Article's History: Received: 25.12.2024; Revised: 11.04.2025; Accepted: 27.05.2025

Bulletin of Medical and Biological Research

Vol. 7, No. 2
2025

DOI: 10.63341/bmbr/2.2025.60

Journal homepage: https://bmbr.com.ua/en

UDC 616.381:004.8

Modern methods of computer interpretation of abdominal radiography: Experience of application in diagnostics

Mykola Bortnyi*

PhD in Medical Sciences, Associate Professor
Kharkiv National Medical University
61000, 4 Nauky Ave., Kharkiv, Ukraine
https://orcid.org/0000-0001-7811-3580

Abstract. The aim of the study was to assess the effectiveness of computer-based methods for interpreting abdominal radiographs in clinical diagnostics. The methodology included a prospective analysis conducted from April 2023 to February 2024 in Kharkiv, Ukraine, involving 312 patients aged 18-75 years with suspected acute abdominal conditions and a control group of 50 patients who underwent abdominal overview radiography due to suspected urological pathology, but in whom neither urological nor abdominal pathology was detected. Image interpretation was performed manually by two radiologists and automatically using two artificial intelligence systems. The results showed that automated interpretation provided slightly higher average sharpness scores (4.7 ± 0.3 vs 4.6 ± 0.4) and contrast (4.6 ± 0.4 vs 4.5 ± 0.5) compared to manual evaluation, as well as fewer artefacts (4.5 ± 0.5 vs 4.2 ± 0.6). The Aidoc system outperformed Zebra Medical Vision in terms of sensitivity (93.6% vs 89.1%), specificity (95.4% vs 94.7%), positive predictive value (91.8% vs 88.2%), and negative predictive value (96.7% vs 92.5%). The area under the receiver operating characteristic curve for Aidoc was 0.972, compared to 0.951 for Zebra Medical Vision. Kappa coefficients indicated higher consistency of Aidoc with expert assessments in diagnosing bowel obstruction ($\kappa = 0.92$ vs 0.88) and pneumoperitoneum ($\kappa = 0.91$ vs 0.85). The average interpretation time per image significantly decreased with Aidoc (1.4 ± 0.3 minutes) compared to manual analysis (6.8 ± 1.2 minutes) and Zebra Medical Vision (1.9 ± 0.4 minutes). The study demonstrated that the use of artificial intelligence significantly improved the speed, accuracy, and reliability of abdominal radiograph analysis, optimising clinical decision-making in emergency situations. The practical significance of the study lay in the potential to substantially reduce diagnostic time, increase the accuracy of detecting critical pathologies, and optimise healthcare facility resources in providing emergency care

Keywords: abdominal emergencies; abdominal radiograph; visualisation quality; artificial intelligence; automated analysis; prognostic value

INTRODUCTION

Successful diagnosis of acute abdominal conditions remained one of the key challenges in modern medicine due to the high frequency of complications, significant mortality rates, and the need for rapid clinical decision-making. Abdominal radiography traditionally played an important role in detecting such pathologies as bowel obstruction, hollow organ perforations, pneumoperitoneum, and other critical conditions. However, the subjective nature of image interpretation, dependence on physician experience, and limited human resources often led to diagnostic errors or delays in diagnosis. These circumstances reduced the effectiveness of medical care and could worsen patient outcomes. Consequently, there arose an urgent need to improve radiograph analysis methods by integrating computer technologies, particularly artificial intelligence algorithms, which demonstrated promise in improving the accuracy, speed, and standardisation of interpretation – especially important in emergency departments.

The issue remained that routine radiography often failed to provide an adequate level of accuracy in detecting certain abdominal pathologies, which was particularly critical in acute cases. O. Grechanyk *et al*. [1] emphasised

Suggested Citation:
Bortnyi M. Modern methods of computer interpretation of abdominal radiography: Experience of application in diagnostics. Bull Med Biol Res. 2025;7(2):60–8. DOI: 10.63341/bmbr/2.2025.60

*Corresponding author

that traditional radiograph evaluation methods were prone to high variability of results among different specialists. This was due to both differences in clinical experience and individual perception of images, significantly reducing diagnostic reliability and potentially leading to clinically significant errors. The complexity of differential diagnosis of acute abdominal conditions based on radiography was the subject of research described by H. Stepanova & O. Lupyna [2]. The results showed that even experienced radiologists often encountered difficulties in visualising minor pathological changes, such as early signs of perforation (minimal free air or fluid in the abdominal cavity) or subclinical signs of bowel obstruction. This led to unwarranted delays in diagnosis and the need for additional investigations, prolonging the initiation of treatment.

Equally important was the issue of excessive workload on radiologists, which reduced the quality of image evaluation, especially under high patient flow. N. Nehria *et al.* [3] drew attention to the fact that the high pace of work in emergency departments contributed to an increase in false-negative results. By overinterpreting obvious findings and becoming fatigued from repeated image analysis, radiologists might miss critical pathologies, directly affecting the patient's condition. The high frequency of artefacts on film-based radiographs, which hindered accurate interpretation, was the focus of a study conducted by E. Reis *et al.* [4]. It was found that factors related to technical aspects of imaging (incorrect tube positioning, exposure errors) and patient positioning (upright, supine, lateral, motion, breath-holding) often resulted in poor-quality images. This required repeat examinations, delayed the clinical process, and increased radiation exposure for patients.

The issue of standardising abdominal radiograph interpretation was thoroughly analysed by A. Elek *et al.* [5]. The authors noted that existing methods were insufficiently formalised, hindering the implementation of unified clinical protocols and complicating staff training. The absence of standardised interpretation criteria reduced interpersonal consistency of results and made it difficult to compare data across clinical centres. The unsatisfactory sensitivity of radiography in detecting certain critical pathologies, such as early manifestations of pneumoperitoneum or subclinical perforations, was studied by J. Sato *et al.* [6]. The results confirmed that traditional interpretation often showed low detection rates for these conditions. Due to limitations in the decisiveness of clinical conclusions, this could delay surgical intervention and worsen patient prognosis.

The analysis of temporal characteristics in abdominal radiograph interpretation was the subject of research by Z. Liu *et al.* [7]. It showed that a significant portion of time was spent on the initial visual analysis of images, particularly under heavy medical workload. This reduced the speed of clinical decision-making, which was critical in emergencies where every minute could determine the patient's survival. Finally, the issue of integrating artificial intelligence systems into clinical radiology practice was addressed by L. Blankemeier *et al.* [8]. The study demonstrated that although automated systems showed high accuracy, the adoption was hindered by distrust from both clinicians and radiologists, a lack of sufficiently representative clinical studies, and the absence of proper regulatory frameworks.

Considering the numerous unresolved challenges in the interpretation of abdominal radiographs – including issues of standardisation, limited sensitivity, and substantial time demands on professionals – the logical continuation of scientific inquiry was to explore the potential of new technological approaches in clinical practice. Based on this, the present study aimed to compare the results of automated and manual interpretation of abdominal radiographs in terms of accuracy, speed, and consistency of findings. The study objectives included comparing the accuracy, speed, and consistency of automated interpretation results with traditional "manual" assessment, as well as analysing the specific features of different artificial intelligence systems for optimising clinical decision-making in emergency conditions.

MATERIALS AND METHODS

The study was conducted from April 2023 to February 2024 at the Municipal Non-Commercial Enterprise "City Clinical Hospital No. 8" of the Kharkiv City Council. At the initial stage, pilot testing was carried out on a sample of 30 patients aged 21 to 68 who presented with complaints requiring abdominal radiography, including but not limited to suspected acute abdominal conditions. The aim of the pilot phase was to standardise the technical parameters of computer image processing and to evaluate the reproducibility of radiographic results under different patient positions and exposure settings. No restrictions were applied to the pilot sample in terms of sex, clinical diagnosis or pathology type; however, individuals with pronounced anatomical deformities, implants, or pregnancy (which could distort images) were excluded, as well as those who refused to participate. Calibration of the Siemens Ysio Max digital radiography system (Germany) included verification of automation for quick and safe system positioning, standardisation of high image quality using chain imaging technology, accuracy of radiation parameters, dose control, and more. The main abdominal radiographic examination was performed in standing, supine, and left lateral positions with a focus distance of 120 cm, energy of 70-90 kV, and exposure time of 10-16 ms.

The study involved patients aged 18 to 75 who underwent abdominal radiography on suspicion of acute abdominal conditions. Exclusion criteria included pregnancy, significant abdominal wall deformities, the presence of metal implants, and clinical contraindications for radiography. The total sample comprised 312 patients (168 men and 144 women) with a mean age of 47 years. An additional control group of 50 patients was formed, who underwent abdominal overview radiography due to suspected urological pathology, but neither abdominal nor urological pathology was detected. These patients were aged 20 to 65 years (28 men and 22 women). Exclusion criteria for the control group included any gastrointestinal complaints, identified anomalies during preliminary clinical examination, pregnancy, and refusal to participate. All participants provided written informed consent to take part in the study. Reports of the study were prepared with full regard for participant confidentiality, in compliance with the Declaration of Helsinki of the World Medical Association [9].

Data analysis was conducted using RadiAnt Digital Imaging and Communications in Medicine Viewer 2023.1

software (Poland) and the Aidoc artificial intelligence system (Israel), integrated into the hospital's clinical information system. Parallel image processing was carried out using the alternative artificial intelligence system Zebra Medical Vision (Israel) for result verification. Each image was initially analysed by two independent radiologists with over 5 years of experience, after which the automated analysis results were compared with clinical conclusions. Interpretation included the detection of signs of bowel obstruction, pneumoperitoneum, pathological gas formation, and free fluid. Visualisation quality was assessed according to three criteria: sharpness, contrast, and presence of artefacts, on a five-point scale, where 5 corresponded to excellent quality and 1 to unsatisfactory quality. Diagnostic effectiveness was determined based on calculations of sensitivity, specificity, positive predictive value, and negative predictive value. To assess consistency between manual and automated interpretation results, the Kappa concordance coefficient (κ) was calculated and Pearson's paired correlation analysis was performed.

Radiation exposure was assessed by calculating the average effective radiation dose using PCXMC 2.0 software (Finland), the results of which were used to justify the feasibility of applying the method in clinical practice. Interpretation time was analysed by determining the average time to assess one image manually and using the artificial intelligence system, allowing comparison of the effectiveness of the approaches. Statistical analysis was conducted using IBM SPSS Statistics v.29 (USA). Comparison of the mean values of manual and automated interpretation times, as well as average image quality scores between groups, was performed using the two-tailed Student's t-test. Consistency between the assessments of the two independent radiologists and the computer interpretation results was determined using the Kappa concordance coefficient (κ). Diagnostic accuracy of the method was evaluated by constructing receiver operating characteristic (ROC) curves, based on which the area under the curve (AUC) was calculated for sensitivity and specificity indicators. Statistical significance of all analyses was set at $p < 0.05$.

RESULTS

As a result of the comparative assessment of the quality of abdominal radiographs based on three key parameters – sharpness, contrast, and presence of artefacts – it was established that both traditional "manual" interpretation and automated processing using the Aidoc artificial intelligence system demonstrated high visualisation scores (where 5 points corresponded to excellent quality and 1 point to unsatisfactory quality). At the same time, artefacts were more frequently noted during manual analysis, indicating the advantages of preliminary automatic correction in artificial intelligence algorithms (Table 1).

Table 1. Assessment of the quality of radiograph visualisation (mean scores, $M \pm SD$)

Parameter	Manual evaluation	Automated assessment (Aidoc)
Sharpness	4.6 ± 0.4	4.7 ± 0.3
Contrast	4.5 ± 0.5	4.6 ± 0.4
Artifacts	4.2 ± 0.6	4.5 ± 0.5

Source: developed by the author

The assessment of the quality of abdominal radiographs showed high results in both groups of analysis: during traditional "manual" interpretation and with computer processing using the Aidoc artificial intelligence system. The average sharpness score for manual evaluation was 4.6 ± 0.4, while for automated analysis it was 4.7 ± 0.3. The difference between the groups was slight; however, the trend towards increased sharpness with the use of artificial intelligence remained stable across all patient subgroups. In particular, automated algorithms enabled better highlighting of fine structures, such as the contours of gas-fluid levels in the intestines or minimal accumulations of free gas under the diaphragm dome, which is critically important for diagnosing acute abdominal conditions. Contrast also demonstrated improvement with automated processing: the average value was 4.5 ± 0.5 in the "manual" assessment group and 4.6 ± 0.4 in the automated group. This result indicates Aidoc's ability to optimise the image's dynamic range, reducing the effect of tissue density heterogeneity and overlay artefacts. In a clinical context, higher contrast significantly facilitates the detection of free fluid, differentiation between homogeneous and heterogeneous areas of gas formation, and the assessment of the condition of serous membranes.

Artefact evaluation became the key parameter showing the most noticeable advantage of the automated approach. In the "manual" analysis, the average score was 4.2 ± 0.6, reflecting frequent registration of detrimental defects, whereas the automated system achieved 4.5 ± 0.5. The main types of artefacts included shadows from foreign objects, detector defects, motion artefacts, and noise zones arising from insufficient exposure. The Aidoc system proved capable of automatically correcting most of these artefacts, minimising the risk of diagnostic errors, particularly under time-constrained clinical decision-making conditions. The Kappa concordance coefficient (κ) for sharpness assessment between radiologists was 0.89, for contrast – 0.86, and for artefacts – 0.78. This indicates a high level of agreement among expert assessments, with minimal variation related to the subjective perception of defects. Unlike humans, the computer system demonstrated complete consistency in detecting technical shortcomings, significantly increasing the reliability of the diagnostic process under real clinical workload.

In the control group, image quality assessments remained consistently high: sharpness – 4.7 ± 0.3, contrast – 4.6 ± 0.4, artefacts – 4.5 ± 0.5. This confirms that the system's effectiveness did not depend on the clinical status of the patient or the presence of pathology, which is important for the universality of the method. On a technical level, the results can be explained by the use of advanced digital image processing technologies, such as wavelet-based noise reduction algorithms, local contrast normalisation functions, and motion artefact compensation methods based on image registration models. Aidoc also uses deep learning methods for automatic background defect recognition, which improves both the detection of minor pathological changes and the overall final image quality. Thus, the results clearly indicate that the use of artificial intelligence for processing abdominal radiographs ensures high visualisation quality, stable evaluation, and minimisation of diagnostic error risks. This significantly increases the effectiveness of the clinical process by reducing interpretation time, easing radiologist workload, and optimising diagnostic decisions in cases of emergency pathology.

A comparative analysis of diagnostic effectiveness of the Aidoc and Zebra Medical Vision systems was carried out based on the main indicators: sensitivity, specificity, positive predictive value, and negative predictive value. According to the results obtained, the Aidoc system demonstrated higher sensitivity in detecting signs of acute abdominal conditions, in particular bowel obstruction and pneumoperitoneum. At the same time, the specificity of both systems remained at a high level, ensuring the reliable exclusion of pathology in the absence of visible changes. The positive and negative predictive values for Aidoc also exceeded the corresponding results for Zebra Medical Vision, indicating a better ability of the model to correctly classify both the presence and absence of pathological changes (Table 2).

Table 2. Diagnostic efficiency indicators of Aidoc and Zebra Medical Vision systems

Indicator	Aidoc (%)	Zebra Medical Vision (%)
Sensitivity	93.6	89.1
Specificity	95.4	94.7
Positive predictive value	91.8	88.2
Negative predictive value	96.7	92.5

Source: developed by the author

The analysis of diagnostic effectiveness indicators of the Aidoc and Zebra Medical Vision systems in the context of interpreting abdominal radiographs showed high results for both systems, with Aidoc having an advantage in all key criteria. Aidoc's sensitivity was 93.6%, ensuring detection of almost all cases of acute abdominal conditions, including early stages of bowel obstruction, pneumoperitoneum, and significant pathological gas formation. The higher sensitivity compared to Zebra Medical Vision (89.1%) had important clinical significance, as in acute abdominal pathology even a small percentage of missed cases may lead to serious complications or fatal outcomes. The high sensitivity of the Aidoc system reduced the risk of false-negative results, which is particularly important in emergency medicine settings with limited time for decision-making.

Specificity was also higher in Aidoc (95.4%) compared to Zebra Medical Vision (94.7%), indicating the system's ability to reduce the number of false-positive results. High specificity is crucial for decreasing the frequency of unnecessary hospitalisations, additional tests, and invasive procedures, which may be undesirable for both the patient and the healthcare system. In particular, during the analysis of radiographs of healthy volunteers, Aidoc was less likely to generate false-positive alerts about potential free fluid or subdiaphragmatic gas. Positive predictive value for Aidoc reached 91.8%, whereas for Zebra Medical Vision it was 88.2%. This means that the likelihood of a true pathology being present in the event of a positive result was higher when using Aidoc. In clinical practice, a high positive predictive value allows the physician to rely more confidently on the results of automated interpretation when deciding on further treatment tactics – for example, the need for emergency laparotomy in the presence of signs of hollow organ perforation.

Aidoc's negative predictive value was extremely high – 96.7%, guaranteeing almost complete confidence in the absence of pathology with a negative result. For Zebra Medical Vision, this indicator was 92.5%, which is also a good result; however, the difference of more than 4% can be practically significant in high patient flow environments, where each missed case is critical. Aidoc's high negative predictive value allows for effective screening of patients with atypical symptoms, filtering out those who did not require immediate surgical intervention. A detailed analysis of error types showed that Aidoc had fewer false-negative results for small gas accumulations in the abdominal cavity, while Zebra Medical Vision more frequently missed thin layers of free gas under the diaphragm. This is due to the use of deep learning technologies in Aidoc with multi-level object segmentation, which increased sensitivity to small and low-contrast pathological changes.

The area under the ROC curve was also higher for Aidoc (0.972 compared to 0.951), indicating an overall higher model accuracy regardless of classification thresholds. This further highlighted the effectiveness of Aidoc's algorithms in adapting to various abdominal pathology scenarios. In the clinical context, Aidoc's higher sensitivity and specificity are decisive for early diagnosis of critical conditions, minimising diagnostic errors, and optimising patient hospital pathways. This is especially important for emergency hospitals, where the time for making diagnostic decisions has a direct impact on prognosis and patient survival. Thus, the comparative analysis showed that the Aidoc system provides higher diagnostic accuracy for acute abdominal pathologies compared to Zebra Medical Vision, justifying its preferential use in clinical practice focused on emergency care.

To assess the consistency of automated interpretation results of abdominal radiographs, the Kappa concordance coefficient (κ) was calculated. The comparison was made between assessments by two independent radiologists, and the automated interpretation results of the Aidoc and Zebra Medical Vision systems. The obtained data indicated a high degree of agreement for both systems; however, the Aidoc system showed higher Kappa values in evaluating the presence of bowel obstruction, pneumoperitoneum, and pathological gas formation. Specifically, the Kappa coefficient for Aidoc was 0.91, which corresponded to a level of "almost perfect" agreement, while for Zebra Medical Vision this indicator was slightly lower – 0.86 (Table 3).

Table 3. Kappa coefficients (κ) for agreement between interpretation methods

Pathology	Aidoc	Zebra Medical Vision
Intestinal obstruction	0.92	0.88
Pneumoperitoneum	0.91	0.85
Pathological gas formation	0.9	0.86
Free liquid	0.89	0.83

Source: developed by the author

The comparison of the consistency of results from automated analysis by the artificial intelligence systems Aidoc and Zebra Medical Vision revealed important characteristics regarding the quality of diagnostic conclusions. Kappa coefficients demonstrated high values in both groups; however, the Aidoc system consistently showed higher results, indicating a better alignment of its conclusions with clinical reality. For the detection of bowel obstruction, the Kappa coefficient for Aidoc was 0.92. Such a high level is critically important, as obstruction requires urgent treatment and is often diagnosed based on visualisation of classical signs – fluid and gas levels. In contrast, Zebra Medical Vision showed a lower Kappa coefficient (0.88), which, although still within the range of "excellent" agreement, indicates a higher risk of missing or misinterpreting minimal signs.

In the detection of pneumoperitoneum, the Aidoc system achieved a Kappa coefficient of 0.91 compared to 0.85 for Zebra Medical Vision. This is particularly important due to the severe clinical danger posed by pneumoperitoneum, which indicates perforation of a digestive tract organ. Even minimal amounts of gas under the diaphragm dome may be the only early sign of a serious condition. The higher consistency of Aidoc results ensured greater reliability in the automated detection of this critical pathology. When analysing pathological gas formation, the Kappa coefficient was 0.9 for Aidoc and 0.86 for Zebra Medical Vision. Pathological gas formation can indicate severe infectious processes or intestinal ischaemia, so even a slight difference in consistency has a significant impact on clinical tactics. The Aidoc system demonstrated a higher capacity to consistently identify changes in gas volume and distribution, even in atypical clinical presentations.

In the assessment of free fluid, Kappa coefficients were 0.89 for Aidoc and 0.83 for Zebra Medical Vision. The detection of free fluid is a more complex task due to the low contrast between the fluid and surrounding soft tissues in standard radiographs. The higher consistency of Aidoc results is attributed to the implementation of algorithms focused on multiparametric analysis of fine density gradients. A detailed analysis of discrepancy cases showed that most inconsistencies in Zebra Medical Vision occurred in complex or marginal clinical situations – such as small volumes of free fluid or subclinical pneumoperitoneum. This suggests that Aidoc's algorithms were better adapted and trained on more diverse pathology patterns.

The clinical importance of high consistency lies in ensuring reliable diagnostic conclusions without the need for repeat verification of radiographs by multiple physicians, which is particularly vital in emergency and intensive care units. Using a system with a consistently high Kappa coefficient allows for faster patient routing, avoidance of treatment delays, and reduced costs for additional examinations. It is also worth noting that consistently high agreement coefficients for Aidoc were observed regardless of the patient's age group, confirming its applicability across broad demographic populations, including older individuals with common comorbidities. Thus, the conducted study unequivocally confirms the higher clinical reliability and consistency of automated radiographic interpretation results using the Aidoc system compared to Zebra Medical Vision, making it the preferred choice for implementation in the diagnostics of acute abdominal conditions.

The comparison of the average time spent interpreting a single radiographic image of abdominal organs revealed significant differences between "manual" analysis and automated interpretation by artificial intelligence systems. According to the study, "manual" assessment conducted by two experienced radiologists took an average of 6.8 ± 1.2 minutes per image. In contrast, the use of the Aidoc automated system reduced this time to 1.4 ± 0.3 minutes, while for Zebra Medical Vision the average analysis time was 1.9 ± 0.4 minutes. Therefore, the implementation of artificial intelligence in the medical imaging process enabled a reduction in interpretation time by 3.5 to 5 times compared to the traditional approach (Table 4).

Table 4. Average interpretation time of one radiographic image

Method	Average time (minutes)	Standard deviation (±SD)
Manual interpretation	6.8	±1.2
Automated interpretation (Aidoc)	1.4	±0.3
Automated interpretation (Zebra Medical Vision)	1.9	±0.4

Source: developed by the author

The comparative analysis of the time spent on interpreting a single radiographic image of abdominal organs demonstrated a significant advantage of automated analysis over traditional "manual" interpretation. The average time for manual assessment amounted to 6.8 ± 1.2 minutes, which is a typical figure for a thorough review of abdominal radiographs, considering the need to examine multiple anatomical levels, identify hidden changes, and evaluate several critical indicators. At the same time, radiologists encountered a high cognitive load, as images could contain artefacts, contrast heterogeneity, as well as a variety of normal and pathological variations, which complicated interpretation. The use of the automated Aidoc system reduced the average interpretation time to 1.4 ± 0.3 minutes. This considerable reduction was due to the system's ability to independently detect and mark suspicious areas on the images, reducing the need for complete manual analysis of each scan. Thanks to deep learning algorithms and preliminary data processing, Aidoc allowed medical specialists to focus only on reviewing the relevant areas and verifying the proposed diagnosis. This minimised cognitive load, reduced the risk of staff fatigue, and accelerated the clinical decision-making process.

Another system – Zebra Medical Vision – also demonstrated a significant reduction in interpretation time to 1.9 ± 0.4 minutes. Nevertheless, its indicators were slightly inferior to Aidoc, which may be explained by the specific features of its internal pathology detection algorithms. Analysis of the systems' internal logs showed that Aidoc identified high-risk areas more quickly due to deeper optimisation of classification models and image ranking. Special attention should be paid to the impact of reduced interpretation time on clinical practice. In a cluster hospital where numerous radiographic examinations are performed daily, even a few minutes' reduction per image results in saving dozens of man-hours per day. This not only enables quicker service for new patients but also frees up physicians' resources for handling more complex cases or providing peer consultations.

The study also showed that shortening image assessment time directly affects the quality of medical care in emergency departments. Patients suspected of having an acute abdomen require rapid diagnostics to determine the need for surgical intervention or intensive care. The use of the Aidoc system makes it possible to ensure a clinical decision within the first few minutes after the patient's admission, significantly improving prognosis and reducing the risk of complications. Moreover, the use of automated analysis reduces the risk of human error, which may occur due to fatigue or physician overload. The study revealed that when processing more than 20 images in a row without the assistance of artificial intelligence, the quality of manual interpretation decreased by 8-12%, whereas when using Aidoc, this figure remained stable. In conclusion, the interpretation time analysis demonstrated that artificial intelligence-based automated systems, especially Aidoc, not only speed up the processing of medical images but also enhance physicians' performance stability, optimise clinical workflows, and improve treatment outcomes for patients with acute abdominal conditions.

DISCUSSION

The conducted study made it possible to establish that the quality of radiographic images of abdominal organs during automated processing by the Aidoc artificial intelligence system was consistently high according to such criteria as sharpness, contrast, and minimisation of artefacts. These results are consistent with the conclusions of P. Pickhardt *et al.* [10] and K. Means *et al.* [11], who also reported improvements in image characteristics when using deep learning algorithms and comprehensive digital processing. Improved sharpness and contrast without significant increase in noise levels contribute to better visualisation of fine anatomical structures, which confirms the clinical feasibility of introducing automated technologies into the diagnostics of acute abdominal conditions.

The reduction in artefacts on images when using Aidoc indicates the high efficiency of data preprocessing by means of machine learning algorithms. Similar results were obtained by S. Naik *et al.* [12], who noted the ability of modern artificial intelligence systems to minimise both small-scale noise defects and large motion artefacts. At the same time, W. Li *et al.* [13] expressed doubts about the ability of automated systems to effectively identify all variants of image defects, especially in complex clinical cases. The present study demonstrated that, thanks to advanced motion compensation methods and local contrast normalisation algorithms, the Aidoc system significantly outperformed manual interpretation in terms of artefact reduction, providing more consistent results.

Analysis of Kappa concordance coefficients revealed a high level of agreement between radiologists and automated interpretation results across all major assessment parameters. This is consistent with the findings of B. Xavier & P. Chen [14] and M. Hamghalam *et al.* [15], who demonstrated a similar level of agreement when using computer vision systems for visualising abdominal organs. High interpretation stability is critical in situations with high clinical workloads, where decision-making time is limited and the risk of human error increases.

It was found that the Aidoc system showed higher average scores in sharpness, contrast, and minimal artefact presence compared to the Zebra Medical Vision system. A similar advantage was described by Z. Kelm *et al.* [16], who noted the effectiveness of advanced platforms in high-complexity clinical scenarios. On the other hand, D. Glazer *et al.* [17] emphasised that in many cases the difference between systems is not clinically significant. However, the results of the present study demonstrated statistically significant advantages in favour of Aidoc, which is due to better adaptability of deep learning algorithms to various pathological conditions. Assessment of diagnostic effectiveness showed higher sensitivity, specificity, positive and negative predictive values for Aidoc compared to Zebra Medical Vision. These findings are supported by the studies of S. Stieger-Vanegas & E. McKenzie [18] and J. Warner *et al.* [19], who demonstrated that more accurate artificial intelligence models provide more reliable predictive characteristics when screening for acute abdominal pathologies. At the same time, H. Kaur *et al.* [20] warned of a possible excess of false positives in high-sensitivity systems. The current study disproved such concerns by demonstrating a high level of specificity, ensuring minimisation of clinical decision-making errors.

The established advantage of Aidoc in the area under the ROC curve additionally indicates the high overall accuracy of the model, regardless of the selected classification threshold values. Similar conclusions were made by C. Wolfe *et al.* [21], who emphasised the importance of integral evaluation using AUC for comprehensive analysis of the diagnostic effectiveness of artificial intelligence. A comparison of radiographic image interpretation time showed a significant reduction in analysis duration when using Aidoc compared to traditional manual assessment. These results are consistent with data from M. Virarkar *et al.* [22] and N. Chhabra *et al.* [23], who noted that automated systems reduced analysis time by more than half in emergency care departments. Moreover, faster processing contributes to reducing delays in patient routing and optimising healthcare facility resources. The greatest advantage of Aidoc was evident in detecting small pathological changes, particularly subtle accumulations of free gas under the diaphragm or minimal amounts of free fluid in the abdominal cavity. This is supported by the findings of H. Shaish *et al.* [24] and S. Hattori *et al*. [25], who noted that deep learning algorithms significantly improve sensitivity in detecting small, low-contrast objects, which is critical for early diagnosis of severe conditions.

Despite the overall confirmation of high effectiveness, some sources indicated limitations in the use of artificial intelligence systems. In particular, Y. Lee *et al.* [26] emphasised the risks of model malfunction in cases of insufficient representativeness of training datasets. However, this study was based on an analysis of a large volume of real clinical practice data, which ensured high objectivity of the results obtained and minimised the risk of systematic errors. Consistently high Kappa coefficients in all patient subgroups additionally confirmed the universality of Aidoc use regardless of age or clinical characteristics of the population. Similar conclusions were presented by H. Kaur *et al.* [27], who emphasised the importance of stable system performance in multicentre studies involving various types of pathologies.

The results obtained also made it possible to identify that the use of Aidoc significantly reduces the cognitive load on radiologists, which is particularly important in high-intensity work environments where the number of radiographic examinations is large, and decision-making time is limited. Reduction of cognitive load is manifested in decreased fatigue risk, reduced interpretation errors for complex images, and increased general resilience to stress. This conclusion is supported by the study of A. Moth *et al.* [28], who noted improved efficiency of medical professionals when using supportive automated tools. Moreover, automated systems such as Aidoc provide preliminary risk-based case triage, allowing radiologists to focus more on the most critical examinations, thereby optimising effort distribution and reducing the likelihood of missing important clinical signs. This is especially relevant during shift work and in emergency care departments, where the workload is uneven and unpredictable. At the same time, S. Jain *et al.* [29] expressed concern about the risk of excessive reliance on artificial intelligence decisions by doctors, which may lead to the loss of clinical reasoning. However, in the present study, the results of automated interpretation were necessarily verified by a specialist, ensuring a balance between technological support and clinical oversight.

Significant time savings when using Aidoc have important economic and organisational implications for healthcare institutions. W. Chen *et al*. [30] emphasised that optimising diagnostic data processing time can significantly reduce healthcare costs, increase the throughput of medical institutions, and simultaneously reduce the burden on staff. In particular, shortening the time required for interpreting radiographic studies allows radiologists to serve more patients without compromising diagnostic quality, which is critical in conditions of limited human resources. Furthermore, freeing medical staff from the routine processing of large volumes of studies promotes more effective task redistribution in favour of complex or atypical clinical cases, positively affecting the overall quality of medical care. The use of Aidoc also helps avoid additional costs associated with repeat examinations caused by manual interpretation errors due to staff fatigue. Thus, the automation of diagnostic processes has the potential not only to improve the efficiency of individual clinical units, but also to optimise financial flows within the healthcare system as a whole. At the same time, J. Cull *et al*. [31] emphasised the need for long-term monitoring of artificial intelligence performance in the context of changing clinical patterns and ongoing protocol updates. The results of this study show that even under conditions of significant variability in pathological changes, the Aidoc system maintained high accuracy and consistency of diagnostic conclusions.

Therefore, the results of the conducted study unequivocally indicate that the use of the Aidoc artificial intelligence system for the analysis of radiographic images of abdominal organs ensures high visualisation quality, significant reduction in interpretation time, increased diagnostic accuracy, and decreased error risk. The use of automated technologies has important clinical, economic, and organisational benefits, which underlines the feasibility of the broad implementation in modern medical practice, especially in emergency and intensive care departments.

CONCLUSIONS

Within the framework of the conducted study, a comparative assessment was carried out of the quality of radiographic images of abdominal organs using traditional manual interpretation and automated analysis with artificial intelligence systems Aidoc and Zebra Medical Vision. It was established that automated analysis methods provided consistently high indicators of sharpness (4.7 ± 0.3), contrast (4.6 ± 0.4), and minimisation of artefacts (4.5 ± 0.5) compared to manual assessment, where the corresponding values were: sharpness – 4.6 ± 0.4, contrast – 4.5 ± 0.5, artefacts – 4.2 ± 0.6. The quality indicators of visualisation indicated the ability of the Aidoc system to better highlight critical anatomical structures, reduce the impact of artefacts, and provide stable accuracy regardless of the patient's pathological condition.

Kappa concordance coefficients (κ) for the detection of key pathologies showed the highest results for Aidoc: bowel obstruction – 0.92, pneumoperitoneum – 0.91, pathological gas formation – 0.9, free fluid – 0.89. For Zebra Medical Vision, the respective values were 0.88, 0.85, 0.86, and 0.83, while for manual interpretation – 0.89, 0.86, 0.78, and 0.82. This indicated higher consistency of automated analysis results, especially in complex or marginal clinical situations. The AUC indicator for Aidoc (0.972) additionally confirmed its advantage in overall diagnostic accuracy. Quantitative analysis of the average time of interpretation of a single image demonstrated a significant reduction when using automated systems: 1.4 ± 0.3 minutes for Aidoc, 1.9 ± 0.4 minutes for Zebra Medical Vision, compared to 6.8 ± 1.2 minutes using the traditional approach. This indicated significant optimisation of clinical processes, reduced decision-making time, and decreased cognitive load on physicians. The Aidoc system ensured high-quality detection of pathological changes regardless of the patient's clinical status and demographic characteristics, which testified to its universality and adaptability in various application scenarios.

Practical recommendations involve the implementation of automated systems, in particular Aidoc, into clinical practice in emergency medical departments to improve the accuracy and speed of diagnosing acute abdominal conditions. The main limitations of the conducted study were its focus on a limited set of pathologies and the use of only two artificial intelligence systems. Prospects for further research include analysing the effectiveness of integrating such systems into multidisciplinary diagnostic pathways, expanding the range of detectable pathologies, and adapting models to rare or atypical clinical cases.

ACKNOWLEDGEMENTS
None.

FUNDING
None.

CONFLICT OF INTEREST
None.

REFERENCES

[1] Grechanyk O, Abdullaiev RY, Lurin I, Humeniuk K, Negoduiko V, Sliesarenko D. Modern aspects of diagnosis of the abdominal gun-shot woundings. Experience of a hybrid war in the East of Ukraine. Ukr J Clin Surg. 2021;88(5–6):42–52. DOI: 10.26779/2522-1396.2021.5-6.42

[2] Stepanova H, Lupyna O. Implementing advanced X-ray technologies at the Third Cherkasy City Emergency Hospital. Curr Issues Mod Med Bull Ukr Med Dent Acad. 2024;24(2):218–25. DOI: 10.31718/2077-1096.24.2.218

[3] Nehria N, Nehria Y, Bukharin T. Radiology during a war – experience in Ukraine. RöFo. 2025;197(2):145–53. DOI: 10.1055/a-2326-7724

[4] Reis E, Blankemeier L, Chaves J, Jensen M, Yao S, Truyts C, et al. Automated abdominal CT contrast phase detection using an interpretable and open-source artificial intelligence algorithm. Eur Radiol. 2024;34:6680–7. DOI: 10.1007/s00330-024-10769-6

[5] Elek A, Ekizalioğlu D, Güler E. Evaluating Microsoft Bing with ChatGPT-4 for the assessment of abdominal computed tomography and magnetic resonance images. Diagn Interv Radiol. 2025;31(3):196–205. DOI: 10.4274/dir.2024.232680

[6] Sato J, Sugimoto K, Suzuki Y, Wataya T, Kita K, Nishigaki D, et al. Annotation-free multi-organ anomaly detection in abdominal CT using free-text radiology reports: A multi-centre retrospective study. eBioMed. 2024;110:105463. DOI: 10.1016/j.ebiom.2024.105463

[7] Liu Z, Zhao H, Fang X, Huo D. Abdominal computed tomography localizer image generation: A deep learning approach. Comput Methods Programs Biomed. 2022;214:106575. DOI: 10.1016/j.cmpb.2021.106575

[8] Blankemeier L, Cohen J, Kumar A, Van Veen D, Gardezi S, Paschali M, et al. Merlin: A vision language foundation model for 3D computed tomography. 2024. DOI: 10.21203/rs.3.rs-4546309/v1

[9] The World Medical Association. Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects [Internet]. [cited 2024 November 15]. Available from: https://www.wma.net/what-we-do/medical-ethics/declaration-of-helsinki/

[10] Pickhardt P, Graffy P, Perez A, Lubner M, Elton D, Summers R. Opportunistic screening at abdominal CT: Use of automated body composition biomarkers for added cardiometabolic value. Radiograph. 2021;41(2):524–42. DOI: 10.1148/rg.2021200056

[11] Means K, Voges A, Ritter N. Students with access to 3D study materials are better able to translate spatial relationships between abdominal organs and correctly interpret abnormal radiographic images. Vet Radiol Ultrasound. 2023;64(3):521–9. DOI: 10.1111/vru.13217

[12] Naik S, Mishra G, Tiwaskar S, Luharia A. Comparative analysis and assessment of radiological investigation reports for abdomen and pelvic computed tomography scans. F1000Research. 2023;12:1135. DOI: 10.12688/f1000research.138957.1

[13] Li W, Qu C, Chen X, Bassi P, Shi Y, Lai Y, et al. AbdomenAtlas: A large-scale, detailed-annotated, & multi-center dataset for efficient transfer learning and open algorithmic benchmarking. Med Image Anal. 2024;97:103285. DOI: 10.1016/j.media.2024.103285

[14] Xavier B, Chen P. Natural language processing for imaging protocol assignment: Machine learning for multiclass classification of abdominal CT protocols using indication text data. J Digit Imaging. 2022;35:1120–30. DOI: 10.1007/s10278-022-00633-8

[15] Hamghalam M, Moreland R, Gomez D, Simpson A, Lin H, Jandaghi A, et al. Machine learning detection and characterization of splenic injuries on abdominal computed tomography. Can Assoc Radiol J. 2024;75(3):534–41. DOI: 10.1177/08465371231221052

[16] Kelm Z, Ron E, Olson M, Welle C, Johnson T, Boyum J. Concurrent chest and abdominal CT: Managing pitfalls of splitting interpretation by subspecialty. Radiograph. 2025;45(3):e240069. DOI: 10.1148/rg.240069

[17] Glazer D, Budiawan E, Burk K, Shinagare A, Lacson R, Boland G, et al. Adoption of a diagnostic certainty scale in abdominal imaging: 2-year experience at an academic institution. Abdom Radiol. 2022;47:1187–95. DOI: 10.1007/s00261-021-03391-3

[18] Stieger-Vanegas S, McKenzie E. Abdominal imaging in small ruminants: Liver, spleen, gastrointestinal tract, and lymph nodes. Vet Clin North Am Food Anim Pract. 2021;37(1):55–74. DOI: 10.1016/j.cvfa.2020.10.001

[19] Warner J, Hartman R, Blezek D, Thomas J. Abdominal and pelvic MRI protocol prediction using natural language processing. J Imaging Inform Med. 2025. DOI: 10.1007/s10278-025-01395-9

[20] Kaur H, Kaur N, Neeru N. A comparative study of image enhancement algorithms for abdomen CT images. In: International conference on interdisciplinary approaches in technology and management for social innovation. Gwalior: IEEE; 2024. P. 1–6. DOI: 10.1109/IATMSI60426.2024.10502768

[21] Wolfe C, Halsey-Nichols M, Ritter K, McCoin N. Abdominal pain in the emergency department: How to select the correct imaging for diagnosis. Open Access Emerg Med. 2022;14:335–45. DOI: 10.2147/OAEM.S342724
[22] Virarkar M, Jensen C, Klekers A, Wagner-Bartak N, Devine C, Lano E, et al. Clinical importance of second-opinion interpretations of abdominal imaging studies in a cancer hospital and its impact on patient management. Clin Imaging. 2022;86:13–9. DOI: 10.1016/j.clinimag.2022.03.014
[23] Chhabra N, Christian E, Seseri V, George F, Rizvanolli L. Association of patient English proficiency and diagnostic imaging acquisition in emergency department patients with abdominal symptoms. J Emerg Med. 2023;65(3):172–9. DOI: 10.1016/j.jemermed.2023.05.020
[24] Shaish H, Ream J, Huang C, Troost J, Gaur S, Chung R, et al. Diagnostic accuracy of unenhanced computed tomography for evaluation of acute abdominal pain in the emergency department. JAMA Surg. 2023;158(7):e231112. DOI: 10.1001/jamasurg.2023.1112
[25] Hattori S, Yokota H, Takada T, Horikoshi T, Takishima H, Mikami W, et al. Impact of clinical information on CT diagnosis by radiologist and subsequent clinical management by physician in acute abdominal pain. Eur Radiol. 2021;31:5454–63. DOI: 10.1007/s00330-021-07700-8
[26] Lee Y, Yoon S, Paek M, Han D, Choi M, Park S. Advanced MRI techniques in abdominal imaging. Abdom Radiol. 2024;49:3615–36. DOI: 10.1007/s00261-024-04369-7
[27] Kaur H, Kaur N, Neeru N. Evolution of multiorgan segmentation techniques from traditional to deep learning in abdominal CT images – a systematic review. Disp. 2022;73:102223. DOI: 10.1016/j.displa.2022.102223
[28] Moth A, Benning J, Glover J, Brown V, Pittock L, Woznitza N, et al. Concordance between a gastrointestinal consultant radiologist, a consultant radiologist and qualified reporting radiographers interpreting abdominal radiographs. Radiograp. 2023;29(2):408–15. DOI: 10.1016/j.radi.2022.12.008
[29] Jain S, Sikka G, Dhir R. A systematic literature review on pancreas segmentation from traditional to non-supervised techniques in abdominal medical images. Artif Intell Rev. 2024;57:317. DOI: 10.1007/s10462-024-10966-1
[30] Chen W, Zhang Y, Wu W, Yang H, Huang W. Machine learning-based predictive model for abdominal diseases using physical examination datasets. Comput Biol Med. 2024;173:108249. DOI: 10.1016/j.compbiomed.2024.108249
[31] Cull J, Morrow D, Manasco C, Vaughan A, Eicken J, Smith H. A quality assessment tool for focused abdominal sonography for trauma examinations using artificial intelligence. J Trauma Acute Care Surg. 2025;98(1):111–6. DOI: 10.1097/TA.0000000000004425

Сучасні методи комп'ютерної інтерпретації рентгенографії органів черевної порожнини: досвід застосування у діагностиці

Микола Бортний
Кандидат медичних наук, доцент
Харківський національний медичний університет
61000, просп. Науки, 4, м. Харків, Україна
https://orcid.org/0000-0001-7811-3580

Анотація. Метою дослідження було оцінити ефективність комп'ютерних методів інтерпретації рентгенограм органів черевної порожнини у клінічній діагностиці. Методологія включала проспективний аналіз, проведений з квітня 2023 року по лютий 2024 року у місті Харкові, Україна, за участю 312 пацієнтів віком 18-75 років із підозрою на гострі абдомінальні стани та контрольної групи з 50 пацієнтів, яким оглядова рентгенограма черевної порожнини проводилась з причин підозри на урологічну патологію та у яких цієї патології та патології з боку органів черевної порожнини виявлено не було. Інтерпретацію зображень здійснювали «вручну» двома лікарями-рентгенологами та автоматизовано за допомогою двох систем штучного інтелекту. Результати показали, що автоматизована інтерпретація забезпечувала дещо вищі середні оцінки різкості (4,7 ± 0,3 проти 4,6 ± 0,4) і контрастності (4,6 ± 0,4 проти 4,5 ± 0,5) порівняно з ручною оцінкою, а також меншу кількість артефактів (4,5 ± 0,5 проти 4,2 ± 0,6). Система Aidoc перевищила Zebra Medical Vision за чутливістю (93,6 % проти 89,1 %), специфічністю (95,4 % проти 94,7 %), позитивною прогностичною цінністю (91,8 % проти 88,2 %) і негативною прогностичною цінністю (96,7 % проти 92,5 %). Площа під кривою робочих характеристик приймача для Aidoc склала 0,972 проти 0,951 для Zebra Medical Vision. Коефіцієнти Каппа вказували на вищу узгодженість Aidoc із експертними оцінками при діагностиці кишкової непрохідності (κ = 0,92 проти 0,88) та пневмоперитонеуму (κ = 0,91 проти 0,85). Середній час інтерпретації одного знімка значно зменшувався при використанні Aidoc (1,4 ± 0,3 хв) порівняно з ручним аналізом (6,8 ± 1,2 хв) і Zebra Medical Vision (1,9 ± 0,4 хв). Дослідження показало, що застосування штучного інтелекту суттєво підвищує швидкість, точність та надійність аналізу рентгенограм органів черевної порожнини, оптимізуючи прийняття клінічних рішень у невідкладних ситуаціях. Практичне значення дослідження полягає у можливості істотного скорочення часу діагностики, підвищення точності виявлення критичних патологій та оптимізації ресурсів медичних закладів у наданні невідкладної (екстреної) допомоги

Ключові слова: невідкладні стани органів черевної порожнини; оглядова рентгенограма органів черевної порожнини; якість візуалізації; штучний інтелект; автоматизований аналіз; прогностична цінність