Evaluation of Different Scoring Systems in Predicting the Severity of Acute Pancreatitis

Rohan J. Harsoda[1], Sharma Vipin Jaishree[2], Krishna Prasad G.V.[3]

[1]Department of Surgery, Military Hospital, Ahmedabad, Gujrat, India.
[2, 3] Department of Anaesthesiology, Military Hospital Kirkee, Pune, Maharashtra, India.

ABSTRACT

BACKGROUND
Accurate prediction of the severity of acute pancreatitis will help in identifying patients at increased risk for morbidity and mortality. We wanted to evaluate the different scoring systems in predicting the severity of acute pancreatitis.

METHODS
This cross-sectional study was undertaken in the Department of Surgery at a zonal hospital between April 2013 and December 2014.

RESULTS
40 patients were selected and enrolled in the study as per the selection criteria. 20 (50 %) patients had fair outcome and 20 (50 %) had a poor outcome. Accuracy of different scoring systems in predicting patient outcome ranged from 45 % (48-hr APACHE II) to 62.5 % (Goris MOF at baseline and 48 hr). Baseline Goris MOF was 70 % sensitive and 55 % specific in prediction of poor outcome. It had an accuracy of 62.5 % in prediction of outcome. 48-hr Goris MOF was 80 % sensitive and 45 % specific in predicting the outcome. Baseline APACHE II scores were below the cut-off level in all the patients. 48-hr APACHE II scores were 5 % sensitive and 100% specific for prediction of outcome. Ranson score > 3 was 25 % sensitive and 90 % specific in the prediction of outcome. Balthazar score > 6 was 65 % sensitive and 55 % specific in prediction of outcome. Ranson score was found to have a limited sensitivity for different outcomes (ranging from 21.1 % to 50 %) but was found to have a high specificity (83.8 % to 90 %).

CONCLUSIONS
Goris scoring system (at 48 hrs) was found to be highly sensitive to different poor outcomes as well as duration of hospital stay. It also correlated with Balthazar scoring system, which was also highly sensitive to different poor outcomes studied.

KEYWORDS
Acute Pancreatitis, Prediction, Scoring System, APACHE II, Goris MOF, Ranson's Score, Balthazar Score

Corresponding Author:
Dr. Krishna Prasad G.V.,
72/1 New Map, Near Symbiosis,
Range Hills, Kirkee – 411020,
Pune, Maharashtra, India.
E-mail: drkaypee99@yahoo.com

DOI: 10.18410/jebmh/2020/537

How to Cite This Article:
Harsoda RJ, Jaishree SV, Krishna Prasad GV. Evaluation of different scoring systems in predicting the severity of acute pancreatitis. J Evid Based Med Healthc 2020; 7(45), 2604-2610. DOI: 10.18410/jebmh/2020/537

Submission 17-06-2020,
Peer Review 23-06-2020,
Acceptance 22-07-2020,
Published 09-11-2020.

BACKGROUND

Acute pancreatitis is a protean disease capable of wide clinical variation, ranging from mild discomfort to apocalyptic prostration. It is a disease of variable severity in which some patients experience mild, self-limited attacks while others manifest a severe, highly morbid, and frequently lethal attack.[1] The ability to predict its severity can help identify patients at increased risk for morbidity and mortality, thereby assisting in appropriate early triage to intensive care units and selection of patients for specific interventions.[2] A multitude of predictive models have been developed to predict the severity of acute pancreatitis based upon clinical, laboratory, and radiologic risk factors, various severity grading systems, and serum markers.[3] Some of these can be performed on admission to assist in triage of patients, while others can only be obtained after the first 48 to 72 hours or later. However, these predictive models have low specificity (i.e., high false positive rates), which, when coupled with the low prevalence of severe acute pancreatitis (15 to 25 percent), results in low positive predictive values. Future predictive models will need to incorporate additional factors (e.g. biomarkers, genetic polymorphisms and mutations, and proteomic and metabolomic patterns) and methods of analyses.[4] The Atlanta Classification has been considered the global standard tool for the assessment of acute pancreatitis severity since its establishment in 1992. However, as time goes on, some of the definitions in the original Atlanta Classification has been proved to be confusing, especially its definition of "severity". In 2012, the Atlanta classification was revised with an emphasis on persistent organ failure. Multi-factorial scoring systems, including Ranson's and Acute Physiology and Chronic Health Evaluation (APACHE)-II scores have been used since the 1970s for assessment of the severity of acute pancreatitis.[5] Keeping in view the relevance of utility of different scoring systems in specific clinical settings, the present study was proposed to compare and evaluate the usefulness of four different severity scoring systems for acute pancreatitis namely Ranson's criteria, Acute Physiology and Chronic Health Evaluation [APACHE] II, Balthazar CT Severity Index and Organ Failure scale in our set up. Therefore, the current study aimed to evaluate the different scoring systems in predicting the severity of acute pancreatitis.

METHODS

This cross-sectional study was undertaken in the Department of Surgery at a zonal hospital between the period April 2013 and December 2014. Patient selection for the study involved the suspected acute pancreatitis pathology falling in the following inclusion and exclusion criteria. Inclusion Criteria included adult patients aged > 18 years, patients must have clinical signs and symptoms suggestive of pancreatitis (epigastric pain, radiation to back), the levels of serum amylase > 3 times normal range and radiologic evidence. All the patients who were unable to complete all the investigations, or expired within 24 hours of admission or not willing to participate in study were excluded from the study. A total 40 eligible patients of acute pancreatitis were enrolled in the study. Approval for conducting the study was obtained from Institutional Ethical Committee. Informed consent was obtained from all the patients enrolled in the study.

Case Definition

Severe pancreatitis was defined as the presence of organ failure and / or local pancreatic complications, complemented by unfavourable prognostic signs i.e., Ranson's score > 3, APACHE II score > 8, Goris > 1, Balthazar score > 6. Local pancreatic complications were defined as the development of a pseudocyst, abscess or parenchymal necrosis (more than 30 % or more than 3 cm of necrosis).

Balthazar's CT Severity Index Score

Grading of Pancreatitis

A: Normal pancreas: 0.
B: Enlargement of pancreas: 1.
C: Inflammatory changes in pancreas and peripancreatic fat: 2.
D: Ill-defined single fluid collection: 3.
E: Two or more poorly defined fluid collections: 4.

Pancreatic Necrosis

None: 0
Less than / equal to 30 %: 2
> 30 - 50 %: 4
> 50 %: 6
Total Score: Out of 10.
Total score < 5 indicated mild acute pancreatitis; score >5 indicated SAP.

Ranson's Scoring Criteria[6]

Every patient was assessed separately with Ranson's Scoring system on admission and after 48 hrs. According to the Ranson's Scoring Criteria, the parameters were as follows:

Non-Gallstone Pancreatitis	
At Admission	**After 48 hours**
1. Age in years > 55 years 2. White blood cell count > 16000 cells / mm^3 3. Blood glucose > 10 mmol / L (> 200 mg / dL) 4. Serum AST > 250 IU / L 5. Serum LDH > 350 IU / L	1. Serum calcium < 2.0 mmol / L (< 8.0 mg / dL) 2. Haematocrit fall > 10% 3. Oxygen (hypoxemia PaO_2< 60 mmHg) 4. BUN increased by 1.8 or more mmol / L (5 or more mg / dL) after IV fluid hydration 5. Base deficit (negative base excess) > 4 mEq / L 6. Sequestration of fluids > 6 L
For Gallstone Pancreatitis	
At Admission	**After 48 hours**
1. Age in years > 70 years 2. White blood cell count > 18000 cells / mm^3 3. Blood glucose > 12.2 mmol / L (> 220 mg / dL) 4. Serum AST > 250 IU / L 5. Serum LDH > 400 IU / L	1. Serum calcium < 2.0 mmol / L (< 8.0 mg / dL) 2. Haematocrit fall > 10% 3. Oxygen (hypoxemia PaO_2 < 60 mmHg) 4. BUN increased by 0.7 or more mmol / L (2 or more mg / dL) after IV fluid hydration 5. Base deficit (negative base excess) > 5 mEq / L 6. Sequestration of fluids > 4 L

Table 1. Ranson's Scoring Criteria.[6]

	Normal organ function 0 point	Organ Dysfunction 1 point	Organ failure 2 points
Lung	No mechanical ventilation	Mechanical ventilation with PEEP < 10 and FiO_2 < 0.4	Mechanical ventilation with PEEP > 10 and FiO2 > 0.4
Heart	Normal blood pressure (BPsys)	BPsys>100 mmHg with low dose of vasoactive drugs	Periods with BPsys < 100 mmHg and / or high dose of vasoactive drugs
Kidney	Serum creatinine < 2 Mg / dl (<150 µmol/l)	Serum creatinine > 2 mg / dl (> 150 µmol / l)	Haemodialysis or peritoneal dialysis
Liver	Normal SGOT and bilirubin	SGOT > 25 units / l; bilirubin > 2 mg / dl (> 31 µmol / l)	SGOT > 55 units / l; bilirubin > 6 mg / dl (>100 µmol / l)
Blood	Normal counts	Leukocytes > 30,000; platelets < 50,000	Leukocytes >60,000 or < 2,500
GI tract	Normal	Stress ulcer, Acalculous cholecystitis	Bleeding ulcer; Necrotizing enterocolitis and / or pancreatitis; perforation of gallbladder
CNS	Normal	Diminished responsiveness	Severely disturbed responsiveness; Diffuse neuropathy.

Table 2. Goris Multi Organ Failure (MOF) Score.[7]

A score < 3 indicated mild acute pancreatitis; Score > 3 indicated SAP.

Total Score Persistent / Progressive / Transient Organ Failure, during the First Week

Predicted Severity Score

No organ failure - 0.
Transient - 1.
Persistent - 2.
Progressive - 3.

For the purpose of present study, the actual severity of the acute pancreatitis was adjudged on the basis of following criteria –

Poor Outcome - Death during hospital stay, need for surgery, > 2 organ failures, duration of ICU stay > 7 days and total duration of hospital stay > 15 days.

Fair Outcome - Discharged alive, no need for surgery, < 2 organ failures, duration of ICU stay < 7 days and total duration of hospital stay < 15 days.

Statistical Analysis

Statistical analysis will be done using Statistical Package for Social Sciences version 15.0 or above. Chi-square test, independent samples and paired "t"-test shall be used to compare and evaluate the data.

RESULTS

The present study was carried out with an aim to evaluate the different scoring systems in predicting the severity of acute pancreatitis in patients of acute pancreatitis admitted in a zonal hospital. Table 1 shows distribution of cases according to severity as per the criteria used in the study.

Sl. No	Severity	No. of Cases (%)
1	Fair Outcome	20 (50)
2	Poor Outcome	20 (50)

Table 1. Distribution of Cases According to Actual Severity

As per the criteria used for the study, out of 40 patients enrolled in the study, a total of 20 (50 %) patients had fair outcome while remaining 20 (50 %) had a poor outcome.

Table 2 shows the association of baseline haematological, biochemical and gaseous assessment with outcome and 48-hr haematological, biochemical and gaseous assessment with outcome.

None of the scoring systems at chosen cut-off values showed a significant association with the outcome (p > 0.05).

Sl. No	Variable	Fair Outcome (n = 20)		Poor Outcome (n = 20)		Significance of Association	
		Mean	SD	Mean	SD	t	p
1	S. amylase	952.85	1075.61	1037.90	1187.89	-0.237	0.814
2	Haematocrit	38.55	4.32	38.60	4.63	-0.035	0.972
3	TLC ('000)	11.57	3.87	11.96	3.68	-0.331	0.743
4	Platelet Count (lakhs)	1.97	0.40	1.98	0.39	-0.064	0.949
5	Glucose	130.45	62.05	180.30	78.75	-2.224	0.032
6	LDH	313.50	223.50	282.60	128.76	0.536	0.595
7	AST	42.80	29.79	106.55	100.89	-2.710	0.010
8	PaO_2	102.43	2.77	99.35	10.14	1.309	0.199
9	Arterial pH	7.43	0.05	7.40	0.03	1.841	0.073
10	S. Sodium	138.00	3.51	136.75	4.30	1.007	0.320
11	S. Potassium	4.27	0.53	4.08	0.47	1.196	0.239
12	S. Creatinine	0.97	0.23	1.02	0.35	-0.529	0.600
13	BUN	32.45	8.97	31.60	9.04	0.298	0.767
14	Serum Bilirubin	1.77	2.13	1.71	1.35	0.115	0.909
15	ALT	49.25	26.80	88.00	69.92	-2.314	0.026
16	Alkaline Phosphatase	216.15	124.48	221.90	114.79	-0.152	0.880

Association of 48-hr Haematological, Biochemical and Gaseous Assessment with Outcome

Sr. No	Variable	Fair Outcome (n = 20)		Poor Outcome (n = 20)		Significance of Association	
		Mean	SD	Mean	SD	t	p
1	Haematocrit	38.85	2.85	38.20	3.47	0.647	0.522
2	TLC ('000)	11.4	3.4	11.8	4.8	-0.296	0.769
3	Platelet Count (lakhs)	2.22	0.53	2.16	0.52	0.374	0.711
4	Glucose	115.80	42.39	142.05	59.77	-1.602	0.117
5	LDH	243.65	122.63	264.70	171.39	-0.447	0.658
6	Serum Bilirubin	1.32	0.97	1.28	0.61	0.138	0.891
7	ALT	56.05	58.78	65.80	55.77	-0.538	0.594
8	AST	43.55	32.07	64.00	37.01	-1.867	0.070
9	Alkaline Phosp	212.35	137.02	211.15	107.94	0.031	0.976
10	PaO_2	103.11	1.74	100.51	10.19	1.125	0.268
11	Arterial pH	7.42	0.03	7.40	0.03	2.215	0.033
12	S. Sodium	140.40	3.50	138.30	4.35	1.761	0.086
13	S. Potassium	3.98	0.39	4.28	0.69	-1.701	0.097
14	S. Creatinine	0.89	0.16	1.01	0.39	-1.317	0.196
15	BUN	30.85	4.85	31.92	13.34	-0.336	0.739
16	Calcium	8.87	0.93	8.12	1.54	1.862	0.070
17	Base deficient	2.13	1.49	2.20	1.18	-0.165	0.870

Table 2. Association of Baseline Haematological, Biochemical and Gaseous Assessment with Outcome

Sl. No	Variable	Fair Outcome (n = 20)		Poor Outcome (n = 20)		Significance of Association	
		No.	%	No.	%	X^2	p
1	Baseline Goris >1	9	45	14	70	2.558	0.110
2	48 hr Goris >1	11	55	16	80	2.849	0.09
3	Baseline APACHE :	0	0	0	0	-	-
4	48 hr APACHE >	3	15	1	5	1.111	0.292
5	Ranson's score ≥	2	10	5	25	1.558	0.212
6	Balthazar Score ≥	9	45	13	65	1.616	0.204

Table 3. Evaluation of Different Scoring Systems against Outcome

Table 3; statistically, no significant difference in mean scores of different scoring systems was observed for the two outcomes except for Balthazar score which was seen

to be significantly higher in cases with poor outcome as compared to those with fair outcome (p > 0.05).

Only two deaths took place. None of the scoring systems showed a significant association with the event of death (p > 0.05).

Sl. No	Variable	Fair Outcome (n = 20)		Poor Outcome (n = 20)		Significance of Association	
		Mean	SD	Mean	SD	t	P
1	Goris at baseline	0.65	0.81	1.00	0.92	-1.277	0.209
2	Goris at 48 hr	1.25	1.21	2.00	1.41	-1.803	0.079
3	Balthazar Score	4.70	1.95	6.90	2.63	-3.003	0.005
4	APACHE at Baseline	1.80	2.12	1.35	1.53	0.770	0.446
5	APACHE at 48 hr	1.80	2.12	1.35	1.53	0.770	0.446
6	Ranson at baseline	0.70	0.86	1.10	0.97	-1.378	0.176
7	Ranson at 48 hr	0.35	0.49	0.80	1.06	-1.729	0.092

Evaluation of Different Scoring Systems against Outcome Death

Sr. no	Variable	Alive (n = 38)		Death (n = 2)		Significance of Association Fisher exact Test (p)
		No.	%	No.	%	
1	Baseline Goris>	21	55.3	0	0	0.180
2	48 hr Goris > 1	22	57.9	2	100	0.532
3	Baseline APACHE > 8	-	-	-	-	-
4	48 hr APACHE > 8	3	7.9	1	50	0.207
5	Ranson's Score ≥ 3	6	15.8	1	50	0.347
6	Balthazar Score ≥ 6	19	50	2	100	0.495

Table 3a. Evaluation of Mean Scores of Different Scoring Systems against Outcome

Sl. No.	Variable	No Need for Surgery (n = 35)		Need for Surgery (n = 5)		Significance of Association Fisher Exact Test (p)
		No.	%	No.	%	
1	Baseline Goris > 1	20	57.1	3	60	1.0
2	48 hr Goris > 1	22	62.9	5	100	0.154
3	Baseline APACHE > 8	-	-	-	-	-
4	48 hr APACHE > 8	3	8.6	1	20	0.427
5	Ranson's score ≥ 3	5	14.3	2	40	0.204
6	Balthazar score ≥ 6	17	48.6	5	100	0.053

Table 4a. Evaluation of Different Scoring Systems against Outcome Need for Surgery

Sl. No.	Variable	ICU Stay ≤ 7 Days (n = 37)		ICU Stay > 7 Days (n = 3)		Significance of Association Fisher Exact Test (p)
		No.	%	No.	%	
1	Baseline Goris > 1	21	56.8	2	66.7	1.00
2	48 hr Goris > 1	24	64.9	3	100	0.538
3	Baseline APACHE > 8	-	-	-	-	-
4	48 hr APACHE > 8	4	10.8	0	0	1.00
5	Ranson's Score ≥ 3	6	16.2	1	33.3	0.448
6	Balthazar Score ≥ 6	21	56.8	1	33.3	0.579

Table 4b. Evaluation of Different Scoring Systems against Outcome ICU Stay > 7 Days

Table 4; only five patients needed surgery. None of the scoring systems showed a significant association with the event of need for surgery (p > 0.05).

Sr. No.	Variable	Hospital Stay ≤ 15 days (n = 21)		Hospital Stay > 15 days (n = 19)		Significance of Association Fisher Exact Test (p)
		No.	%	No.	%	
1	Baseline Goris >1	11	52.4	12	63.2	0.538
2	48 hr Goris > 1	12	57.1	15	78.9	0.186
3	Baseline APACHE > 8	-	-	-	-	-
4	48 hr APACHE> 8	3	14.3	1	5.3	0.607
5	Ranson's score ≥ 3	3	14.3	4	21.1	0.689
6	Balthazar Score ≥ 6	10	47.6	12	63.2	0.360

Table 4c. Evaluation of Different Scoring Systems against Outcome Hospital Stay > 15 days

Only three patients needed ICU stay > 7 days. None of the scoring systems showed a significant association with ICU stay > 7 days (p > 0.05).

Almost half (n = 19) patients needed hospital stay > 15 days. None of the scoring systems showed a significant association with hospital stay > 15 days (p > 0.05).

Sl. No	Variable	≤ 2 Organ Failures (n = 35)		> 2 Organ Failures (n = 5)		Significance of Association Fisher Exact Test (p)
		No.	%	No.	%	
1	Baseline Goris > 1	21	60	2	40	0.634
2	48 hr Goris > 1	22	62.9	5	100	0.154
3	Baseline APACHE > 8	-	-	-	-	-
4	48 hr APACHE > 8	3	8.6	1	20	0.427
5	Ranson's Score ≥ 3	5	14.3	2	40	0.204
6	Balthazar Score ≥ 6	18	51.4	4	80	0.355

Evaluation of Ranson's Criteria against other Scoring SYSTEMS

Sl. No	Variable	Ranson's Criteria Category				Significance of Association Fisher Exact Test (p)
		Mild (n = 33)		Severe (n = 7)		
		No.	%	No.	%	
1	Baseline Goris > 1	19	57.6	4	57.1	1.0
2	48 hr Goris > 1	22	66.7	5	71.4	1.0
3	Baseline APACHE > 8	-	-	-	-	-
4	48 hr APACHE > 8	3	9.1	1	14.3	0.552
5	Balthazar Score ≥ 6	17	51.5	5	71.4	0.427

Table 5. Evaluation of Different Scoring Systems against Outcome > 2 Organ Failure

Table 5; A total of 5 (12.5 %) patients had > 2 organ failures. None of the scoring systems showed a significant association with > 2 organ failures (p > 0.05).

On evaluating Ranson's criteria against other scoring systems, statistically no significant association was observed between Ranson's criteria and other scoring systems (p > 0.05).

Table 6; on evaluating baseline Goris multi organ failure criteria against other scoring systems, statistically no

significant association was observed between baseline Goris MOF criteria and other scoring systems ($p > 0.05$) except with APACHE > 8 which was found to be inversely associated with MOF ($p = 0.026$).

On evaluating 48-hr Goris multi-organ failure criteria against their scoring systems, statistically no significant association was observed between 48-hr Goris MOF criteria and other scoring systems ($p > 0.05$) except for Balthazar score with which it had a significant positive association ($p = 0.001$).

Table 7; the accuracy of different scoring systems based on the results obtained in present study is depicted in table above. With respect to sensitivity for prediction of actual poor outcome (as defined in present study), 48 hr. Goris MOF was most sensitive (80 %) followed by baseline Goris MOF (70 %), Balthazar score (65 %) and Ranson score (25 %). APACHE II score at 48 hr. (5 %) and APACHE II score at baseline (0 %) were the worst performers as far as sensitivity was concerned.

Sl. No	Variable	Ranson's Criteria Category				Significance of Association
		Mild (n = 17)		Severe (n = 23)		Fisher Exact Test (p)
		No.	%	No.	%	
1	Baseline APACHE > 8	-	-	-	-	-
2	48 hr APACHE > 8	4	23.5	0	0	0.026
3	Ranson's score ≥ 3	3	17.6	4	17.4	1.0
4	Balthazar Score ≥ 6	8	47.1	14	60.9	0.523

Evaluation of 48-hr Goris Multi Organ failure Criteria against other Scoring Systems

Sl. No	Variable	Ranson's Criteria Category				Significance of Association
		Mild (n = 13)		Severe (n = 27)		Fisher Exact Test (p)
		No.	%	No.	%	
1	Baseline APACHE > 8	-	-	-	-	-
2	48 hr APACHE > 8	2	15.4	2	7.4	0.584
3	Ranson's score ≥ 3	2	15.4	5	18.5	1.0
4	Balthazar score ≥ 6	2	15.4	20	74.1	0.001

Table 6. Evaluation of Baseline Goris Multi Organ Failure Criteria against Other Scoring Systems

Sl. No	System	TP	FP	FN	TN	Accuracy	% Accuracy
1	Baseline Goris > 1	14	9	6	11	25	62.5
2	48 hr Goris > 1	16	11	4	9	25	62.5
3	Baseline APACHE > 8	0	0	20	20	20	50
4	48 hr APACHE > 8	1	3	19	17	18	45
5	Ranson's score ≥ 3	5	2	15	18	23	57.5
6	Balthazar score ≥ 6	13	9	7	11	24	60

Table 7. Accuracy of Different Scoring Systems for Prediction of Poor Outcome as Observed in the Present Study

DISCUSSION

There have been a number of scoring and classification systems proposed for assessment of acute pancreatitis that a clinician is often confused.[8] Hence there is always a need for an objective criterion through which a clinician could assess the severity of the disease and assess its complications.

In present study, we made an attempt to evaluate two such scoring systems, viz. Ranson's criteria and Goris Multi Organ Failure (MOF) against some of the other commonly used scoring systems, viz. APACHE II scoring system and Balthazar scoring system apart from evaluating the efficacy of each of different scoring systems to predict severity of acute pancreatitis against the actual outcome. For this purpose, an observational study was carried out in which a total of 40 patients with acute pancreatitis were enrolled. The severity of acute pancreatitis was done using APACHE-II, Balthazar scoring systems, Ranson's scoring and Goris multiorgan failure scoring systems. All these scoring systems have been used for assessment of severity of acute pancreatitis.[9-12]

In present study, the 48-hr baseline haematological, biochemical and gaseous parameters, did not show a significant difference between two outcome groups for any of the parameters except arterial pH which was found to be significantly lower among cases with poor outcome as compared to that in cases with fair outcome. Contrary to findings in present study Ranson criteria focuses on haematological, biochemical and gaseous parameters as a predictor of severity of acute pancreatitis.[13]

In present study, APACHE II scores were found to have lower accuracy (50 % and 45 % for baseline and 48-hr APACHE II scores) as compared to Goris and Ranson scores (62.5 % and 57.5 % respectively). Our results are in line with the study conducted by Chatzi Costas et al. (2002) who also observed APACHE II scoring system to be less accurate as compared to Ranson scoring system for the prediction of severity of acute pancreatitis.[14]

Most of the studies in literature have focused on sensitivity of different scoring systems against generalized outcomes. Chatzi Costas et al. (2003)[15] in their study comparing APACHE II / III, 48-hr Ranson score and 72 hr Balthazar score computed only likelihood ratios of positive test and found Balthazar score to have the highest likelihood ratio of positive test. However, in present study we found that Goris scoring systems at baseline or at 48 hr had higher sensitivity as compared to Balthazar score for poor outcome as well as for different specific outcomes.

In present study too we found that Balthazar score > 6 was able to predict the death, need for surgery and > 2 organ failures with a sensitivity ranging from 80 to 100 %. In terms of prediction of different adverse events in present study 48 hr BORIS score was found to have a sensitivity ranging from 78.9 % to 100 % for all the five specific outcomes in study.[16,17]

In present study, APACHE II score at admission were the worst predictors, however, Suvarna et al. (2011)[18] reported them to have high accuracy and superiority over other systems like Ranson's criteria.

In present study, we found that the mutual association of different scoring systems was poor except for a significant association between 48-hr Goris and Balthazar CT scoring systems ($p = 0.001$), in fact these two systems

had the highest sensitivity for different adverse / poor outcomes and thus indicated a good correlation.

The findings in present study stressed on the need of using different scoring systems for clinical decision making and improvising and determining the direction of patient management. The utility of these scoring systems is in fact not in forecasting a poor and adverse outcome but in the determining the course of management in order to avert these poor and adverse outcomes. In present study, we found that these scoring systems helped to improvise the patient management and reverted the adverse outcome although this process ended up in a prolonged ICU / hospital stay for the patient but helped to reduce the more severe and serious outcomes. Balthazar score and Goris MOF score at 48 hours were found to be two scoring systems that were highly sensitive to different poor outcomes vis-a-vis the severity of acute pancreatitis. They provide us ample inputs to strengthen our facilities in order to reduce the adverse outcomes even further.

More studies on the issue are recommended in same or similar settings in order to test the repeatability of results in similar situations. Continuous studies on the issue are recommended with changing settings to evaluate the relevance and contemporariness of different scoring systems and their ability to avert adverse outcome.

CONCLUSIONS

There is no independent association of baseline or 24-hr. haemodynamic or vital parameters with outcome. Baseline elevated glucose, AST and ALT levels were significantly associated with a poor outcome. And, except for 48-hr. arterial pH levels, none of the laboratory, haematological or gaseous parameters were significantly associated with poor outcome. The accuracy of different scoring systems in the prediction of patient outcome ranged from 45 % (48-hr. APACHE II) to 62.5 % (Goris MOF at baseline and 48 hr.). The baseline Goris MOF was 70 % sensitive and 55 % specific in the prediction of poor outcome. It had an accuracy of 62.5 % in the prediction of outcome. Similarly, 48-hr. Goris MOF was 80 % sensitive and 45 % specific in the prediction of outcome. The baseline APACHE II scores were below the cut-off level in all the patients. And, 48-hr. APACHE II scores were 5 % sensitive and 100 % specific for prediction of outcome. Ranson's score > 3 was 25 % sensitive and 90% specific in prediction of outcome. Ranson score was found to have a limited sensitivity for different outcomes (ranging from 21.1 % to 50 %) but was found to have a high specificity (83.8 % to 90 %), thus indicating that the criteria used in study was a strict one and needs to be relaxed in order to enhance the sensitivity. Whereas, Balthazar score > 6 was 65 % sensitive and 55 % specific in prediction of outcome. For the outcome ICU stay > 7 days, 48-hr. Goris score was 100 % sensitive but only 35.1 % specific while baseline Goris score was 66. 7 % sensitive and 43.2 % specific. All the other scoring systems had sensitivity below 50 %. For the outcome hospital stay > 15 days, 48 hr. Goris score was most sensitive (78.9 %) but had limited specificity (42.9 %). Next to 48 hr. Goris score was the Balthazar score which had a sensitivity of 63.2 % and specificity of 52.4 %. A significant association between 48-hr Goris and Balthazar score was observed. In summary, Goris scoring system (at 48 hr.) was found to be highly sensitive to different poor outcomes as well as duration of hospital stay. It also correlated with Balthazar scoring system which was also highly sensitive to different poor outcomes studied.

Data sharing statement provided by the authors is available with the full text of this article at jebmh.com.
Financial or other competing interests: None.
Disclosure forms provided by the authors are available with the full text of this article at jebmh.com.

REFERENCES

[1] Bhatia M, Wong FL, Cao Y, et al. Pathophysiology of acute pancreatitis. Pancreatology 2005;5(2-3):132-144.
[2] Corfield AP, Williamson RC, McMahon MJ, et al. Prediction of severity in acute pancreatitis: prospective comparison of three prognostic indices. The Lancet 1985;326(8452):403-407.
[3] Blum T, Maisonneuve P, Lowenfels AB, et al. Fatal outcome in acute pancreatitis: its occurrence and early prediction. Pancreatology 2001;1(3):237-241.
[4] Mounzer R, Langmead CJ, Wu BU, et al. Comparison of existing clinical scoring systems to predict persistent organ failure in patients with acute pancreatitis. Gastroenterology 2012;142(7):1476-1482.
[5] Souza GD, Souza LR, Cuenca RM, et al. Understanding the international consensus for acute pancreatitis: classification of Atlanta 2012. ABCD Arquivos Brasileiros de Cirurgia Digestiva (São Paulo) 2016;29(3):206-210.
[6] Lalithkumar J, Chitra T, Kodieswaran N. Comparative study between BISAP and Ranson's score in predicting severity of acute pancreatitis. Int Arc Intg Med 2016;3(9):23-33.
[7] Lefering R, Goris JR, van Nieuwenhoven EJ, et al. Revision of the multiple organ failure score. Langenbeck's Archives of Surgery 2002;387(1):14-20.
[8] Yadav D. Acute pancreatitis: too many classifications - What is a clinician or researcher to do? Clinical Gastroenterology and Hepatology 2014;12(2):317-319.
[9] De Beaux AC, Palmer KR, Carter DC. Factors influencing morbidity and mortality in acute pancreatitis: an analysis of 279 cases. Gut 19951;37(1):121-126.
[10] Ranson JH, Rifkind KM, Roses DF, et al. Prognostic signs and the role of operative management in acute pancreatitis. Surg Gynecol Obstet 1974;139(1):69-81.
[11] Khan AA, Parekh D, Cho Y, et al. Improved prediction of outcome in patients with severe acute pancreatitis by the APACHE II score at 48 hours after hospital admission compared with the APACHE II score at admission. Archives of Surgery 2002;137(10):1136-1140.

[12] Balthazar EJ. Acute pancreatitis: assessment of severity with clinical and CT evaluation. Radiology 2002;223(3):603-613.
[13] Munoz A, Katerndahl DA. Diagnosis and management of acute pancreatitis. American Family Physician 2000;62(1):164-174.
[14] Chatzicostas C, Roussomoustakaki M, Vlachonikolis IG, et al. Comparison of Ranson, APACHE II and APACHE III scoring systems in acute pancreatitis. Pancreas 2002;25(4):331-335.
[15] Chatzicostas C, Roussomoustakaki M, Vardas E, et al. Balthazar computed tomography severity index is superior to Ranson criteria and APACHE II and III scoring systems in predicting acute pancreatitis outcome. Journal of Clinical Gastroenterology 2003;36(3):253-260.
[16] Leung TK, Lee CM, Lin SY, et al. Balthazar computed tomography severity index is superior to Ranson criteria and APACHE II scoring system in predicting acute pancreatitis outcome. World Journal of Gastroenterology 2005;11(38):6049-6052.
[17] Fan JY, Huang ZW, Guo J. Value of four scoring systems for predicting prognosis of severe acute pancreatitis. Zhong xi yijie he xuebao: Journal of Chinese Integrative Medicine 2009;7(1):34-40.
[18] Suvarna R, Pallipady A, Bhandary N, et al. The clinical prognostic indicators of acute pancreatitis by APACHE II scoring. Journal of Clinical and Diagnostic Research 2011;5(3):459-463.