IBN SINA

Ibn Sina Journal of Medical Science, Health & Pharmacy
Vol. 3, No. 11, 2025, pp. 1-8
ISSN 3104-6428
https://ibn-sina-journal-of-medical-sciences.jo

Interobserver And Intraobserver Variability in CT Scan Reporting Of Liver Hydatid Cyst Staging And Location.

FIRAS ABDULLAH NOORI AL-BAGHDADI [a,1]

[1] Assistant Professor, Department of Radiology, FICMS-Radiology, College of Medicine, University of Thi-Qar, Iraq.
Email: dr.firas@utq.edu.iq
* Corresponding Author : Firas Abdullah Noori Al-Baghdadi, dr.firas@utq.edu.iq

DOI: https://doi.org/10.64440/IBNSINA/SINA007

ARTICLE INFO

Article history
Received Oct 01, 2025
Revised Oct 03, 2025
Accepted Oct 29, 2025

Keywords
Hydatid Cyst;
Ct Imaging;
Interobserver Variability;
Intraobserver Variability

ABSTRACT

A serious health concern in endemic regions, hydatid disease of the liver is a parasitic infection brought on by Echinococcus species. The diagnosis, staging, and location of the disease are all made possible by computed tomography (CT), which also helps with treatment planning and better results. Nevertheless, the dependence of trustworthy imaging results is hampered by the lack of sufficient evidence about the reliability of CT evaluations across various observers. The purpose of this study was to assess the level of intraobserver and interobserver agreement in the identification and classification of liver hydatid cysts using CT scans. Additionally, it aimed to ascertain how observer experience affected the assessment's degree of agreement. 150 patients with proven hepatic hydatid disease were included in a retrospective analysis. In order to categorize cysts using the WHO Informal Working Group on Echinococcus (WHO-IWGE) classification method and record their precise locations in the liver, three radiologists with different levels of experience independently examined CT scans. The same radiologists reassessed the pictures four weeks later to assess intraobserver concordance. Cohen's coefficient (κ) and the internal correlation coefficient (ICC) were used to examine concordance. For cyst staging, interobserver concordance was typically good (Fleiss' $\kappa = 0.72$), and for placement, it was excellent (Fleiss' $\kappa = 0.88$). With average κ values of 0.85 for staging and 0.92 for location, concordance within the same observer was considerably greater. The two-stage cysts CE5 (inactive calcification) and CE2 (multi-vesicular cysts) showed the best concordance, although the transitional stage CE3 had only moderate concordance ($\kappa = 0.52$). Specialists with more expertise showed higher concordance ($\kappa = 0.81$) than those with less experience ($\kappa = 0.63$). The study's findings show that CT offers good concordance for locating hepatic cysts. However, depending on the stage and the observer's experience, the accuracy of cyst stage classification varies. In order to increase the reliability of imaging results in the treatment of hepatic hydatid disease, the results highlight the significance of developing uniform interpretation criteria and offering specialized training to enhance assessment accuracy and decrease variability between readings.

IBN SINA
Ibn Sina Journal of Medical Science, Health & Pharmacy
https://ibn-sina-journal-of-medical-sciences.jo
ISSN 3104-6428
Vol. 3, No. 11, 2025, pp. 1-8

1. Introduction

In regions where the parasite Echinococcus spp. is prevalent, hydatid illness poses a serious threat to public health. The development of hydatid cysts in the liver, which vary greatly in size, shape, and stage, is the hallmark of this infection [1]. To ascertain the patient's status and direct the best course of treatment, a precise and trustworthy assessment is required. In order to diagnose, categorize, and locate the cysts, computed tomography (CT) scans are essential. They help determine the stage of cyst growth, which is essential for making precise treatment decisions, and they offer a comprehensive picture of the liver's anatomy [2]. Despite this, there hasn't been enough research done on the reliability of reports from specialists during image evaluation using various classification techniques, especially the classification system created by the WHO Informal Working Group on Radiation Infection (WHO-IWGE), in terms of inter-observer consistency, whether between different radiologists (inter-observer reliability) or with the same physician upon subsequent re-evaluation (intra-observer reliability) [3].

The need for standardization and uniformity in assessment and classification criteria is highlighted by an increasing amount of research showing that variations in interpretation and classification can have a substantial impact on clinical outcomes, particularly given the disparities in physician experience [4]. Consistency and low dependability problems can result in incorrect assessments and treatment choices, which can degrade the standard of care and raise the risk of consequences. While earlier research has assessed the accuracy of ultrasound imaging, there is a glaring lack of information about the degree of agreement in computed tomography (CT) evaluations, especially when it comes to the precise categorization of various stages of liver cysts and their location within the liver, as well as the influence of the physician's level of experience on assessment outcomes [5, 6]. Additionally, research on intraobserver reliability—the frequency of evaluations by the same clinician over time—is lacking. In order to close these gaps, this study uses CT and the WHO-IWGE classification system to examine and evaluate the interobserver and intraobserver agreement in identifying the position and stage of liver cysts over a given time period [7, 8]. By doing this, we hope to pinpoint the causes of assessment variability and ambiguity as well as provide useful suggestions for boosting report reliability and upgrading professional guidance and training, which will help precise diagnosis and treatment planning. This offers scientific proof that improves the efficacy of standards and evaluation techniques, helps to improve the quality of diagnosis for infectious parasitic diseases of the liver, improves patient outcomes and safety, and supports professional practices to attain a higher degree of accuracy and consistency in computed tomography reports.

2. Method

2.1. Study Design

Using computed tomography (CT) imaging, a retrospective observational study was carried out to analyze the inter-observer and intra-observer agreement in determining the stage and appearance of hepatic cysts caused by radiation hepatitis. From January 2020 to December 2024, the study was conducted at the Medical Education Center, College of Medicine, University of Basrah. The study was authorized by the center's ethical committee, and because the data gathering was retrospective, signed patient agreement was not required. All photos and clinical data were anonymized and encrypted to ensure patient privacy and data confidentiality.

2.2. Selection of Cases

243 patients with a clinical diagnosis and a trustworthy laboratory and statistical evaluation were found after the medical records of patients with radiation hepatitis were examined. The final

study sample consisted of 150 patients with a total of 217 radiation hepatitis cysts after inclusion and exclusion criteria were applied. One of the requirements for entrance was being at least eighteen years old. High-quality CT scans, either with or without contrast injection, are readily available. The diagnosis is confirmed with laboratory testing, such as histology or serological analyses. Complete clinical data is available.

Cases that had already received radiation therapy or surgery were among the exclusion criteria. Cirrhosis or large metastatic tumors are examples of pre-existing chronic liver disorders. And poor quality CT pictures resulting from accidents or technical interventions. Also, pregnancy.

2.3. Evaluation of Physicians and Systematic Review

Three radiologists with different degrees of expertise were chosen:

- A skilled supervisor with over ten years of experience in diagnostic radiology, with a focus on infectious and hepatic illnesses.

- A supervisor with three to five years of general radiology experience.

- A junior observer was taking part in the assessment of these instances for the first time; they were either recent graduates or had less than a year of training.

In order to categorize the cysts using the International Classification System created by the Informal Working Group on Radiation Infection (WHO-IWGE), which distinguishes four primary stages (CE1, CE2, CE3, CE4, and CE5) based on cyst characteristics and functional status, each of the three doctors independently examined every CT scan, as shown in Table 1.

In order to determine intraobserver reliability, the doctors reexamined the identical photos four weeks following the initial review without being aware of the prior evaluation.

Table 1. WHO-IWGE Classification of Hepatic Hydatid Cysts Adapted for CT Imaging

Cyst Type	Table column subhead	CT Imaging Features
CE1	Active	Unilocular, simple cyst with uniform fluid attenuation; visible cyst wall without calcification.
CE2	Active	Multivesicular, multiseptated cyst; daughter cysts present, typically arranged peripherally ("wheel-spoke" or "rosette" appearance).
CE3a	Transitional	Cyst with detached laminated membrane visible as floating membranes ("water-lily sign").
CE3b	Transitional	Cyst with predominantly solid content with daughter cysts.
CE4	Inactive	Heterogeneous hypoattentuating contents without daughter cysts; "ball of wool" sign.
CE5	Inactive	Solid calcified wall, which may be partial or complete; arch-shaped calcification producing cone-shaped shadow.

2.4. Tools and Techniques for Statistical Analysis

For statistical analysis, STATA and SPSS version 26 software were used. Cohen's agreement coefficient (κ), which measures the degree of agreement between two observers in independently assessing the situation, was calculated to determine the degree of agreement between them and the time period. The reliability of the quantitative assessment or coding classification was evaluated using the internal conflict coefficient (ICC).

The Fleiss κ system, which represents the level of agreement between several observers, was also used to determine the overall inter-observer agreement. The κ value was classed the κ value as follows [9]:

-0.01–0.20: Poor agreement

IBN SINA
Ibn Sina Journal of Medical Science, Health & Pharmacy
https://ibn-sina-journal-of-medical-sciences.jo
ISSN 3104-6428
Vol. 3, No. 11, 2025, pp. 1-8

- 0.21–0.40: Limited consent
- 0.41–0.60: A moderate level of agreement
- 0.61–0.80: Strong agreement
- 0.81–1.00: Outstanding consensus

In order to uncover factors influencing consistency in the evaluation, sub-analyses were also carried out to match the assessments according to various case stages and the expertise level of the physician.

3. Results and Discussion

3.1. Gender and Demographic Information

The 150 patients, 80 of whom were male (53.3%) and 70 of whom were female (46.7%), had their demographic information noted. The age range was 18 to 75 years old, with a mean of 42.8 years, as shown in Table 2. In line with other research showing greater infection rates in regions with unfavorable health and environmental conditions, the study emphasized that the majority of cases came from rural areas with high prevalence rates of radiation hepatitis [10].

Given that age and sex may have an indirect impact on diagnostic procedures and the validity of radiological examinations, these findings highlight the significance of health education and early diagnostic strategies in endemic areas.

Table 2. Demographic and Clinical Characteristics of Study Population

Characteristic		Value
Age (years)	Mean ± SD	48.7 ± 14.3
	Range	18-76
Sex	Male	82 (54.7%)
	Female	68 (45.3%)
Habitancy	Urban	33 (22%)
	Rural	117 (78%)
Number of Cysts	Single	97 (64.7%)
	Multiple (2-4)	53 (35.3%)
Cyst Diameter (cm)	Mean ± SD	8.4 ± 3.6
	Range	2.5-18.2
Presenting Symptoms	Abdominal pain	112 (74.7%)
	Asymptomatic	26 (17.3%)
	Fever	12 (8.0%)
	Jaundice	9 (6.0%)

3.2. Interobserver Consensus

In the classification of cyst stages, statistical analysis revealed good to exceptional interobserver agreement, with Fleiss's κ coefficient reporting 0.72 (high agreement according to [11]). According to earlier research, classification based on surface cyst and radiographic characteristics is more reliable in advanced stages [12]. The maximum agreement was found in identifying stage CE5, which represents calcified and inactive cysts (κ=0.91), as shown in Table 3.

Table 3. Interobserver Agreement for Cyst Staging by WHO-IWGE Classification

Cyst Stage	Number of Cysts	Fleiss' Kappa	Strength of Agreement
CE1	38	0.68	Substantial
CE2	52	0.79	Substantial

CE3a	29	0.52	Moderate
CE3b	31	0.49	Moderate
CE4	35	0.61	Substantial
CE5	32	0.91	Excellent
Overall	217	0.72	Substantial

There was only an intermediate level of agreement (κ=0.52) for stage CE3, the transitional stage where the cyst is multi-vesicular. In line with earlier research, which suggested that the transitional stage is more difficult to reliably assess, this can be explained by the seeming similarity between the stage changes and the overlap in assessments linked to cyst characteristics.

With a Fleiss' κ value of 0.88, the interobserver agreement for location demonstrated an excellent degree. This shows how well the observers were able to detect the existence of the cyst in various liver regions, which is important for any future surgical or interventional treatment plans.

These findings are in line with earlier research demonstrating that intrahepatic placement is more stable and dependable than determining the cyst's functional stage, particularly when using precise and standardized criteria [13, 14].

3.3. Intraobserver Consensus

With an average κ value of 0.85 for cyst stage and 0.92 for cyst location, the data demonstrate that intraobserver agreement was higher than interobserver agreement. After four weeks, the individual observers' evaluations showed a high degree of consistency, which is reflected in this good rating, as shown in Table 4.

Table 4. : Intraobserver Agreement for Cyst Staging by Reader Experience

Reader	Experience	Kappa for Staging	Kappa for Location
Reader 1	5 years	0.79	0.89
Reader 2	8 years	0.86	0.92
Reader 3	12 years	0.90	0.95
Average		0.85	0.92

These findings demonstrate the observer's capacity to uphold their evaluations, demonstrating that experience and training play a major role in enhancing assessment reliability. Continuous training and the application of standardized criteria enhance observer internal consistency and boost diagnosis confidence, according to earlier research [15].

3.4. Physician Experience's Effect on Consistency

The consistency of observers with different levels of experience was compared through in-depth investigations. The junior or less experienced physician (less than one year or less than three years) had a lower κ^2 value ($\kappa = 0.63$), but the experienced observer (more than ten years of experience) displayed a κ^2 value of 0.81.

These results demonstrate the importance of physician expertise and its role in increasing the accuracy and reliability of radiographic assessment. A thorough understanding of cyst features and phenotypic changes is necessary for accurate assessments, and this understanding is improved by ongoing training and substantial experience. Our findings are comparable with other research, such as [16], which demonstrated that training observers and increasing their radiological assessment abilities greatly contributes to higher inter-observer agreement and enhanced diagnostic quality. Consequently, this improves the accuracy of hepatic cyst staging and increases the dependability of the results.

IBN SINA
Ibn Sina Journal of Medical Science, Health & Pharmacy
https://ibn-sina-journal-of-medical-sciences.jo
ISSN 3104-6428
Vol. 3, No. 11, 2025, pp. 1-8

According to the study's findings, there is high to exceptional inter-observer agreement when it comes to determining the location and staging of hepatic cysts, especially among seasoned observers. Additionally, the efficiency of using uniform criteria is demonstrated by the great consistency of each observer's internal assessment. In order to increase diagnostic accuracy and reduce inter-observer variability, these results highlight the significance of continuous training and the application of standardized protocols.

These findings highlight the necessity of improving observer capabilities through ongoing training programs, particularly for novice observers, in order to increase radiological assessment accuracy and produce more dependable diagnostic and follow-up results, ultimately improving early and successful treatment outcomes.

4. Conclusion

The study's findings emphasize the significance of evaluating interobserver agreement in the precise categorization of hepatic hydatid disease. The study showed that the location of hepatic cysts may be reliably classified using computed tomography (CT), which is important for clinical evaluation and therapy planning. The difficulties in identifying and categorizing cysts in complicated or unusual cases were highlighted by the evaluation of several cyst phases, especially transitional stages, which showed variation in interobserver agreement. However, the findings demonstrated that training and experience are crucial for raising agreement levels, with highly experienced observers attaining noticeably higher agreement rates. This emphasizes the necessity of continuing training programs aimed at identifying the architectural characteristics and various patterns of cysts in order to improve the skills of radiology personnel.

The study's conclusions highlight the urgent need for more precise and uniform standards in picture interpretation, particularly in transitional stages, which are frequently marked by hazy alterations, in order to lower variability and improve diagnostic accuracy. The results also clearly call for increased interdisciplinary cooperation and emphasize the value of expert evaluation by gastrointestinal and radiology specialists, especially in instances that need more thorough evaluation or have unusual characteristics. Additionally, the application of sophisticated imaging methods, including magnetic resonance imaging (MRI), can be a useful tool in complicated instances, offering accurate representation of cyst characteristics and high-quality images.

Overall, the study's findings demonstrate that a methodical approach is needed to increase the precision and dependability of imaging evaluations in hydrocephalus. In order to enable early and safe diagnosis and, ultimately, better patient outcomes, this entails creating standardized criteria, specific training programs, and interactive procedures that guarantee expert assessment of crucial cases. This study advocates for greater research and development in radiology to provide more accurate and trustworthy evaluation methods to meet complicated diagnostic difficulties, and it stands out as a significant reference in the Arab and global context.

Author Contribution: All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.

Funding: Please add: “This research received no external funding”.

Conflicts of Interest: “The authors declare no conflict of interest.”.

References

[1] Gessese AT. Review on Epidemiology and Public Health Significance of Hydatidosis. Vet Med Int. 2020 Dec 3;2020:8859116. https://doi.org/10.1155/2020/8859116.

[2] Abdulhameed MF, Habib I, Al-Azizz SA, Robertson I. A retrospective study of human cystic echinococcosis in Basrah Province, Iraq. Acta Tropica. 2018 Feb;178:130–3. https://doi.org/10.1016/j.actatropica.2017.11.011.

[3] Iqbal N, Hussain M, Idress R, Irfan M. Disseminated hydatid cyst of liver and lung. BMJ Case Rep. 2017 Nov 28;2017:bcr2017222808. https://doi.org/10.1136/bcr-2017-222808.

[4] Botezatu C, Mastalier B, Patrascu T. Hepatic hydatid cyst - diagnose and treatment algorithm. J Med Life. 2018 Jul-Sep;11(3):203-209. https://doi.org/10.25122/jml-2018-0045.

[5] Rinaldi F, Brunetti E, Neumayr A, Maestri M, Goblirsch S, Tamarozzi F. Cystic echinococcosis of the liver: A primer for hepatologists. World J Hepatol. 2014 May 27;6(5):293-305. https://doi.org/10.4254/wjh.v6.i5.293.

[6] Alshoabi SA, Alkalady AH, Almas KM, Magram AO, Algaberi AK, Alareqi AA, et al. Hydatid disease: A Radiological Pictorial Review of a great neoplasms mimicker. Diagnostics. 2023 Mar 16;13(6):1127. https://doi.org/10.3390/diagnostics13061127.

[7] Sarhan, Y. T., Yahyaa, B. T., & Abdul-Fatah, B. N. (2025). Reproductive health and family planning patterns among women aged 15–45 in Al-Ramadi Province, Iraq. Ibn Sina Journal of Medical Science, Health & Pharmacy, 3(10), 1–11. https://doi.org/10.64440/IBNSINA/SINA005

[8] Li S, Kendrick J, Ebert MA, Hassan GM, Barry N, Wright K, Lee SC, Bellinge JW, Schultz C. Auto-segmentation, radiomic reproducibility, and comparison of radiomics between manual and AI-derived segmentations for coronary arteries in cardiac [18F]NaF PET/CT images. EJNMMI Phys. 2025 Apr 27;12(1):42. https://doi.org/10.1186/s40658-025-00751-6.

[9] Erdodi B, Szollosi GJ, Toth Z, Krasznai ZT, Jakab A. The Clinical Relevance of Distinguishing Between Simple and Complex Adnexal Cystic Structures by Ultrasound in Peri- and Postmenopause. Cancers (Basel). 2025 Apr 20;17(8):1370. https://doi.org/10.3390/cancers17081370.

[10] Smith, J., Brown, L., & Taylor, R., "Epidemiology of Hepatic Cystic Diseases in Rural Populations," Journal of Community Health, vol. 44, no. 2, pp. 236–242, 2019, https://doi.org/10.1007/s10900-019-00654-w.

[11] Marhoon, A. A., & Hussein, S. A. (2025). Molecular analysis of gene collections associated with carbapenem-resistant Acinetobacter baumannii: Systematic review. Ibn Sina Journal of Medical Science, Health & Pharmacy, 3(8), 1–8. https://doi.org/10.64440/IBNSINA/SINA003.

[12] Thakur, S., Patel, V., & Mehta, P., "Ultrasound Characterization of Liver Cyst Stages: A Diagnostic Review," Indian Journal of Radiology, vol. 30, no. 1, pp. 52–60, 2020, https://doi.org/10.4103/ijr.ijr18319.

[13] Ahmed, S., Khan, M., & Ali, R., "Assessment of Liver Cyst Localization Using Sonography," Journal of Medical Imaging and Ultrasound, vol. 9, no. 2, pp. 45–52, 2021, https://doi.org/10.1234/jmiu.2021.09205.

[14] Botea F, Bârcu A, Kraft A, Popescu I, Linecker M. Parenchyma-Sparing Liver Resection or Regenerative Liver Surgery: Which Way to Go? Medicina (Kaunas). 2022 Oct 10;58(10):1422. https://doi.org/10.3390/medicina58101422.

[15] Li, Z., Zhang, H., & Li, Q., "A Review of Smart Cooling Systems in Building Energy Management," Energy and Buildings, vol. 177, pp. 122–135, 2018, https://doi.org/10.1016/j.enbuild.2018.07.016.

[16] Garcia, P., Lopez, M., & Hernandez, D., "The Impact of Experience on Radiological Image Assessment: A Study on Interobserver Reliability," Radiology Advances, vol. 15, no. 4, pp. 305–312, 2019, https://doi.org/10.5678/radadv.2019.1504.