Abdominal Radiology
https://doi.org/10.1007/s00261-025-05296-x

REVIEW

Comparison of the predictive value of different assessment times in the severity and prognostic outcomes of CTSI in patients with acute pancreatitis: a systematic review and meta-analysis

Shan Huang[1] · **Xinwei Liu**[2] · **Longyan Zhu**[3] · **Kun Ai**[1]

Received: 1 September 2025 / Revised: 4 November 2025 / Accepted: 7 November 2025

Abstract

Background Acute pancreatitis (AP) necessitates accurate severity and prognosis assessment. While the Computed Tomography Severity Index (CTSI) is widely used, its time-dependent predictive value remains unclarified. This meta-analysis compares the predictive efficacy of CTSI across assessment times for AP outcomes.

Methods Systematic searches were performed in PubMed, Cochrane, Embase, and Web of Science. Literature was collated by using EndNote 20. Methodological quality was assessed using QUADAS-2. Bivariate meta-analyses were conducted with Stata 16 and Meta-Disc 1.4.

Results Analysis of 28 studies ($n=5,419$) revealed significant time-dependent variations: CTSI achieved optimal sensitivity for severity prediction at ≤48 h (0.84, 95%CI 0.78–0.89) with specificity 0.79 (0.76–0.82), and peak sensitivity for organ failure at ≤48 h (0.90, 0.82–0.95) though with moderate specificity (0.52, 0.48–0.57). Mortality prediction showed the highest sensitivity at ≤72 h (0.89, 0.74–0.97), despite suboptimal specificity (0.65, 0.61–0.68). Pancreatic necrosis detection demonstrated superior accuracy at >72 h (sensitivity 0.91, 0.84–0.95; specificity 0.82, 0.79–0.85). The pooled area under the receiver operating characteristic curve values with 95% CI for overall predictive performance were: severity, 0.85 (0.82–0.88); organ failure, 0.86 (0.82–0.88); mortality, 0.85 (0.81–0.88); and pancreatic necrosis, 0.94 (0.99−0.96).

Conclusion CTSI exhibits distinct temporal predictive patterns: ≤48 h assessments optimise early evaluation of severity/organ failure, ≤72 h scans best predict mortality, while delayed imaging (>72 h) maximises pancreatic necrosis accuracy at the expense of clinical timeliness. Future research must standardise imaging timepoints, integrate CTSI with physiological biomarkers, and develop dynamic assessment models to resolve discrepancies in anatomical-pathophysiological prognostic factors in AP management.

Keywords Acute pancreatitis · Computed tomography severity index · Prognosis · Meta-analysis

✉ Kun Ai
aikun650@qq.com

Shan Huang
501837343@qq.com

Xinwei Liu
1109627843@qq.com

Longyan Zhu
1679978423@qq.com

[1] College of Acupuncture-Moxibustion, Tuina and Rehabilitation, Hunan University of Chinese Medicine, Changsha, China

[2] Department of Critical Care Medicine, The Second Xiangya Hospital of Central South University, Changsha, China

[3] Institute of Translational Medicine, Nanchang University, Nanchang, China

Introduction

Acute pancreatitis (AP) is an inflammatory disorder of the pancreas characterised by premature activation of pancreatic enzymes, leading to autodigestion, oedema, haemorrhage, and pancreatic tissue necrosis [1]. Globally, AP affects 13–45 individuals per 100,000 annually, with mortality rates reaching 1.6 per 100,000 [2, 3]. Approximately 15–20% of patients progress to severe acute pancreatitis (SAP), defined by persistent organ failure (≥48 h) under the Revised Atlanta Classification [4], and SAP mortality remains alarmingly high (10–30%) [5]. Adverse outcomes such as pancreatic necrosis, systemic inflammatory response syndrome (SIRS), and multi-organ failure are key

Published online: 20 November 2025

determinants of poor prognosis [6]. For instance, pancreatic necrosis occurs in 20–40% of SAP cases. Patients with pancreatic necrosis account for over 70% of all AP-related mortality cases, with infected necrosis being the primary cause of death [7, 8]. Organ failure increases the risk of mortality by 5- to 10-fold [9]. The guidelines [10, 11] recommend that prognosis be determined by assessing organ failure and SIRS within the 48 h after admission. These underscore the critical need for early risk stratification to guide timely interventions, optimize resource allocation, and enhance survival [12].

Several prognostic tools have been developed to assess the severity of AP and its outcomes. The Acute Physiology and Chronic Health Evaluation-II (APACHE-II) score, a multi-organ dysfunction assessment, provides dynamic risk stratification, with an area under the receiver-operating characteristic (ROC) curve (AUC) of 0.82–0.88 for mortality prediction [13]. However, its reliance on 24-hour data collection limits practicality in rapid triage [14]. The Ranson Criteria, introduced in 1974, evaluates 11 clinical and laboratory parameters within 48 h, demonstrating moderate accuracy (AUC: 0.70–0.85) but requiring prolonged data collection [15]. In contrast, imaging-based systems, such as the Balthazar Score and its evolution, the Computed Tomography Severity Index (CTSI), provide morphologic insights into pancreatic inflammation and necrosis. CTSI integrates pancreatic inflammation grading (0–4) and necrosis extent (0–6), demonstrating moderate predictive accuracy for AP severity assessment [16, 17]. However, CTSI's reliance on contrast-enhanced computed tomography (CECT) raises concerns about radiation exposure and renal risks [18, 19]. Though validated in critical care, the Sequential Organ Failure Assessment (SOFA) score lacks specificity for AP. Despite advancements, consensus on the optimal timing for CTSI assessment remains elusive, as early imaging may underestimate necrosis, while delayed scans risk clinical deterioration [20, 21].

As for the temporal relationship between CTSI assessment and its predictive accuracy, evidence-based clinical guidelines [11] recommend that initial CT assessment be performed in AP patients to confirm severity, with the optimal timing no earlier than 72 to 96 h post-symptom onset (GRADE 1 C, strong consensus). Existing studies have shown that CTSI, performed 72 h after admission of patients with AP, demonstrated a strong ability to predict pancreatic necrosis (AUC = 0.879; 95% CI, 0.77–0.96) [22]. However, because pancreatic morphological changes may manifest late, earlier CT scanning may have limited utility [23]. By contrast, emerging evidence [24] suggests that prompt CT evaluation (within 48 h of symptom onset) enables quantitative assessment of extrapancreatic necrosis volume and pancreatic tissue damage, thereby preventing missed opportunities for critical early intervention windows. This temporal discordance underscores the need for standardised assessment windows, particularly given the emergence of quantitative imaging biomarkers that map time-dependent diagnostic trajectories [24]. Our systematic review aims to resolve these discrepancies by synthesising evidence on optimal CTSI evaluation timelines for predicting multidimensional outcomes.

Methods

This systematic review adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [25]. The study protocol was prospectively registered with the International Prospective Register of Systematic Reviews (PROSPERO) under registration (Registration ID: CRD42024574245).

Search strategy

We conducted comprehensive literature searches in PubMed, Cochrane Library, Embase, and Web of Science from the database inception through March 1st, 2025. Only literature published after 2000 was included, with full search syntax detailed in Table S1.

Inclusion and exclusion criteria

The following inclusion criteria were defined: (1) Studies in which patients were definitely diagnosed with AP. The diagnosis of AP required at least 2 of the following 3 criteria: typical acute epigastric pain with or without back radiation; a serum amylase or lipase level at least 3 times the upper limit of normal; and imaging findings consistent with AP. (2) Studies reporting CTSI or Modified CTSI (MCTSI). (3) Studies on at least one of the following outcomes: severity, mortality, organ failure, and pancreatic necrosis. (3) Studies on at least one of the following outcomes: severity, mortality, organ failure, and pancreatic necrosis. AP severity was classified as mild, moderate, or severe according to the 2012 Revised Atlanta Classification. Data on SAP were included in this study. Deaths occurred during hospitalisation. According to the modified Marshall scoring system, organ failure was defined as three organ systems with scores of 2 or more: respiratory, cardiovascular, and renal. Pancreatic necrosis was determined by CECT, which showed absent pancreatic parenchymal enhancement, and was confirmed. (4) Studies mentioning the time of CT evaluation. (5) Studies providing sensitivity, specificity, the gold standard, and the number of negative and positive cases.

Exclusion criteria comprised: (1) duplicate publications or irrelevant content; (2) preclinical studies using animal/cell models; (3) non-original research (reviews, meta-analyses, conference abstracts, case reports, guidelines, or commentaries); (4) insufficient data for constructing 2 × 2 contingency tables (true positives, false positives, true negatives, false negatives); (5) inaccessible full-text articles.

Literature screening

In this study, two researchers independently used EndNote 20 to screen the literature. A third researcher makes the judgment in the event of any dispute. The screening process mainly consisted of two steps. First, an initial screening was performed by the titles and abstracts of the retrieved studies to exclude those that did not meet the criteria. Second, after the initial screening, the literature was then screened in full text, and the full texts of the screened articles were downloaded. The final inclusion of studies was decided based on the screening results.

Data extraction

Two researchers independently collected data from eligible studies. Disagreements were resolved by discussion with a third researcher. For studies that met the inclusion criteria, data extracted mainly included the first author, year of publication, study design, country, type of evaluation, time of evaluation, diagnostic criteria, aetiology, age, gender (female/male), sample size, and outcome.

Quality assessment

Two researchers used the revised Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) [26] to independently assess the methodological quality and potential bias of the selected studies. If necessary, a third researcher will join to make a final judgment. Disagreements in the quality assessments were resolved through discussion. The QUADAS-2 tool consists of 4 main components: patient selection, index test, reference standard, flow, and timing. Risk of bias includes patient selection, index test, reference standard, flow, and timing. Clinical applicability includes patient selection, index tests, and reference standards. Each item was categorised as high (risk), low (risk), or unclear (risk).

Statistical analysis

Stata 16 and Meta-Disc 1.4 were used for the meta-analysis. Sensitivity, specificity, the gold standard, and the number of negative and positive cases were provided in the included studies. Then, true positives, false positives, true negatives, and false negatives were calculated from a 2 × 2 table of diagnostic tests. If only the ROC curves were available, Origin software (version 2021) was used to extract the sensitivity and specificity at the optimal thresholds, which were then used to calculate true positives, false positives, true negatives, and false negatives. Diagnostic accuracy was evaluated using sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), diagnostic odds ratio (DOR), ROC curve, area under the curve (AUC), and 95% confidence interval (CI) under a random-effects model. Publication bias was assessed using the Deeks funnel plot asymmetry test. The stability of the results was assessed by sensitivity and specificity analyses. Possible sources of heterogeneity were explored by meta-regression and subgroup analyses. Heterogeneity was assessed using the I2 statistic and the corresponding chi-square test, with a random-effects model required for I2 values greater than 50% and a fixed-effects model for values below 50%. $I2 > 50\%$ and $p < 0.1$ indicated significant heterogeneity.

Results

Search results

The systematic search across four databases initially yielded 2,028 records. After duplicate removal and title/abstract screening, 629 full-text reports were assessed for eligibility. Of these, 28 studies [14, 16, 22, 27–51] met the inclusion criteria for meta-analysis. The primary reasons for exclusion during full-text review were irretrievable data and unreported outcome indicators. The detailed selection process is illustrated in the PRISMA flow diagram (Fig. 1).

Characteristics of the included studies

A total of 28 studies were included in this study, which contained data on 5,419 patients with AP. The included studies were published between 2003 and 2023. The 28 studies included 14 retrospective studies [27–31, 33, 41–45, 47, 49, 50], 12 prospective studies [14, 16, 22, 34–37, 39, 40, 46, 48, 51], and 2 cohort studies [32, 47]. 8 studies [29–32, 41, 43, 44, 47] were from China, 9 studies [22, 34, 35, 37–40, 42, 46] from India, 3 studies [28, 33, 36] from Turkey, 2 studies [14, 48] from the USA, 1 studies [27] from France, 1 studies [16] from Spain, 1 studies [50] from Belgium, 1 studies [51] from Greece, 1 studies [49] from Japan, and 1 studies [45] from Korea. 18 CTSI-assessed studies [14, 16, 22, 27, 28, 32, 34, 36, 43–51] and 10 MCTSI-assessed studies [29–31, 33, 35, 37, 39–42] were included. Of these, 3 studies [48–50] were assessed within 24 h, 5 studies [14,

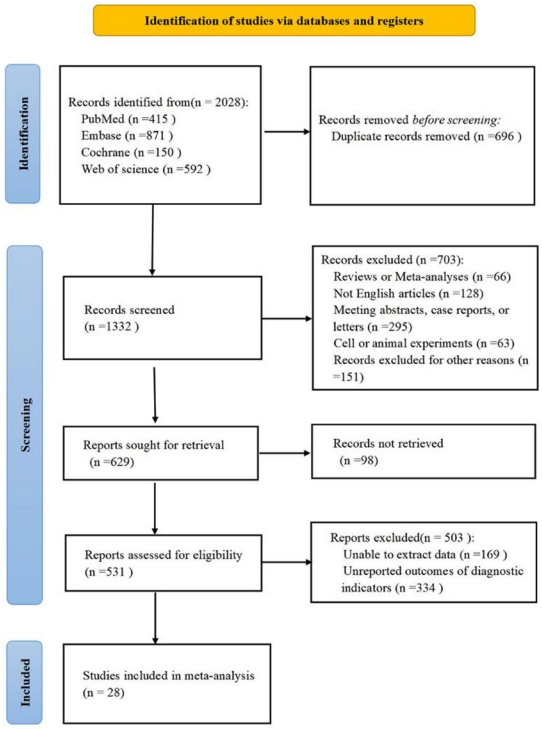

Fig. 1 Flowchart of study selection

28, 29, 32, 35] within 48 h, 7 studies [16, 33, 34, 41, 43, 47, 51] within 72 h, and 13 studies [22, 27, 30, 31, 36–40, 42, 44–46] after 72 h. The age range of patients was 53.1 ± 27.7 years. 18 studies [14, 22, 29, 30, 32, 33, 35, 37, 38, 41, 43–46, 48–51] reported severity, 18 studies [14, 16, 22, 28, 30, 33, 34, 37, 38, 40–42, 44–46, 48–50] reported mortality, 18 studies [16, 22, 27, 29–32, 34, 36, 37, 39, 40, 42, 43, 45–47, 51] reported organ failure, and 7 studies [14, 22, 29, 30, 39, 45, 46] reported pancreatic necrosis. The main characteristics of the included studies are shown in Table 1.

Quality assessment of the selected studies

QUADAS-2 was used to assess methodological quality. For risk of bias, on patient selection, there were 7 studies [14, 16, 22, 27, 29, 32] with low risk, 2 studies [41, 43] with high risk, and 19 studies [28, 30, 31, 33, 34, 36–40, 42, 44–51] with unclear risk. On the index test, there were 18 studies [14, 22, 27, 31, 32, 34, 35, 38, 41–48, 50, 51] with low risk, 5 studies [28, 37, 39, 40, 49] with high risk, and 5 studies [16, 29, 30, 33, 36] with unclear risk. On reference standard, there were 12 studies [14, 22, 27, 31, 32, 38, 40, 41, 43, 44, 48, 50] with low risk, 4 studies [16, 36, 45, 47] with high risk, and 12 studies [28–30, 33–35, 37, 39, 42, 46, 49, 51] with unclear risk. On flow and timing, there were 14 studies [14, 16, 22, 27–29, 31, 32, 34, 35, 41, 43, 47, 48] with low risk, 8 studies [30, 38, 40, 42, 45, 49–51] with high risk, and 6 studies [33, 36, 37, 39, 44, 46] with unclear risk. For clinical applicability, regarding patient selection, except for two studies [42, 44] with high risk, the remaining 26 studies [14, 16, 22, 27–40, 42, 44–51] had a low risk. On the index test, 25 studies [14, 22, 23, 26, 28–37, 39–41, 43–50] had a low risk, 1 study [43] had a high risk, and 2 studies [16, 39] had an unclear risk. Of the reference standard, 24 studies [14, 22, 27, 29, 31–44, 46–51] had a low risk, 1 study [28] had a high risk, and 3 studies [16, 30, 45] had an unclear risk. The specific evaluation results are shown in Fig. 2.

Outcome indicators

Predictive performance for severity

Subgroup analyses were performed according to the CTSI evaluation time, divided into four time periods: ≤24 h, ≤48 h, ≤72 h, and >72 h. A total of 18 studies provided data on the prediction of AP severity by CTSI, with a pooled sensitivity of 0.76 (95% CI: 0.67–0.83), a pooled specificity of 0.80 (95% CI: 0.73–0.86), and a pooled AUC of 0.85 (95% CI: 0.82–0.88); pooled PLR was 3.90 (95% CI: 2.88–5.28); pooled NLR was 0.29 (95% CI: 0.22–0.40); and pooled DOR was 13.26 (95% CI: 8.60–20.46) (Fig. 3A; Table 2).

In the study on severity, CTSI evaluation time ≤ 24 h ($n = 3$) had a sensitivity of 0.81 (95% CI: 0.72–0.88) and a specificity of 0.76 (95% CI: 0.70–0.81); evaluation time ≤ 48 h ($n = 4$) had a sensitivity of 0.82 (95% CI: 0.69–0.90) and a specificity of 0.78 (95% CI: 0.68–0.85); evaluation time ≤ 72 h ($n = 4$) had a sensitivity of 0.56 (95% CI: 0.48–0.64) and a specificity of 0.86 (95% CI: 0.84–0.88); and evaluation time > 72 h ($n = 7$) had a sensitivity of 0.69 (95% CI: 0.56–0.79) and a specificity of 0.83 (95% CI: 0.73–0.90). The evaluation time ≤ 48 h had the highest sensitivity for predicting severity, while the highest specificity was observed in the ≤ 72 h subgroup (Table 2). There was a statistically significant difference in the prediction of severity at ≤ 24 h compared with ≤ 72 h ($p < 0.01$), and at ≤ 48 h compared with ≤ 72 h and > 72 h ($p < 0.05$).

Predictive performance for mortality

A total of 18 studies provided data on the prediction of mortality by CTSI, with a pooled sensitivity of 0.84 (95% CI: 0.71–0.92), pooled specificity of 0.71 (95% CI: 0.59–0.81), pooled AUC of 0.85 (95% CI: 0.81–0.88); pooled PLR was 2.93 (95% CI: 2.06–4.16); pooled NLR was 0.22 (95% CI: 0.12–0.42); and pooled DOR was 13.34 (95% CI: 6.07, 29.31) (Fig. 3B; Table 2).

Table 1 Basic characteristics of the included studies

No.	Author	Year	Study design	Country	Evaluation	Evaluation time	Diagnostic criteria	Etiology	Age (Mean±SD)	Gender (Female/Male)	Sample size	Outcome
1	Tortum et al. [28]	2023	Retrospective	Turkey	CTSI	≤48 h	Serum amylase/lipase levels	–	60.50±6.75	162/78	240	Mortality
2	Luo et al. [29]	2023	Retrospective	China	MCTSI	≤48 h	The 2012 AP Grading and Classification System	Gallstone 43.09%; Hyperlipidemic 22.04%; Alcoholic 13.82%; Unknown 21.05%	47.45±2.74	90/214	304	Severity, organ failure, pancreatic necrosis
3	Zhao et al. [30]	2022	Retrospective	China	MCTSI	>72 h	The guidelines for the Diagnosis and Treatment of Acute Pancreatitis revised in 2019 by the Pancreatic Surgical Science Section of the Chinese Medical Association Surgery Branch; the modified Marshall scoring system; death during hospitalization	Gallstone 54.23%; Hyperlipidemic 24.30%; Alcoholic 4.58%; Other 21.48%	54.05±22.96	123/161	284	Severity, mortality, organ failure, pancreatic necrosis
4	Alberti et al. [16]	2021	Prospective	Spain	CTSI	≤72 h	2012 revised Atlanta; the modified Marshall scoring system; CECT	Gallstone 47.7%; Hyperlipidemic 39.6%; Alcoholic 17.4%; Idiopathic 22.1%; Post ERCP 4%; Other 6%	–	57/92	149	Mortality, organ failure
5	Liu et al. [31]	2021	Retrospective	China	MCTSI	>72 h	2012 revised Atlanta	Gallstone 45.8%; Hyperlipidemic 40.6%; Alcoholic 3.2%; Post ERCP 5.8%; Other 4.5%	48.70±14.70	47/108	155	Organ failure
6	Çakar et al. [36]	2020	Prospective	Turkey	CTSI	>72 h	2012 revised Atlanta	Gallstone 51.5%; Hyperlipidemic 15.1%; Alcoholic 6.1%; Post ERCP 5.1%; Unknown 22.2%	51.50±11.50	37/62	99	Organ failure
7	Gezer et al. [33]	2020	Retrospective	Turkey	MCTSI	≤72 h	2012 revised Atlanta	Gallstone 57.5%; Alcoholic 7.6%; Hyperlipidemic 8.8%; Other 5.1%; Unknown 21%	55.00±17.00	46/34	80	Severity, mortality
8	Peng et al. [32]	2020	Cohort	China	CTSI	≤48 h	2012 revised Atlanta; the modified Marshall scoring system	Gallstone 43.7%; Hyperlipidemic 22.0%; Alcoholic 7.4%; Trauma 0.6%; Postoperative state 1.6%; Unknown 24.6%	50.00±16.00	113/196	309	Severity, organ failure

Table 1 (continued)

No.	Author	Year	Study design	Country	Evaluation	Evaluation time	Diagnostic criteria	Etiology	Age (Mean±SD)	Gender (Female/Male)	Sample size	Outcome
9	Chitrambalam et al. [34]	2020	Prospective	India	CTSI	≤72 h	2012 revised Atlanta	–	38.68±11.94	12/108	120	Mortality, organ failure
10	Chatterjee et al. [35]	2020	Prospective	India	MCTSI	≤48 h	2012 revised Atlanta; the modified Marshall scoring system	Alcoholic 58.6%; Gallstones 20.7%; Hyperlipidemic 1.1%; Drug-induced 5.7%; Hypercalcemia 3.4%; Pancreas division 3.4%; Uncontrolled diabetes, 3.4%; Idiopathic 1.1%; Pancreatic duct obstruction 1.1%; Hydatid cysts 1.1%	37.72±12.6	12/75	87	Severity
11	Jain et al. [37]	2019	Prospective	India	MCTSI	>72 h	2012 revised Atlanta; the modified Marshall scoring system	Alcoholic 98%; Idiopathic 2%	42.06±13.27	0/50	50	Severity, mortality, organ failure
12	Hagjer & Kumar [22]	2018	Prospective	India	CTSI	>72 h	2012 revised Atlanta; the modified Marshall scoring system; CECT	Alcoholic 45%; Gallstone 40%; Idiopathic 15%	37.17±11.77	19/41	60	Severity, mortality, organ failure, pancreatic necrosis
13	Vaidya et al. [38]	2018	Cohort	India	CTSI	>72 h	2012 revised Atlanta; the modified Marshall scoring system	Alcoholic 84%; Gallstone 13%; Idiopathic 3%	40.16	6/49	55	Severity, mortality
14	Kumar & Griwan [39]	2018	Prospective	India	MCTSI	>72 h	2012 revised Atlanta; the modified Marshall scoring system	Gallstone 74%; Alcoholic 18%; Traumatic 2%; Idiopathic 6%	48.42±15.25	33/17	50	Organ failure, pancreatic necrosis
15	Sahu et al. [40]	2017	Prospective	India	MCTSI	>72 h	2012 revised Atlanta; the modified Marshall scoring system; CECT	Alcoholic 50.0%; Gallstone 25.0%; Idiopathic 22.0%; Other 3.0%	36.60±9.80	24/36	60	Mortality, organ failure
16	Yang et al. [41]	2016	Retrospective	China	MCTSI	≤72 h	Chinese guideline for diagnosis and treatment of acute pancreatitis; Atlanta Classification of Acute Pancreatitis	–	44.00±17.75	142/184	326	Severity, mortality
17	Qiu et al. [43]	2015	Retrospective	China	CTSI	≤72 h	2012 revised Atlanta	Gallstone 65%; Alcoholic 5.2%; Hyperlipidemic 14.2%; Unknown and idiopathic 15.6%	57.00±17.30	441/468	909	Severity, organ failure

Table 1 (continued)

No.	Author	Year	Study design	Country	Evaluation	Evaluation time	Diagnostic criteria	Etiology	Age (Mean±SD)	Gender (Female/Male)	Sample size	Outcome
18	Meyrignac et al. [27]	2015	Retrospective	France	CTSI	>72 h	2012 revised Atlanta	Gallstone 34%; Alcoholic 22%; Ampullary tumors 8%; Post ERCP 5%; Other 1%; Unknown 26%	55.30±19.00	104/160	264	Organ failure
19	Sharma et al. [42]	2015	Retrospective	India	MCTSI	>72 h	2012 revised Atlanta; the modified Marshall scoring system	Alcoholic 50.5%; Gallstone 34.3%; Idiopathic 11.4%; Other 3.8%	40.60±12.99	40/65	105	Mortality, organ failure
20	Xu et al. [44]	2014	Retrospective	China	CTSI	>72 h	the International Classification of Diseases, Ninth Revision, Clinical Modification code for AP (577.0); the Atlanta	Gallstone 35.8%; Alcoholic 27.6%; Gallstone+alcohol 21.4%; Post ERCP 6.6%; Other or unknown 8.6%	51.20±12.30	61/196	257	Severity, mortality
21	Chen et al. [47]	2013	Retrospective	China	CTSI	≤72 h	the Atlanta	Gallstone 66.0%; Alcoholic 6.8%; Hyperlipidemia 10.1%; Idiopathic 16.1%; Other 1.0%	53.60±16.60	222/275	497	Organ failure
22	Khanna et al. [46]	2013	Prospective	India	CTSI	>72 h	the Atlanta	Gallstone 61.1%; Alcoholic 18%; Idiopathic 12.5%; Hyperlipidemia 2.8%; post ERCP 2.8%; Traumatic 2.8%	40.50±14.50	35/37	72	Severity, mortality, organ failure, pancreatic necrosis
23	Park et al. [45]	2013	Retrospective	Korea	CTSI	>72 h	CECT; organ failure (transient or persistent) and/or local complications; the most extreme laboratory value or clinical measurement	Alcoholic 49.8%; Gallstone 29.0%; Hyperlipidemic 1.3%; Idiopathic 16.2%; Cancer 3.6%	52.00±17.00	87/216	303	Severity, mortality, organ failure, pancreatic necrosis
24	Bollen et al. [48]	2012	Prospective	USA	CTSI	≤24 h	2012 revised Atlanta; the modified Marshall scoring system	Gallstones 30%; miscellaneous (e.g., hypertriglyceridemia, hereditary, and post-endoscopic retrograde cholangiopancreatography) 24%; alcohol 21%; idiopathic 17%; drug-induced 8%	56.00±17.50	66/84	159	Severity, mortality
25	Imamura et al. [49]	2010	Retrospective	Japan	CTSI	≤24 h	the Atlanta	Alcoholic 42%; Gallstone 23%; Idiopathic 20%; Postoperative 5%	53.00±17.00	37/85	122	Severity, mortality

Table 1 (continued)

No.	Author	Year	Study design	Country	Evaluation	Evaluation time	Diagnostic criteria	Etiology	Age (Mean ± SD)	Gender (Female/Male)	Sample size	Outcome
26	Papachristou et al. [14]	2010	Prospective	USA	CTSI	≤48 h	CECT; the presence of organ failure for more than 48 h	Gallstone 36%; Idiopathic 27%; Alcoholic 14%; post ERCP 14%; Hyperlipidemia 4%; Other 5%	52.35 ± 18.75	90/95	185	Severity, mortality, pancreatic necrosis
27	De Waele et al. [50]	2007	Retrospective	Belgium	CTSI	≤24 h	the International Classification of Diseases, Ninth Revision, Clinical Modification code for AP (577.0).	Alcoholic 50%; Gallstone 20%; Hyperlipidemia (15%); Traumatic 5%; Unknown 10%	50.00 ± 17.70	12/28	40	Severity, mortality
28	Chatzicostas et al. [51]	2003	Prospective	Greece	CTSI	≤72 h	the Atlanta	Gallstone 60.3%; Alcoholic 5.1%; Others (hyperlipidemia, pancreatic malignancy, pancreas division) 6.4%; Unknown 28.2%	63.80 ± 17.00	36/42	78	Severity, organ failure

CTSI computed tomography severity index, *MCTSI* modified computed tomography severity index, *SD* standard difference

In the study on mortality, it was shown that the sensitivity of CTSI evaluation time ≤ 72 h ($n = 4$) was 0.90 (95% CI:0.72–0.97) with a specificity of 0.56 (95% CI:0.34–0.76); the sensitivity of evaluation time ≤ 48 h ($n = 2$) was 0.85 (95% CI:0.70–0.94) and a specificity of 0.77 (95% CI:0.73–0.82); the sensitivity for evaluation time ≤ 24 h ($n = 3$) was 0.81 (95% CI:0.63–0.93) and a specificity of 0.66 (95% CI:0.60–0.71); the sensitivity was 0.79 (95% CI:0.48–0.94) and the specificity was 0.74 (95% CI:0.72–0.77) for evaluation time > 72 h ($n = 9$). The highest sensitivity for predicting mortality was observed at evaluation times ≤ 72 h, and the highest specificity was observed at evaluation times ≤ 48 h (Table 2). No statistically significant difference in mortality prediction was found in patients with AP in either comparison ($p > 0.05$).

Predictive performance for organ failure

A total of 18 studies provided data on the prediction of organ failure by CTSI, with a pooled sensitivity of 0.84 (95% CI: 0.70–0.92), a pooled specificity of 0.75 (95% CI: 0.63–0.83), a pooled AUC of 0.86 (95% CI: 0.82–0.88); pooled PLR was 3.30 (95% CI: 2.30–4.72); pooled NLR was 0.22 (95% CI: 0.12–0.40); and pooled DOR was 14.99 (95% CI: 7.48, 30.04) (Fig. 3C; Table 2).

In studies on organ failure, it was shown that the sensitivity of CTSI evaluation time ≤ 48 h ($n = 2$) was 0.90 (95% CI:0.82–0.95), with a specificity of 0.52 (95% CI:0.48–0.57); the sensitivity of evaluation time > 72 h ($n = 11$) was 0.86 (95% CI:0.63–0.95) and specificity was 0.73 (95% CI:0.60–0.82); and the sensitivity for assessment time ≤ 72 h ($n = 5$) was 0.74 (95% CI:0.56–0.86) and specificity was 0.84 (95% CI:0.61–0.95). Organ failure outcomes were not analysed for the ≤ 24 h subgroup because no studies reported CTSI assessments at this early time point. Among the results, evaluation ≤ 48 h had the highest sensitivity for predicting organ failure, and evaluation ≤ 72 h had the highest specificity (Table 2). There was a statistically significant difference in the prediction of organ failure at ≤ 48 h compared with ≤ 72 h and > 72 h ($p < 0.05$).

Predictive performance for pancreatic necrosis

With 7 studies providing data on the prediction of pancreatic necrosis by CTSI, the pooled sensitivity was 0.92 (95% CI: 0.80–0.97), the pooled specificity was 0.84 (95% CI: 0.70–0.92), the pooled AUC was 0.94 (95% CI: 0.92–0.96); pooled PLR was 5.57 (95% CI: 3.04–10.20); pooled NLR was 0.10 (95% CI: 0.04–0.24); and pooled DOR was 56.75 (95% CI: 23.79, 135.37) (Fig. 3D; Table 2).

In the study on pancreatic necrosis, it was shown that the sensitivity of CTSI evaluation time > 72 h ($n = 5$) was

Fig. 2 Risk of bias graph (**A**); Risk of bias summary (**B**)

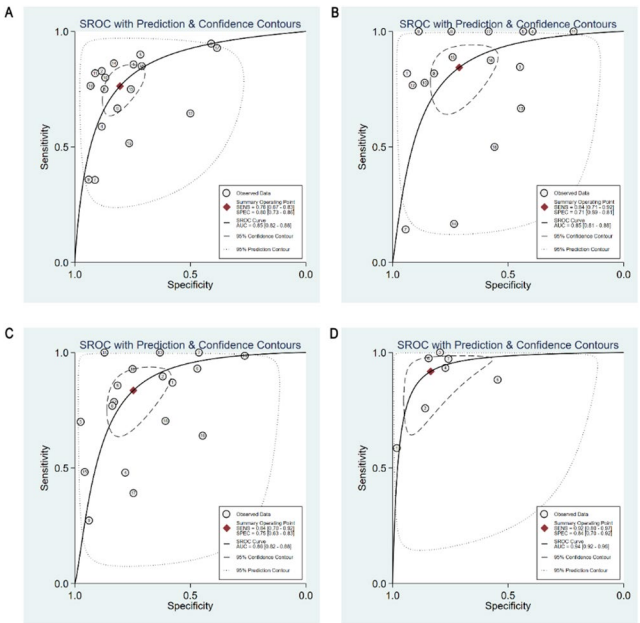

Fig. 3 SROC curve of the CTSI for predicting severity (**A**), mortality (**B**), organ failure (**C**), and pancreatic necrosis (**D**) in AP

0.93 (95% CI:0.81–0.98) with a specificity of 0.78 (95% 0.68–0.86); the sensitivity of evaluation time≤48 h ($n=2$) was 0.80 (95% CI:0.68–0.89) and specificity was 0.90 (95% CI:0.87–0.93). Similarly, no pancreatic necrosis data were available for ≤24 h analyses since no studies performed CTSI evaluations at this timepoint. That is, evaluation time>72 h had the highest sensitivity for predicting pancreatic necrosis, and evaluation time≤48 h had the highest specificity (Table 2). There was a statistically significant difference in the prediction of pancreatic necrosis at ≤48 h compared to >72 h ($p<0.05$).

Subgroup analysis of CTSI versus MCTSI

Subgroup analyses comparing the diagnostic performance of CTSI and MCTSI across outcomes are presented in Table 3. For severity prediction, MCTSI demonstrated higher sensitivity (0.82 vs. 0.74) and diagnostic odds ratio (DOR 17.62 vs. 12.31) than CTSI, with comparable pooled AUC (0.87 vs. 0.84). Mortality prediction revealed superior sensitivity for MCTSI (0.87 vs. 0.82) and significantly higher AUC (0.89 vs. 0.82). Notably, MCTSI achieved higher specificity in pancreatic necrosis detection (0.91 vs. 0.78) despite lower sensitivity (0.73 vs. 0.97), reflecting its omission of contrast-dependent necrosis grading. For organ failure, MCTSI maintained robust early sensitivity (≤48 h: 0.90) comparable to CTSI. Critically, MCTSI showed non-inferior pooled accuracy for key outcomes (severity AUC 0.87, mortality AUC 0.89), while enabling reduced contrast exposure—a finding supporting its safety advantages in clinical implementation.

Meta-regression

Univariate meta-regression analyses were conducted to investigate potential sources of heterogeneity in severity, mortality, and organ failure. For severity, country could be a significant influence on the presence of heterogeneity in sensitivity ($p<0.05$), while age could be a potential source of heterogeneity in specificity ($p<0.05$) (Fig. 4A); for mortality and organ failure, year, country, sample size, evaluation, and age were not found to be possible factors contributing to heterogeneity (Fig. 4B, C).

Sensitivity analysis and publication bias

After conducting separate sensitivity analyses of the CTSI to predict AP severity, mortality, organ failure, and pancreatic necrosis (Fig. S1−S4), the results showed that the sensitivities were relatively stable. No significant changes were observed when each study was excluded individually; the results indicated that sensitivities remained relatively stable. This suggests that any particular study did not significantly affect the overall results.

Deeks' funnel plot asymmetry test was used to assess the publication bias of more than 10 studies. The results of the Deeks' funnel plot asymmetry test were not significant for studies on AP severity ($p=0.397$) and organ failure ($p=0.757$), indicating no significant publication bias ($p>0.05$). For the study on mortality, there was a possibility of publication bias ($p=0.009$, $p<0.05$) (Fig. S5).

Table 2 Results of univariate meta–regression and subgroup analyses of CTSI predicting severity, mortality, organ failure, and pancreatic necrosis

Outcome	Evaluation time	Studies (*n*)	Sensitivity(95%CI)	Specificity (95%CI)	PLR(95%CI)	NLR(95%CI)	DOR(95%CI)
Severity	≤24 h	3	0.81(0.72–0.88)	0.76(0.70–0.81)	2.86(1.32–6.19)	0.27(0.17–0.43)	13.98(6.00–32.58)
	≤48 h	4	0.82(0.69–0.90)	0.78(0.68–0.85)	3.67(2.76–4.87)	0.23(0.14–0.39)	15.90(9.90–25.53)
	≤72 h	4	0.56(0.48–0.64)	0.86(0.84–0.88)	4.78(1.79–12.78)	0.31(0.12–0.80)	17.33(7.68–39.10)
	>72 h	7	0.69(0.56–0.79)	0.83(0.73–0.90)	4.09(2.35–7.13)	0.38(0.25–0.56)	10.88(4.43–26.73)
	Combined	18	0.76(0.67–0.83)	0.80(0.73–0.86)	3.90(2.88–5.28)	0.29(0.22–0.40)	13.26(8.60–20.46)
Mortality	≤24 h	3	0.81(0.63–0.93)	0.66(0.60–0.71)	2.22(1.32–3.76)	0.35(0.09–1.40)	6.15(1.21–31.25)
	≤48 h	2	0.85(0.70–0.94)	0.77(0.73–0.82)	5.37(0.44–66.15)	0.19(0.09–0.38)	58.63(21.90–156.92)
	≤72 h	4	0.90(0.72–0.97)	0.56(0.34–0.76)	2.05(1.28–3.28)	0.17(0.06–0.51)	11.83(3.46–40.38)
	>72 h	9	0.79(0.48–0.94)	0.74(0.72–0.77)	3.40(1.90–6.10)	0.27(0.09–0.78)	12.50(3.27–47.78)
	Combined	18	0.84(0.71–0.92)	0.71(0.59–0.81)	2.93(2.06–4.16)	0.22(0.12–0.42)	13.34(6.07–29.31)
Organ failure	≤48 h	2	0.90(0.82–0.95)	0.52(0.48–0.57)	1.89(1.61–2.20)	0.20(0.11–0.37)	10.00(4.89–20.46)
	≤72 h	5	0.74(0.56–0.86)	0.84(0.61–0.95)	4.74(1.85–12.10)	0.31(0.19–0.51)	15.23(5.69–40.73)
	>72 h	11	0.86(0.63–0.95)	0.73(0.60–0.82)	3.12(2.15–4.53)	0.20(0.07–0.55)	15.64(4.98–49.11)
	Combined	18	0.84(0.70–0.92)	0.75(0.63–0.83)	3.30(2.30–4.72)	0.22(0.12–0.40)	14.99(7.48–30.04)
Pancreatic necrosis	≤48 h	2	0.80(0.68–0.89)	0.90(0.87–0.93)	10.90(1.30–91.79)	0.14(0.01–3.05)	83.60(30.71–227.58)
	>72 h	5	0.93(0.81–0.98)	0.78(0.68–0.86)	4.32(2.82–6.64)	0.08(0.03–0.26)	51.24(13.75–190.93)
	Combined	7	0.92(0.80–0.97)	0.84(0.70–0.92)	5.57(3.04–10.20)	0.10(0.04–0.24)	56.75(23.79–135.37)

CTSI computed tomography severity index, *PLR* Positive Likelihood Ratio, *NLR* Negative Likelihood Ratio, *DOR* Diagnostic Odds Ratio

Table 3 Sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, diagnostic odds ratio, and AUC of CTSI or MCTSI for predicting severity, mortality, organ failure, and pancreatic necrosis

Outcome	Evaluation type	Studies (*n*)	Sensitivity(95%CI)	Specificity (95%CI)	PLR(95%CI)	NLR(95%CI)	DOR(95%CI)	AUC(95%CI)
Severity	CTSI	12	0.74 (0.65–0.81)	0.81 (0.72–0.88)	3.91 (2.63–5.83)	0.32 (0.23–0.43)	12.31 (6.83–22.19)	0.84 (0.80–0.87)
	MCTSI	6	0.82 (0.60–0.93)	0.80 (0.64–0.90)	4.04 (2.51–6.48)	0.23 (0.11–0.48)	17.62 (10.62–29.24)	0.87 (0.84–0.90)
Mortality	CTSI	12	0.82 (0.59–0.94)	0.70 (0.54–0.82)	2.72 (1.76–4.21)	0.25 (0.10–0.63)	10.68 (3.58–31.85)	0.82 (0.78–0.85)
	MCTSI	6	0.87 (0.75–0.93)	0.74 (0.56–0.86)	3.30 (1.86–5.87)	0.18 (0.09–0.35)	18.34 (7.01–47.97)	0.89 (0.86–0.91)
Organ failure	CTSI	11	0.78 (0.57–0.90)	0.78 (0.62–0.88)	3.53 (2.10–5.94)	0.29 (0.15–0.56)	12.28 (5.10–29.59.10.59)	0.85 (0.81–0.88)
	MCTSI	7	0.90 (0.73–0.97)	0.69 (0.53–0.81)	2.86 (1.89–4.33)	0.15 (0.06–0.40)	19.24 (6.62–55.90)	0.85 (0.82–0.88)
Pancreatic necrosis	CTSI	4	0.97 (0.91–0.99)	0.78 (0.74–0.82)	3.86 (2.27–6.54)	0.06 (0.02–0.22)	65.42 (12.24–349.66.24.66)	
	MCTSI	3	0.73 (0.61–0.82)	0.91 (0.89–0.94)	8.23 (3.13–21.69)	0.32 (0.18–0.56)	37.30 (13.90–100.09.90.09)	

**CTSI* computed tomography severity index, *MCTSI* modified computed tomography severity index, *PLR* positive likelihood ratio, *NLR* negative likelihood ratio, *DOR* diagnostic odds ratio, *AUC* area under the curve

Discussion

This meta-analysis evaluated the predictive performance of the CTSI at different assessment times for severity and prognostic outcomes in AP. Our findings demonstrate that CTSI exhibits time-dependent variations in its predictive efficacy across distinct clinical endpoints, underscoring the importance of optimising imaging timing for specific outcomes. The analysis of 28 studies ($n=5,419$) revealed that CTSI assessment at ≤48 h showed optimal sensitivity (0.84) and specificity (0.79) for predicting AP severity, while assessment at ≤72 h performed best for mortality (sensitivity: 0.89; specificity: 0.65). Organ failure prediction peaked at ≤48 h (sensitivity: 0.90), whereas pancreatic necrosis was most accurately identified at >72 h (sensitivity: 0.91; specificity: 0.82). These results highlight the dynamic relationship between imaging timing and pathophysiological progression in AP.

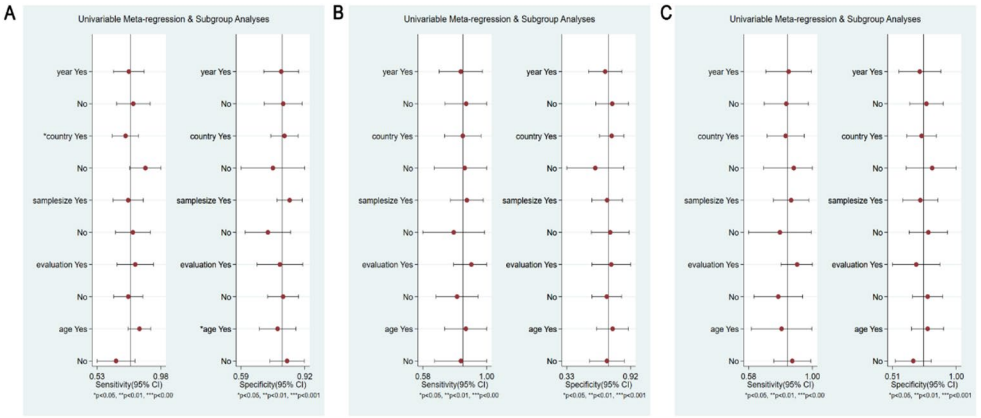

Fig. 4 Univariate meta-regression and subgroup analyses of CTSI predicting severity (**A**), mortality (**B**), and organ failure (**C**) in AP

CTSI at ≤ 48 h for severity prediction

The optimal performance of CTSI at ≤48 h for severity prediction aligns with critical pathophysiological transitions in AP. During this window, SIRS typically peaks, driving early organ failure [4]. While pancreatic necrosis may not yet be radiologically evident [52], CTSI effectively captures peripancreatic complications detectable within 48 h [53], including acute fluid collections (0–30 Hounsfield units) that correlate with evolving SAP risk [54]. This temporal association explains CTSI's balanced sensitivity (0.84) and specificity (0.79) at ≤48 h, outperforming ≤24-hour assessments (sensitivity 0.81, specificity 0.76) where inflammatory evolution remains incomplete. Notably, CTSI's superiority over clinical scores, such as APACHE-II (AUC 0.78 vs. CTSI 0.84) [55], stems from its direct visualisation of local complications. However, its predictive capacity remains secondary to biochemical markers of persistent SIRS at 48 h, which guide current severity classification [11]. This discrepancy may be due to: (1) CTSI's inability to quantify dynamic cytokine cascades driving organ failure [4]; (2) up to 25% of SAP cases lacking radiologically detectable necrosis [55]; and (3) the poor correlation between necrosis development and organ failure timing [55]. Comparative analyses reveal CTSI's practical advantage over Ranson's criteria (AUC 0.94) [14], which requires non-routine laboratory parameters. While no single tool perfectly predicts SAP, CTSI's operational feasibility and spatial resolution make it valuable for early triage when combined with inflammatory biomarkers (e.g., CRP, interleukin-6) [46]. Nevertheless, its reduced sensitivity at ≤72 h (0.56) suggests limited utility beyond the 48-hour window for severity assessment, likely because necrosis manifests later [52].

Mortality prediction: ≤72 h as the optimal window

Our meta-analysis identified a ≤72-hour CTSI assessment as the most sensitive time window (pooled sensitivity 0.89) for mortality prediction in AP. This temporal advantage likely stems from CTSI's capacity to capture two critical mortality determinants: [1] early pancreatic necrosis patterns detectable by 72-hour CECT (sensitivity approaching 100% for necrosis detection at 4 days [56]), and [2] evolving organ failure dynamics particularly persistent organ failure (>48 h duration) which increases mortality risk by 3.8-fold [57]. Notably, the observed decline in specificity (0.65 vs. 0.77 at ≤48 h) suggests compromised differentiation power in later phases, when secondary complications emerge. This finding aligns with SAP management guidelines, which emphasise 72-hour reassessment windows [58], particularly for patients with persistent organ failure and infected necrosis, a subgroup exhibiting mortality rates of 28–39% in our included studies. Cumulative evidence indicates that CTSI ranks lower in predictive accuracy for AP-related mortality compared to Ranson's criteria, APACHE-II, BISAP, and SOFA [41, 46, 58, 59]. Recent proposals for multidimensional risk stratification [14] that incorporate genetic variants (e.g., PRSS1 mutations), metabolic factors (BMI>30), and serial biomarker panels (IL-6, D-dimer) may enhance predictive models. Machine learning approaches integrating CTSI with these variables could potentially overcome the current specificity ceiling observed in our analysis.

Organ failure: early imaging imperative

The meta-analysis revealed that CTSI assessed within ≤48 h demonstrated superior sensitivity (0.90) but limited specificity (0.52) for predicting organ failure in AP patients. This temporal pattern aligns with the pathophysiological progression of AP, in which early pancreatic necrosis and peripancreatic inflammation (detected by CT) often precede systemic manifestations of organ dysfunction [58, 59]. The observed sensitivity declines beyond 72 h (0.74), with improved specificity (0.69), likely reflecting the delayed onset of definitive systemic complications that correlate better with imaging findings. Current guidelines [60] emphasize avoiding premature severity stratification within

the first 48 h, a recommendation supported by our findings regarding CTSI's evolving predictive characteristics. While CTSI outperforms clinical scoring systems like APACHE-II (cumbersome 12-parameter protocol) [61] and Ranson criteria (limited specificity for persistent organ failure) [60] in early detection, its moderate specificity underscores the need for serial assessments. Notably, the strong association between pancreatic necrosis volume (threshold > 100 mL) and organ failure [24, 31] underscores CTSI's unique advantage in quantifying the anatomical correlates of dysfunction. This temporal precision makes CTSI particularly valuable for initial risk stratification, as early reversal of organ failure within 48 h correlates with an improved prognosis [57]. However, clinicians should interpret ≤ 48-hour CTSI results cautiously in cases of isolated peripancreatic necrosis, as these may disproportionately influence scores despite uncertain clinical significance [62].

Pancreatic necrosis: delayed imaging advantage

Our meta-analysis reveals CTSI achieves peak diagnostic accuracy for pancreatic necrosis (sensitivity 0.91, specificity 0.82) when assessed > 72 h after symptom onset, substantially outperforming evaluations at ≤ 48 h (sensitivity 0.80, specificity 0.90). This temporal divergence aligns with the pathophysiological trajectory of necrotic evolution. Pancreatic necrosis typically becomes radiographically discernible around day 4 post-onset, with progression continuing over the subsequent 48–72 h as mixed parenchymal-peripancreatic necrosis [55, 63]. Early CT imaging (≤ 3–4 days) demonstrates reduced accuracy in delineating necrosis, while delayed scans (> 5 days) are critical for assessing necrotic demarcation and fluid collections in non-responsive patients [64]. These findings corroborate guideline-endorsed protocols [11] that recommend an optimal initial CT timing of 72–96 h. Our > 72-hour peak accuracy finding robustly supports the guideline's 72–96 h window [11]. Crucially, necrosis becomes radiographically discernible around day 4 (~ 96 h), suggesting imaging closer to 96 h within this window maximises accuracy. This aligns with delayed scanning (> 5 d) being critical for non-responsive patients [64], thereby specifically endorsing the 96-hour endpoint for definitive necrosis assessment in this cohort. CTSI appeared to maintain superior accuracy in predicting pancreatic necrosis compared to other clinical scoring systems. Papachristou et al. [14]demonstrated that the AUC of the CTSI (0.98; 95% CI: 0.94–1.00) was significantly superior to those of the Ranson (AUC = 0.85), BISAP (AUC = 0.78), and APACHE-II (AUC = 0.72) scores. However, CTSI's inability to quantify systemic inflammation necessitates complementary early clinical assessments (e.g., Ranson's or BISAP) for comprehensive risk stratification [61]. While APACHE-II and SOFA prioritise critical care monitoring over necrosis prediction [46, 61, 65], CTSI remains the imaging gold standard within its validated temporal window [14]. This evidence supports a staged diagnostic approach: delayed CT (> 72 h) optimises confirmation of necrosis, while early clinical scores guide initial management. Future studies should investigate hybrid models integrating temporal CTSI data with serial biomarker profiles to enhance phase-specific predictive accuracy.

Addressing safety concerns with MCTSI

While the timing of CTSI assessment is critical for accuracy, its reliance on CECT protocols raises concomitant concerns about cumulative radiation exposure and contrast-induced nephropathy, particularly in patients requiring repeated imaging [66]. These risks underscore the need for alternative scoring approaches that maintain diagnostic efficacy while mitigating safety issues. In this context, MCTSI, which omits necrosis grading and permits non-contrast or abbreviated CECT protocols [67], emerges as a promising strategy. Our meta-analysis included 10 studies using MCTSI (representing 35.7% of included studies), and its pooled diagnostic performance for severity (AUC 0.87, 95% CI 0.84–0.90) and mortality prediction (AUC 0.89, 95% CI 0.86–0.92) was comparable to that of CTSI. Crucially, by eliminating the need for contrast-dependent necrosis assessment, MCTSI reduces the requirement for repeated CECT scans, thereby lowering cumulative radiation dose and minimising the risk of contrast-induced nephropathy, especially in high-risk populations such as those with renal impairment [68]. This advantage is particularly relevant for early assessment windows (e.g., ≤ 48 h), where MCTSI's sensitivity for organ failure (0.90) remains robust. *Future studies should standardise MCTSI protocols to further validate its role in risk reduction.*

Limitations and future directions

A key strength of this meta-analysis is its comprehensive quantitative synthesis of evidence across multiple CTSI assessment time windows for diverse AP outcomes, providing novel insights into time-dependent predictive performance. However, several limitations warrant consideration. First, the temporal classification of CTSI assessments (e.g., ≤ 24 h, ≤ 48 h) inherently introduces temporal heterogeneity, as studies within the same category may differ in the exact timing of imaging (e.g., scans labeled "≤ 72 h" could span 48–72 h). This imprecise classification may dilute time-specific accuracy estimates, particularly for outcomes such as pancreatic necrosis, where delayed imaging (> 72 h)

demonstrated superior predictive performance but limits early clinical utility. Second, while CTSI effectively stratifies anatomical severity, its moderate specificity for predicting organ failure (especially at ≤48 h) highlights challenges in distinguishing transient imaging abnormalities from those that progress to clinical dysfunction. This anatomical-physiological discordance is further evident in mortality prediction, where CTSI's omission of dynamic physiological parameters (e.g., hemodynamic status) poorly aligns with modern multiparameter prognostic frameworks. Third, despite standardized time categories, inter-study variations in imaging protocols (e.g., contrast timing, slice thickness) and CTSI interpretation thresholds may introduce unmeasured heterogeneity. Finally, the analysis did not evaluate synergistic use of CTSI with biochemical markers (e.g., CRP, procalcitonin) or clinical scores (e.g., BISAP), which could enhance early risk stratification.

Future research should prioritize three directions: (1) Prospectively validating standardized imaging timepoints (e.g., 24 h±4 h, 72 h±12 h) to minimize temporal misclassification and establish clinically actionable accuracy thresholds, requiring dedicated studies to explore optimal timing further; (2) Developing integrated models combining CTSI with physiologic parameters (e.g., SOFA scores) and biomarkers to improve specificity for organ failure and mortality; (3) Investigating dynamic CTSI trajectories via serial imaging to better correlate with evolving clinical courses. Additionally, comparative studies evaluating CTSI against emerging imaging biomarkers or revised scoring systems are warranted.

Conclusion

CTSI exhibits distinct temporal predictive patterns: it demonstrates optimal sensitivity for severity assessment and organ failure prediction when performed early, within 48 h of admission; scans within 72 h best predict mortality, while delayed imaging (>72 h) maximizes pancreatic necrosis accuracy at the expense of clinical timeliness. Future research must standardize imaging timepoints, integrate CTSI with physiological biomarkers, and develop dynamic assessment models to resolve discrepancies in anatomical-pathophysiological prognostic factors in AP.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00261-025-05296-x.

Author contributions Shan Huang: Conceptualization, Methodology, Software, Writing- Original draft, Data Curation. Xinwei Liu: Software, Data Curation, Writing- Reviewing and Editing. Longyan Zhu: Formal analysis, Visualization, Investigation. Kun Ai: Validation, Resources, Supervision. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript. Shan Huang and Xinwei Liu have contributed equally to this work and share first authorship.

Data availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

References

1. Boxhoorn L, Voermans RP, Bouwense SA, Bruno MJ, Verdonk RC, Boermeester MA, van Santvoort HC, Besselink MG. Acute pancreatitis. *The Lancet*. 2020;396(10252):726–734.
2. Iannuzzi JP, King JA, Leong JH, Quan J, Windsor JW, Tanyingoh D, Coward S, Forbes N, Heitman SJ, Shaheen AA, Swain M, Buie M, Underwood FE, Kaplan GG. Global Incidence of Acute Pancreatitis Is Increasing Over Time: A Systematic Review and Meta-Analysis. *Gastroenterology*. 2022;162(1):122–134.
3. Li CL, Jiang M, Pan CQ, Li J, Xu LG. The global, regional, and national burden of acute pancreatitis in 204 countries and territories, 1990–2019. *BMC Gastroenterol*. 2021;21(1):332.
4. Banks PA, Bollen TL, Dervenis C, Gooszen HG, Johnson CD, Sarr MG, Tsiotos GG, Vege SS. Classification of acute pancreatitis–2012: revision of the Atlanta classification and definitions by international consensus. *Gut*. 2013;62(1):102–111.
5. Petrov MS, Shanbhag S, Chakraborty M, Phillips AR, Windsor JA. Organ failure and infection of pancreatic necrosis as determinants of mortality in patients with acute pancreatitis. *Gastroenterology*. 2010;139(3):813–820.
6. Mofidi R, Duff MD, Wigmore SJ, Madhavan KK, Garden OJ, Parks RW. Association between early systemic inflammatory response, severity of multiorgan dysfunction and death in acute pancreatitis. *Br J Surg*. 2006;93(6):738–744.
7. Xiao AY, Tan ML, Wu LM, Asrani VM, Windsor JA, Yadav D, Petrov MS. Global incidence and mortality of pancreatic diseases: a systematic review, meta-analysis, and meta-regression of population-based cohort studies. *Lancet Gastroenterol Hepatol*. 2016;1(1):45–55.
8. Podda M, Pellino G, Di Saverio S, Coccolini F, Pacella D, Cioffi SPB, Virdis F, Balla A, Ielpo B, Pata F, Poillucci G, Ortenzi M, Damaskos D, De Simone B, Sartelli M, Leppaniemi A, Jayant K, Catena F, Giuliani A, Di Martino M, Pisanu A. Infected pancreatic necrosis: outcomes and clinical predictors of mortality. A post hoc analysis of the MANCTRA-1 international study. *Updates Surg*. 2023;75(3):493–522.
9. Schepers NJ, Bakker OJ, Besselink MG, Ahmed Ali U, Bollen TL, Gooszen HG, van Santvoort HC, Bruno MJ. Impact of characteristics of organ failure and infected necrosis on mortality in necrotising pancreatitis. *Gut*. 2019;68(6):1044–1051.
10. Arvanitakis M, Dumonceau JM, Albert J, Badaoui A, Bali MA, Barthet M, Besselink M, Deviere J, Oliveira Ferreira A, Gyökeres T, Hritz I, Hucl T, Milashka M, Papanikolaou IS, Poley JW, Seewald S, Vanbiervliet G, van Lienden K, van Santvoort H, Voermans R, Delhaye M, van Hooft J. Endoscopic management of acute necrotizing pancreatitis: European Society of Gastrointestinal Endoscopy (ESGE) evidence-based multidisciplinary guidelines. *Endoscopy*. 2018;50(5):524–546.
11. IAP/APA evidence-based guidelines for the management of acute pancreatitis. *Pancreatology*. 2013;13(4 Suppl 2):e1-15.

12. Wu BU, Johannes RS, Sun X, Tabak Y, Conwell DL, Banks PA. The early prediction of mortality in acute pancreatitis: a large population-based study. *Gut*. 2008;57(12):1698–1703.
13. Venkatesh NR, Vijayakumar C, Balasubramaniyan G, Chinnakkulam Kandhasamy S, Sundaramurthi S, G SS, Srinivasan K. Comparison of Different Scoring Systems in Predicting the Severity of Acute Pancreatitis: A Prospective Observational Study. *Cureus*. 2020;12(2):e6943.
14. Papachristou GI, Muddana V, Yadav D, O'Connell M, Sanders MK, Slivka A, Whitcomb DC (2010) Comparison of BISAP, Ranson's, APACHE-II, and CTSI scores in predicting organ failure, complications, and mortality in acute pancreatitis. Am J Gastroenterol 105(2):435–441
15. Ranson JH, Rifkind KM, Roses DF, Fink SD, Eng K, Spencer FC. Prognostic signs and the role of operative management in acute pancreatitis. *Surg Gynecol Obstet*. 1974;139(1):69–81.
16. Alberti P, Pando E, Mata R, Vidal L, Roson N, Mast R, Armario D, Merino X, Dopazo C, Blanco L, Caralt M, Gomez C, Balsells J, Charco R. Evaluation of the modified computed tomography severity index (MCTSI) and computed tomography severity index (CTSI) in predicting severity and clinical outcomes in acute pancreatitis. *J Dig Dis*. 2021;22(1):41–48.
17. Balthazar EJ, Robinson DL, Megibow AJ, Ranson JH. Acute pancreatitis: value of CT in establishing prognosis. *Radiology*. 1990;174(2):331–336.
18. Aycock RD, Westafer LM, Boxen JL, Majlesi N, Schoenfeld EM, Bannuru RR. Acute Kidney Injury After Computed Tomography: A Meta-analysis. *Ann Emerg Med*. 2018;71(1):44–53.e44.
19. Wrazidlo R, Walder L, Estler A, Gutjahr R, Schmidt B, Faby S, Fritz J, Nikolaou K, Horger M, Hagen F. Radiation Dose Reduction in Contrast-Enhanced Abdominal CT: Comparison of Photon-Counting Detector CT with 2nd Generation Dual-Source Dual-Energy CT in an oncologic cohort. *Acad Radiol*. 2023;30(5):855–862.
20. Balthazar EJ. Complications of acute pancreatitis: clinical and CT evaluation. *Radiol Clin North Am*. 2002;40(6):1211–1227.
21. Thoeni RF. The revised Atlanta classification of acute pancreatitis: its importance for the radiologist and its effect on treatment. *Radiology*. 2012;262(3):751–764.
22. Hagjer S, Kumar N. Evaluation of the BISAP scoring system in prognostication of acute pancreatitis - A prospective observational study. *Int J Surg*. 2018;54(Pt A):76–81.
23. Banks PA, Freeman ML. Practice guidelines in acute pancreatitis. *Am J Gastroenterol*. 2006;101(10):2379–2400.
24. Porões F, Vietti Violi N, Piazza G, Uldry E, Lázaro-Fontanet E, Gaspar-Figueiredo S, Hilbert T, Ledoux JB, Denys A, Schütz F, Schmidt S. Quantitative T2 Mapping of Acute Pancreatitis. *J Magn Reson Imaging*. 2024;60(6):2683–2691.
25. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, Cohen JF, Deeks JJ, Gatsonis C, Hooft L, Hunt HA, Hyde CJ, Korevaar DA, Leeflang MMG, Macaskill P, Reitsma JB, Rodin R, Rutjes AWS, Salameh JP, Stevens A, Takwoingi Y, Tonelli M, Weeks L, Whiting P, Willis BH. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. *Jama*. 2018;319(4):388–396.
26. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529–536.
27. Meyrignac O, Lagarde S, Bournet B, Mokrane FZ, Buscail L, Rousseau H, Otal P. Acute Pancreatitis: Extrapancreatic Necrosis Volume as Early Predictor of Severity. *Radiology*. 2015;276(1):119–128.
28. Fatma Tortum ET, Aydın Fahri, Özdal Emine, Tatlısu Kübra (2023) The relationship of biochemical parameters and radiological parameters in the evaluation of the clinical severity of acute pancreatitis in the emergency department – a retrospective analysis. European Journal of Clinical and Experimental Medicine 21(2):277–282
29. Luo X, Wang J, Wu Q, Peng P, Liao G, Liang C, Yang H, Huang J, Qin M. A modified Ranson score to predict disease severity, organ failure, pancreatic necrosis, and pancreatic infection in patients with acute pancreatitis. *Front Med (Lausanne)*. 2023;10:1145471.
30. Zhao Y, Xia W, Lu Y, Chen W, Zhao Y, Zhuang Y. Predictive value of the C-reactive protein/albumin ratio in severity and prognosis of acute pancreatitis. *Front Surg*. 2022;9:1026604.
31. Liu N, He J, Hu X, Xu SF, Su W, Luo JF, Wang QF, Guo F. Acute necrotising pancreatitis: measurements of necrosis volume and mean CT attenuation help early prediction of organ failure and need for intervention. *Eur Radiol*. 2021;31(10):7705–7714.
32. Peng R, Zhang L, Zhang ZM, Wang ZQ, Liu GY, Zhang XM. Chest computed tomography semi-quantitative pleural effusion and pulmonary consolidation are early predictors of acute pancreatitis severity. *Quant Imaging Med Surg*. 2020;10(2):451–463.
33. Gezer NS, Bengi G, Baran A, Erkmen PE, Topalak ÖS, Altay C, Dicle O (1992) Comparison of radiological scoring systems, clinical scores, neutrophil-lymphocyte ratio and serum C-reactive protein level for severity and mortality in acute pancreatitis. Rev Assoc Med Bras 66(6):762–770
34. Tharun GanapaThy ChiTrambalam Sr, jeyakumar Sundaraj, koShy maThew paniCker, ramyaSree paladuGu. A longitudinal study on comparison of haematocrit, glasgow score and computed tomography abdomen as prognostic markers in assessment of severity in acute pancreatitis. *Journal of Clinical and Diagnostic Research*. 2020;14(12):1–4.
35. Chatterjee R, Parab N, Sajjan B, Nagar VS. Comparison of Acute Physiology and Chronic Health Evaluation II, Modified Computed Tomography Severity Index, and Bedside Index for Severity in Acute Pancreatitis Score in Predicting the Severity of Acute Pancreatitis. *Indian J Crit Care Med*. 2020;24(2):99–103.
36. Çakar İ, Keven A, Eseroğlu E, Çubuk SM. Role of extrapancreatic necrosis volume in determining early prognosis in patients with acute pancreatitis. *Abdom Radiol (NY)*. 2020;45(5):1507–1516.
37. Jain D, Bhaduri G, Jain P. DIFFERENT SCORING SYSTEMS IN ACUTE ALCOHOLIC PANCREATITIS: WHICH ONE TO FOLLOW? AN ONGOING DILEMA. *Arq Gastroenterol*. 2019;56(3):280–285.
38. Vaidya Y, R. Vaithianathan, and R. Manickam. Comparative Evaluation of the BISAP Score with CT Severity Index in Predicting the Severity of Acute Pancreatitis. *Indian Journal of Surgery*. 2018;80(4):353–358.
39. Harshit Kumar A, Singh Griwan M. A comparison of APACHE II, BISAP, Ranson's score and modified CTSI in predicting the severity of acute pancreatitis based on the 2012 revised Atlanta Classification. *Gastroenterol Rep (Oxf)*. 2018;6(2):127–131.
40. Sahu B, Abbey P, Anand R, Kumar A, Tomer S, Malik E. Severity assessment of acute pancreatitis using CT severity index and modified CT severity index: Correlation with clinical outcomes and severity grading as per the Revised Atlanta Classification. *Indian J Radiol Imaging*. 2017;27(2):152–160.
41. Yang L, Liu J, Xing Y, Du L, Chen J, Liu X, Hao J. Comparison of BISAP, Ranson, MCTSI, and APACHE II in Predicting Severity and Prognoses of Hyperlipidemic Acute Pancreatitis in Chinese Patients. *Gastroenterol Res Pract*. 2016;2016:1834256.
42. Sharma V, Rana SS, Sharma RK, Kang M, Gupta R, Bhasin DK. A study of radiological scoring system evaluating extrapancreatic inflammation with conventional radiological and clinical scores in predicting outcomes in acute pancreatitis. *Ann Gastroenterol*. 2015;28(3):399–404.
43. Qiu L, Sun RQ, Jia RR, Ma XY, Cheng L, Tang MC, Zhao Y. Comparison of Existing Clinical Scoring Systems in Predicting

Severity and Prognoses of Hyperlipidemic Acute Pancreatitis in Chinese Patients: A Retrospective Study. *Medicine (Baltimore)*. 2015;94(23):e957.
44. Xu H, Ebner L, Jiang S, Wu Y, Christe A, Zhang S, Zhang X, Luo Z, Tian F. Retrocrural space involvement on computed tomography as a predictor of mortality and disease severity in acute pancreatitis. *PLoS One*. 2014;9(9):e107378.
45. Park JY, Jeon TJ, Ha TH, Hwang JT, Sinn DH, Oh TH, Shin WC, Choi WC. Bedside index for severity in acute pancreatitis: comparison with other scoring systems in predicting severity and organ failure. *Hepatobiliary Pancreat Dis Int*. 2013;12(6):645–650.
46. Khanna AK, Meher S, Prakash S, Tiwary SK, Singh U, Srivastava A, Dixit VK (2013) Comparison of Ranson, Glasgow, MOSS, SIRS, BISAP, APACHE-II, CTSI Scores, IL-6, CRP, and Procalcitonin in Predicting Severity, Organ Failure, Pancreatic Necrosis, and Mortality in Acute Pancreatitis. HPB Surg 2013:367581
47. Chen L, Lu G, Zhou Q, Zhan Q. Evaluation of the BISAP score in predicting severity and prognoses of acute pancreatitis in Chinese patients. *Int Surg*. 2013;98(1):6–12.
48. Bollen TL, Singh VK, Maurer R, Repas K, van Es HW, Banks PA, Mortele KJ. A comparative evaluation of radiologic and clinical scoring systems in the early prediction of severity in acute pancreatitis. *Am J Gastroenterol*. 2012;107(4):612–619.
49. Imamura Y, Hirota M, Ida S, Hayashi N, Watanabe M, Takamori H, Awai K, Baba H. Significance of renal rim grade on computed tomography in severity evaluation of acute pancreatitis. *Pancreas*. 2010;39(1):41–46.
50. De Waele JJ, Delrue L, Hoste EA, De Vos M, Duyck P, Colardyn FA. Extrapancreatic inflammation on abdominal computed tomography as an early predictor of disease severity in acute pancreatitis: evaluation of a new scoring system. *Pancreas*. 2007;34(2):185–190.
51. Chatzicostas C, Roussomoustakaki M, Vardas E, Romanos J, Kouroumalis EA. Balthazar computed tomography severity index is superior to Ranson criteria and APACHE II and III scoring systems in predicting acute pancreatitis outcome. *J Clin Gastroenterol*. 2003;36(3):253–260.
52. Buter A, Imrie CW, Carter CR, Evans S, McKay CJ. Dynamic nature of early organ dysfunction determines outcome in acute pancreatitis. *Br J Surg*. 2002;89(3):298–302.
53. Siegelman SS, Copeland BE, Saba GP, Cameron JL, Sanders RC, Zerhouni EA. CT of fluid collections associated with pancreatitis. *AJR Am J Roentgenol*. 1980;134(6):1121–1132.
54. Zaheer A, Singh VK, Qureshi RO, Fishman EK. The revised Atlanta classification for acute pancreatitis: updates in imaging terminology and guidelines. *Abdom Imaging*. 2013;38(1):125–136.
55. Tenner S. Initial management of acute pancreatitis: critical issues during the first 72 hours. *Am J Gastroenterol*. 2004;99(12):2489–2494.
56. Balthazar EJ. Acute pancreatitis: assessment of severity with clinical and CT evaluation. *Radiology*. 2002;223(3):603–613.
57. Johnson CD, Abu-Hilal M. Persistent organ failure during the first week as a marker of fatal outcome in acute pancreatitis. *Gut*. 2004;53(9):1340–1344.
58. Sharma M, Banerjee D, Garg PK. Characterization of newer subgroups of fulminant and subfulminant pancreatitis associated with a high early mortality. *Am J Gastroenterol*. 2007;102(12):2688–2695.
59. Yadav J, Yadav SK, Kumar S, Baxla RG, Sinha DK, Bodra P, Besra RC, Baski BM, Prakash O, Anand A. Predicting morbidity and mortality in acute pancreatitis in an Indian population: a comparative study of the BISAP score, Ranson's score and CT severity index. *Gastroenterol Rep (Oxf)*. 2016;4(3):216–220.
60. Tenner S, Baillie J, DeWitt J, Vege SS (2013) American College of Gastroenterology guideline: management of acute pancreatitis. Am J Gastroenterol 108(9):1400–1415
61. Mounzer R, Langmead CJ, Wu BU, Evans AC, Bishehsari F, Muddana V, Singh VK, Slivka A, Whitcomb DC, Yadav D, Banks PA, Papachristou GI (2012) Comparison of existing clinical scoring systems to predict persistent organ failure in patients with acute pancreatitis. Gastroenterology 142(7):1476–1482
62. Koutroumpakis E, Dasyam AK, Furlan A, Slivka A, Gougol A, Zeh HJ, Lee KK, Zureikat AH, Whitcomb DC, Yadav D, Papachristou GI (2016) Isolated Peripancreatic Necrosis in Acute Pancreatitis Is Infrequent and Leads to Severe Clinical Course Only When Extensive: A Prospective Study From a US Tertiary Center. J Clin Gastroenterol 50(7):589–595
63. Lesniowski B. IL-18, ghrelin, and acute inflammatory phase factors concentration in serum in patients with acute pancreatitis (AP). *Pancreatology*. 2009;9(4):498.
64. van Dijk SM, Hallensleben NDL, van Santvoort HC, Fockens P, van Goor H, Bruno MJ, Besselink MG. Acute pancreatitis: recent advances through randomised trials. *Gut*. 2017;66(11):2024–2032.
65. Cazacu SM, Parscoveanu M, Cartu D, Moraru E, Rogoveanu I, Ungureanu BS, Iordache S, Florescu DN, Iovanescu VF, Dragomir MI. NLR48 is Better Than CRP, and mCTSI, and Similar to BISAP and SOFA Scores for Mortality Prediction in Acute Pancreatitis: A Comparison of 6 Scores. *J Inflamm Res*. 2023;16:4793–4804.
66. Einstein AJ, Henzlova MJ, Rajagopalan S. Estimating risk of cancer associated with radiation exposure from 64-slice computed tomography coronary angiography. *Jama*. 2007;298(3):317–323.
67. Mortele KJ, Wiesner W, Intriere L, Shankar S, Zou KH, Kalantari BN, Perez A, vanSonnenberg E, Ros PR, Banks PA, Silverman SG. A modified CT severity index for evaluating acute pancreatitis: improved correlation with patient outcome. *AJR Am J Roentgenol*. 2004;183(5):1261–1265.
68. Davenport MS, Khalatbari S, Dillman JR, Cohan RH, Caoili EM, Ellis JH. Contrast material-induced nephrotoxicity and intravenous low-osmolality iodinated contrast material. *Radiology*. 2013;267(1):94–105.