the follow-up of nephrolithiasis? Journal of Endourology. 2016; 30(4):366-70

14. Catalano O, Nunziata A, Altei F, Siani A. Suspected ureteral colic: primary helical CT versus selective helical CT after unenhanced radiography and sonography. AJR American Journal of Roentgenology. 2002; 178(2):379-87

15. Cauberg ECC, Nio CY, De La Rosette JMCH, Laguna MP, De Reijke TM. Computed tomography-urography for upper urinary tract imaging: is it required for all patients who present with hematuria? Journal of Endourology. 2011; 25(11):1733-40

16. Chan VO, Buckley O, Persaud T, Torreggiani WC. Urolithiasis: how accurate are plain radiographs? Canadian Association of Radiologists Journal. 2008; 59(3):131-4

17. Chen MY, Zagoria RJ. Can noncontrast helical computed tomography replace intravenous urography for evaluation of patients with acute urinary tract colic? Journal of Emergency Medicine. 1999; 17(2):299-303

18. Cifci E, Coban G, Cicek T, Gonulalan U. The diagnostic value of magnetic resonance urography using a balanced turbo field echo sequence. European Radiology. 2016; 26(12):4624-31

19. Cochon L, Smith J, Baez AA. Bayesian comparative assessment of diagnostic accuracy of low-dose CT scan and ultrasonography in the diagnosis of urolithiasis after the application of the STONE score. Emergency Radiology. 2017; 24(2):177-82

20. de Souza LR, Goldman SM, Faintuch S, Faria JF, Bekhor D, Tiferes DA et al. Comparison between ultrasound and noncontrast helical computed tomography for identification of acute ureterolithiasis in a teaching hospital setting. Sao Paulo Medical Journal. 2007; 125(2):102-7

21. Dillman JR, Kappil M, Weadock WJ, Rubin JM, Platt JF, DiPietro MA et al. Sonographic twinkling artifact for renal calculus detection: correlation with CT. Radiology. 2011; 259(3):911-6

22. Dorio PJ, Pozniak MA, Lee FT, Jr., Kuhlman JE. Non-contrast-enhanced helical computed tomography for the evaluation of patients with acute flank pain. WMJ. 1999; 98(6):30-4

23. Drake T, Jain N, Bryant T, Wilson I, Somani BK. Should low-dose computed tomography kidneys, ureter and bladder be the new investigation of choice in suspected renal colic?: a systematic review. Indian Journal of Urology. 2014; 30(2):137-43

24. Dundee P, Bouchier-Hayes D, Haxhimolla H, Dowling R, Costello A. Renal tract calculi: comparison of stone size on plain radiography and noncontrast spiral CT scan. Journal of Endourology. 2006; 20(12):1005-9

25. Edmonds ML, Yan JW, Sedran RJ, McLeod SL, Theakston KD. The utility of renal ultrasonography in the diagnosis of renal colic in emergency department patients. CJEM Canadian Journal of Emergency Medical Care. 2010; 12(3):201-6

26. Ege G, Akman H, Kuzucu K, Yildiz S. Can computed tomography scout radiography replace plain film in the evaluation of patients with acute urinary tract colic? Acta Radiologica. 2004; 45(4):469-73

27. Eikefjord E, Askildsen JE, Rorvik J. Cost-effectiveness analysis (CEA) of intravenous urography (IVU) and unenhanced multidetector computed tomography (MDCT) for initial investigation of suspected acute ureterolithiasis. Acta Radiologica. 2008; 49(2):222-9

28. Ekici S, Sinanoglu O. Comparison of conventional radiography combined with ultrasonography versus nonenhanced helical computed tomography in evaluation of patients with renal colic. Urological Research. 2012; 40(5):543-7

29. Eray O, Cubuk MS, Oktay C, Yilmaz S, Cete Y, Ersoy FF. The efficacy of urinalysis, plain films, and spiral CT in ED patients with suspected renal colic. American Journal of Emergency Medicine. 2003; 21(2):152-4

30. Eshed I, Witzling M. The role of unenhanced helical CT in the evaluation of suspected renal colic and atypical abdominal pain in children. Pediatric Radiology. 2002; 32(3):205-8

31. Feroze S, Singh B, Gojwari T, Manjeet S, Athar B, Hamid H. Role of non-contrast spiral computerized tomography in acute ureteric colic. Indian Journal of Urology. 2007; 23(2):119-21

32. Foell K, Ordon M, Ghiculete D, Lee JY, Honey RJ, Pace KT. Does baseline radiography of the kidneys, ureters, and bladder help facilitate stone management in patients presenting to the emergency department with renal colic? Journal of Endourology. 2013; 27(12):1425-30

33. Fowler JC, Cutress ML, Abubacker Z, Saleemi MA, Alam A, Shekhdar J et al. Clinical evaluation of ultra-low dose contrast-enhanced CT in patients presenting with acute ureteric colic. British Journal of Medical and Surgical Urology. 2011; 4(2):56-63

34. Fowler KA, Locken JA, Duchesne JH, Williamson MR. US for detecting renal calculi with nonenhanced CT as a reference standard. Radiology. 2002; 222(1):109-13

35. Gaspari RJ, Horst K. Emergency ultrasound and urinalysis in the evaluation of flank pain. Academic Emergency Medicine. 2005; 12(12):1180-4

36. German I, Lantsberg S, Crystal P, Assali M, Rachinsky I, Kaneti J et al. Non contrast computerized tomography and dynamic renal scintigraphy in the evaluation of patients with renal colic: are both necessary? European Urology. 2002; 42(2):188-91

37. Gliga ML, Chirila CN, Podeanu DM, Imola T, Voicu SL, Gliga MG et al. Twinkle, twinkle little stone: an artifact improves the ultrasound performance! Medical Ultrasonography. 2017; 19(3):272-5

38. Graumann O, Osther SS, Spasojevic D, Osther PJ. Can the CT planning image determine whether a kidney stone is radiopaque on a plain KUB? Urological Research. 2012; 40(4):333-7

39. Gurel S, Akata D, Gurel K, Ozmen MN, Akhan O. Correlation between the renal resistive index (RI) and nonenhanced computed tomography in acute renal colic: how reliable is the RI in distinguishing obstruction? Journal of Ultrasound in Medicine. 2006; 25(9):1113-20; quiz 1121-3

40. Hamm M, Lampart L, Knopfle E, Hackel T, Hauser H, Wawroschek F et al. Evaluation of acute flank pain - diagnostic value of unenhanced helical computed tomography. Klinikarzt. 2001; 30(3):72-6

41. Hamm M, Wawroschek F, Weckermann D, Knopfle E, Hackel T, Hauser H et al. Unenhanced helical computed tomography in the evaluation of acute flank pain. European Urology. 2001; 39(4):460-5

42. Haroun AA, Hadidy AM, Mithqal AM, Mahafza WS, Al-Riyalat NT, Sheikh-Ali RF. The role of B-mode ultrasonography in the detection of urolithiasis in patients with acute renal colic. Saudi Journal of Kidney Diseases and Transplantation. 2010; 21(3):488-93

43. Herbst MK, Rosenberg G, Daniels B, Gross CP, Singh D, Molinaro AM et al. Effect of provider experience on clinician-performed ultrasonography for hydronephrosis in patients with suspected renal colic. Annals of Emergency Medicine. 2014; 64(3):269-76

44. Homer JA, Davies-Payne DL, Peddinti BS. Randomized prospective comparison of non-contrast enhanced helical computed tomography and intravenous urography in the diagnosis of acute ureteric colic. Australasian Radiology. 2001; 45(3):285-90

45. Hu H, Hu XY, Fang XM, Chen HW, Yao XJ. Unenhanced helical CT following excretory urography in the diagnosis of upper urinary tract disease: a little more cost, a lot more value. Urological Research. 2010; 38(2):127-33

46. Ibrahim EH, Cernigliaro JG, Bridges MD, Pooley RA, Haley WE. The capabilities and limitations of clinical magnetic resonance imaging for detecting kidney stones: a retrospective study. International Journal of Biomedical Imaging. 2016; 2016:4935656

47. Jackman SV, Potter SR, Regan F, Jarrett TW. Plain abdominal x-ray versus computerized tomography screening: sensitivity for stone localization after nonenhanced spiral computerized tomography. Journal of Urology. 2000; 164(2):308-10

48. Jeng CM, Kung CH, Wang YC, Wu CY, Lee WY, Fan JK et al. Urolithiasis in patients with acute flank pain: comparison of plain abdominal radiography to unenhanced spiral CT. Chinese Journal of Radiology. 2001; 26(6):243-9

49. Johnston R, Lin A, Du J, Mark S. Comparison of kidney-ureter-bladder abdominal radiography and computed tomography scout films for identifying renal calculi. BJU International. 2009; 104(5):670-3

50. Joshi KS, Karki S, Regmi S, Joshi HN, Adhikari SP. Sonography in acute ureteric colic: an experience in Dhulikhel Hospital. Kathmandu University Medical Journal. 2014; 12(45):9-15

51. Jung SI, Kim YJ, Park HS, Jeon HJ, Park HK, Paick SH et al. Sensitivity of digital abdominal radiography for the detection of ureter stones by stone size and location. Journal of Computer Assisted Tomography. 2010; 34(6):879-82

52. Kanno T, Kubota M, Funada S, Okada T, Higashi Y, Yamada H. The utility of the kidneys-ureters-bladder radiograph as the sole imaging modality and its combination with ultrasonography for the detection of renal stones. Urology. 2017; 104:40-4

53. Kanno T, Kubota M, Sakamoto H, Nishiyama R, Okada T, Higashi Y et al. Determining the efficacy of ultrasonography for the detection of ureteral stone. Urology. 2014; 84(3):533-7

54. Kanno T, Kubota M, Sakamoto H, Nishiyama R, Okada T, Higashi Y et al. The efficacy of ultrasonography for the detection of renal stone. Urology. 2014; 84(2):285-8

55. Kennish SJ, Bhatnagar P, Wah TM, Bush S, Irving HC. Is the KUB radiograph redundant for investigating acute ureteric colic in the non-contrast enhanced computed tomography era? Clinical Radiology. 2008; 63(10):1131-5

56. Khan N, Anwar Z, Zafar AM, Ahmed F, Ather MH. A comparison of non-contrast CT and intravenous urography in the diagnosis of urolithiasis and obstruction. African Journal of Urology. 2012; 18(3):108-11

57. Kielar AZ, Shabana W, Vakili M, Rubin J. Prospective evaluation of Doppler sonography to detect the twinkling artifact versus unenhanced computed tomography

for identifying urinary tract calculi. Journal of Ultrasound in Medicine. 2012; 31(10):1619-25

58. Kluner C, Hein PA, Gralla O, Hein E, Hamm B, Romano V et al. Does ultra-low-dose CT with a radiation dose equivalent to that of KUB suffice to detect renal and ureteral calculi? Journal of Computer Assisted Tomography. 2006; 30(1):44-50

59. Korkmaz M, Aras B, Sanal B, Yucel M, Guneyli S, Kocak A et al. Investigating the clinical significance of twinkling artifacts in patients with urolithiasis smaller than 5 mm. Japanese Journal of Radiology. 2014; 32(8):482-6

60. Kravchick S, Stepnov E, Lebedev V, Linov L, Leibovici O, Ben-Horin CL et al. Non-contrast computerized tomography (NCCT) and dynamic renal scintigraphy (DRS) in the patients with refractory renal colic. European Journal of Radiology. 2006; 58(2):301-6

61. Lee DH, Chang IH, Kim JW, Chi BH, Park SB. Usefulness of nonenhanced computed tomography for diagnosing urolithiasis without pyuria in the emergency department. BioMed Research International. 2015; 2015:810971

62. Leo MM, Langlois BK, Pare JR, Mitchell P, Linden J, Nelson KP et al. Ultrasound vs. computed tomography for severity of hydronephrosis and its importance in renal colic. Western Journal of Emergency Medicine. 2017; 18(4):559-68

63. Levine JA, Neitlich J, Verga M, Dalrymple N, Smith RC. Ureteral calculi in patients with flank pain: correlation of plain radiography with unenhanced helical CT. Radiology. 1997; 204(1):27-31

64. Lew HM, Seow JH, Hewavitharana CP, Burrows S. Alternatives to the baseline KUB for CTKUB-detected calculi: evaluation of CT scout and average and maximum intensity projection images. Abdominal Radiology. 2017; 42(5):1459-63

65. Lin K, Dowling S. Ultrasonography versus computed tomography for initial investigation of suspected nephrolithiasis. CJEM Canadian Journal of Emergency Medical Care. 2016; 18(4):315-8

66. Lisanti CJ, Toffoli TJ, Stringer MT, DeWitt RM, Schwope RB. CT evaluation of the upper urinary tract in adults younger than 50 years with asymptomatic microscopic hematuria: is IV contrast enhancement needed? American Journal of Roentgenology. 2014; 203(3):615-9

67. Liu W, Esler SJ, Kenny BJ, Goh RH, Rainbow AJ, Stevenson GW. Low-dose nonenhanced helical CT of renal colic: assessment of ureteric stone detection and measurement of effective dose equivalent. Radiology. 2000; 215(1):51-4

68. Longo J, Akbar SA, Schaff T, Jafri ZH, Jackson RE. A prospective comparative study of non-contrast helical computed tomography and intravenous urogram for the assessment of renal colic. Emergency Radiology. 2001; 8(5):285-92

69. Lorberboym M, Kapustin Z, Elias S, Nikolov G, Katz R. The role of renal scintigraphy and unenhanced helical computerized tomography in patients with ureterolithiasis. European Journal of Nuclear Medicine. 2000; 27(4):441-6

70. MacEjko A, Okotie OT, Zhao LC, Liu J, Perry K, Nadler RB. Computed tomography-determined stone-free rates for ureteroscopy of upper-tract stones. Journal of Endourology. 2009; 23(3):379-82

71. Malaki M. The comparison of ultrasound and non-contrast helical computerized tomography for children nephrolithiasis detection. Urology Annals. 2014; 6(4):309-13

72. Marumo K, Horiguchi Y, Nakagawa K, Oya M, Ohigashi T, Asakura H et al. Significance and diagnostic accuracy of renal calculi found by ultrasonography in patients with asymptomatic microscopic hematuria. International Journal of Urology. 2002; 9(7):363-7; discussion 367

73. Masch WR, Cohan RH, Ellis JH, Dillman JR, Rubin JM, Davenport MS. Clinical effectiveness of prospectively reported sonographic twinkling artifact for the diagnosis of renal calculus in patients without known urolithiasis. American Journal of Roentgenology. 2016; 206(2):326-31

74. Matani YS, Al-Ghazo MA. Role of helical nonenhanced computed tomography in the evaluation of acute flank pain. Asian Journal of Surgery. 2007; 30(1):45-51

75. May PC, Haider Y, Dunmire B, Cunitz BW, Thiel J, Liu Z et al. Stone-mode ultrasound for determining renal stone size. Journal of Endourology. 2016; 30(9):958-62

76. Meagher T, Sukumar VP, Collingwood J, Crawley T, Schofield D, Henson J et al. Low dose computed tomography in suspected acute renal colic. Clinical Radiology. 2001; 56(11):873-6

77. Melnikow J, Xing G, Cox G, Leigh P, Mills L, Miglioretti DL et al. Cost analysis of the STONE randomized trial: can health care costs be reduced one test at a time? Medical Care. 2016; 54(4):337-42

78. Mendelson RM, Arnold-Reed DE, Kuan M, Wedderburn AW, Anderson JE, Sweetman G et al. Renal colic: a prospective evaluation of non-enhanced spiral CT versus intravenous pyelography. Australasian Radiology. 2003; 47(1):22-8

79. Mermuys K, De Geeter F, Bacher K, Van De Moortele K, Coenegrachts K, Steyaert L et al. Digital tomosynthesis in the detection of urolithiasis: diagnostic performance and dosimetry compared with digital radiography with MDCT as the reference standard. American Journal of Roentgenology. 2010; 195(1):161-7

80. Middleton WD, Dodds WJ, Lawson TL, Foley WD. Renal calculi: sensitivity for detection with US. Radiology. 1988; 167(1):239-44

81. Miller OF, Rineer SK, Reichard SR, Buckley RG, Donovan MS, Graham IR et al. Prospective comparison of unenhanced spiral computed tomography and intravenous urogram in the evaluation of acute flank pain. Urology. 1998; 52(6):982-7

82. Mitterberger M, Aigner F, Pallwein L, Pinggera GM, Neururer R, Rehder P et al. Sonographic detection of renal and ureteral stones. Value of the twinkling sign. International Brazilian Journal of Urology. 2009; 35(5):532-9; discussion 540-1

83. Mitterberger M, Pinggera GM, Maier E, Neuwirt H, Neururer R, Pallwein L et al. Value of 3-dimensional transrectal/transvaginal sonography in diagnosis of distal ureteral calculi. Journal of Ultrasound in Medicine. 2007; 26(1):19-27

84. Moak JH, Lyons MS, Lindsell CJ. Bedside renal ultrasound in the evaluation of suspected ureterolithiasis. American Journal of Emergency Medicine. 2012; 30(1):218-21

85. Mos C, Holt G, Iuhasz S, Mos D, Teodor I, Halbac M. The sensitivity of transabdominal ultrasound in the diagnosis of ureterolithiasis. Medical Ultrasonography. 2010; 12(3):188-97

86. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual. London. National Institute for Health and Care Excellence, 2014. Available

from: http://www.nice.org.uk/article/PMG20/chapter/1%20Introduction%20and%20overview

87. NHS Improvement. Reference costs 2016/17: highlights, analysis and introduction to the data. London. 2017. Available from: https://improvement.nhs.uk/resources/reference-costs/

88. Niall O, Russell J, MacGregor R, Duncan H, Mullins J. A comparison of noncontrast computerized tomography with excretory urography in the assessment of acute flank pain. Journal of Urology. 1999; 161(2):534-7

89. Nishiura JL, Neves RF, Eloi SR, Cintra SM, Ajzen SA, Heilberg IP. Evaluation of nephrolithiasis in autosomal dominant polycystic kidney disease patients. Clinical Journal of the American Society of Nephrology. 2009; 4(4):838-44

90. O'Kane D, Papa N, Manning T, Quinn J, Hawes A, Smith N et al. Contemporary accuracy of digital abdominal x-ray for follow-up of pure calcium urolithiasis: is there still a role? Journal of Endourology. 2016; 30(8):844-9

91. Olcott EW, Sommer FG, Napel S. Accuracy of detection and measurement of renal calculi: in vitro comparison of three-dimensional spiral CT, radiography, and nephrotomography. Radiology. 1997; 204(1):19-25

92. Oner S, Oto A, Tekgul S, Koroglu M, Hascicek M, Sahin A et al. Comparison of spiral CT and US in the evaluation of pediatric urolithiasis. Journal Belge de Radiologie. 2004; 87(5):219-23

93. Palmer JS, Donaher ER, O'Riordan MA, Dell KM. Diagnosis of pediatric urolithiasis: role of ultrasound and computerized tomography. Journal of Urology. 2005; 174(4 Pt 1):1413-6

94. Park SJ, Yi BH, Lee HK, Kim YH, Kim GJ, Kim HC. Evaluation of patients with suspected ureteral calculi using sonography as an initial diagnostic tool: how can we improve diagnostic accuracy? Journal of Ultrasound in Medicine. 2008; 27(10):1441-50

95. Passerotti C, Chow JS, Silva A, Schoettler CL, Rosoklija I, Perez-Rossello J et al. Ultrasound versus computerized tomography for evaluating urolithiasis. Journal of Urology. 2009; 182(4 Suppl):1829-34

96. Patlas M, Farkas A, Fisher D, Zaghal I, Hadas-Halpern I. Ultrasound vs CT for the detection of ureteric stones in patients with renal colic. British Journal of Radiology. 2001; 74(886):901-4

97. Pepe P, Motta L, Pennisi M, Aragona F. Functional evaluation of the urinary tract by color-Doppler ultrasonography (CDU) in 100 patients with renal colic. European Journal of Radiology. 2005; 53(1):131-5

98. Pfister SA, Deckart A, Laschke S, Dellas S, Otto U, Buitrago C et al. Unenhanced helical computed tomography vs intravenous urography in patients with acute flank pain: accuracy and economic impact in a randomized prospective trial. European Radiology. 2003; 13(11):2513-20

99. Pichler R, Skradski V, Aigner F, Leonhartsberger N, Steiner H. In young adults with a low body mass index ultrasonography is sufficient as a diagnostic tool for ureteric stones. BJU International. 2012; 109(5):770-4

100. Poletti PA, Platon A, Rutschmann OT, Verdun FR, Schmidlin FR, Iselin CE et al. Abdominal plain film in patients admitted with clinical suspicion of renal colic: should it be replaced by low-dose computed tomography? Urology. 2006; 67(1):64-8

101. Quirke M, Divilly F, O'Kelly P, Winder S, Gilligan P. Imaging patients with renal colic: a comparative analysis of the impact of non-contrast helical computed tomography versus intravenous pyelography on the speed of patient processing in the Emergency Department. Emergency Medicine Journal. 2011; 28(3):197-200

102. Rajaie Esfahani M, Momeni A. Comparison of ultrasonography and intravenous urography in the screening and diagnosis of hematuria causes. Urology Journal. 2006; 3(1):54-60; discussion 60

103. Ray AA, Ghiculete D, Pace KT, Honey RJ. Limitations to ultrasound in the detection and measurement of urinary tract calculi. Urology. 2010; 76(2):295-300

104. Rengifo Abbad D, Rodriguez Caravaca G, Barreales Tolosa L, Villar del Campo MC, Martel Villagran J, Trapero Garcia MA. Diagnostic validity of helical CT compared to ultrasonography in renal-ureteral colic. Archivos Españoles de Urología. 2010; 63(2):139-44

105. Resorlu M, Abdulmajed MI, Resorlu EB, Ates C, Uysal F, Adam G et al. The accuracy of urinary ultrasound in the diagnosis of urinary stone disease in patients with acute flank pain: is it influenced by the time of ultrasound performance during the day or week? Wiener Klinische Wochenschrift. 2015; 127(11-12):445-50

106. Richards JR, Christman CA. Intravenous urography in the emergency department: when do we need it? European Journal of Emergency Medicine. 1999; 6(2):129-33

107. Riddell J, Case A, Wopat R, Beckham S, Lucas M, McClung CD et al. Sensitivity of emergency bedside ultrasound to detect hydronephrosis in patients with computed tomography-proven stones. Western Journal of Emergency Medicine. 2014; 15(1):96-100

108. Ripolles T, Agramunt M, Errando J, Martinez MJ, Coronel B, Morales M. Suspected ureteral colic: plain film and sonography vs unenhanced helical CT. A prospective study in 66 patients. European Radiology. 2004; 14(1):129-36

109. Ripolles T, Martinez-Perez MJ, Vizuete J, Miralles S, Delgado F, Pastor-Navarro T. Sonographic diagnosis of symptomatic ureteral calculi: usefulness of the twinkling artifact. Abdominal Imaging. 2013; 38(4):863-9

110. Rosen CL, Brown DF, Sagarin MJ, Chang Y, McCabe CJ, Wolfe RE. Ultrasonography by emergency physicians in patients with suspected ureteral colic. Journal of Emergency Medicine. 1998; 16(6):865-70

111. Rosser CJ, Zagoria R, Dixon R, Scurry WC, Bare RL, McCullough DL et al. Is there a learning curve in diagnosing urolithiasis with noncontrast helical computed tomography? Canadian Association of Radiologists Journal. 2000; 51(3):177-81

112. Rowland JL, Kuhn M, Bonnin RL, Davey MJ, Langlois SL. Accuracy of emergency department bedside ultrasonography. Emergency Medicine. 2001; 13(3):305-13

113. Ryu JA, Kim B, Jeon YH, Lee J, Lee JW, Jeon SS et al. Unenhanced spiral CT in acute ureteral colic: a replacement for excretory urography? Korean Journal of Radiology. 2001; 2(1):14-20

114. Sade R, Ogul H, Eren S, Levent A, Kantarci M. Comparison of ultrasonography and low-dose computed tomography for the diagnosis of pediatric urolithiasis in the emergency department. The Eurasian Journal of Medicine. 2017; 49(2):128-31

115. Sarofim M, Teo A, Wilson R. Management of alternative pathology detected using CT KUB in suspected ureteric colic. International Journal of Surgery. 2016; 32:179-82

116. Sattar A, Ahmed I, Anjum S, Chishti AT. Comparative study of non-contrast-enhanced spiral CT scan to ultrasonography in the diagnosis of acute renal colic. Medical Forum Monthly. 2011; 22(12):24-6

117. Schwartz G, Lipschitz S, Becker JA. Detection of renal calculi: the value of tomography. American Journal of Roentgenology. 1984; 143(1):143-5

118. Selberherr A, Hormann M, Prager G, Riss P, Scheuba C, Niederle B. "Silent" kidney stones in "asymptomatic" primary hyperparathyroidism-a comparison of multidetector computed tomography and ultrasound. Langenbeck's Archives of Surgery. 2017; 402(2):289-93

119. Semins MJ, Feng Z, Trock B, Bohlman M, Hosek W, Matlaga BR. Evaluation of acute renal colic: a comparison of non-contrast CT versus 3-T non-contrast HASTE MR urography. Urolithiasis. 2013; 41(1):43-6

120. Sen V, Imamoglu C, Kucukturkmen I, Degirmenci T, Bozkurt IH, Yonguc T et al. Can Doppler ultrasonography twinkling artifact be used as an alternative imaging modality to non-contrast-enhanced computed tomography in patients with ureteral stones? A prospective clinical study. Urolithiasis. 2017; 45(2):215-19

121. Sheafor DH, Hertzberg BS, Freed KS, Carroll BA, Keogan MT, Paulson EK et al. Nonenhanced helical CT and US in the emergency evaluation of patients with renal colic: prospective comparison. Radiology. 2000; 217(3):792-7

122. Shokeir AA, Abdulmaaboud M. Prospective comparison of nonenhanced helical computerized tomography and Doppler ultrasonography for the diagnosis of renal colic. Journal of Urology. 2001; 165(4):1082-4

123. Smith-Bindman R, Aubin C, Bailitz J, Bengiamin RN, Camargo CA, Jr., Corbo J et al. Ultrasonography versus computed tomography for suspected nephrolithiasis. New England Journal of Medicine. 2014; 371(12):1100-10

124. Sternberg KM, Eisner B, Larson T, Hernandez N, Han J, Pais VM. Ultrasonography significantly overestimates stone size when compared to low-dose, noncontrast computed tomography. Urology. 2016; 95:67-71

125. Sudah M, Vanninen RL, Partanen K, Kainulainen S, Malinen A, Heino A et al. Patients with acute flank pain: comparison of MR urography with unenhanced helical CT. Radiology. 2002; 223(1):98-105

126. Thomson JM, Glocer J, Abbott C, Maling TM, Mark S. Computed tomography versus intravenous urography in diagnosis of acute flank pain from urolithiasis: a randomized study comparing imaging costs and radiation dose. Australasian Radiology. 2001; 45(3):291-7

127. Ulusan S, Koc Z, Tokmak N. Accuracy of sonography for detecting renal stone: comparison with CT. Journal of Clinical Ultrasound. 2007; 35(5):256-61

128. Unal D, Yeni E, Karaoglanoglu M, Verit A, Karatas OF. Can conventional examinations contribute to the diagnostic power of unenhanced helical computed tomography in urolithiasis? Urologia Internationalis. 2003; 70(1):31-5

129. Uraiqat A, Al Khateeb M, Al Shishani J. Non-enhanced spiral CT versus excretory urography in acute renal colic. Qatar Medical Journal. 2007; 16(2):21-4

130. Valencia V, Moghadassi M, Kriesel DR, Cummings S, Smith-Bindman R. Study of tomography of nephrolithiasis evaluation (STONE): methodology, approach and rationale. Contemporary Clinical Trials. 2014; 38(1):92-101

131. Vallone G, Napolitano G, Fonio P, Antinolfi G, Romeo A, Macarini L et al. US detection of renal and ureteral calculi in patients with suspected renal colic. Critical Ultrasound Journal. 2013; 5(Suppl 1):S3

132. Van Appledorn S, Ball AJ, Patel VR, Kim S, Leveillee RJ. Limitations of noncontrast CT for measuring ureteral stones. Journal of Endourology. 2003; 17(10):851-4; discussion 854

133. Van Beers BE, Dechambre S, Hulcelle P, Materne R, Jamart J. Value of multislice helical CT scans and maximum-intensity-projection images to improve detection of ureteral stones at abdominal radiography. American Journal of Roentgenology. 2001; 177(5):1117-21

134. Vieweg J, Teh C, Freed K, Leder RA, Smith RH, Nelson RH et al. Unenhanced helical computerized tomography for the evaluation of patients with acute flank pain. Journal of Urology. 1998; 160(3 Pt 1):679-84

135. Viprakasit DP, Sawyer MD, Herrell SD, Miller NL. Limitations of ultrasonography in the evaluation of urolithiasis: a correlation with computed tomography. Journal of Endourology. 2012; 26(3):209-13

136. Vrtiska TJ, Hattery RR, King BF, Charboneau JW, Smith LH, Williamson B, Jr. et al. Role of ultrasound in medical management of patients with renal stone disease. Urologic Radiology. 1992; 14(3):131-8

137. Wang JH, Lin WC, Wei CJ, Chang CY. Diagnostic value of unenhanced computerized tomography urography in the evaluation of acute renal colic. Kaohsiung Journal of Medical Sciences. 2003; 19(10):503-9

138. Wang JH, Shen SH, Huang SS, Chang CY. Prospective comparison of unenhanced spiral computed tomography and intravenous urography in the evaluation of acute renal colic. Journal of the Chinese Medical Association. 2008; 71(1):30-6

139. Wang LJ, Ng CJ, Chen JC, Chiu TF, Wong YC. Diagnosis of acute flank pain caused by ureteral stones: value of combined direct and indirect signs on IVU and unenhanced helical CT. European Radiology. 2004; 14(9):1634-40

140. Wang RC, Rodriguez RM, Fahimi J, Hall MK, Shiboski S, Chi T et al. Derivation of decision rules to predict clinically important outcomes in acute flank pain patients. American Journal of Emergency Medicine. 2017; 35(4):554-63

141. Watkins S, Bowra J, Sharma P, Holdgate A, Giles A, Campbell L. Validation of emergency physician ultrasound in diagnosing hydronephrosis in ureteric colic. Emergency Medicine Australasia. 2007; 19(3):188-95

142. Westergreen-Thorne M, Lee SY, Babawale K, Lovegrove C, Brewer J, Shrotri N. Comparing the diagnostic accuracy of ultrasound in the community and in the hospital setting for urinary calculi: a retrospective cohort study. Journal of Clinical Urology. 2017; 10(2):133-36

143. Winkel RR, Kalhauge A, Fredfeldt KE. The usefulness of ultrasound colour-Doppler twinkling artefact for detecting urolithiasis compared with low dose nonenhanced computerized tomography. Ultrasound in Medicine and Biology. 2012; 38(7):1180-7

144. Wong SK, Ng LG, Tan BS, Cheng CW, Chee CT, Chan LP et al. Acute renal colic: value of unenhanced spiral computed tomography compared with intravenous urography. Annals of the Academy of Medicine, Singapore. 2001; 30(6):568-72

145. Yap WW, Belfield JC, Bhatnagar P, Kennish S, Wah TM. Evaluation of the sensitivity of scout radiographs on unenhanced helical CT in identifying ureteric calculi: a large

UK tertiary referral centre experience. British Journal of Radiology. 2012; 85(1014):800-6

146. Yavuz A, Ceken K, Alimoglu E, Kabaalioglu A. The reliability of color doppler "twinkling" artifact for diagnosing millimetrical nephrolithiasis: comparison with B-Mode US and CT scanning results. Journal of Medical Ultrasonics. 2015; 42(2):215-22

147. Yilmaz S, Sindel T, Arslan G, Ozkaynak C, Karaali K, Kabaalioglu A et al. Renal colic: comparison of spiral CT, US and IVU in the detection of ureteral calculi. European Radiology. 1998; 8(2):212-7

148. Zilberman DE, Tsivian M, Lipkin ME, Ferrandino MN, Frush DP, Paulson EK et al. Low dose computerized tomography for detection of urolithiasis-its effectiveness in the setting of the urology clinic. Journal of Urology. 2011; 185(3):910-4

Appendices

Appendix A: Review protocols

Table 6: Review protocol: What is the clinical and cost effectiveness of performing imaging for diagnosing renal and ureteric stones?

Field	Content
Review question	What is the clinical and cost effectiveness of performing imaging for diagnosing renal and ureteric stones?
Type of review question	Diagnostic review A review of health economic evidence related to the same review question was conducted in parallel with this review. For details see the health economic review protocol for this NICE guideline.
Objective of the review	To evaluate the diagnostic test accuracy of imaging techniques in diagnosing renal and ureteric stones.
Eligibility criteria – population / disease / condition / issue / domain	People (adults, children and young people) with suspected (or under investigation for) renal and ureteric stones
Eligibility criteria – intervention(s) / exposure(s) / prognostic factor(s)	○ Plain abdominal radiograph (conventional, KUB) ○ Ultrasound ○ Magnetic resonance imaging
Eligibility criteria – comparator(s) / control or reference (gold) standard	• Non contrast computed tomography
Outcomes and prioritisation	• Specificity • Sensitivity • Positive Predictive Value • Negative Predictive Value
Eligibility criteria – study design	Cross-sectional studies, cohort studies, Case–control and case series studies should be included only if there is no other evidence (as they are biased).
Other inclusion exclusion criteria	Bladder stones Open surgery for renal (kidney and ureteric) stones Laparoscopic nephrolithotomy and pyelolithotomy Non-English language studies
Proposed sensitivity / subgroup analysis, or meta-regression	Strata: • Adults (≥16 years) • Children and young people (<16 years) • Pregnant women
Selection process – duplicate screening / selection / analysis	Studies are sifted by title and abstract. Potentially significant publications obtained in full text are then assessed against the inclusion criteria specified in this protocol.
Data management (software)	• Sensitivity and specificity are calculated using Cochrane Review Manager (RevMan5). • Diagnostic meta-analyses are conducted using WinBUGS14 and graphically presented using RevMan5. • Endnote for bibliography, citations, sifting and reference management

Information sources – databases and dates	Clinical search databases to be used: Medline, Embase, Cochrane Library Date: all years Health economics search databases to be used: Medline, Embase, NHSEED, HTA Date: Medline, Embase from 2014 NHSEED, HTA – all years Language: Restrict to English only Supplementary search techniques: backward citation searching Key papers: Not known
Identify if an update	Not applicable
Author contacts	https://www.nice.org.uk/guidance/indevelopment/gid-ng10033
Highlight if amendment to previous protocol	For details please see section 4.5 of Developing NICE guidelines: the manual.
Search strategy – for one database	For details please see appendix B
Data collection process – forms / duplicate	A standardised evidence table format will be used, and published as appendix D of the evidence report.
Data items – define all variables to be collected	For details please see evidence tables in Appendix D (clinical evidence tables) or H (health economic evidence tables).
Methods for assessing bias at outcome / study level	Standard study checklists are used to critically appraise individual studies. For details please see section 6.2 of Developing NICE guidelines: the manual. The risk of bias is evaluated for each outcome on a study using the QUADAS-2 checklist.
Criteria for quantitative synthesis	For details please see section 6.4 of Developing NICE guidelines: the manual.
Methods for quantitative analysis – combining studies and exploring (in)consistency	For details please see the separate Methods report for this guideline.
Meta-bias assessment – publication bias, selective reporting bias	For details please see section 6.2 of Developing NICE guidelines: the manual. [Consider exploring publication bias for review questions where it may be more common, such as pharmacological questions, certain disease areas, etc. Describe any steps taken to mitigate against publication bias, such as examining trial registries.]
Confidence in cumulative evidence	For details please see sections 6.4 and 9.1 of Developing NICE guidelines: the manual.
Rationale / context – what is known	For details please see the introduction to the evidence review.
Describe contributions of authors and guarantor	A multidisciplinary committee developed the evidence review. The committee was convened by the National Guideline Centre (NGC) and chaired by Andrew Dickinson in line with section 3 of Developing NICE guidelines: the manual. Staff from NGC undertook systematic literature searches, appraised the evidence, conducted meta-analysis and cost-effectiveness analysis where appropriate, and drafted the evidence review in collaboration with the committee. For details please see Developing NICE guidelines: the manual.

Sources of funding / support	NGC is funded by NICE and hosted by the Royal College of Physicians.
Name of sponsor	NGC is funded by NICE and hosted by the Royal College of Physicians.
Roles of sponsor	NICE funds NGC to develop guidelines for those working in the NHS, public health and social care in England.
PROSPERO registration number	Not registered

1 **Table 7: Health economic review protocol**

Review question	**All questions – health economic evidence**
Objectives	To identify economic studies relevant to any of the review questions.
Search criteria	• Populations, interventions and comparators must be as specified in the individual review protocol above. • Studies must be of a relevant economic study design (cost-utility analysis, cost-effectiveness analysis, cost-benefit analysis, cost-consequences analysis, comparative cost analysis). • Studies must not be a letter, editorial or commentary, or a review of economic evaluations. (Recent reviews will be ordered although not reviewed. The bibliographies will be checked for relevant studies, which will then be ordered.) • Unpublished reports will not be considered unless submitted as part of a call for evidence. • Studies must be in English.
Search strategy	An economic study search will be undertaken using population-specific terms and an economic study filter – see Appendix G *[in the Full guideline]*.
Review strategy	Studies not meeting any of the search criteria above will be excluded. Studies published before 2002, abstract-only studies and studies from non-OECD countries or the USA will also be excluded. Each remaining study will be assessed for applicability and methodological limitations using the NICE economic evaluation checklist which can be found in Appendix G of the 2014 NICE guidelines manual.[86] **Inclusion and exclusion criteria** • If a study is rated as both 'Directly applicable' and with 'Minor limitations' then it will be included in the guideline. An economic evidence table will be completed and it will be included in the economic evidence profile. • If a study is rated as either 'Not applicable' or with 'Very serious limitations' then it will usually be excluded from the guideline. If it is excluded then an economic evidence table will not be completed and it will not be included in the economic evidence profile. • If a study is rated as 'Partially applicable', with 'Potentially serious limitations' or both then there is discretion over whether it should be included. **Where there is discretion** The health economist will make a decision based on the relative applicability and quality of the available evidence for that question, in discussion with the Committee if required. The ultimate aim is to include economic studies that are helpful for decision-making in the context of the guideline and the current NHS setting. If several studies are considered of sufficiently high applicability and methodological quality that they could all be included, then the health economist, in discussion with the Committee if required, may decide to include only the most applicable studies and to selectively exclude the remaining studies. All studies excluded on the basis of applicability or

	methodological limitations will be listed with explanation as excluded economic studies in Appendix M. The health economist will be guided by the following hierarchies. *Setting:* • UK NHS (most applicable). • OECD countries with predominantly public health insurance systems (for example, France, Germany, Sweden). • OECD countries with predominantly private health insurance systems (for example, Switzerland). • Studies set in non-OECD countries or in the USA will have been excluded before being assessed for applicability and methodological limitations. *Economic study type:* • Cost-utility analysis (most applicable). • Other type of full economic evaluation (cost-benefit analysis, cost-effectiveness analysis, cost-consequences analysis). • Comparative cost analysis. • Non-comparative cost analyses including cost-of-illness studies will have been excluded before being assessed for applicability and methodological limitations. *Year of analysis:* • The more recent the study, the more applicable it will be. • Studies published in 2002 or later but that depend on unit costs and resource data entirely or predominantly from before 2002 will be rated as ‘Not applicable’. • Studies published before 2002 will have been excluded before being assessed for applicability and methodological limitations. *Quality and relevance of effectiveness data used in the economic analysis:* • The more closely the clinical effectiveness data used in the economic analysis matches with the outcomes of the studies included in the clinical review the more useful the analysis will be for decision-making in the guideline.

Appendix B: Literature search strategies

The literature searches for this review are detailed below and complied with the methodology outlined in Developing NICE guidelines: the manual 2014, updated 2017 https://www.nice.org.uk/guidance/pmg20/resources/developing-nice-guidelines-the-manual-pdf-72286708700869

For more detailed information, please see the Methodology Review. [Add cross reference]

B.1 Clinical search literature search strategy

Searches were constructed using a PICO framework where population (P) terms were combined with Intervention (I) and in some cases Comparison (C) terms. Outcomes (O) are rarely used in search strategies for interventions as these concepts may not be well described in title, abstract or indexes and therefore difficult to retrieve. Search filters were applied to the search where appropriate.

Table 8: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline (OVID)	1946 – 29 November 2017	Exclusions Randomised controlled trials Systematic review studies Observational studies

Database	Dates searched	Search filter used
		Diagnostic tests studies
Embase (OVID)	1974 – 29 November 2017	Exclusions Randomised controlled trials Systematic review studies Observational studies Diagnostic tests studies
The Cochrane Library (Wiley)	Cochrane Reviews to 2017 Issue 11 of 12 CENTRAL to 2017 Issue 10 of 12 DARE, and NHSEED to 2015 Issue 2 of 4 HTA to 2016 Issue 4 of 4	None

1 **Medline (Ovid) search terms**

1.	exp urolithiasis/
2.	(nephrolitiasis or nephrolith or nephroliths or urolithias?s or ureterolithias?s).ti,ab.
3.	((renal or kidney* or urinary or ureter* or urethra*) adj3 (stone* or calculi or calculus or calculosis or lithiasis or c?olic*)).ti,ab.
4.	stone disease*.ti,ab.
5.	((calculi or calculus or calcium oxalate or cystine) adj3 (crystal* or stone* or lithiasis)).ti,ab.
6.	or/1-5
7.	letter/
8.	editorial/
9.	news/
10.	exp historical article/
11.	Anecdotes as Topic/
12.	comment/
13.	case report/
14.	(letter or comment*).ti.
15.	or/7-14
16.	randomized controlled trial/ or random*.ti,ab.
17.	15 not 16
18.	animals/ not humans/
19.	exp Animals, Laboratory/
20.	exp Animal Experimentation/
21.	exp Models, Animal/
22.	exp Rodentia/
23.	(rat or rats or mouse or mice).ti.
24.	or/17-23
25.	6 not 24
26.	limit 25 to English language
27.	exp Tomography/
28.	tomograph*.ti,ab.
29.	(NCCT or CT or UHCT).ti,ab.
30.	((CAT or body) adj2 scan*).ti,ab.

31.	or/27-30
32.	Radiography/
33.	Radiography, Abdominal/
34.	Urography/
35.	(radiograph* or x ray* or xray* KUB or urograph*).ti,ab.
36.	or/32-35
37.	Ultrasonography/
38.	(ultrasonograph* or ultrasound or ultrasonic or sonograph* or echograph* or echotomograph*).ti,ab.
39.	(ultra adj2 (sound or sonic)).ti,ab.
40.	(sound* adj2 (wave* or frequenc*)).ti,ab.
41.	(US adj3 imag*).ti,ab.
42.	or/37-41
43.	Magnetic Resonance Imaging/
44.	((magnetic or nuclear) adj2 resonance adj3 imag*).ti,ab.
45.	(MRI or NMR or NMRI or fMRI or MR).ti,ab.
46.	or/43-45
47.	31 or 36 or 42 or 46
48.	26 and 47
49.	exp "sensitivity and specificity"/
50.	(sensitivity or specificity).ti,ab.
51.	((pre test or pretest or post test) adj probability).ti,ab.
52.	(predictive value* or PPV or NPV).ti,ab.
53.	likelihood ratio*.ti,ab.
54.	likelihood function/
55.	((area under adj4 curve) or AUC).ti,ab.
56.	(receive* operat* characteristic* or receive* operat* curve* or ROC curve*).ti,ab.
57.	(diagnos* adj3 (performance* or accurac* or utilit* or value* or efficien* or effectiveness)).ti,ab.
58.	gold standard.ab.
59.	or/49-58
60.	randomized controlled trial.pt.
61.	controlled clinical trial.pt.
62.	randomi#ed.ti,ab.
63.	placebo.ab.
64.	randomly.ti,ab.
65.	Clinical Trials as topic.sh.
66.	trial.ti.
67.	or/60-66
68.	Meta-Analysis/
69.	exp Meta-Analysis as Topic/
70.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
71.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
72.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.

73.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
74.	(search* adj4 literature).ab.
75.	(medline or pubmed or cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or science citation index or bids or cancerlit).ab.
76.	cochrane.jw.
77.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
78.	or/68-77
79.	Epidemiologic studies/
80.	Observational study/
81.	exp Cohort studies/
82.	(cohort adj (study or studies or analys* or data)).ti,ab.
83.	((follow up or observational or uncontrolled or non randomi#ed or epidemiologic*) adj (study or studies or data)).ti,ab.
84.	((longitudinal or retrospective or prospective or cross sectional) and (study or studies or review or analys* or cohort* or data)).ti,ab.
85.	Controlled Before-After Studies/
86.	Historically Controlled Study/
87.	Interrupted Time Series Analysis/
88.	(before adj2 after adj2 (study or studies or data)).ti,ab.
89.	or/79-88
90.	exp case control study/
91.	case control*.ti,ab.
92.	or/90-91
93.	89 or 92
94.	Cross-sectional studies/
95.	(cross sectional and (study or studies or review or analys* or cohort* or data)).ti,ab.
96.	or/94-95
97.	89 or 96
98.	89 or 92 or 96
99.	59 or 67 or 78 or 98
100.	48 and 99

Embase (Ovid) search terms

1.	exp urolithiasis/
2.	(nephrolitiasis or nephrolith or nephroliths or urolithias?s or ureterolithias?s).ti,ab.
3.	((renal or kidney* or urinary or ureter* or urethra*) adj3 (stone* or calculi or calculus or calculosis or lithiasis or c?olic*)).ti,ab.
4.	stone disease*.ti,ab.
5.	((calculi or calculus or calcium oxalate or cystine) adj3 (crystal* or stone* or lithiasis)).ti,ab.
6.	or/1-5
7.	letter.pt. or letter/
8.	note.pt.
9.	editorial.pt.
10.	case report/ or case study/
11.	(letter or comment*).ti.

12.	or/7-11
13.	randomized controlled trial/ or random*.ti,ab.
14.	12 not 13
15.	animal/ not human/
16.	nonhuman/
17.	exp Animal Experiment/
18.	exp Experimental Animal/
19.	animal model/
20.	exp Rodent/
21.	(rat or rats or mouse or mice).ti.
22.	or/14-21
23.	6 not 22
24.	limit 23 to English language
25.	exp *tomography/
26.	tomograph*.ti,ab.
27.	(NCCT or CT or UHCT).ti,ab.
28.	((CAT or body) adj2 scan*).ti,ab.
29.	or/25-28
30.	*radiography/
31.	*abdominal radiography/
32.	*urography/
33.	(radiograph* or x ray* or xray* KUB or urograph*).ti,ab.
34.	or/30-33
35.	*echography/
36.	(ultrasonograph* or ultrasound or ultrasonic or sonograph* or echograph* or echotomograph*).ti,ab.
37.	(ultra adj2 (sound or sonic)).ti,ab.
38.	(sound* adj2 (wave* or frequenc*)).ti,ab.
39.	(US adj3 imag*).ti,ab.
40.	or/35-39
41.	*nuclear magnetic resonance imaging/
42.	((magnetic or nuclear) adj2 resonance adj3 imag*).ti,ab.
43.	(MRI or NMR or NMRI or fMRI or MR).ti,ab.
44.	or/41-43
45.	29 or 34 or 40 or 44
46.	24 and 45
47.	exp "sensitivity and specificity"/
48.	(sensitivity or specificity).ti,ab.
49.	((pre test or pretest or post test) adj probability).ti,ab.
50.	(predictive value* or PPV or NPV).ti,ab.
51.	likelihood ratio*.ti,ab.
52.	((area under adj4 curve) or AUC).ti,ab.
53.	(receive* operat* characteristic* or receive* operat* curve* or ROC curve*).ti,ab.
54.	(diagnos* adj3 (performance* or accurac* or utilit* or value* or efficien* or effectiveness)).ti,ab.
55.	diagnostic accuracy/

56.	diagnostic test accuracy study/
57.	gold standard.ab.
58.	or/47-57
59.	random*.ti,ab.
60.	factorial*.ti,ab.
61.	(crossover* or cross over*).ti,ab.
62.	((doubl* or singl*) adj blind*).ti,ab.
63.	(assign* or allocat* or volunteer* or placebo*).ti,ab.
64.	crossover procedure/
65.	single blind procedure/
66.	randomized controlled trial/
67.	double blind procedure/
68.	or/59-67
69.	systematic review/
70.	meta-analysis/
71.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
72.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
73.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
74.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
75.	(search* adj4 literature).ab.
76.	(medline or pubmed or cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or science citation index or bids or cancerlit).ab.
77.	cochrane.jw.
78.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
79.	or/69-78
80.	Clinical study/
81.	Observational study/
82.	family study/
83.	longitudinal study/
84.	retrospective study/
85.	prospective study/
86.	cohort analysis/
87.	follow-up/
88.	cohort*.ti,ab.
89.	87 and 88
90.	(cohort adj (study or studies or analys* or data)).ti,ab.
91.	((follow up or observational or uncontrolled or non randomi#ed or epidemiologic*) adj (study or studies or data)).ti,ab.
92.	((longitudinal or retrospective or prospective or cross sectional) and (study or studies or review or analys* or cohort* or data)).ti,ab.
93.	(before adj2 after adj2 (study or studies or data)).ti,ab.
94.	or/80-86,89-93
95.	exp case control study/
96.	case control*.ti,ab.
97.	or/95-96

98.	94 or 97
99.	cross-sectional study/
100.	(cross sectional and (study or studies or review or analys* or cohort* or data)).ti,ab.
101.	or/99-100
102.	94 or 101
103.	94 or 97 or 101
104.	58 or 68 or 79 or 103
105.	46 and 104

Cochrane Library (Wiley) search terms

#1.	MeSH descriptor: [Urolithiasis] explode all trees
#2.	(nephrolitiasis or nephrolith or nephroliths or urolithias?s or ureterolithias?s):ti,ab
#3.	((renal or kidney* or urinary or ureter* or urethra*) near/3 (stone* or calculi or calculus or calculosis or lithiasis or c?olic*)):ti,ab
#4.	stone disease*:ti,ab
#5.	((calculi or calculus or calcium oxalate or cystine) near/3 (crystal* or stone* or lithiasis)):ti,ab
#6.	(or #1-#5)
#7.	MeSH descriptor: [Tomography] explode all trees
#8.	tomograph*:ti,ab
#9.	(NCCT or CT or UHCT):ti,ab
#10.	((CAT or body) near/2 scan*):ti,ab
#11.	(or #7-#10)
#12.	MeSH descriptor: [Radiography] this term only
#13.	MeSH descriptor: [Radiography, Abdominal] this term only
#14.	MeSH descriptor: [Urography] explode all trees
#15.	(radiograph* or x ray* or xray or KUB or urograph*):ti,ab
#16.	(or #12-#15)
#17.	MeSH descriptor: [Ultrasonography] this term only
#18.	(ultrasonograph* or ultrasound or ultrasonic or sonograph* or echograph* or echotomograph*):ti,ab
#19.	(ultra near/2 (sound or sonic)):ti,ab
#20.	(sound* near/2 (wave* or frequenc*)):ti,ab
#21.	(US near/3 imag*):ti,ab
#22.	(or #17-#21)
#23.	MeSH descriptor: [Magnetic Resonance Imaging] this term only
#24.	((magnetic or nuclear) near/2 resonance near/3 imag*):ti,ab
#25.	(MRI or NMR or NMRI or fMRI or MR):ti,ab
#26.	(or #23-#25)
#27.	#11 or #16 or #22 or #26
#28.	#6 and #27

B.2 Health Economics literature search strategy

Health economic evidence was identified by conducting a broad search relating to renal and ureteric stones population in NHS Economic Evaluation Database (NHS EED – this ceased to be updated after March 2015) and the Health Technology Assessment database (HTA) with no date restrictions. NHS EED and HTA databases are hosted by the Centre for

Research and Dissemination (CRD). Additional searches were run on Medline and Embase for health economics studies.

Table 9: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline	2014 – 9 March 2018	Exclusions Health economics studies
Embase	2014 – 9 March 2018	Exclusions Health economics studies
Centre for Research and Dissemination (CRD)	HTA - Inception – 9 March 2018 NHSEED - Inception to March 2015	None

Medline (Ovid) search terms

1.	exp urolithiasis/
2.	(nephrolitiasis or nephrolith or nephroliths or urolithias?s or ureterolithias?s).ti,ab.
3.	((renal or kidney* or urinary or ureter* or urethra*) adj3 (stone* or calculi or calculus or calculosis or lithiasis or c?olic*)).ti,ab.
4.	stone disease*.ti,ab.
5.	((calculi or calculus or calcium oxalate or cystine) adj3 (crystal* or stone* or lithiasis)).ti,ab.
6.	or/1-5
7.	letter/
8.	editorial/
9.	news/
10.	exp historical article/
11.	Anecdotes as Topic/
12.	comment/
13.	case report/
14.	(letter or comment*).ti.
15.	or/7-14
16.	randomized controlled trial/ or random*.ti,ab.
17.	15 not 16
18.	animals/ not humans/
19.	exp Animals, Laboratory/
20.	exp Animal Experimentation/
21.	exp Models, Animal/
22.	exp Rodentia/
23.	(rat or rats or mouse or mice).ti.
24.	or/17-23
25.	6 not 24
26.	limit 25 to English language
27.	Economics/
28.	Value of life/
29.	exp "Costs and Cost Analysis"/
30.	exp Economics, Hospital/

31.	exp Economics, Medical/
32.	Economics, Nursing/
33.	Economics, Pharmaceutical/
34.	exp "Fees and Charges"/
35.	exp Budgets/
36.	budget*.ti,ab.
37.	cost*.ti.
38.	(economic* or pharmaco?economic*).ti.
39.	(price* or pricing*).ti,ab.
40.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
41.	(financ* or fee or fees).ti,ab.
42.	(value adj2 (money or monetary)).ti,ab.
43.	or/27-42
44.	26 and 43

Embase (Ovid) search terms

1.	exp urolithiasis/
2.	(nephrolitiasis or nephrolith or nephroliths or urolithias?s or ureterolithias?s).ti,ab.
3.	((renal or kidney* or urinary or ureter* or urethra*) adj3 (stone* or calculi or calculus or calculosis or lithiasis or c?olic*)).ti,ab.
4.	stone disease*.ti,ab.
5.	((calculi or calculus or calcium oxalate or cystine) adj3 (crystal* or stone* or lithiasis)).ti,ab.
6.	or/1-5
7.	letter.pt. or letter/
8.	note.pt.
9.	editorial.pt.
10.	case report/ or case study/
11.	(letter or comment*).ti.
12.	or/7-11
13.	randomized controlled trial/ or random*.ti,ab.
14.	12 not 13
15.	animal/ not human/
16.	nonhuman/
17.	exp Animal Experiment/
18.	exp Experimental Animal/
19.	animal model/
20.	exp Rodent/
21.	(rat or rats or mouse or mice).ti.
22.	or/14-21
23.	6 not 22
24.	limit 23 to English language
25.	health economics/

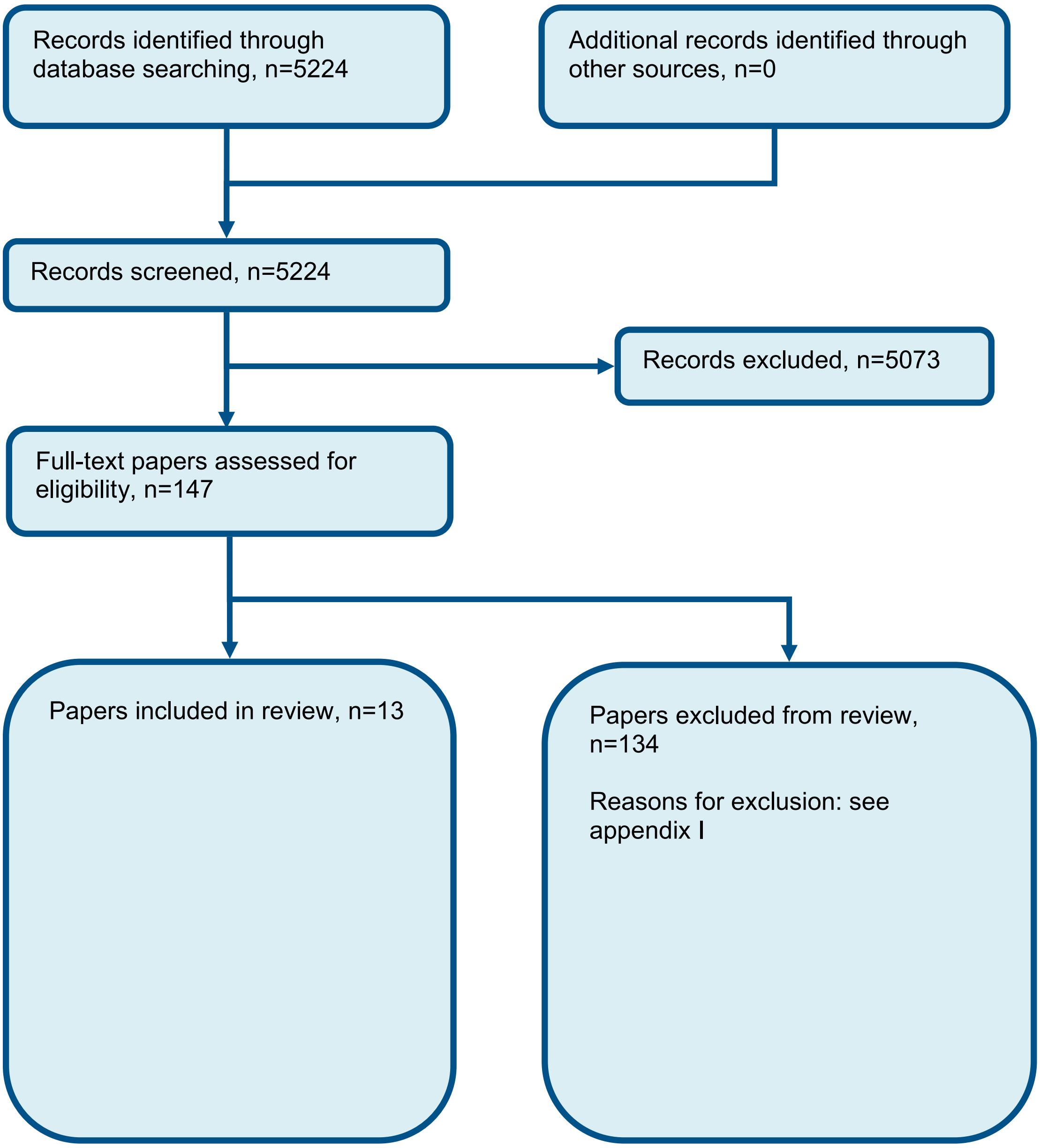
26.	exp economic evaluation/
27.	exp health care cost/
28.	exp fee/
29.	budget/
30.	funding/
31.	budget*.ti,ab.
32.	cost*.ti.
33.	(economic* or pharmaco?economic*).ti.
34.	(price* or pricing*).ti,ab.
35.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
36.	(financ* or fee or fees).ti,ab.
37.	(value adj2 (money or monetary)).ti,ab.
38.	or/25-37
39.	24 and 38

NHS EED and HTA (CRD) search terms

#1.	MeSH DESCRIPTOR urolithiasis EXPLODE ALL TREES
#2.	(((nephrolitiasis or nephrolith or urolithiasis)))
#3.	((((renal or kidney or urinary or ureteric or ureteral or ureter or urethra*) adj2 (stone* or calculi or calculus or calculosis or lithiasis or colic))))
#4.	((stone disease*))
#5.	((((calculi or calculus) adj2 (stone* or lithiasis))))
#6.	(#1 OR #2 OR #3 OR #4 OR #5)

Appendix C: Clinical evidence selection

Figure 1: Flow chart of clinical study selection for the review of What is the clinical and cost effectiveness of performing imaging for diagnosing renal and ureteric stones?



Appendix D: Clinical evidence tables

Reference	Chan 2008
Study type	Retrospective cohort study
Study methodology	Data source: retrospective review of imaging Recruitment: consecutive patients being investigated for suspected urolithiasis
Number of patients	n = 100
Patient characteristics	Age, mean (SD): 45 (SD not reported) Gender (male to female ratio): 63:37 Ethnicity: Not reported Setting: Not reported Country: Ireland Inclusion criteria: Patients presenting with acute renal colic who had both a KUB and UHCT within a 3 hour period. Exclusion criteria: Not reported
Target condition(s)	Urolithiasis
Index test(s) and reference standard	<u>Index test</u> KUB using a standard KUB protocol. <u>Reference standard</u> CT. All CT imaging was obtained with a spiral CT scanner using a standard low dose protocol (100 mAs, 120 kVp, 1.4 pitch, single breath-hold) extending from the top of the kidneys to the base of the bladder. UHCT and KUB pairs were divided into two separate groups and read by two radiologists in consensus who were experienced in genitourological radiology. The UHCT and KUB subsets were reviewed at separate time intervals to prevent internal bias. All revisions were done on a picture archiving and communications system (PACS). Time between measurement of index test and reference standard: 3 hours

<table>
<tr><th>Reference</th><th>Chan 2008</th></tr>
<tr><td></td><td></td></tr>
<tr><td>2×2 table</td><td>
<table>
<tr><td></td><td>Reference standard +</td><td>Reference standard −</td><td>Total</td></tr>
<tr><td>Index test +</td><td>39</td><td>2</td><td>41</td></tr>
<tr><td>Index test −</td><td>20</td><td>39</td><td>59</td></tr>
<tr><td>Total</td><td>59</td><td>41</td><td>100</td></tr>
</table>
</td></tr>
<tr><td>Statistical measures</td><td><u>Index text: KUB</u>
Sensitivity: 66.1% (52.61, 77.92)
Specificity: 95.12% (83.47, 99.40)
PPV: 95.12% (83.29, 98.71)
NPV: 66.1% (57.56, 73.71)
PLR: 13.55 (3.46, 53.01)
NLR: 0.36 (0.25, 0.51)</td></tr>
<tr><td>Source of funding</td><td>Not reported</td></tr>
<tr><td>Limitations</td><td>Risk of bias: None
Indirectness: None</td></tr>
<tr><td>Comments</td><td></td></tr>
</table>

1

Reference	Cifci 2016
Study type	Prospective cohort study
Study methodology	Data source: patients admitted to the urology department Recruitment: consecutive patients being investigated for suspected urolithiasis
Number of patients	n = 159
Patient characteristics	Age, mean (SC): 41.5 years (13 years) Gender (male to female ratio): 120:39 Ethnicity: Not reported Setting: Not reported

Reference	Cifci 2016
	Country: Turkey Inclusion criteria: patients who were admitted to the urology department with suspected acute urinary calculi (flank pain, hematuria or patients with a history of urinary calculi) Exclusion criteria: Images that did not meet the required conditions (inadequate anatomy and image quality) were excluded from the study. Patients with MRU images performed more than 3 hours after CT examination were also excluded from the study, due to a potential change in the stone localization. Additionally, patients who were not able to hold their breath, yielding images with respiratory, motion or volume artifacts, were excluded from the study.
Target condition(s)	Ureteral calculi
Index test(s) and reference standard	<u>Index test</u> MRU. MRU scans were performed on a 1.5-Tesla (T) unit (Intera, Gyroscan, Philips Medical Systems, The Netherlands) and a four-channel phased-array body coil was used, from the top of the kidneys to the pubic symphysis. Axial, coronal and sagittal B-TFE (TR: 3.5–5, TE: 1.5–2.5, slice thickness: +4 mm and gap: -2.5 mm) images were taken while breath was held from the top of the kidneys to the pubic symphysis. <u>Reference standard</u> CT. Abdominal non-enhanced CT (Philips, MX 8000, The Netherlands) was performed from the top of the kidneys to the pubic symphysis, while breath was held, at 80–120 kV, 100–120 mAs, and 5a -mm slice thickness with 3-mm reconstruction. All patients were hydrated a minimum of 1 hour before CT and MRU imaging. No contrast media was used. Time between measurement of index test and reference standard: 1-3 hours
2×2 table observer 1	(see table below)
Statistical measures	<u>Index text: KUB</u> Sensitivity: 65.88% (54.80, 75.82) Specificity: 95.95% (88.61, 99.16) PPV: 94.92% (85.91, 98.28) NPV: 71.00% (64.48, 76.75) PLR: 16.25 (5.31, 49.75) NLR: 0.36 (0.26, 0.48)

2×2 table observer 1:

	Reference standard +	Reference standard −	Total
Index test +	56	3	59
Index test −	29	71	100
Total	85	74	159

Reference	Cifci 2016			
2×2 table observer 2		Reference standard +	Reference standard −	Total
	Index test +	61	0	61
	Index test −	24	74	98
	Total	85	74	159
Statistical measures	Index text: KUB Sensitivity: 71.76% (60.96, 81.00) Specificity: 100.00% (95.14, 100.00) PPV: 100.00 NPV: 75.51% (68.72, 81.23) PLR: - NLR: 0.28 (0.20, 0.40)			
Source of funding	No funding			
Limitations	Risk of bias: reference standard Indirectness: history of stones			
Comments				

Reference	De Souza 2007
Study type	Prospective cohort study
Study methodology	Data source: patients referred from the emergency department Recruitment: consecutive patients being evaluated for acute flank pain
Number of patients	n = 52

Reference	De Souza 2007
Patient characteristics	Age, mean (SD): Not reported Gender (male to female ratio): Not reported Ethnicity: Not reported Setting: Not reported Country: Brazil Inclusion criteria: Not reported Exclusion criteria: known renal diseases or imaging signs of pyelonephritis, chronic renal insufficiency, nephrocalcinosis and staghorn calculus.
Target condition(s)	Ureteral stone
Index test(s) and reference standard	<u>Index test</u> US performed transabdominally, after ingestion of water. Sonography was performed by senior radiology residents and immediately checked by attending radiologists, using a Philips SD800 scanner (Philips Medical Systems, Eindhoven, Netherlands) with a convex (curved phased array) transducer (2-5 MHz) and transducer frequencies selected to optimize the imaging of the kidneys, ureters and bladder. The US diagnosis of ureteral calculi required the demonstration of an intraluminal hyperechoic structure causing acoustic shadowing. The presence of collecting system dilatation was also evaluated. No patient underwent transvaginal or transrectal sonographic examination. <u>Reference standard</u> CT. NCT scans were acquired after US examination, on a Tomoscan EV-EV1 (Philips Medical Systems, Eindhoven, Netherlands) using Secura Release 1.3 software. The scan parameters included helical data acquisition, with section thickness 3-5 mm, using 120 kV and 200 mAs and a pitch of 1-1.5. Images were obtained during apnea, from the top of the kidneys to the base of the bladder, and no contrast medium was used. The NCT images were interpreted by a senior resident, using an electronic workstation (Philips), and subsequently reviewed by three experienced abdominal radiologists in a blinded manner. Time between measurement of index test and reference standard: both were performed within 8 hours of the onset of colic (average of 4 hours between the two tests)
2×2 table	

	Reference standard +	Reference standard −	Total
Index test +	9	0	9
Index test −	31	12	43

Reference	**De Souza 2007**				
	Total	40	12	52	
Statistical measures	Index text: KUB Sensitivity: 22.50% (10.84, 38.45) Specificity: 100.00% (73.54, 100.00) PPV: 100.00% NPV: 27.91% (24.67, 31.39) PLR: - NLR: 0.78 (0.66, 0.92)				
Source of funding	None				
Limitations	Risk of bias: index test, reference standard Indirectness: none				
Comments					

Reference	**Haroun 2010**
Study type	Retrospective cohort study
Study methodology	Data source: Not reported Recruitment: Not reported
Number of patients	n = 156
Patient characteristics	Age, mean (SD): male patients 51 (16) years; female patients 46 (18) years Gender (male to female ratio): 102:54 Ethnicity: Not reported Setting: Not reported Country: Jordan Inclusion criteria: patients who underwent UHCT scan and US for suspicion of urolithiasis

<table>
<tr><th>Reference</th><th>Haroun 2010</th></tr>
<tr><td></td><td>Exclusion criteria: Not reported</td></tr>
<tr><td>Target condition(s)</td><td>Urolithiasis</td></tr>
<tr><td>Index test(s) and reference standard</td><td><u>Index test</u>
Ultrasound examinations were performed by the trans-abdominal approach for all patients, after ensuring a full urinary bladder, using 3.5 or 5MHz probes. The kidneys were evaluated in the longitudinal and transverse projections. Whenever possible, the course of ureters was also followed down to the urinary bladder with special attention to the uretero-vesical junction. The urinary bladder was also examined in both planes.

<u>Reference standard</u>
CT. UHCT scan was performed with a Somatom Plus 4 machine (Siemens, Germany). The images were obtained with the patient in supine position during breath-hold plus quiet breathing. The explored area extended from the upper poles of both kidneys down to pubic symphysis using five mm collimation with a table speed of 7.5 mm/second giving a pitch of 1.5:1 The images were obtained with a 0.75-second gantry rotation using 120 KVp and 206 mA giving 155 mAs. Multiplanar reformation (MPR) in coronal oblique direction was used when the location of stone was uncertain. CT scan images were reported by consultant radiologists on hard copy films.

Time between measurement of index test and reference standard: Not reported</td></tr>
<tr><td>2×2 table</td><td><table>
<tr><th></th><th>Reference standard +</th><th>Reference standard −</th><th>Total</th></tr>
<tr><td>Index test +</td><td>34</td><td>9</td><td>43</td></tr>
<tr><td>Index test −</td><td>25</td><td>88</td><td>113</td></tr>
<tr><td>Total</td><td>59</td><td>97</td><td>156</td></tr>
</table></td></tr>
<tr><td>Statistical measures</td><td><u>Index text: KUB</u>
Sensitivity: 57.63% (44.07, 70.39)
Specificity: 90.72% (83.12, 95.67)
PPV: 79.07% (66.14, 87.96)
NPV: 77.88% (72.19, 82.68)
PLR: 6.21 (3.21, 12.01)
NLR: 0.47 (0.34, 0.63)</td></tr>
<tr><td>Source of funding</td><td>Not reported</td></tr>
<tr><td>Limitations</td><td>Risk of bias: index test, reference test, flow and timing
Indirectness: none</td></tr>
<tr><td>Comments</td><td></td></tr>
</table>

Reference	Kanno 2017
Study type	Retrospective cohort study
Study methodology	Data source: database review Recruitment: Not reported
Number of patients	n = 822
Patient characteristics	Age, median: 60 years Gender (male to female ratio): 553:269 Ethnicity: Not reported Setting: Not reported Country: Japan Inclusion criteria: Indications for imaging included acute flank pain, hematuria, or a history of urinary stones as previously described. Whereas new patients routinely underwent US for screening of the urinary tract at our institution, we also performed both KUB and NCCT in patients with acute flank pain and suspected urolithiasis to get information such as stone size, mean stone density, and skin-to-stone distance, except for patients who underwent NCCT in other hospitals and were referred to our institution. Exclusion criteria: Patients with solitary kidney, staghorn calculi, and urinary diversion were excluded from our analysis.
Target condition(s)	Renal stones
Index test(s) and reference standard	<u>Index test</u> US. All US examinations were performed by 7 experienced sonographers who are specialized in handling urologic US. Echogenic foci (with or without acoustic shadowing) that were seen in the renal pelvis or calices on US were diagnosed as renal stones, because small stones may not cast an acoustic shadow as described previously KUB. No further details reported. <u>Reference standard</u>

Reference	Kanno 2017			
	CT. NCCT (Toshiba Aquilion ONE 640, Toshiba Medical Systems, Tokyo, Japan) was performed from the upper abdomen to the pelvis with images reconstructed at 1- to 2-mm intervals. US of the kidneys and bladder was performed using grayscale sonography (Toshiba SSA550A) with a 3.5-MHz convex transducer. Time between measurement of index test and reference standard: same day			
2×2 table - US		Reference standard +	Reference standard −	Total
	Index test +	882	206	1088
	Index test −	112	444	556
	Total	994	650	1644
2×2 table - KUB		Reference standard +	Reference standard −	Total
	Index test +	488	6	494
	Index test −	506	644	1150
	Total	994	650	1644
Statistical measures	<u>Index text: US</u> Sensitivity: 88.73% (86.60, 90.63) Specificity: 68.31% (64.58, 71.87) PPV: 81.07% (79.24, 82.77) NPV: 79.86% (76.77, 82.63) PLR: 2.80 (2.50, 3.14) NLR: 0.16 (0.14, 0.20) <u>Index text: KUB</u> Sensitivity: 49.09% (45.94, 52.25) Specificity: 99.08% (98.00, 99.66) PPV: 98.79% (97.34, 99.45) NPV: 56.00% (54.48, 57.51) PLR: 53.19 (23.92, 118.25) NLR: 0.51 (0.48, 0.55)			
Source of funding	Not reported			
Limitations	Risk of bias: patient selection Indirectness: history of stones			
Comments				

Reference	**Kanno 2014a #1274**
Study type	Retrospective cohort study
Study methodology	Data source: database review Recruitment: Not reported
Number of patients	n = 428
Patient characteristics	Age, median: Not reported Gender (male to female ratio): Not reported Ethnicity: Not reported Setting: Not reported Country: Japan Inclusion criteria: Indications for imaging included acute flank pain, haematuria, or a history of urinary stones Exclusion criteria: Patients with solitary kidney, staghorn calculi, and urinary diversion were excluded
Target condition(s)	Renal stones
Index test(s) and reference standard	<u>Index test</u> US. US was performed using gray scale sonography (SSA550A; Toshiba) with a 3.5-MHz convex transducer. All US examinations were performed by 4 experienced sonographers who are specialized in handling urologic US. Echogenic foci (with or without acoustic shadowing) that were seen in the renal pelvis or calices on US were diagnosed as renal stones because small stones may not cast an acoustic shadow. <u>Reference standard</u> CT. NCCT (Aquilion ONE 640; Toshiba) was performed from the upper abdomen to the pelvis with images reconstructed at 1 or 2 mm intervals. Time between measurement of index test and reference standard: same day

Reference	**Kanno 2014a #1274**			
2×2 table – individual stone		Reference standard +	Reference standard −	Total
	Index test +	285	80	365
	Index test −	76	412	488
	Total	361	492	853
Statistical measures	Index text: US – individual stones Sensitivity: 78.95% (74.37, 83.04) Specificity: 83.74 % (80.18, 86.89) PPV: 78.08% (74.33, 81.43) NPV: 84.43% (81.56, 86.92) PLR: 4.86 (3.95, 5.97) NLR: 0.25 (0.21, 0.31)			
Source of funding	No financial interests			
Limitations	Risk of bias: reference standard Indirectness: history of stones			
Comments				

Reference	**Kanno 2014**
Study type	Retrospective cohort study
Study methodology	Data source: database review Recruitment: Not reported
Number of patients	n = 428
Patient characteristics	Age, median: Not reported Gender (male to female ratio): Not reported Ethnicity: Not reported Setting: Not reported

<table>
<tr><td>Reference</td><td>Kanno 2014</td></tr>
<tr><td></td><td>Country: Japan

Inclusion criteria: Indications for imaging included acute flank pain, haematuria, or a history of urinary stones

Exclusion criteria: Patients with solitary kidney, staghorn calculi, and urinary diversion were excluded</td></tr>
<tr><td>Target condition(s)</td><td>Ureteric stones</td></tr>
<tr><td>Index test(s) and reference standard</td><td><u>Index test</u>
US. US was performed using gray scale sonography (SSA550A; Toshiba) with a 3.5-MHz convex transducer. All US examinations were performed by 4 experienced sonographers who are specialized in handling urologic US. Echogenic foci (with or without acoustic shadowing) that were seen in the renal pelvis or calices on US were diagnosed as renal stones because small stones may not cast an acoustic shadow.

<u>Reference standard</u>
CT. NCCT (Aquilion ONE 640; Toshiba) was performed from the upper abdomen to the pelvis with images reconstructed at 1 or 2 mm intervals.

Time between measurement of index test and reference standard: same day</td></tr>
<tr><td>2×2 table – individual stone</td><td><table>
<tr><td></td><td>Reference standard +</td><td>Reference standard −</td><td>Total</td></tr>
<tr><td>Index test +</td><td>98</td><td>73</td><td>171</td></tr>
<tr><td>Index test −</td><td>17</td><td>668</td><td>685</td></tr>
<tr><td>Total</td><td>115</td><td>741</td><td>856</td></tr>
</table></td></tr>
<tr><td>Statistical measures</td><td><u>Index text: US</u>
Sensitivity: 85.22% (77.39, 91.15)
Specificity: 90.15% (87.77, 92.20)
PPV: 57.31% (51.59, 62.84)
NPV: 97.52% (96.20, 98.39)
PLR: 8.65 (6.87, 10.89)
NLR: 0.16 (0.11, 0.25)</td></tr>
<tr><td>Source of funding</td><td>No financial interests</td></tr>
</table>

Reference	**Kanno 2014**
Limitations	Risk of bias: reference standard Indirectness: none
Comments	

Reference	**Kielar 2012**
Study type	Prospective cohort study
Study methodology	Data source: patients from an Emergency Department Recruitment: Not reported
Number of patients	n = 55
Patient characteristics	Age, mean (range): 49 years (28–81 years) Gender (male to female ratio): 38:17 Ethnicity: Not reported Setting: Not reported Country: Canada Inclusion criteria: People with flank pain in whom an unenhanced CT of the abdomen and pelvis was requested Exclusion criteria: Not reported
Target condition(s)	Urolithiasis
Index test(s) and reference standard	Index test US. The patient directly underwent a limited sonographic scan of the kidneys, ureters, and bladder (iU22; Philips Healthcare, Bothell, WA). This examination was performed for research purposes and was not considered the usual standard of care. It was performed with a standard ultrasound unit, which is always situated in the emergency radiology department, in a dedicated room next to the emergency radiology CT scanner. The examination was performed by a trained sonographer using a curved low-frequency probe (2–5 MHz) and a high pulse repetition frequency, with the machine's scale in the range of 60 to 70 cm/s. The pulse repetition frequency is defined as the number of pulses sent per second Reference standard

Reference	Kielar 2012
	CT. All CT scans were performed on a 64-slice CT scanner (Lightspeed VCT; GE Healthcare, Milwaukee, WI). All scans were performed with a low-dose, unenhanced “renal colic protocol.” The images were sent to a picture archiving and communication system at the original axial 1.25mmslice thickness in addition to the reconstructed5-mm axial images and2-mmcoronal images. The1.25-mm raw data were reviewed to eliminate the possibility of missing small calculi because of volume averaging. The post processing techniques do not expose the patient to any additional radiation. Time between measurement of index test and reference standard: US was performed right after the CT scan. No further details.
2×2 table – individual stone	(see 2×2 table below)

2×2 table – individual stone

	Reference standard +	Reference standard −	Total
Index test +	74	8	82
Index test −	40	5	45
Total	114	13	127

Statistical measures	<u>Index text: US</u> Sensitivity: 64.91% (55.41, 73.62) Specificity: 38.46% (13.86, 68.42) PPV: 90.24% (85.50, 93.55) NPV: 11.11% (5.67, 20.62) PLR: 1.05 (0.67, 1.66) NLR: 0.91 (0.44, 1.90)
Source of funding	No financial interests
Limitations	Risk of bias: patient selection Indirectness: none
Comments	

Reference	Levine 1997
Study type	Retrospective cohort study
Study methodology	Data source: retrospective review of records Recruitment: Not reported
Number of patients	n = 178

Reference	**Levine 1997**
Patient characteristics	Age, mean (range): 49 years (28–81 years) Gender (male to female ratio): 38:17 Ethnicity: Not reported Setting: Not reported Country: USA Inclusion criteria: People with acute flank pain who had undergone plain abdominal radiography before CT Exclusion criteria: Not reported
Target condition(s)	Urolithiasis
Index test(s) and reference standard	<u>Index test</u> Plain abdominal radiograph. No details reported <u>Reference standard</u> CT. All CT examinations were performed with a HiSpeed Advantage scanner (GE Medical Systems, Milwaukee, Wis). Images were obtained from the superior poles of the kidneys to the bladder base. Helical data acquisition was used with 5mm thick sections and a pitch of 1. Image clusters of 15-20 sections were obtained during suspended respiration. Time between measurement of index test and reference standard: 4 hours
2×2 table – individual stone	(see below)

	Reference standard +	Reference standard −	Total
Index test +	39	25	70
Index test −	27	60	87
Total	72	85	157

Reference	**Levine 1997**
Statistical measures	Index text: US Sensitivity: 59.09% (46.29, 71.05) Specificity: 70.59% (59.71, 79.98) PPV: 60.94% (51.47, 69.64) NPV: 68.97% (61.72, 75.39) PLR: 2.01 (1.37, 2.95) NLR: 0.58 (0.42, 0.80)
Source of funding	Not reported
Limitations	Risk of bias: reference standard Indirectness: none
Comments	

Reference	**Passerotti 2009**
Study type	Prospective cohort study
Study methodology	Data source: emergency department Recruitment: Consecutive patients
Number of patients	n = 50
Patient characteristics	Age, mean (range): 13.1 (2–18 years) Gender (male to female ratio): 25:25 Ethnicity: Not reported Setting: Not reported Country: USA Inclusion criteria: People younger than 18 years who presented to the emergency department or the urology clinic with signs, symptoms or a history suggestive of urolithiasis.

Reference	Passerotti 2009
	Exclusion criteria: Not reported
Target condition(s)	Urinary calculi
Index test(s) and reference standard	Index test US. Renal US using a standard departmental protocol with a Sequoia™ 512 or an IU22 machine (Philips, Bothell, Washington) with a variable combination of curved and vector array transducers, and frequencies best suited to the patient. Stones were defined as echogenic foci with posterior acoustic shadowing, confirmed in 2 planes and measured in 3 planes. Reference standard CT. Noncontrast CT of the abdomen and pelvis using 5 mm cuts with reconstruction at 1.5 mm was done according to standard departmental renal stone protocol using 1 of 3 scanners, including a LightSpeed™ Ultra 16, a LightSpeed CT 32 or a Somatom® Sensation 64. No patient required sedation. Peak kV and mA settings were device specific, and based on patient weight and age. Images were obtained in the axial plane and reconstructed in the coronal and sagittal planes. Stones were defined as hyper dense foci seen in the kidneys, ureters or bladder. Time between measurement of index test and reference standard: 0.5-8 hours

2×2 table – individual stone

	Reference standard +	Reference standard −	Total
Index test +	26	0	26
Index test −	8	16	24
Total	34	16	50

Statistical measures	Index text: US Sensitivity: 76.47% (58.83, 89.25) Specificity: 100.00% (79.41, 100.00) PPV: 100.00% NPV: 66.67% (52.188, 78.57) PLR: - NLR: 0.24 (0.13, 0.43)
Source of funding	Not reported
Limitations	Risk of bias: index test, reference standard

Reference	**Passerotti 2009**
	Indirectness: history of stones
Comments	

1

Reference	**Resorlu 2015**
Study type	Retrospective cohort study
Study methodology	Data source: retrospective analysis of records Recruitment: not reported
Number of patients	n = 500
Patient characteristics	Age, mean (SD): 49.8 (16.9) years Gender (male to female ratio): 297:203 Ethnicity: Not reported Setting: Not reported Country: Turkey Inclusion criteria: People who underwent an initial urinary US, followed by urinary NCCT as part of their investigation for acute flank pain during working hours (between 08:00 a.m. and 05:00p.m.—Mondays to Fridays) Exclusion criteria: Patients who passed their stone in the interval between US and NCCT and those requiring stone treatment in terms of shock wave lithotripsy (SWL) or endoscopic intervention (ureterorenoscopic stone extraction) were excluded from the study.
Target condition(s)	Urinary stones
Index test(s) and reference standard	Index test US. Urinary US was performed by radiologists with grayscale ultrasound machines (Toshiba® Aplio XG and General Electric Logiq 9) using two convex transducers with 3.5, 4.0 MHz frequency. The presence of stone was defined as an echogenic image with or without posterior acoustic shadowing, clearly located within the urinary tract. Reference standard

Reference	Resorlu 2015
	CT. All NCCT were performed with a 4-multidetector CT scanner (Toshiba® Asteion TSX-021B) without intravenous or oral contrast medium. Slices of 3-mm thickness with 1-mm reconstructed intervals were obtained, beginning from the superior aspect of the kidneys to the inferior aspect of the pubic symphysis. Stones were defined as hyperdense foci seen in the renal pelvis, calices, or ureters Time between measurement of index test and reference standard: 10 days

2×2 table

	Reference standard +	Reference standard −	Total
Index test +	91	42	133
Index test −	111	256	367
Total	202	298	500

Statistical measures	<u>Index text: US</u> Sensitivity: 45.05% (38.06, 52.19) Specificity: 85.91% (81.43, 89.65) PPV: 68.42% (61.16, 74.88) NPV: 69.75% (66.88, 72.49) PLR: 3.20 (2.32, 4.40) NLR: 0.64 (0.56, 0.73)
Source of funding	Not reported
Limitations	Risk of bias: index test, reference standard Indirectness: none
Comments	

Reference	Semins 2013
Study type	Prospective cohort study
Study methodology	Data source: emergency department Recruitment: not reported
Number of patients	n = 22

Reference	Semins 2013
Patient characteristics	Age, median (range): Not reported Gender (male to female ratio): Not reported Ethnicity: Not reported Setting: Not reported Country: USA Inclusion criteria: male or female patients with acute renal colic, age greater than 18 years, and ability to complete all necessary components of study Exclusion criteria: inability to give informed consent, contraindication to MRI, inability or unwillingness to comply with completion of studies, and women who were pregnant or in whom pregnancy status could not be determined
Target condition(s)	Obstructing stone
Index test(s) and reference standard	Index test MRU. Non-contrast HASTE (Half-Fourier single shot turbo spin-echo) MRU with a 3-T scanner Reference standard CT. Non-contrast CT. No further details. Time between measurement of index test and reference standard: MRU was performed immediately after CT
2×2 table	(see below)

2×2 table

	Reference standard +	Reference standard −	Total
Index test +	16	0	16
Index test −	3	3	6
Total	19	3	22

Reference	**Semins 2013**
Statistical measures	Index text: US Sensitivity: 84.21% (60.42, 96.62) Specificity: 100% (29.24, 100.00) PPV: 100.00% NPV: 50.00% (26.15, 73.85) PLR: NLR: 0.16 (0.06, 0.45)
Source of funding	Not reported
Limitations	Risk of bias: patient selection, index test, reference standard Indirectness: population
Comments	

Reference	**Sternberg 2016**
Study type	Retrospective cohort study
Study methodology	Data source: retrospective review of records from three institutions (University of Vermont Medical Center, Massachusetts General Hospital, Dartmouth-Hitchcock Medical Center) Recruitment: not reported
Number of patients	n = 155
Patient characteristics	Age, mean (range): Not reported Gender (male to female ratio): Not reported Ethnicity: Not reported Setting: Not reported Country: Lebanon Inclusion criteria: adult patients >18 years of age. Only formal radiologic US, not bedside-US, were included.

<table>
<tr><th>Reference</th><th>Sternberg 2016</th></tr>
<tr><td></td><td>Exclusion criteria: images were obtained >1 day apart, if imaging was of poor quality for interpretation, and/or if staghorn calculi were present</td></tr>
<tr><td>Target condition(s)</td><td>Urinary calculi</td></tr>
<tr><td>Index test(s) and reference standard</td><td><u>Index test</u>
US. No further details.

<u>Reference standard</u>
CT. A standard protocol was followed using abdominal windows and zooming in to best visualize the stone of interest. Three measurements were made (length, width, height) using axial, sagittal, and coronal sections.

Time between measurement of index test and reference standard: 1 day</td></tr>
<tr><td>2×2 table</td><td>
<table>
<tr><td></td><td>Reference standard +</td><td>Reference standard −</td><td>Total</td></tr>
<tr><td>Index test +</td><td>79</td><td>2</td><td>81</td></tr>
<tr><td>Index test −</td><td>58</td><td>16</td><td>74</td></tr>
<tr><td>Total</td><td>137</td><td>18</td><td>155</td></tr>
</table>
</td></tr>
<tr><td>Statistical measures</td><td><u>Index text: US</u>
Sensitivity: 57.66% (48.94, 66.05)
Specificity: 88.89% (65.29, 98.62)
PPV: 97.53% (91.39, 99.32)
NPV: 21.62% (17.62, 26.25)
PLR: 5.19 (1.39, 19.32)
NLR: 0.48 (0.37, 0.61)</td></tr>
<tr><td>Source of funding</td><td>Not reported</td></tr>
<tr><td>Limitations</td><td>Risk of bias: patient selection, index test, reference standard
Indirectness: none</td></tr>
<tr><td>Comments</td><td>One of the authors is owner of the Ravine Group, and a Consultant for Boston Scientific, Bard, Retrophin, and Olympus.</td></tr>
</table>

Appendix E: Coupled sensitivity and specificity forest plots and sROC curves

E.1 Coupled sensitivity and specificity forest plots

Figure 2: Sensitivity and specificity of index test ultrasound for condition renal or ureteric stones in adults

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
de Souza 2007	9	0	31	12	0.23 [0.11, 0.38]	1.00 [0.74, 1.00]
Haroun 2010	34	9	25	88	0.58 [0.44, 0.70]	0.91 [0.83, 0.96]
Kanno 2014a	285	80	76	412	0.79 [0.74, 0.83]	0.84 [0.80, 0.87]
Kanno 2014b	98	17	73	668	0.57 [0.50, 0.65]	0.98 [0.96, 0.99]
Kanno 2017	882	206	112	444	0.89 [0.87, 0.91]	0.68 [0.65, 0.72]
Resorlu 2015	91	42	111	256	0.45 [0.38, 0.52]	0.86 [0.81, 0.90]
Sternberg 2016	79	2	58	16	0.58 [0.49, 0.66]	0.89 [0.65, 0.99]

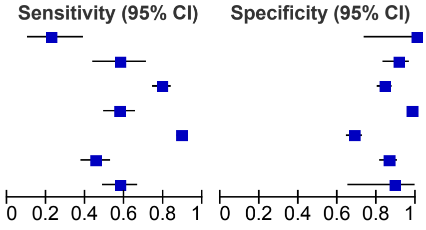


Figure 3: Sensitivity and specificity of index test plain abdominal radiograph for condition renal or ureteric stones in adults

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Chan 2008	39	2	20	39	0.66 [0.53, 0.78]	0.95 [0.83, 0.99]
Kanno 2017	488	6	506	644	0.49 [0.46, 0.52]	0.99 [0.98, 1.00]
Levine 1997	39	25	27	60	0.59 [0.46, 0.71]	0.71 [0.60, 0.80]

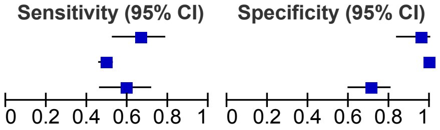


Figure 4: Sensitivity and specificity of index test MRI for condition renal or ureteric stones in adults

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Cifci 2016 (observer 1)	56	3	29	71	0.66 [0.55, 0.76]	0.96 [0.89, 0.99]
Cifi 2016 (observer 2)	61	0	24	74	0.72 [0.61, 0.81]	1.00 [0.95, 1.00]
Semins 2013	16	0	3	3	0.84 [0.60, 0.97]	1.00 [0.29, 1.00]

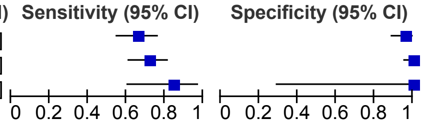
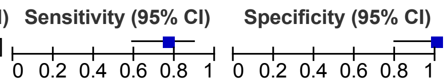


Figure 5: Sensitivity and specificity of index test ultrasound for condition renal or ureteric stones in children

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Passerotti 2009	26	0	8	16	0.76 [0.59, 0.89]	1.00 [0.79, 1.00]



E.2 ROC curves

Figure 6: Pooled with prediction region: ultrasound

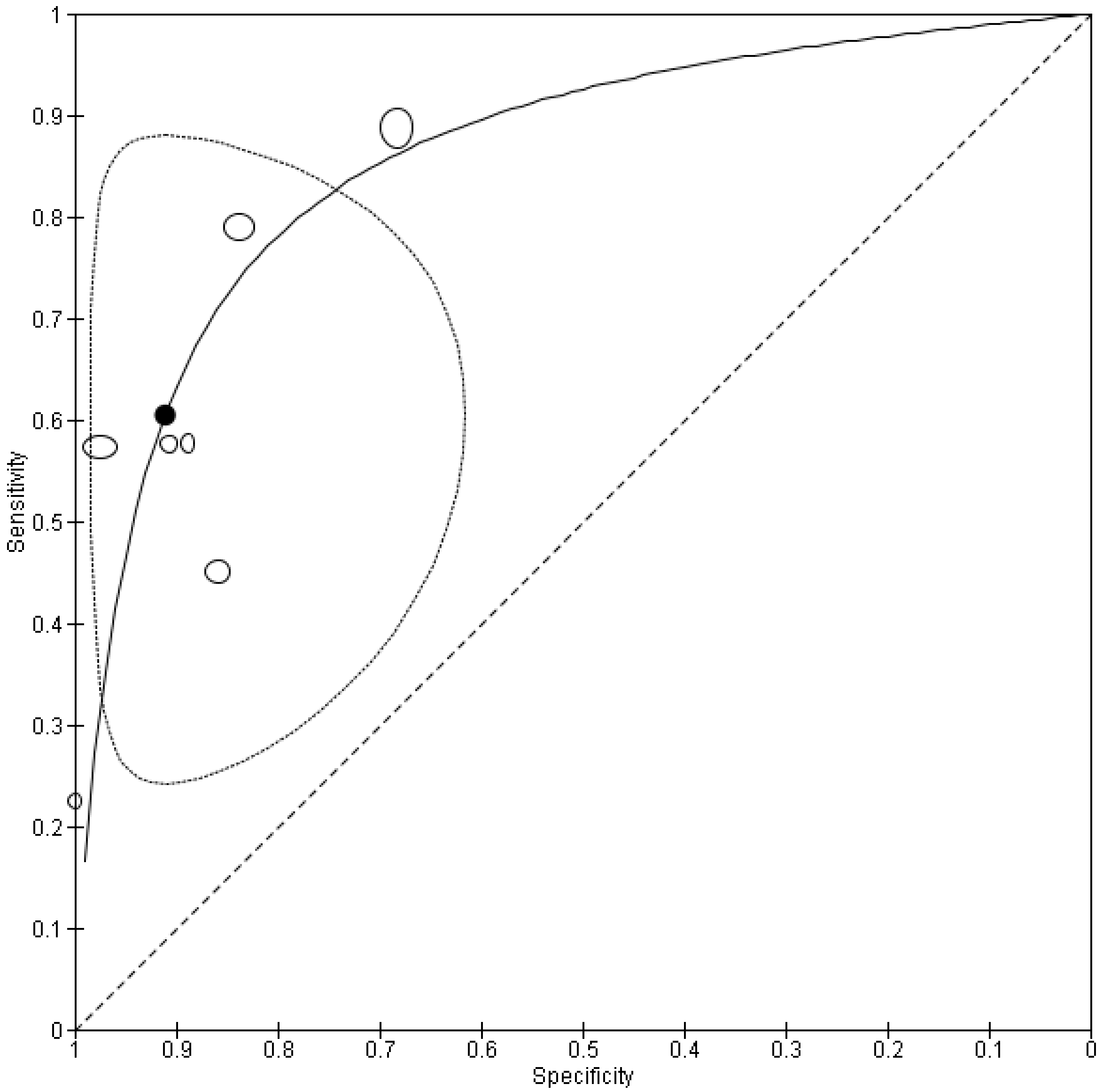
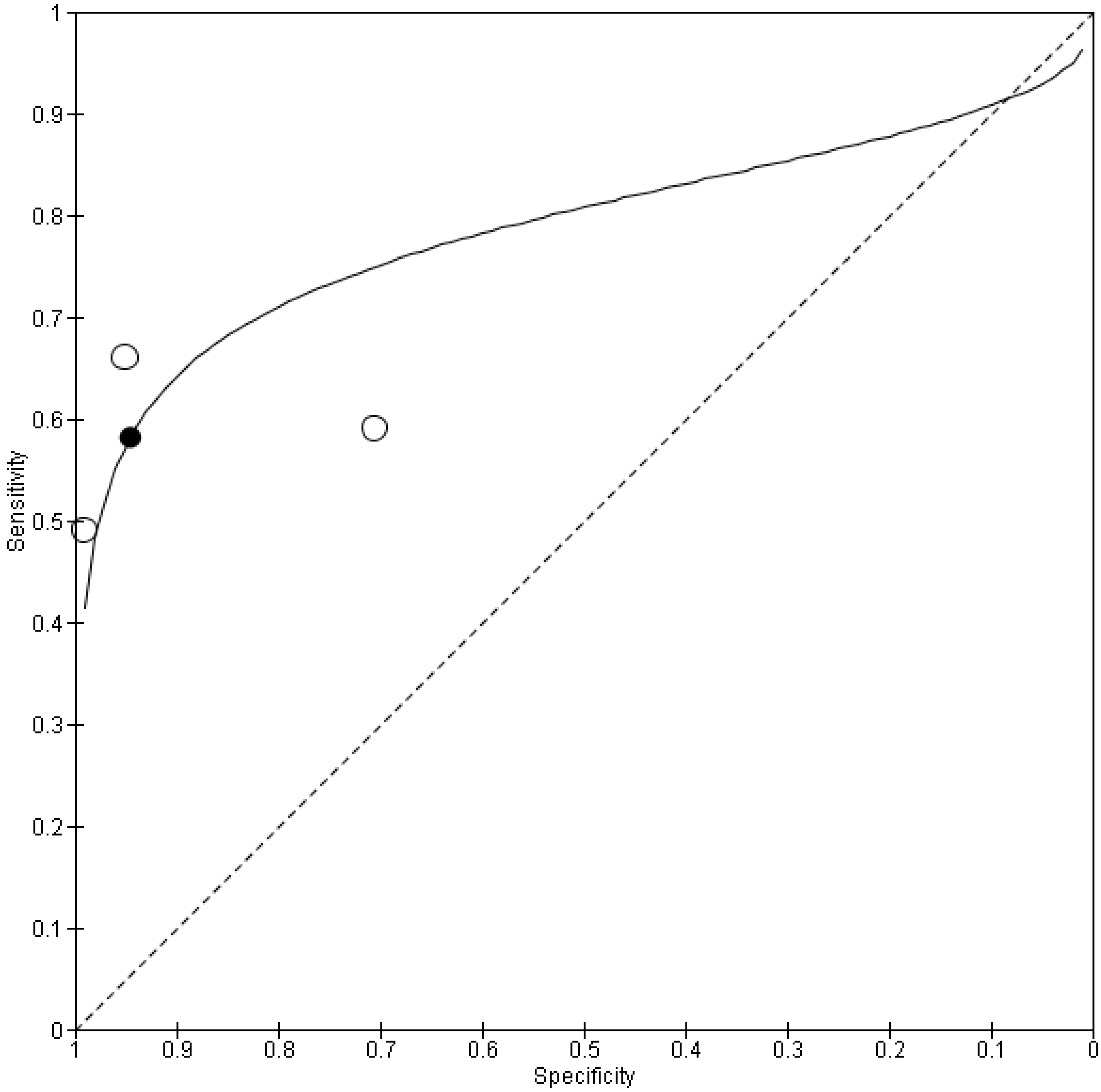
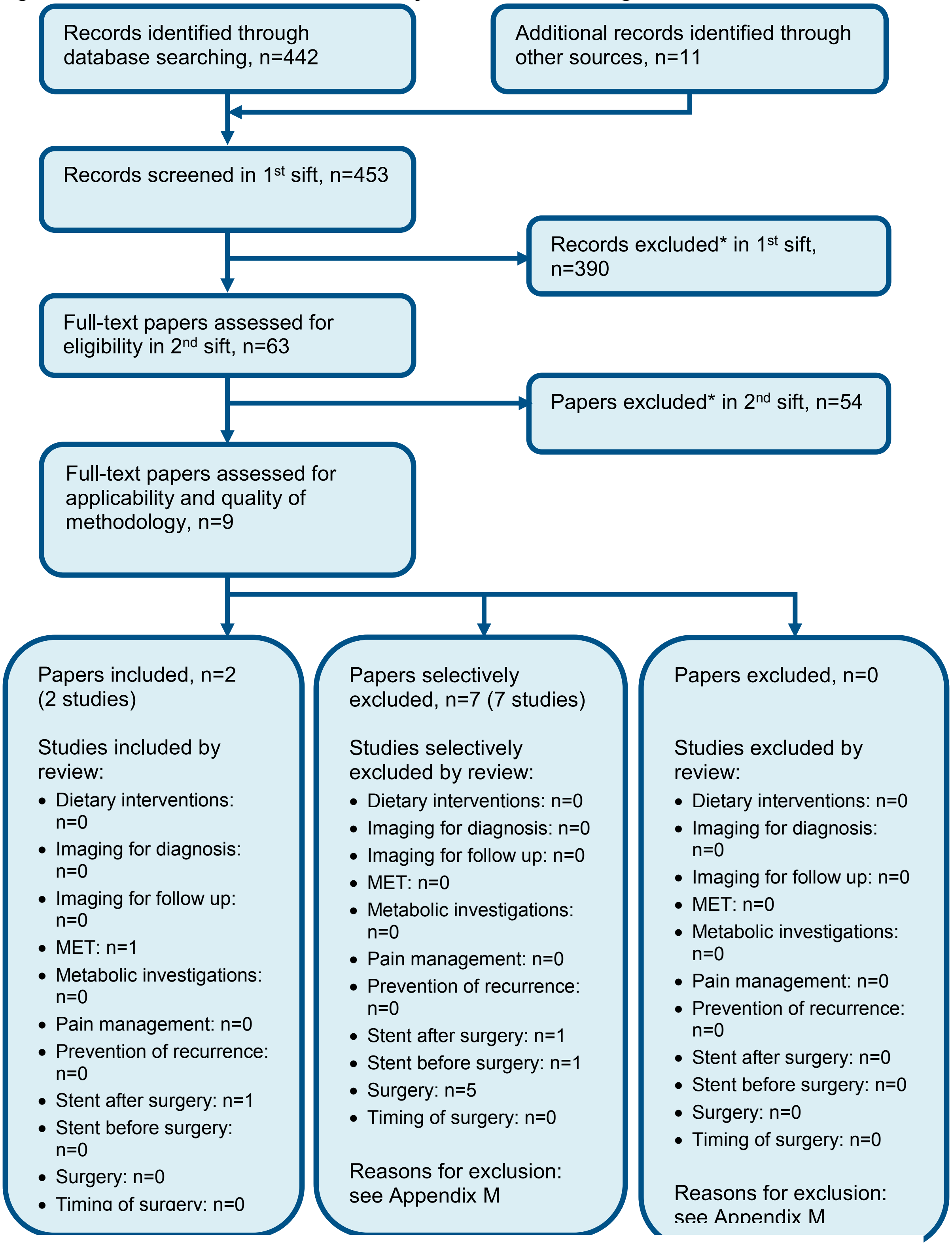


Figure 7: Pooled with prediction region: plain abdominal radiograph



Appendix F: Health economic evidence selection

Figure 8: Flow chart of economic study selection for the guideline



* Non-relevant population, intervention, comparison, design or setting; non-English language

Appendix G: Health economic evidence tables

None

Appendix H: Excluded studies

H.1 Excluded clinical studies

Table 10: Studies excluded from the clinical review

Reference	Reason for exclusion
Abdel-Gawad 2014 [1]	Incorrect reference standard
Abdel-Gawad 2016 [2]	Incorrect reference standard
Ahmad 2014 [3]	No outcomes
Ahn 2002[4]	Incorrect population
Albani 2007 [5]	Incorrect study design
Andresen 1997 [6]	Incorrect reference standard
Arif 2013 [7]	Incorrect reference standard
Assi 2000 [8]	Incorrect reference standard
Ather 2004 [9]	Incorrect reference standard
Ben Nakhi 2010 [10]	Incorrect index test
Blandino 2004 [11]	Incorrect reference standard
Bozdar 2016 [12]	No outcomes
Cabrera 2016 [13]	Incorrect index test
Catalano 2002 [14]	Incorrect reference standard
Cauberg 2011 [15]	Incorrect population
Chen 1999 [17]	Incorrect reference standard
Cochon 2017 [19]	Incorrect study design
Dillman 2011 [21]	Incorrect index test
Dorio 1999 [22]	Not available
Drake 2014 [23]	Systematic review checked for references
Dundee 2006 [24]	No outcomes
Edmonds 2010 [25]	Incorrect population
Ege 2004 [26]	Incorrect population
Eikefjord 2008 [27]	Incorrect reference standard
Ekici 2012 [28]	Incorrect reference standard
Eray 2003 [29]	Incorrect reference standard
Eshed 2002 [30]	Incorrect study design
Feroze 2007 [31]	Incorrect reference standard
Foell 2013 [32]	Incorrect population
Fowler 2011 [33]	Incorrect reference standard
Fowler 2002 [34]	Incorrect reference standard
Gaspari 2005 [35]	Incorrect target condition
German 2002 [36]	No outcomes
Gliga 2017 [37]	Incorrect reference standard
Graumann 2012 [38]	Incorrect reference standard
Gurel 2006 [39]	Incorrect index test
Hamm 2001 [40]	Incorrect reference standard
Hamm 2001 [41]	Not in English
Herbst 2014 [43]	Incorrect target condition

Reference	Reason for exclusion
Homer 2001 [44]	Incorrect index test
Hu 2010 [45]	Incorrect reference standard
Ibrahim 2016 [46]	Unclear population?
Jackman 2000 [47]	Incorrect population
Jeng 2001 [48]	Not in English
Johnston 2009 [49]	Incorrect index test
Joshi 2014 [50]	Incorrect reference standard
Jung 2010 [51]	Incorrect population
Kennish 2008 [55]	Incorrect study design
Khan 2012 [56]	No outcomes
Kluner 2006 [58]	Incorrect reference standard
Korkmaz 2014 [59]	Incorrect population
Kravchick 2006 [60]	Incorrect reference standard
Lee 2015 [61]	Incorrect study design
Leo 2017 [62]	Incorrect target condition
Lew 2017 [64]	Incorrect population
Lin 2016 [65]	Incorrect reference standard
Lisanti 2014 [66]	Incorrect target condition
Liu 2000 [67]	Not available
Longo 2001 [68]	Incorrect reference standard
Lorberboym 2000 [69]	No outcomes
MacEjko 2009 [70]	Incorrect study design
Malaki 2014 [71]	Incorrect population
Marumo 2002 [72]	Incorrect study design
Masch 2016 [73]	Incorrect index test
Matani 2007 [74]	Incorrect index test
May 2016 [75]	Incorrect population
Meagher 2001 [76]	Incorrect population
Melnikow 2016 [77]	Incorrect study design
Mendelson 2003 [78]	Incorrect index test
Mermuys 2010 [79]	No outcomes
Middleton 1988 [80]	Incorrect reference standard
Miller 1998 [81]	Incorrect reference standard
Mitterberger 2009 [82]	Incorrect population
Mitterberger 2007 [83]	Incorrect reference standard
Moak 2012	Incorrect index test
Moesbergen 2011 [84]	Incorrect population
Mos 2010 [85]	Incorrect population
Niall 1999 [88]	Incorrect reference standard
Nishiura 2009 [89]	Incorrect population
O'Kane 2016 [90]	Incorrect population
Olcott 1997 [91]	Incorrect population
Oner 2004 [92]	Incorrect reference standard
Palmer 2005 [93]	Incorrect population
Park 2008 [94]	Incorrect reference standard

Reference	Reason for exclusion
Patlas 2001 [96]	Incorrect reference standard
Pepe 2005 [97]	Incorrect reference standard
Pfister 2003 [98]	Incorrect reference standard
Pichler 2012 [99]	Incorrect study design
Poletti 2006 [100]	Incorrect study design
Quirke 2011 [101]	Incorrect study design
Rajaie 2006 [102]	Incorrect index test
Ray 2010 [103]	No outcomes
Rengifo 2010 [104]	Not in English
Richards 1999 [106]	Incorrect index test
Riddell 2014 [107]	Incorrect population
Ripolles 2004 [108]	Incorrect reference standard
Ripolles 2013 [109]	Incorrect reference standard
Rosen 1998 [110]	Incorrect reference standard
Rosser 2000 [111]	Incorrect reference standard
Rowland 2001 [112]	Incorrect target condition
Ryu 2001 [113]	Incorrect reference standard
Sade 2017 [114]	No outcomes
Sarofim 2016 [115]	Incorrect reference standard
Sattar 2011 [116]	Not available
Schwartz 1984 [117]	Incorrect index text and reference standard
Selberherr 2017 [118]	Incorrect population
Sen 2017 [120]	Incorrect population
Sheafor 2000 [121]	Incorrect reference standard
Shokeir 2001 [122]	Incorrect reference standard
Smith-Bindman 2014 [123]	Incorrect reference standard
Sudah 2002 [125]	Incorrect reference standard
Thomson 2001 [126]	Incorrect index test
Ulusan 2007[127]	No usable outcomes
Unal 2003 [128]	Incorrect reference standard
Uraiqat 2007 [129]	Incorrect reference standard
Valencia 2014 [130]	Incorrect study design
Vallone 2013 [131]	Incorrect population
Van Appledorn 2003 [132]	Incorrect population
Van Beers 2001 [133]	Incorrect reference standard
Vieweg 1998 [134]	Incorrect reference standard
Viprakasit 2012 [135]	Time between tests, no usable outcomes
Vrtiska 1992 [136]	Incorrect population
Wang 2003 [137]	Incorrect reference standard
Wang 2008 [138]	Incorrect reference standard
Wang 2004 [139]	Incorrect reference standard
Wang 2017 [140]	Incorrect study design
Watkins 2007 [141]	Incorrect target condition
Westergreen 2017 [142]	No outcomes
Winkel 2012 [143]	Incorrect index test

Reference	Reason for exclusion
Wong 2001 [144]	Incorrect index test
Yap 2012 [145]	Incorrect population
Yavuz 2015 [146]	Incorrect index test
Yilmaz 1998 [147]	Incorrect reference standard
Zilberman 2011 [148]	Incorrect study design

H.2 Excluded health economic studies

None