BRAZILIAN JOURNAL OF EMERGENCY MEDICINE *VOLUME 05 | NÚMERO 1 / 29-39* **DOI:** 10.5935/2764-1449.20250005

AI diagnostic accuracy for abdominal free fluid in emergency: a meta-analysis

Acurácia diagnóstica da IA para líquido abdominal livre em emergência: uma meta-análise

João Martins da Fonseca[1], Sarah Verdan Moreira[1], Sanda Kolenda Zloic[3], Karabo Kago Marole[4], Gianluca Capello Ingold[5], Emma Finnegan[6], Marcia Harumy Yoshikawa[7], Júlia Azevedo Miranda[8], Silvija Daugėlaitė[9], Prathyusha Songa[10], Marco Aurélio Soato Ratti[11]

[1]Hospital Geral de Salvador, Salvador (BA), Brazil.
[2]Hospital Universitário da Universidade Federal de Juiz de Fora, Minas Gerais (MG), Brazil.
[3]Special Hospital Agram in Zagreb, Zagreb, Croatia.
[4]St George's University School of Medicine, True Blue, Grenada.
[5]Hospital Universitario Austral, Pilar, Argentina.
[6]School of Medicine, Trinity College Dublin, Ireland.
[7]Brigham and Women's Hospital, Harvard Medical School, USA.
[8]Institute of Radiology, Hospital das Clínicas, Faculty of Medicine, University of São Paulo.
[9]Radvili kis Hospital, Lithuania.
[10]Dr Nimmy's Skin Clinic, Hyderabad, India.
[11]Fleury Group, Musculoskeletal Radiology, São Paulo (SP), Brazil.

To cite this article: J.M. Fonseca; Moreira S.V.; Zloic S.K.; Marole K.K.; Ingold G.C.; Finnegan E.; Yoshikawa M.H.; Miranda J.A.; Daugėlaitė S.; Songa P.; Ratti M.A.S. AI Diagnostic Accuracy for Abdominal Free Fluid in Emergency: A Meta-Analysis. Brazilian Journal of Emergency Medicine 2025; 5: 29-39.

ABSTRACT

Objective: This systematic review and meta-analysis aimed to assess the diagnostic accuracy of artificial intelligence (AI) models in detecting abdominal free fluid via ultrasonography in emergency and critical care settings. Given the increasing demand for rapid and accurate assessment in trauma care, AI-based models could improve diagnostic efficiency, particularly in point-of-care settings. **Methods:** A comprehensive literature search was conducted across PubMed, Cochrane Library, and Embase up until July 2024, adhering to PRISMA-DTA guidelines. Observational and randomized studies reporting diagnostic accuracy outcomes (sensitivity, specificity) of AI models for abdominal free fluid detection using ultrasonography in emergency cases were included. Key data extracted included study characteristics, patient demographics, AI model details, and diagnostic outcomes. Study quality was assessed using the QUADAS-2 tool. Meta-analyses using random-effects models calculated pooled sensitivity, specificity, and the summary receiver operating characteristic (SROC) curve. Heterogeneity was evaluated with the I statistic, and a leave-one-out sensitivity analysis assessed result robustness. **Results:** Six studies involving over 2,000 participants met inclusion criteria. Pooled sensitivity was 0.916 (95% CI: 0.784-0.970), specificity was 0.941 (95% CI: 0.878-0.972), and

RESUMO

Objetivo: Esta revisão sistemática e meta-análise teve como objetivo avaliar a precisão diagnóstica de modelos de inteligência artificial (IA) na detecção de fluido livre abdominal por ultrassonografia em ambientes de emergência e cuidados críticos. Dada a crescente demanda por avaliações rápidas e precisas em atendimentos de trauma, modelos baseados em IA podem melhorar a eficiência diagnóstica, especialmente em configurações de ponto de cuidado. **Mйtodos:** Foi realizada uma busca abrangente nas bases de dados PubMed, Cochrane Library e Embase até julho de 2024, seguindo as diretrizes PRISMA-DTA. Foram incluídos estudos observacionais e randomizados que relataram resultados de precisão diagnóstica (sensibilidade, especificidade) de modelos de IA para detecção de fluido livre abdominal usando ultrassonografia em casos de emergência. Dados chave extraídos incluíram características dos estudos, demografia dos pacientes, detalhes dos modelos de IA e resultados diagnósticos. A qualidade dos estudos foi avaliada com a ferramenta QUADAS-2. Meta-análises com modelos de efeitos aleatórios calcularam sensibilidade, especificidade e a curva SROC. **Resultados:** Seis estudos com mais de 2.000 participantes foram incluídos. Sensibilidade agrupada

Correspondence to: João Fonseca
E-mail: fonseca.joaomd@gmail.com

the SROC curve indicated an area under the curve (AUC) of 0.965 (95% CI: 0.906-0.979). The leave-one-out analysis confirmed the stability of these results, with no single study disproportionately affecting the estimates. **Conclusion:** AI models demonstrate high diagnostic accuracy in detecting abdominal free fluid via ultrasonography in emergency settings. Despite some variability and heterogeneity, AI has the potential to significantly enhance diagnostic accuracy in trauma and non-trauma care.

Key-words Artificial Intelligence; Focused Assessment with Sonography for Trauma; Abdomen; Emergency Medicine.

foi 0,916 (IC 95%: 0,784-0,970), especificidade foi 0,941 (IC 95%: 0,878-0,972), e a curva SROC teve AUC de 0,965 (IC 95%: 0,906-0,979). A análise leave-one-out confirmou a robustez dos resultados. **Conclusro:** Modelos de IA demonstram alta precisão diagnóstica na detecção de fluido livre abdominal por ultrassonografia em ambientes de emergência, com potencial para aprimorar a precisão diagnóstica no atendimento de trauma e não-trauma.

Palavras-chave: Inteligência Artificial; Abdome; Avaliação Sonográfica Focada no Trauma; Emergências.

INTRODUCTION

Abdominal trauma is a frequent injury[1,2] that can result in active hemorrhage, often due to liver damage or hemodynamic instability, requiring prompt intervention and, in many cases, an emergency laparotomy[3,4]. Such injuries, including liver or spleen ruptures and gastrointestinal perforations, are often challenging to diagnose through physical examination alone[5], as clinical signs typically offer insufficient information to determine the need for surgical intervention[3,6]. Consequently, the evaluation of abdominal trauma, especially blunt abdominal trauma, continues to pose a significant challenge[7].

Prompt imaging assessments are critical for trauma diagnosis, as delays in treatment can significantly increase the risk of mortality[8–10]. For instance, in patients requiring laparotomy, mortality increases by approximately 1% for every 3-minute delay in intervention[3]. Ultrasonography has become a widely utilized imaging modality in trauma care due to its accessibility and its ability to provide rapid point-of-care (POC) assessments at the bedside[11,12]. However, despite considerable technological advancements and expanded use over the past 25 years, ultrasonography still faces several limitations that affect its effectiveness in trauma settings[13].

POC ultrasound in the emergency department is largely based on the focused assessment with sonography in trauma (FAST) examination, which is a non-invasive diagnostic technique widely employed in the evaluation of acute abdominal cases[9,14]. Key regions of the abdomen are systematically examined for the presence of free fluid, which serves as a critical indicator of serious intra-abdominal injuries that may necessitate urgent surgical intervention, such as an emergency laparotomy[3]. Numerous studies have highlighted the importance of FAST in guiding clinical decisions, particularly pediatric cases, pregnancy, and hemodynamically unstable patients[15]. Early detection of abdominal free fluid through this technique is crucial for timely treatment and intervention in a wide range of emergency scenarios, often determining the need for angiography or surgery[3,15].

While experienced physicians can readily identify peritoneal free fluid in ultrasound images, the process can be more time-consuming for novice clinicians, individuals lacking expertise in ultrasound imaging, or non-medical training[9]. To enhance the standardization of care among trauma providers with varying levels of proficiency, artificial intelligence (AI) technologies have been developed to improve the quality of bedside ultrasound image acquisition and interpretation[11,16].

The implementation of AI-based techniques offers the potential to quickly detect and localize abdominal free fluid, significantly reducing examination times and allowing optimized clinical interventions[9]. AI and deep learning (DL) applications have been shown to significantly increase diagnostic accuracy in point-of-care (POC) imaging techniques[3,17]. In fact, deep learning algorithms for POC ultrasound have achieved accuracy rates exceeding 98% in some datasets, surpassing the performance of experienced clinicians[11,18].

Thus, the primary objective of this systematic review and meta-analysis is to evaluate the feasibility and diagnostic accuracy of AI algorithms for the timely detection of abdominal free fluid in ultrasound (USG) images obtained in emergency cases

MATERIAL AND METHOD

Literature Search and Study Selection

This retrospective systematic review and meta-analysis were performed following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic Test

Accuracy (PRISMA-DTA) guidelines[19]. A comprehensive search across the electronic databases PubMed/MEDLINE, Cochrane Library and Embase was performed from inception to July 11th 2024. Studies included for meta-analysis were [1] observational or randomized and [2] reported one of the diagnostic accuracy outcomes for AI models for abdominal free fluid detection in [3] emergency adult cases [4] via ultrasonography. The search strategy incorporated terms related to "Artificial Intelligence" (e.g., "AI," "Deep Learning," "Machine Learning," "CNN"), "Ultrasonography" (e.g., "FAST," "point-of-care ultrasound," "POCUS"), and conditions affecting the abdomen (e.g., "Ascites," "Hemoperitoneum," "Abdominal free fluid," "Abdominal Injuries"). Additionally, we registered the study with PROSPERO prior to the initial literature search (CRD42024568898)

After initial literature search, two authors independently performed the removal of duplicate studies and screening for study inclusion. Discrepancies were resolved by third-party adjudication.

Data Extraction and Quality Assessment

Data from the included studies was extracted. Baseline characteristics such as geographic location, mean age, diagnosis, study design and AI model details were extracted and summarized. Risk of bias assessment was performed independently by two authors. The preferred tool was the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2). Each of the included studies was analyzed in the seven proposed domains by the authors. Discrepancies were resolved by consensus or third-party adjudication. Given the small number of included studies, publication bias analysis was not performed outside of sensitivity analysis.

Endpoints and Statistical Analyses

Meta-analyses were performed for the outcomes of sensitivity and specificity and their corresponding 95% confidence intervals (CI), using random-effects models. The summary receiver operating characteristic (SROC) curve was also drawn with the calculation of the area under the curves (AUC) to reflect the overall diagnostic performance. Heterogeneity was assessed using the I statistic, with Cochran's Q test used to determine significance. A p-value < 0.10 and I > 25% were considered indicative of significant heterogeneity. Leave-one-out sensitivity analyses were conducted by excluding each study one by one to evaluate the robustness of the pooled estimates for sensitivity, specificity, and accuracy. The effect of each exclusion was analyzed and plotted to assess whether any study had a disproportionate influence on the overall results. All analyses were performed using R software (version 4.2.1; The R Foundation), employing the 'meta' and 'metafor' packages.

RESULTS

Study Selection and Characteristics

The initial search yielded 571 records. After duplicate removal, 25 records were excluded. After screening the titles and abstracts of the remaining 546 studies, 13 full-text articles were assessed for eligibility. Ultimately, 6 studies met the inclusion criteria for the meta-analysis [3,7,9–11,14] and were included in the final analysis (Fig. 1). The pooled analysis included total of over 2,000 participants and used AI models. The mean age of participants was approximately 53 years and all the included studies were retrospective. The most commonly used structure were convolutional neural networks (CNNs) and most of the studies were a mix of emergency trauma and non-trauma cases. The main characteristics of the studies are summarized in Table 1.

Quality Assessment and Risk of Bias

Overall risk of bias of the included studies was low. The first domain and its applicability, regarding patient selection, was significant for multiple studies with potential risk of bias due to case-control design adoption and potentially inappropriate exclusions based on image quality. The remaining domains had the majority or totality of studies categorized as low bias risk (Fig. 2).

Diagnostic Accuracy and Heterogeneity

The pooled sensitivity under random-effects model was 0.916 (CI 95% 0.784-0.970, I = 99%) (Fig. 3). The pooled specificity under random-effects model was 0.941 (CI 95% 0.878-0.972, I = 97%) (Fig. 3). The summary ROC curve under random-effects model was 0.965 (CI 95% 0.906 - 0.979) (Fig. 4).

Leave-One-Out-Analysis

The leave-one-out sensitivity analysis demonstrates that the overall meta-analytic results are robust, with no individual study exerting an excessive influence on the summary estimates of sensitivity, specificity, or AUC. The small range of fluctuation in these values indicates that the pooled diagnostic accuracy measures are reliable and not dependent on any single study. The model remains consistent even when studies with large sample sizes or extreme values are excluded (Fig. 5).

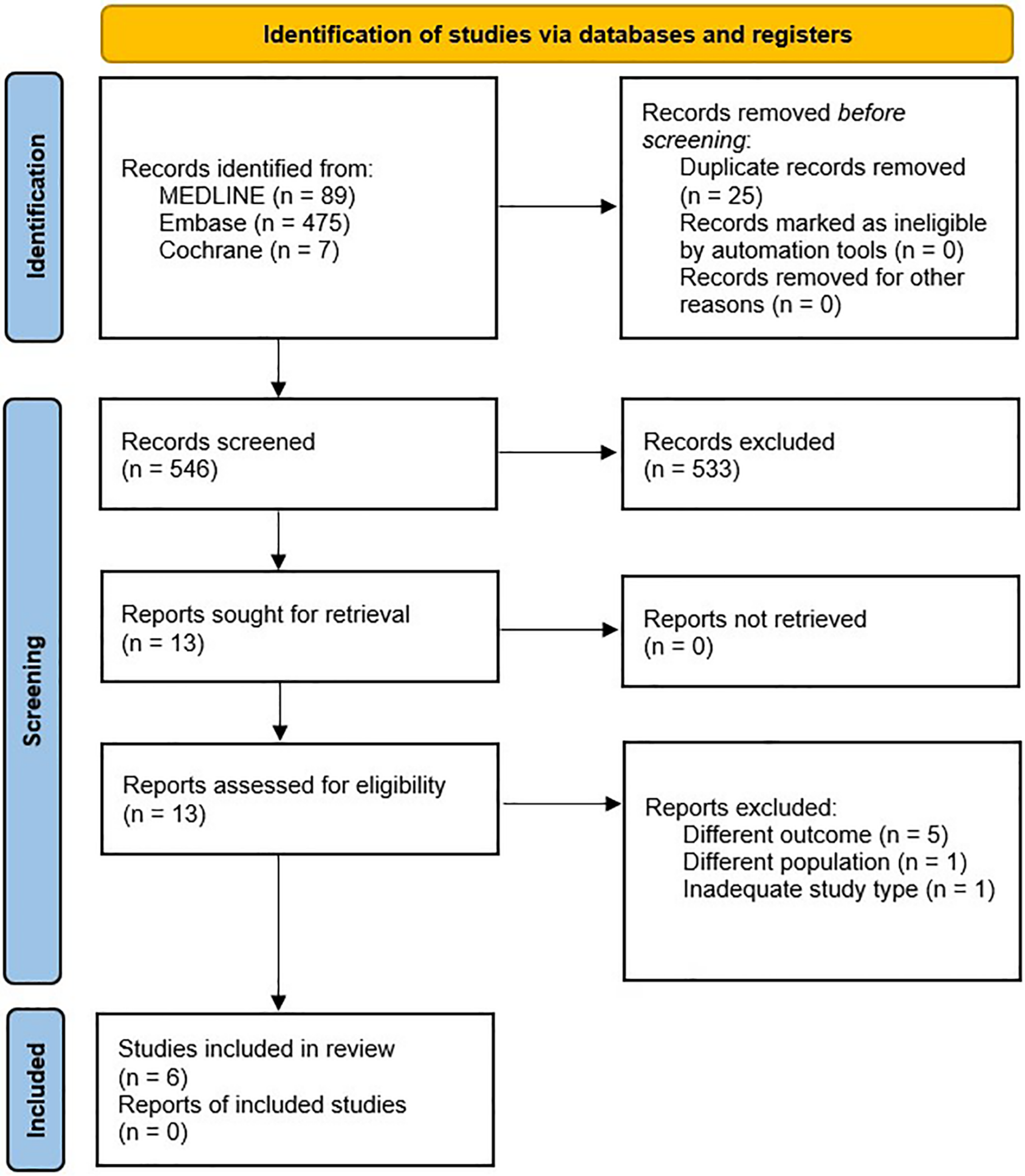

Figura 1. PRISMA 2020 flow diagram of the study.

DISCUSSION

In this systematic review and meta-analysis of six studies and over 2,000 patients, we evaluated the diagnostic accuracy of AI models for detection of abdominal free fluid with ultrasonography in emergency cases. We found a pooled sensitivity of 0.916 (CI 95% 0.784-0.970), pooled specificity of 0.941 (CI 95% 0.878-0.972) and summary ROC curve of 0.965 (CI 95% 0.906 - 0.979) across the included studies.

Overall heterogeneity was substantial for included studies. The leave-one-out analysis revealed minimal changes in diagnostic measures when individual studies were excluded and the overall quality assessment of the included studies showed a low risk of bias outside the patient selection domains.

The increasing demands of healthcare are prominently featured in the emergency department context, where fast and accurate interpretation and decision-making is of great

Table 1. Baseline characteristics of included studies.

Author and year	Total Sample	Age† (years)	Location	Scenario	Study Design	Structure	System	Reference Standard
Levy 2023	109 patients / 6,608 images	N/A	United States	Emergency Trauma cases	Retrospective Cohort	A-CNN B-CNN C-CNN D-CNN	A-DenseNet121 B-Inception v3 C-ResNet50 D-vgg11_bn	Experts
Lin 2022	A-845 patients / 3,192 images B-845 patients / 2,778 images	45.4	China	Emergency and Teaching Ascites cases	Retrospective Cohort	A (Ascites 1)- CNN B (Ascites 2)- CNN	A-U-net B-U-net	Experts
Cheng 2021	A-396 patients / 11,574 images B-396 patients/ 809 images	59.9	Taiwan	Emergency cases	Retrospective Cohort	A (By frame)- CNN B (By 1s majority voting)- CNN	A-ResNet50-V2 B-ResNet50-V2	Experts
Leo 2023	A-94 patients / 94 videos B-94 patients / 94 videos C-94 patients / 94 videos D-94 patients / 94 videos	38.7	United States	Emergency cases	Retrospective Cohort	A-CNN B-CNN C-CNN D-CNN	A-Yolo V3 B- U-net C-MaskRCNN D-ResNet	Experts
Sjogren 2016	20 patients/ 1264 frames	47,4	United States	Emergency Cases	Retrospective Pilot Study	-SVM	MATLAB	Experts
Jeong 2023	864 patients / 2200 images	58	Republic of Korea	Emergency Trauma Cases	Retrospective Cohort	-DL	AutoML	Experts

†mean or median; NA = Not Available; CNN = Convolutional Neural Network; SVM = Support Vector Machine; ML = Machine Learning; DL = Deep Learning.

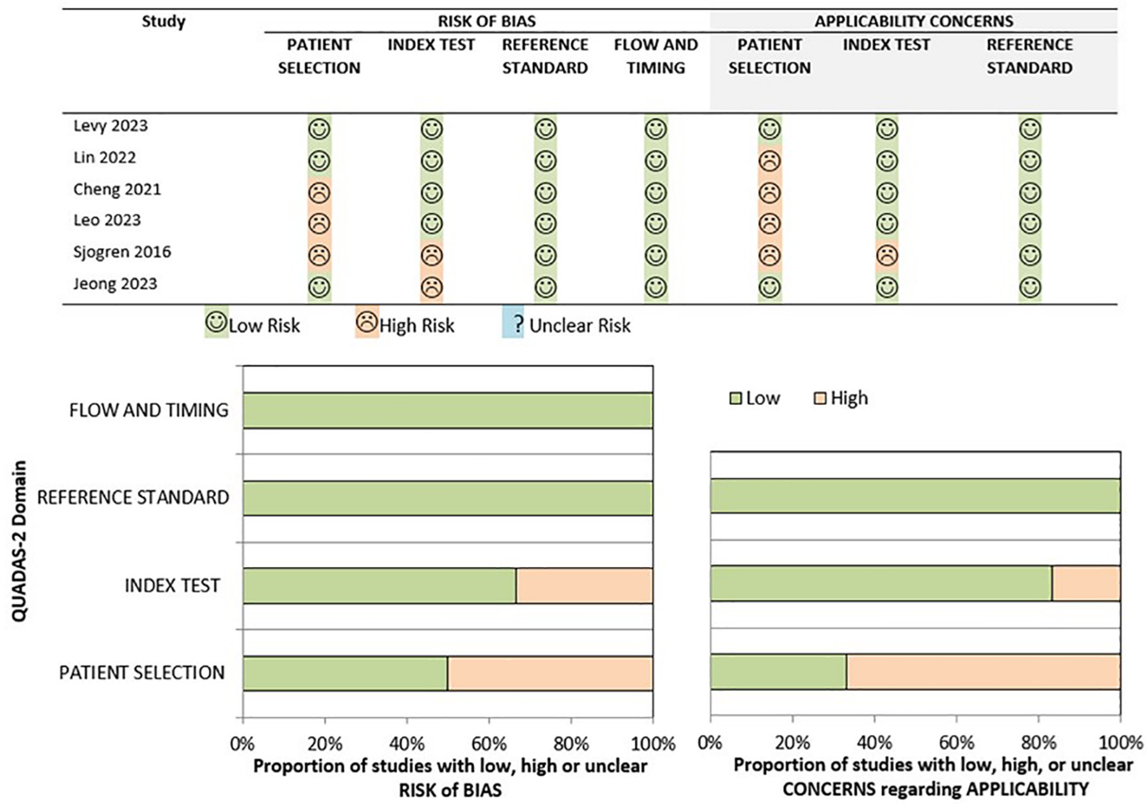

Figure 2. QUADAS-2 assessment of included studies. Risk of bias analysis using the QUADAS-2 tool. Most studies showed a low risk of bias in all domains, with the exception of the patient selection domain, which had a higher risk due to case-control design and potentially inappropriate exclusions based on image quality.

importance[20]. A systematic review by Boonstra and Laven highlights the usefulness of AI-based tools in the ED as a way to cope with an overcrowded emergency case load and mitigate human error[21]. In emergency radiology, AI can provide support to radiologists with patient positioning, imaging acquisition, reconstruction, interpretation and timely structuring of reports[22].

The results of this meta-analysis are comparable to those evidenced by other AI-based tools in emergency neurological and orthopedic cases in terms of accuracy[23–25]. Prior meta-analyses have also been performed for AI tools in emergency cases, but mostly for orthopedic trauma, with similar endpoint results[26,27].

However, abdominal pathologies pose a significant difficulty for validation of AI algorithms due to the complexity of cases and imaging features[22]. Nevertheless, in the emergency scenario, a CNN-based study with conventional radiography achieved sensitivity and specificity > 0.90 to diagnose small-bowel obstruction[28] and Park et al. reported a trained model with similar accuracy to our pooled summary when evaluating acute appendicitis diagnosis via CT scans[29]. Specifically related to abdominal ultrasonography AI models, prior studies have also reported satisfactory accuracy results, but mostly in non-emergency liver pathologies[30–33]. Therefore, to the best of our knowledge, this is the first systematic quantitative synthesis of evidence for detection of abdominal free fluid with sonography in the emergency scenario.

Most of the included studies in this review used CNNs as the blueprint for the models as they are considered today to be the state-of-the-art imaging analysis structure considering they do not necessarily require hand-crafted feature extraction nor structure segmentation by experts[34]. CNNs are a subclass of artificial intelligence and deep learning that consist of layered-shaped networks of assimilation and processing that transform imaging volume into output class scores[35]. The usual pattern of a CNN-based study design includes a computer vision task, data acquisition, data processing, structure selection and validation[35]. However,

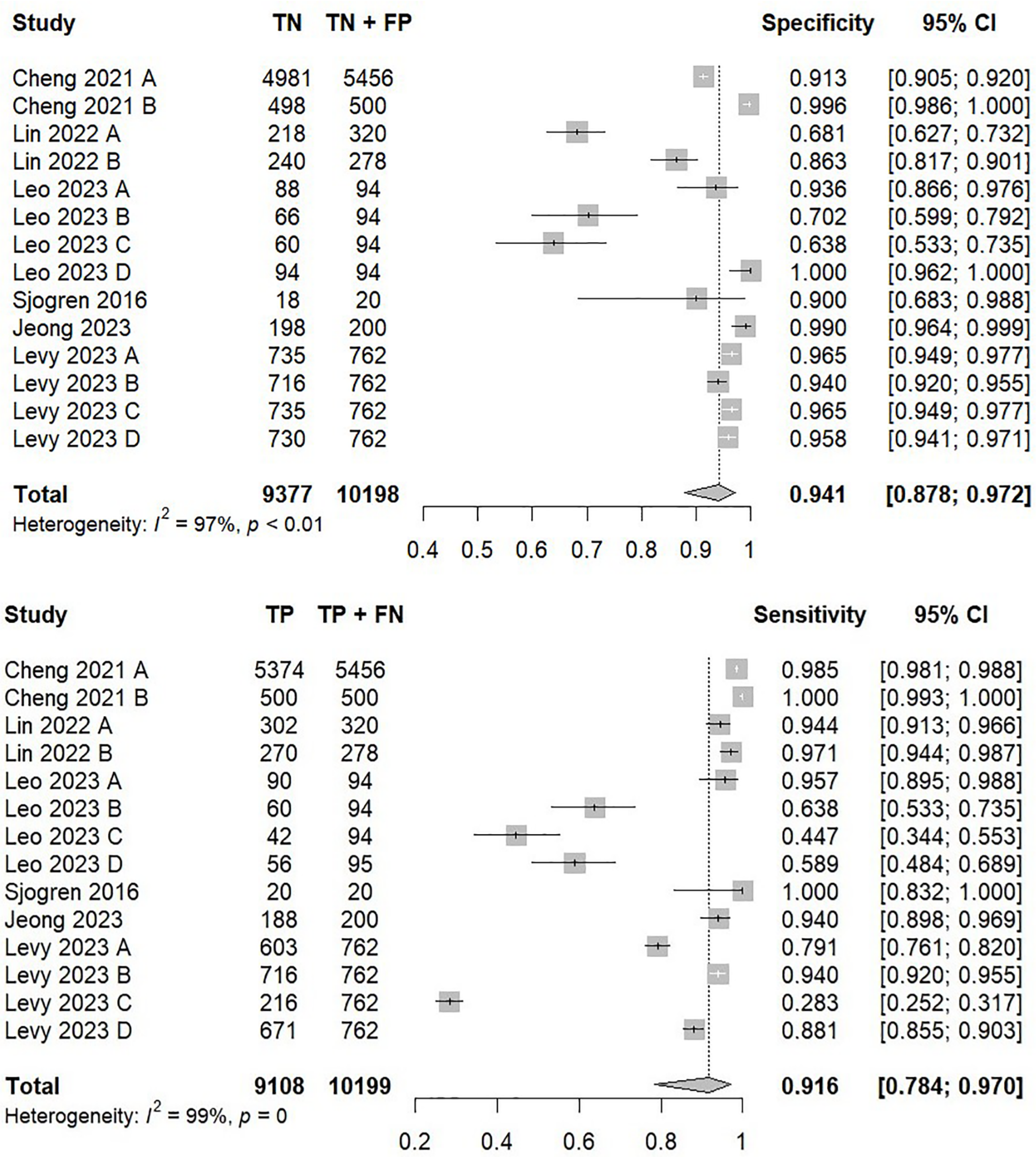

Figure 3. Forest plots for pooled diagnostic sensitivity and specificity of included studies. Forest plot showing the pooled sensitivity and specificity of AI models for detecting abdominal free fluid in emergency cases. Pooled sensitivity was 0.916 (95% CI: 0.784–0.970), and pooled specificity was 0.941 (95% CI: 0.878–0.972). Both plots demonstrate significant diagnostic accuracy across studies. CI = confidence interval; TN = True Negative; FP = False Positive; TP = True Positive; FN = False Negative.

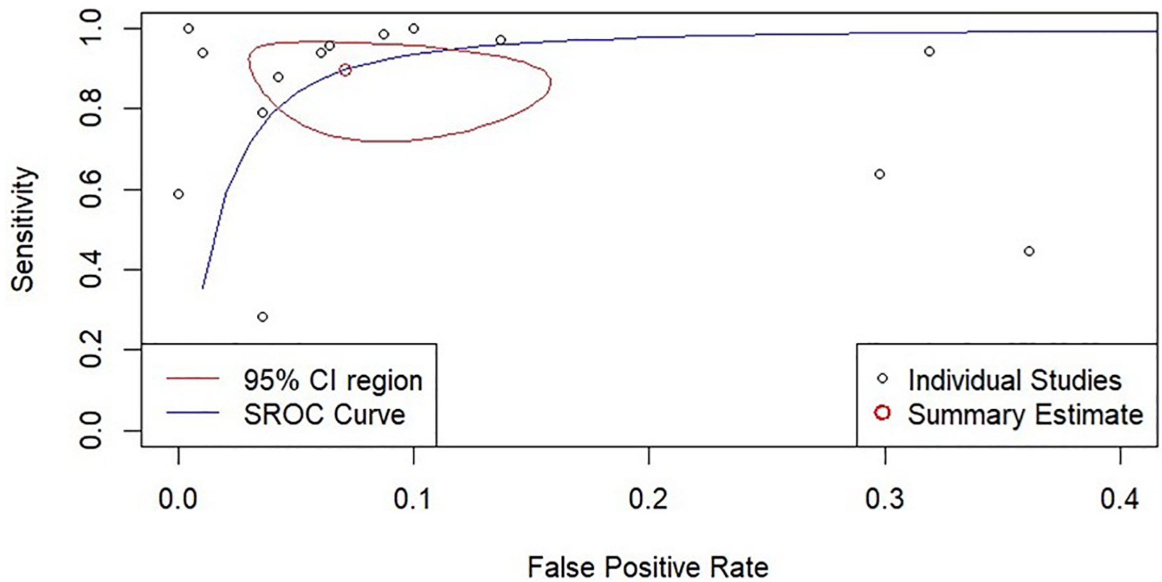

Figure 4. Forest plot for bivariate diagnostic accuracy showing summary receiver operating characteristic (SROC) curve. SROC curve illustrating the overall diagnostic performance of AI models in detecting abdominal free fluid. The area under the curve (AUC) was 0.965 (95% CI: 0.906–0.979), indicating high diagnostic accuracy in emergency settings. CI = confidence interval.

the model has its limitations, as the reasoning behind the algorithm's decision-making remains largely unclear. Additionally, the absence of large datasets and the need for data augmentation observed in the included studies to prevent overfitting have not yet adequately addressed this concern.[34].

Additionally, despite the substantial accuracy observed in the pooled analysis and potential use of the applied systems for real time aide to health care providers, individual concerns on the model's applicability have also been raised by the individual studies. Leo et al., highlights the amount of free fluid and poor imaging quality to be particularly troublesome for the model reported[14]. Lin emphasized that the used model has substantially more prone to error when analyzing small ascites areas[3]. Variability of sonography machines and geographic restrictions to generalizability of results were also reported by most of the included studies.

Our study is not without limitations. The small number of included studies limited a more nuanced evaluation of publication bias, such as through funnel plots and meta-regression analysis. Additionally, many studies did not specify the nature of the emergency cases evaluated, precluding subgroup analyses of trauma versus non-trauma cases. Furthermore, the considerable variation in sample sizes across the included studies increased the risk of one study to disproportionately skew the results. To mitigate this potential bias, a leave-one-out sensitivity analysis was conducted to assess the robustness of our findings. Another source of potential bias derived from the significant heterogeneity observed in most plots, which may be attributed to the wide range of pathologies, differing diagnostic criteria and sonography sites, and model structures employed in each study. Given these complexities, we adopted a random-effects approach to provide a more accurate estimation of the pooled data.

CONCLUSION

In conclusion, despite its limitations, our study suggests that AI models provide reliable diagnostic accuracy for abdominal free fluid via ultrasonography in emergency cases, though further studies are needed to address specific subgroups.

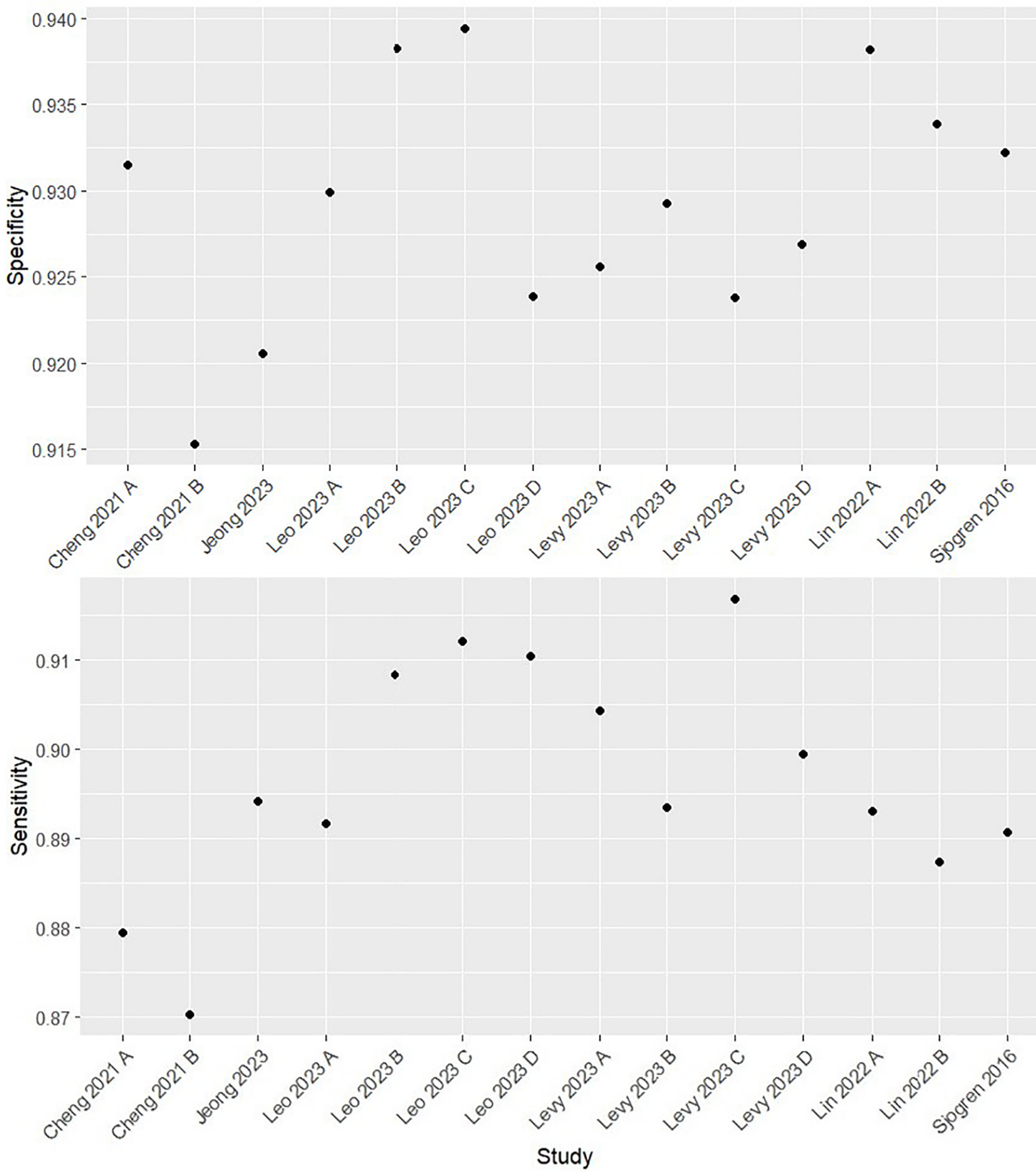

Figure 5. Forest plot of leave one out sensitivity analysis for sensitivity and specificity of included studies. Sensitivity analysis evaluating the stability of the meta-analytic results. The study name in the x-axis represents the excluded study and the diagnostic measure (black dots) represents the summary result of the remaining studies. CI = confidence interval.

REFERENCES

1. Armstrong LB, Mooney DP, Paltiel H, et al. Contrast enhanced ultrasound for the evaluation of blunt pediatric abdominal trauma. *J Pediatr Surg*. 2018;53(3):548-552. doi:10.1016/j.jpedsurg.2017.03.042
2. Simel DL. Does This Adult Patient Have a Blunt Intra-abdominal Injury? *JAMA*. 2012;307(14):1517. doi:10.1001/jama.2012.422
3. Lin Z, Li Z, Cao P, et al. Deep learning for emergency ascites diagnosis using ultrasonography images. *J Appl Clin Med Phys*. 2022;23(7). doi:10.1002/acm2.13695

4. Lv F, Tang J, Luo Y, et al. Contrast-enhanced ultrasound imaging of active bleeding associated with hepatic and splenic trauma. *Radiol Med*. 2011;116(7):1076-1082. doi:10.1007/s11547-011-0680-y
5. Brown MA, Casola G, Sirlin CB, Patel NY, Hoyt DB. Blunt Abdominal Trauma: Screening US in 2,693 Patients. *Radiology*. 2001;218(2):352-358. doi:10.1148/radiology.218.2.r01fe42352
6. Fang JF, Wong YC, Lin BC, Hsu YP, Chen MF. The CT Risk Factors for the Need of Operative Treatment in Initially Hemodynamically Stable Patients After Blunt Hepatic Trauma. *The Journal of Trauma: Injury, Infection, and Critical Care*. 2006;61(3):547-554. doi:10.1097/01.ta.0000196571.12389.ee
7. Cheng CY, Chiu IM, Hsu MY, Pan HY, Tsai CM, Lin CHR. Deep Learning Assisted Detection of Abdominal Free Fluid in Morison's Pouch During Focused Assessment With Sonography in Trauma. *Front Med (Lausanne)*. 2021;8. doi:10.3389/fmed.2021.707437
8. McCarter FD, Luchette FA, Molloy M, et al. Institutional and Individual Learning Curves for Focused Abdominal Ultrasound for Trauma. *Ann Surg*. 2000;231(5):689-700. doi:10.1097/00000658-200005000-00009
9. Jeong D, Jeong W, Lee JH, Park SY. Use of Automated Machine Learning for Classifying Hemoperitoneum on Ultrasonographic Images of Morrison's Pouch: A Multicenter Retrospective Study. *J Clin Med*. 2023;12(12):4043. doi:10.3390/jcm12124043
10. Sjogren AR, Leo MM, Feldman J, Gwin JT. Image Segmentation and Machine Learning for Detection of Abdominal Free Fluid in Focused Assessment With Sonography for Trauma Examinations. *Journal of Ultrasound in Medicine*. 2016;35(11):2501-2509. doi:10.7863/ultra.15.11017
11. Levy BE, Castle JT, Virodov A, et al. Artificial intelligence evaluation of focused assessment with sonography in trauma. *Journal of Trauma and Acute Care Surgery*. 2023;95(5):706-712. doi:10.1097/TA.0000000000004021
12. Morrow D, Cupp J, Schrift D, Nathanson R, Soni NJ. Point-of-Care Ultrasound in Established Settings. *South Med J*. 2018;111(7):373-381. doi:10.14423/SMJ.0000000000000838
13. Lee L, DeCara JM. Point-of-Care Ultrasound. *Curr Cardiol Rep*. 2020;22(11):149. doi:10.1007/s11886-020-01394-y
14. Leo MM, Potter IY, Zahiri M, Vaziri A, Jung CF, Feldman JA. Using Deep Learning to Detect the Presence and Location of Hemoperitoneum on the Focused Assessment with Sonography in Trauma (FAST) Examination in Adults. *J Digit Imaging*. 2023;36(5):2035-2050. doi:10.1007/s10278-023-00845-6
15. Savoia P, Jayanthi SK, Chammas MC. Focused Assessment with Sonography for Trauma (FAST). *J Med Ultrasound*. 2023;31(2):101-106. doi:10.4103/jmu.jmu_12_23
16. Akkus Z, Cai J, Boonrod A, et al. A Survey of Deep-Learning Applications in Ultrasound: Artificial Intelligence–Powered Ultrasound for Improving Clinical Workflow. *Journal of the American College of Radiology*. 2019;16(9):1318-1328. doi:10.1016/j.jacr.2019.06.004
17. Shokoohi H, LeSaux MA, Roohani YH, Liteplo A, Huang C, Blaivas M. Enhanced Point-of-Care Ultrasound Applications by Integrating Automated Feature-Learning Systems Using Deep Learning. *Journal of Ultrasound in Medicine*. 2019;38(7):1887-1897. doi:10.1002/jum.14860
18. Blaivas M, Arntfield R, White M. DIY AI, deep learning network development for automated image classification in a point-of-care ultrasound quality assurance program. *J Am Coll Emerg Physicians Open*. 2020;1(2):124-131. doi:10.1002/emp2.12018
19. Frank RA, Bossuyt PM, McInnes MDF. Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy: The PRISMA-DTA Statement. *Radiology*. 2018;289(2):313-314. doi:10.1148/radiol.2018180850
20. Berlyand Y, Raja AS, Dorner SC, et al. How artificial intelligence could transform emergency department operations. *Am J Emerg Med*. 2018;36(8):1515-1517. doi:10.1016/j.ajem.2018.01.017
21. Boonstra A, Laven M. Influence of artificial intelligence on the work design of emergency department clinicians a systematic literature review. *BMC Health Serv Res*. 2022;22(1):669. doi:10.1186/s12913-022-08070-7
22. Cellina M, Cè M, Irmici G, et al. Artificial Intelligence in Emergency Radiology: Where Are We Going? *Diagnostics*. 2022;12(12):3223. doi:10.3390/diagnostics12123223

23. Jones RM, Sharma A, Hotchkiss R, et al. Assessment of a deep-learning system for fracture detection in musculoskeletal radiographs. *NPJ Digit Med*. 2020;3(1):144. doi:10.1038/s41746-020-00352-w

24. McLouth J, Elstrott S, Chaibi Y, et al. Validation of a Deep Learning Tool in the Detection of Intracranial Hemorrhage and Large Vessel Occlusion. *Front Neurol*. 2021;12. doi:10.3389/fneur.2021.656112

25. Rava RA, Seymour SE, LaQue ME, et al. Assessment of an Artificial Intelligence Algorithm for Detection of Intracranial Hemorrhage. *World Neurosurg*. 2021;150:e209-e217. doi:10.1016/j.wneu.2021.02.134

26. Be uli H, Begagi E, D idi -Krivi A, et al. Sensitivity and specificity of machine learning and deep learning algorithms in the diagnosis of thoracolumbar injuries resulting in vertebral fractures: A systematic review and meta-analysis. *Brain and Spine*. 2024;4:102809. doi:10.1016/j.bas.2024.102809

27. van den Broek MCL, Buijs JH, Schmitz LFM, Wijffels MME. Diagnostic Performance of Artificial Intelligence in Rib Fracture Detection: Systematic Review and Meta-Analysis. *Surgeries*. 2024;5(1):24-36. doi:10.3390/surgeries5010005

28. Cheng PM, Tran KN, Whang G, Tejura TK. Refining Convolutional Neural Network Detection of Small-Bowel Obstruction in Conventional Radiography. *American Journal of Roentgenology*. 2019;212(2):342-350. doi:10.2214/AJR.18.20362

29. Park JJ, Kim KA, Nam Y, Choi MH, Choi SY, Rhie J. Convolutional-neural-network-based diagnosis of appendicitis via CT scans in patients with acute abdominal pain presenting in the emergency department. *Sci Rep*. 2020;10(1):9556. doi:10.1038/s41598-020-66674-7

30. Biswas M, Kuppili V, Edla DR, et al. Symtosis: A liver ultrasound tissue characterization and risk stratification in optimized deep learning paradigm. *Comput Methods Programs Biomed*. 2018;155:165-177. doi:10.1016/j.cmpb.2017.12.016

31. Hassan TM, Elmogy M, Sallam ES. Diagnosis of Focal Liver Diseases Based on Deep Learning Technique for Ultrasound Images. *Arab J Sci Eng*. 2017;42(8):3127-3140. doi:10.1007/s13369-016-2387-9

32. Byra M, Styczynski G, Szmigielski C, et al. Transfer learning with deep convolutional neural network for liver steatosis assessment in ultrasound images. *Int J Comput Assist Radiol Surg*. 2018;13(12):1895-1903. doi:10.1007/s11548-018-1843-2

33. Guo LH, Wang D, Qian YY, et al. A two-stage multi-view learning framework based computer-aided diagnosis of liver tumors with contrast enhanced ultrasound images. *Clin Hemorheol Microcirc*. 2018;69(3):343-354. doi:10.3233/CH-170275

34. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights Imaging*. 2018;9(4):611-629. doi:10.1007/s13244-018-0639-9

35. Soffer S, Ben-Cohen A, Shimon O, Amitai MM, Greenspan H, Klang E. Convolutional Neural Networks for Radiologic Images: A Radiologist's Guide. *Radiology*. 2019;290(3):590-606. doi:10.1148/radiol.2018180547