ACTA RADIOLÓGICA PORTUGUESA
May-August 2026 | Vol 38 nº 2 | e44184 https://doi.org/10.25748/arp.44184

Original Article / Artigo Original

Appropriateness and Diagnostic Yield of Abdominal and Kidney Ultrasound in the Emergency Department

Adequação e Rentabilidade Diagnóstica das Ecografias Abdominais e Renais no Serviço de Urgência

Nuno Lupi Manso[1], Rita Pina Prata[2], Rui Mateus Marques[2]

[1]Serviço de Radiologia, Hospital CUF Tejo, Lisboa, Portugal
[2]Serviço de Radiologia, ULS São José, Lisboa, Portugal

Address

Nuno Lupi Manso
Hospital CUF Tejo
Av. José Manuel de Mello 171
1350-352 Lisboa, Portugal
e-mail: nuno.manso@cuf.pt

Received: 18/11/2025
Accepted: 08/02/2026
Published: 05/03/2026

Abstract

Background: Diagnostic imaging use in emergency departments (ED) is increasing, particularly the use of abdominal and kidney ultrasound, highlighting the need for ongoing evaluation of its appropriateness and diagnostic yield.
Objective: To assess referral reasons, appropriateness, and diagnostic yield of abdominal and kidney US in a tertiary care ED.
Methods: We retrospectively reviewed adult ED patients who underwent urgent abdominal or kidney US between November 2024 and February 2025, excluding point-of-care studies. Appropriateness was determined using American College of Radiology (ACR) Appropriateness Criteria®. Diagnostic yield was defined as at least one positive finding potentially explaining the clinical presentation.
Results: Among 562 US exams, 295 (52.5%) were abdominal and 267 (47.5%) were kidney US. Most referrals were appropriate (abdominal: 88%; kidney: 43%). Abdominal US yielded positive findings in 29.5% of cases, highest for cholestasis (57%) and lowest for appendicitis (18%). Kidney US yielded positive findings in 13.1% of cases, highest for renal colic (37%).
Discussion and conclusion: While most US referrals met ACR criteria, diagnostic yield was modest, particularly for kidney US in acute kidney injury and pyelonephritis. Abdominal US presents a higher yield for biliary pathology but limited yield for appendicitis in adults. More selective use of US, coupled with improved clinical documentation and decision-support tools, may enhance efficiency and diagnostic impact in the ED.

Keywords

Diagnostic ultrasound; Acute disease; Emergency department.

Resumo

Introdução: O uso de exames de imagem no serviço de urgência tem aumentado, nomeadamente as ecografias abdominais e renais, instando à avaliação contínua da sua adequação e rentabilidade.
Objetivo: Avaliar os pedidos clínicos, adequação e rentabilidade diagnóstica da ecografia abdominal e renal num serviço de urgência terciário.
Métodos: Avaliámos retrospetivamente todos os doentes adultos admitidos num serviço de urgência e que realizaram ecografia abdominal ou renal urgente entre novembro de 2024 e fevereiro de 2025, excluindo-se estudos point-of-care. A adequação foi determinada usando os Appropriateness Criteria do American College of Radiology. A rentabilidade diagnóstica foi definida como pelo menos um achado positivo potencialmente explicativo do quadro clínico.
Resultados: Identificaram-se 562 ecografias, sendo 295 (52,5%) abdominais e 267 (47,5%) renais. A maioria dos pedidos foram apropriados (abdominal: 88%; renal: 43%). A ecografia abdominal revelou achados positivos em 29,5% dos casos, com maior rentabilidade na colestase (57%) e mais baixa para apendicite (18%). A ecografia renal revelou achados positivos em 13,1% dos casos, com rentabilidade mais elevada para cólica renal (37%).
Discussão e conclusão: Embora a maioria das ecografias tenham sido considerado apropriadas, a rentabilidade diagnóstica foi modesta, particularmente para a ecografia renal na investigação de lesão renal aguda e pielonefrite. A ecografia abdominal é mais rentável na patologia biliar mas com rentabilidade limitada na investigação de apendicite em adultos. O recurso mais seletivo aos exames ecográficos, juntamente com melhor fundamentação clínica e uso de ferramentas de apoio à decisão poderá aumentar a eficiência e impacto diagnóstico no serviço de urgência.

Palavras-chave

Ecografia; Doença aguda; Serviço hospitalar de emergência.

Introduction

The overuse of diagnostic imaging in the emergency department (ED) setting is a known issue, leading to higher healthcare costs, overdiagnosis and patient harm.[1] Data from the United States of America revealed a significant increase in the use of imaging in the ED over the last two decades even when accounting for increased ED visits[2,3] and there has been an increased interest in the evaluation of the appropriateness and diagnostic yield associated with ultrasound (US) use in the ED.[4,5,6,7]

This increasing interest in the literature may be attributed to ultrasound's greater accessibility, lesser cost and negligible risk of direct patient harm, as it avoids both ionizing radiation and iodinated contrast media. An increasing body of evidence has also suggested ultrasound to be cost-effective in specific

clinical scenarios such as pediatric appendicitis and venous thromboembolism.[7,8]

Despite these favorable attributes, ultrasound still occupies radiologists' time and may unnecessarily delay patient diagnosis.

Therefore, this cross-sectional study aims to:

1. Descriptively explore the reasons for referral.
2. Evaluate the appropriateness of the referrals.
3. Evaluate the diagnostic yield of the US exams.

Methods

This cross-sectional study with retrospective data collection reviewed the electronic health records of all adult patients who underwent urgent abdominal and/or kidney ultrasound – excluding point-of-care ultrasound – during emergency department visits to a tertiary care hospital between November 2024 and February 2025. All ultrasounds were performed by radiology specialists or radiology residents. Study data included reasons for referral and ultrasound imaging findings.

Referral appropriateness was evaluated according to the American College of Radiology (ACR) Appropriateness Criteria®. For all statistical analysis related to appropriateness, each referral was evaluated and considered "appropriate" if deemed either "usually appropriate" or "may be appropriate" and "inappropriate" if it was deemed "usually not appropriate" according to criteria.

Regarding diagnostic yield, an exam was considered to have a positive yield when there was at least one imaging finding that could plausibly explain the patient's presenting symptoms or clinical suspicion, such as detection of hydronephrosis in suspected renal colic.

Continuous variables are expressed as mean ± standard deviation for normally distributed data, or as median with interquartile range for non-normally distributed data. Categorical variables are presented as absolute values and frequencies. To compare diagnostic yield between appropriate and inappropriate exams, Chi-square and/or Fisher's exact test were used as appropriate. A two-tailed p-value < 0.05 was considered statistically significant. All analyses were conducted using IBM SPSS Statistics, version 26.0.

Results

A total of 562 patients were identified, with a mean age of 59.5 years (range 20–98 years), of whom 358 (64%) were female. Among the ultrasounds performed, 295 (52.5%) were abdominal and 267 (47.5%) were kidney ultrasounds.

Abdominal ultrasound

The most frequent reasons for referral were the presence of cholestasis (31%), right upper quadrant pain (18%), acute cholecystitis (17%), hepatic cytolysis (13%), jaundice (9%), right lower quadrant pain (9%) and acute appendicitis (6%). Other less frequent reasons for referral included suspected portal vein thrombosis, increased abdominal perimeter and suspected cirrhosis.

Overall, 53% (n = 156) of abdominal ultrasounds revealed at least one potentially relevant finding. The most frequent findings were gallstones (11%), bile duct dilation (11%), acute cholecystitis (9%), cirrhosis (5%), and acute appendicitis (1%). Rare findings (<1%) included portal vein thrombosis and colitis. These findings are further detailed in table 1(Fig. 1).

Table 1 – Abdominal ultrasound relevant findings and diagnostic yield by reason for referral. RUQ: Right upper quadrant. RLQ: Right lower quadrant.

Referral	N (%) out of all abdominal US	Relevant findings by referral (N,%)	Diagnostic yield (%)
Cholestasis	90 (31%)	Bile duct dilation (12, 13%) Cholecystitis (11, 12%) Gallstones (10, 11%) Pancreatic cancer (3, 3%) Liver nodules (3, 3%)	57%
RUQ pain	51 (18%)	Cholecystitis (10, 20%) Gallstones (10, 20%) Bile duct dilation (6, 12%)	49%
Cholecystitis	49 (17%)	Cholecystitis (17, 35%) Gallstones (6,12%) Bile duct dilation (3, 6%)	35%
Jaundice	26 (9%)	Bile duct dilation (7, 27%) Cholecystitis (2, 8%) Gallstones (2, 8%) Pancreatic cancer (1, 4%)	27%
Appendicitis	17 (6%)	Appendicitis (3, 18%)	18%
Hepatocytolysis	38 (13%)	Gallstones (6, 16%) Bile duct dilation (4, 11%) Cirrhosis (1, 3%) Liver nodules (1, 3%)	16%
RLQ pain	27 (9%)	Appendicitis (3, 11%) Colitis (1, 4%)	15%

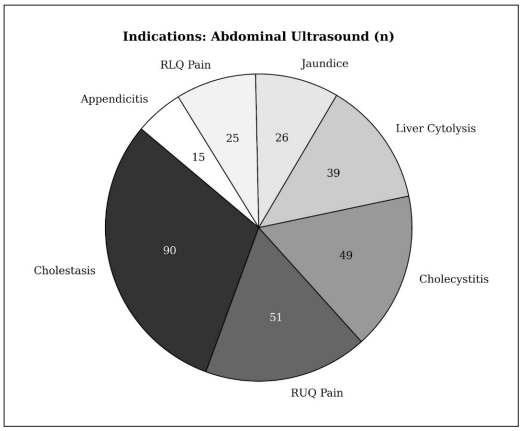


Figure 1 – Distribution of the most frequent clinical indications for abdominal ultrasound. Pie chart showing the absolute frequencies of the primary reasons for referral. RUQ: Right upper quadrant. RLQ: Right lower quadrant.

Referrals were classified as appropriate in 88% of cases, may be appropriate in 7%, and inappropriate in 5%, based on ACR criteria (Fig. 2). Inappropriate referrals were typically associated with non-specific symptoms such as nausea, vomiting, or diarrhea.

As for diagnostic yield, the global yield was estimated to be 29.5%, being significantly greater in appropriate exams (32.4% vs 9%, $\chi 2 = 8.83$, $p=0.003$). The diagnostic yield was greater in the investigation of cholestasis (57%), right upper quadrant pain (49%) and acute cholecystitis (35%). On the other hand, diagnostic yield was lowest for right lower quadrant pain and appendicitis, with a total of 3 positive cases being identified in 17 patients.

Kidney ultrasound

The most frequent reason for referral was the investigation of acute kidney injury (40%). The other reasons for referral

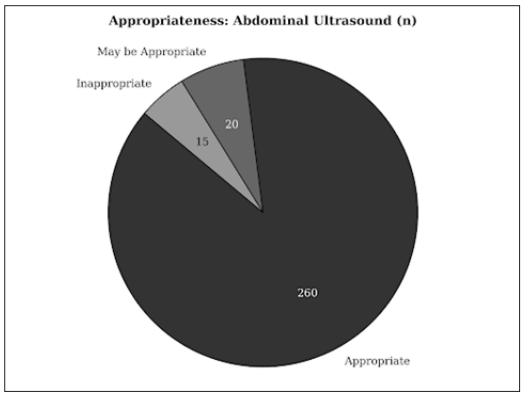


Figure 2 – Appropriateness of abdominal ultrasound referrals. Classification of abdominal ultrasound requests according to the ACR Appropriateness Criteria.

included specific clinical scenarios such as renal colic (19%), acute pyelonephritis (8%), lower back pain (4%), hematuria (4%) and anuria / oliguria (4%). Less common indications included urosepsis, erythrocyturia and leukocyturia.

Regarding the imaging findings, the majority (70%) of exams found no relevant acute findings. The most frequent findings were hydronephrosis (14%), medical nephropathy (8%) and kidney stones (4%). Other far rarer findings included suspicious focal kidney lesions, pyelitis and peri-renal hematoma. These findings are further detailed in table 2 (Fig. 3).

Table 2 – Kidney ultrasound relevant findings and diagnostic yield by reason for referral. AKI: Acute kidney disease.

Referral	N (%) out of all kidney US	Relevant findings by referral (N, %)	Diagnostic yield (%)
Renal colic	51 (19%)	Kidney stones (19, 37%) Hydronephrosis (11, 22%)	37%
AKI	108 (40%)	Hydronephrosis (9, 8%)	8%
Pyelonephritis	21 (8%)	Pyelitis (1, 5%)	3%

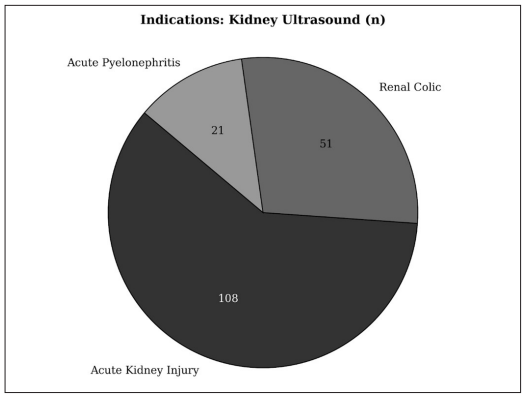


Figure 3 – Distribution of the most frequent clinical indications for kidney ultrasound. Pie chart showing the absolute frequencies of the primary reasons for referral.

Concerning its appropriateness, most referrals were considered either appropriate (43%) or may be appropriate (31%). Inappropriate referrals included clinical scenarios such as acute pyelonephritis and lower back pain with no additional information.

As for diagnostic yield, the global yield was estimated to be 13.1%, being significantly greater in inappropriate exams (19% vs 6%, $\chi 2 = 9.01$, $p=0.003$). The diagnostic yield was greater in the investigation of renal colic (37%) and acute kidney injury (8%). Among 21 suspected acute pyelonephritis cases, only one (5%) demonstrated a positive imaging finding – pyelitis. (Fig. 4)

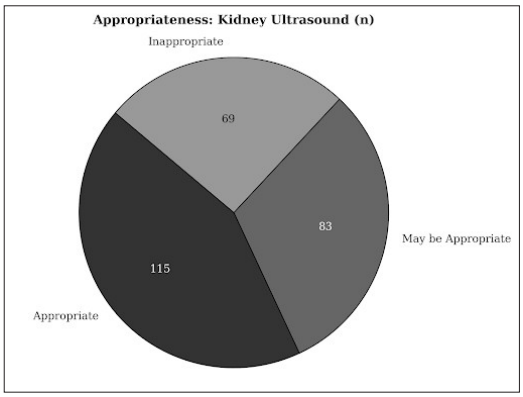


Figure 4 – Appropriateness of kidney ultrasound referrals. Classification of kidney ultrasound requests according to the ACR Appropriateness Criteria.

Discussion

Abdominal ultrasound

While the great majority of abdominal ultrasound referrals were deemed appropriate according to ACR criteria, the overall diagnostic yield was modest (29.5%) rising to 32.4% for appropriate exams. These findings mirror a national study[6] that also revealed increased odds of diagnostic yield for appropriate exams. This discrepancy may be explained by a low pre-test probability when prescribing physicians request the exam, hindering final diagnostic yield.

For instance, in the evaluation of suspected appendicitis, the Alvarado score[9] incorporates multiple clinical elements – such as right lower quadrant pain, fever, and leukocytosis – to estimate diagnostic likelihood. The presence of pain alone may not justify a high probability, and this is reflected in the respective ACR criteria[10], which distinguish between imaging pathways for appendicitis depending on additional clinical factors.

It is therefore essential to incorporate clinical scoring systems and Bayesian reasoning into decision-making for imaging. Prescribing physicians and radiologists alike should consider the cumulative weight of clinical signs and laboratory results when selecting the most appropriate first-line exam.

The leading motives for referral, like right upper and lower quadrant pain, are consistent with prior literature, with appendicitis and cholecystitis being major reasons for ED admission.[11] In biliary pathology, ultrasound is an ideal first-line study considering its high accuracy in detecting gallstones[12] and reasonable sensitivity and specificity in the diagnosis of acute cholecystitis, with a recent systematic review[13] estimating a global diagnostic accuracy of 83%. Our findings lend further support to this thesis, finding the highest diagnostic yields on the evaluation of cholestasis and right upper quadrant pain.

In contrast, ultrasound had a low diagnostic yield in the diagnosis of appendicitis, with only 3 positive cases among 17 patients. Several factors possibly contributed: our study population consisted exclusively of adults, who generally present more technical limitations for ultrasound than children; pre-test probability may have been low in some cases and the absolute number of referrals for appendicitis

was also limited. Additionally, the variability in operator experience – ranging from residents with different levels of training to specialists from various subfields – was not specifically tracked. Despite these findings, previous studies[14] have consistently found reasonable sensitivity and specificity for ultrasound (82% and 86%, respectively), although abdominal CT with contrast had the highest sensitivity and specificity (97% and 96%, respectively).

Kidney ultrasound

Although most kidney ultrasound referrals were considered appropriate, their diagnostic yield was low in several common clinical scenarios. The most frequent indication, acute kidney injury (AKI), resulted in a positive finding – hydronephrosis – in only 8% of the cases.

This is consistent with a large retrospective study[15] of 1471 renal ultrasounds, in which the authors reported urinary tract obstruction in fewer than 1% of patients without a suggestive clinical history. These results support the need for a more selective approach to ultrasound use in AKI, guided by the pre-test probability of an obstructive etiology. When hydronephrosis is present, however, ultrasound has a very high sensitivity in its detection (over 90%).[16]

Notably, the diagnostic yield was paradoxically higher in "inappropriate" referrals. This likely reflects a methodological artifact: in appropriate scenarios such as AKI, the absence of findings is often the expected and desired outcome. Because our study defined diagnostic yield as the presence of a positive finding, appropriately negative ultrasounds in AKI may have contributed to an underestimation of yield in that group.

For suspected renal colic, our study found a low to modest diagnostic yield of 37%. The use of kidney ultrasound in this context is acknowledged as controversial within the respective appropriateness criteria,[17] due to several diagnostic limitations. The ultrasound detection of calculi is limited,[18,19] especially when these are small (less than 5 mm) or impacted in the ureter. Additionally, while being a very sensitive exam for hydronephrosis, hydronephrosis may not have had sufficient time to develop.[20] Therefore, one should consider abdominal CT without intravenous contrast as a competing first-line exam, especially in patients with high BMI or equivocal clinical findings, considering its far higher sensitivity, even when performed with a low-dose protocol.[21]

Suspected acute pyelonephritis was another common but low-yield indication, with only 1 of 21 patients (5%) demonstrating a relevant imaging finding – pyelitis. In uncomplicated cases, imaging is not recommended,[22] suggesting a need for improved guideline adherence. These results point to an opportunity for better alignment between clinical practice and evidence-based recommendations, supported through targeted education for referring clinicians and enhanced radiology oversight.

Limitations

This study has several limitations that should be considered when interpreting our results. Firstly, its retrospective design limited data completeness. Ultrasound referrals that were declined or not recorded by the attending radiologist may have been missed, potentially inducing selection bias. Secondly, while the overall sample size was reasonable, it was insufficient to support more detailed subgroup analysis, such as diagnostic yield by specific diagnoses. Thirdly, as previously discussed, our definition of diagnostic yield was restricted to positive findings. This approach does not capture the clinical utility of negative ultrasound results, which may assist in ruling out disease or reduce diagnostic uncertainty. Additionally, this approach fails to account for false positives, which we were unable to control for due to the retrospective nature of the study. Consequently, categories with a high reported yield may include findings that did not translate into a confirmed diagnosis. This highlights the risk of overestimating a technique's effectiveness when negative utility and diagnostic accuracy (including false-positive rates) are not fully considered. Fourthly, the quality of the referrals was variable. Many referrals lacked sufficient clinical detail, such as accompanying symptoms or previous history, which are crucial for accurate assessment of pre-test probability and exam appropriateness. Fifthly, an exam with a low diagnostic yield is not necessarily inappropriate, as some examinations may have a low yield simply because they are used to investigate rare but clinically critical conditions. Sixthly, there was a lack of control over the experience level of the ultrasound operators. The exams were performed by radiology specialists or residents, but the specific seniority of the residents or the subspecialty of the specialists was not recorded. Moreover, the study did not document whether resident-performed exams were systematically reviewed by a specialist. Lastly, this study was circumscribed to only one emergency department, limiting the generalization of the obtained data.

Further studies / Future work

As different national and international institutions and scientific societies build their own guidelines (such as the ESR's iGuide), prospective studies are warranted to assess their impact on diagnostic accuracy, clinical outcomes, protocol adherence and cost-effectiveness.

Integrating large language model technology in electronic health records could potentially assist physicians by suggesting relevant information and by issuing real-time alerts for the most appropriate imaging and to avoid redundant imaging.

Institutions should also ensure that comprehensive and up-to-date imaging referral guidelines are readily accessible to physicians, both through clinical training and internal documentation.

Conclusion

This study found that while most abdominal and kidney ultrasound referrals in the emergency department were appropriate by ACR criteria, their diagnostic yield was modest, particularly for kidney ultrasound in cases of acute kidney injury and pyelonephritis. Abdominal ultrasound was most useful in evaluating biliary conditions, whereas its utility for appendicitis was limited in the adult population.

Overall, these results advocate more nuanced imaging decisions in the emergency department, balancing clinical judgment, pre-test probability, and guideline-based appropriateness. Improved documentation of clinical reasoning, greater use of diagnostic scoring systems, and closer collaboration between referring physicians and radiologists may enhance both the appropriateness and effectiveness of imaging utilization.

Ethical Disclosure / Divulgações Éticas

Conflicts of interest: The authors have no conflicts of interest to declare.

Conflitos de interesse: Os autores declaram não possuir conflitos de interesse.

Financing Support: This work has not received any contribution, grant or scholarship.

Suporte financeiro: O presente trabalho não foi suportado por nenhum subsídio ou bolsa.

Confidentiality of data: The authors declare that they have followed the protocols of their work center on the publication of data from patients.

Confidencialidade dos dados: Os autores declaram ter seguido os protocolos do seu centro de trabalho acerca da publicação dos dados de doentes.

Protection of human and animal subjects: The authors declare that the procedures followed were in accordance with the regulations of the relevant clinical research ethics committee and with those of the Code of Ethics of the World Medical Association (Declaration of Helsinki).

Protecção de pessoas e animais: Os autores declaram que os procedimentos seguidos estavam de acordo com os regulamentos estabelecidos pelos responsáveis da Comissão de Investigação Clínica e Ética e de acordo com a Declaração de Helsínquia da Associação Médica Mundial.

References

1. Selvarajan SK, Levin DC, Parker L. The increasing use of emergency department imaging in the United States: is it appropriate? American Journal of Roentgenology. 2019;213:W180-W184.

2. Larson DB, Johnson LW, Schnell BM, et al. National trends in CT use in the emergency department: 1995-2007 1. Radiology. 258. Epub ahead of print 2011. Doi: 10.1148/radiol.10100640/-/DC1.

3. Tung M, Sharma R, Hinson JS, et al. Factors associated with imaging overuse in the emergency department: a systematic review. American Journal of Emergency Medicine. 2018;36:301-9.

4. Francisco MZ, Altmayer S, Carlesso L, et al. Appropriateness and imaging outcomes of ultrasound, CT, and MR in the emergency department: a retrospective analysis from an urban academic center. Emerg Radiol. 2024;31:367-72.

5. Francisco MZ, Altmayer S, Verma N, et al. Appropriateness of computed tomography and ultrasound for abdominal complaints in the emergency department. Curr Probl Diagn Radiol. 2021;50:799-802.

6. Martins R, Raimundo P, Alves P, et al. Appropriateness of radiology test requests by an emergency department: a retrospective study. Acta Med Port. 2020;33:7-14.

7. Nicola R, Dogra V. Ultrasound: the triage tool in the emergency department: using ultrasound first. British Journal of Radiology. 89. Epub ahead of print 2016. Doi: 10.1259/bjr.20150790.

8. Lentz B, Fong T, Rhyne R, et al. A systematic review of the cost-effectiveness of ultrasound in emergency care settings. Ultrasound Journal. 13. Epub ahead of print 1 December 2021. Doi: 10.1186/s13089-021-00216-8.

9. Coleman JJ, Carr BW, Rogers T, et al. The Alvarado score should be used to reduce emergency department length of stay and radiation exposure in select patients with abdominal pain. In: Journal of Trauma and Acute Care Surgery. Lippincott Williams and Wilkins, 2018, pp.946-50.

10. Garcia EM, Camacho MA, Karolyi DR, et al. ACR appropriateness criteria® right lower quadrant pain-suspected appendicitis. Journal of the American College of Radiology. 2018;15:S373-S387.

11. Weiss AJ, Jiang HJ. Most frequent reasons for emergency department visits, 2018. https://hcup-us.ahrq.gov/reports/statbriefs/sb286-ED-Frequent-Conditions-2018.pdf

12. Bennett GL. Evaluating patients with right upper quadrant pain. Radiologic Clinics of North America. 2015;53:1093-130.

13. Huang SS, Lin KW, Liu KL, et al. Diagnostic performance of ultrasound in acute cholecystitis: a systematic review and meta-analysis. World Journal of Emergency Surgery. 18. Epub ahead of print 1 December 2023. Doi: 10.1186/s13017-023-00524-5.

14. Arruzza E, Milanese S, Li LSK, et al. Diagnostic accuracy of computed tomography and ultrasound for the diagnosis of acute appendicitis: a systematic review and meta-analysis. Radiography. 2022;28:1127-41.

15. Podoll A, Walther C, Finkel K. Clinical utility of gray scale renal ultrasound in acute kidney injury. BMC Nephrol. 2013;14:188.

16. Ellenbogen PH, Scheible FW, Talner LB, et al. Sensitivity of gray scale ultrasound in detecting urinary tract obstruction. Am J Roentgenol. 1978;130:731-3.

17. Gupta RT, Kalisz K, Khatri G, et al. ACR Appropriateness criteria® acute onset flank pain-suspicion of stone disease (urolithiasis). Journal of the American College of Radiology. 2023;20:S315-S328.

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

19. Fowler KAB, Locken JA, Duchesne JH, et al. US for detecting renal calculi with nonenhanced CT as a reference standard. Radiology. 2002;222:109-13.

20. Varanelli MJ, Coll DM, Levine JA, et al. Relationship between duration of pain and secondary signs of obstruction of the urinary tract on unenhanced helical CT. AJR. 2001;177.

21. Bhojani N, Paonessa JE, El Tayeb MM, et al. Sensitivity of noncontrast computed tomography for small renal calculi with endoscopy as the gold standard. Urology. 2018;117:36-40.

22. Smith AD, Nikolaidis P, Khatri G, et al. ACR Appropriateness criteria® acute pyelonephritis: 2022 update. Journal of the American College of Radiology. 2022;19:S224-S239.