DOI: 10.1590/0100-6991e-20192285 Original Article

Outcomes prediction score for acute abdomen: a proposal.

Proposta de escore preditor de desfechos para abdome agudo.

Felipe Antônio Cacciatori[1,2]; Arthur Dajori Ronchi[1,3]; Sérgio Emerson Sasso[1]

A B S T R A C T

Objective: to establish a correlation between objective data collected at the initial evaluation and the outcomes in cases of acute abdominal pain. **Methods:** we conducted a retrospective, case-control study in which we reviewed medical records of patients treated at an emergency room in 2016, diagnosed with abdominal and pelvic pain. **Results:** of the 2,352 medical records evaluated, we considered 330 valid for the study. Of these patients, 235 (71.2%) were discharged and the remaining 95 (28.8%) were hospitalized, submitted to surgery, or died. The statistical analysis demonstrated that male gender, age ≥50, axillary temperature >37.3ºC, anemia, leukocytosis >10,000/mm^3, neutrophil count above 7,700/mm^3, lymphocyte count less than 2,000/mm^3 and hyperamylasemia were variables in independently associated with worse outcome. **Conclusion:** the presence of three or more of the evaluated variables greatly increases the chances of a patient suffering the outcomes of surgery or death, the chance of death being greater the greater the number of variables present.

Keywords: Abdominal Pain. Abdomen. Acute. Propensity Score. Surgery Department. Hospital.

INTRODUCTION

The acute abdomen can be defined as the presence of abdominal pain and tenderness[1] of non-traumatic etiology, last for a maximum of five days[2], and which can be caused by more than 42 different diseases[3], many of which requiring surgical treatment. The primary objective in these cases is the early and accurate diagnosis of the etiology, to start the appropriate treatment[4]. However, about 40% of patients seeking medical care complaining of abdominal pain remain undiagnosed[5].

From the days of Hippocrates, without ever losing importance, remains the thought that clinical examination must be sovereign. However, it is also known that only 43% to 59% of diagnoses based on anamnesis and physical examination are correct[2], since the normality of vital signs is frequent, despite a serious disease, especially in elderly or immunocompromised patients. Even the intensity of abdominal pain is not related to the severity of the disease[6]. In this scenario, the good use of complementary methods can help to increase the rate of correct diagnoses or, at least, indicate patients who need further investigation.

The importance of the theme is particularly evident when one considers that, in Brazil, only in the month of January 2018, DATASUS registered 3,419 admissions for abdominal and pelvic pain (ICD10 R10), resulting in 18,553 cumulative days of hospitalization. In the same sample, 329 deaths were recorded, corresponding to a lethality rate of 9.62%[7].

The development of diagnostic algorithms or scores to increase the accuracy in cases of acute abdomen, at present non-existent[2], could ultimately contribute to the reduction of morbidity and mortality from abdominal diseases. The first step towards this goal includes establishing a correlation between the objective data obtained in the initial assessment, including the minimal complementary examinations, with the outcome of patients presenting with acute abdominal pain in the emergency room.

1 - Santa Catarina Extreme South University (UNESC), Medicine Course, Criciúma, SC, Brazil. 2 - São José Hospital, General Surgery Service, Criciúma, SC, Brazil. 3 - São José Hospital, Anesthesiology Service, Criciúma, SC, Brazil.

The objective of this study was and establish a correlation between objective data collected in the initial evaluation of patients treated for acute abdominal pain and their clinical outcomes.

METHODS

We conducted a retrospective, case-control study of patients seen in the emergency room service of a medium and high complexity teaching hospital, located in a county in the south of Santa Catarina, which serves a population of 994,000 inhabitants and that, at the time, performed about 130,000 medical care events per year.

After receiving approval by the Ethics in Research Committees of the Santa Catarina Extreme South University (UNESC) and the São José hospital, with opinions of numbers 2,253,11 0 and 2,278,239, respectively, through the *Plataforma Brasil* system, we gathered all the reports classified as CID10 R10 (abdominal and pelvic pain) of care events performed in 2016, totaling 4,575 records. We randomized the reports using the "Random" command in the Microsoft Excel 2016 software. We then had access to the records stored in the Philips TASY hospital system, from which we collected data consecutively until we obtained the final sample of 330 valid events, with 2,022 exclusions, all due to incomplete medical records. We did not identify repeated events amongst the valid medical records sample. Figure 1 summarizes the steps of the present study.

The tabulation of data occurred in the Microsoft Excel software and two groups were designed: the Control group consisted of patients who were discharged after treatment and, therefore, did not display the outcomes surgery or death, and the Case group, with patients who were admitted, underwent surgery, or died. We performed statistical analysis using the IBM SPSS Statistics Vesion 21.0 software by constructing frequency distributions and comparisons between the dependent and independent variables. The measures of central tendency were the mean and standard deviation, and also the median and interquartile range. We used the Kolmogorov-Smirnov test to determine the normality or not of the comparative data, and then applied the Student's t-test or the Mann-Whitney u-test, respectively. We also used the Pearson's chi-square and likelihood ratio tests, with a complementary evaluation of the residual analysis for these tests, and the Kruskal-Wallis h-test, complemented by the Dunn's *post-hoc* test. As an association measure, we calculated the *odds ratio* -OR-, with a 95% confidence interval, for the variables that displayed relevant results[8,9].

RESULTS

Of the 330 valid events, 235 (71.2%) were discharged after evaluation in the emergency room service. The remaining 95 (28.8%) achieved the primary outcomes of this study and were divided into three subgroups: a) hospitalization without surgery, b) hospitalization with surgery, c) death, as shown in figure 1. The sample consisted mainly of women (64.8%) and the average age was 39.54 (±19.09) years, ranging from four to 90 years.

The comparison between groups, detailed in table 1, showed a significant statistical difference between the mean age, heart rate, axillary temperature, leukocyte, band cells, neutrophils, and serum amylase levels, all of which presented higher values in the case group patients. This finding suggests a worse outcome in patients who manifest these changes.

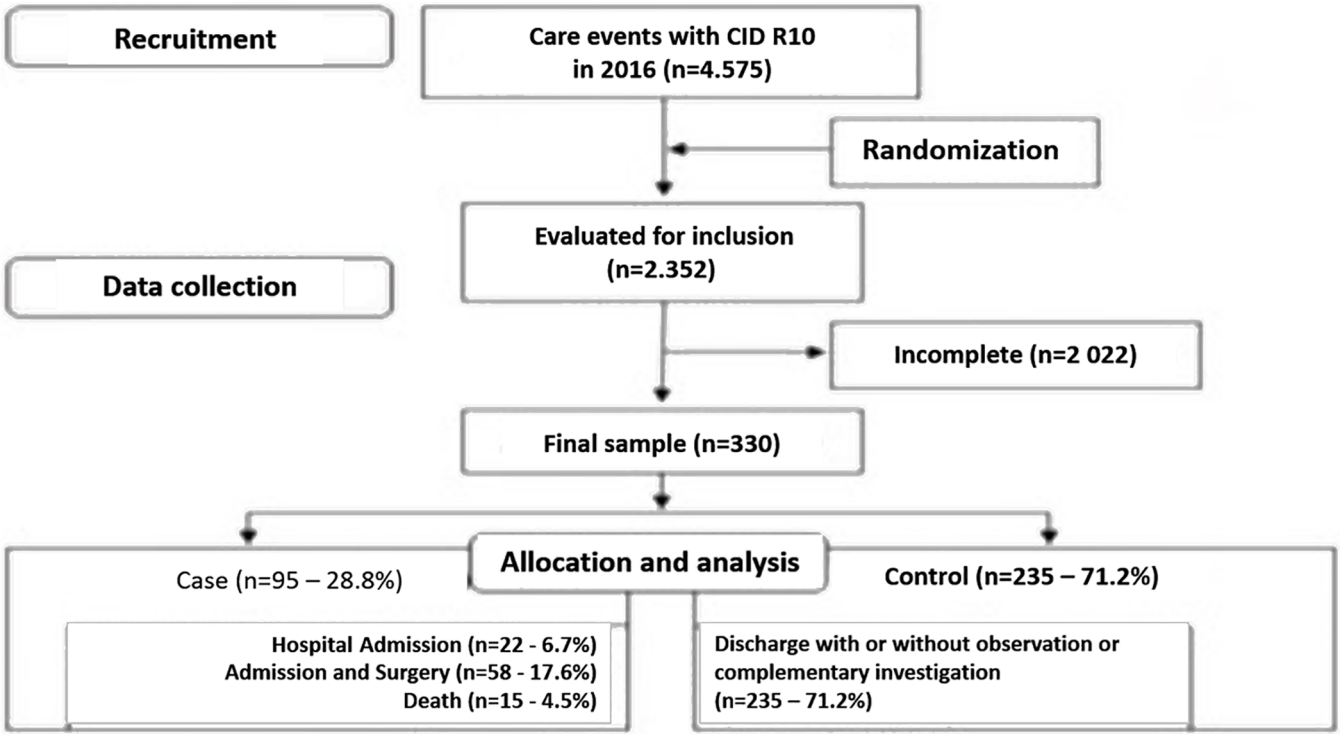

Figure 1. *Study flowchart.*

Table 1. *Characteristics of the groups.*

Variables	Case Group n=95	Control Group n=235	*p*-value
Age (years)*	44.40±19.32	37.57±18.67	0.004‡
Gender**			
Female	41 (43.20)	173 (73.60)[a]	<0.001‡
Male	54 (56.80)[a]	62 (26.40)	
Vital signs*			
Heart rate	90.52±20.66	81.65±13.22	<0.001‡‡
Peripheral saturation of oxygen	97.96±2.36	98.49±1.37	0.450‡
Axillary temperature	37.13±0.95	36.55±0.75	<0.001‡
Systolic blood pressure	118.36±21.28	126.40±17.80	0.001‡
Diastolic blood Pressure	70.66±14.12	78.18±11.39	<0.001‡
Laboratory tests			
Hematocrit*	36.84±6.69	39.31±3.89	<0.001‡‡
Hemoglobin*	12.28±2.43	13.23±1.50	<0.001‡‡
Leukocytes*	12,812.13±6,010.86	10,077.44±4,128.04	<0.001‡
Band cells***	214.00 (133.0-340.5)	170.00 (109.0-221.5)	<0.001‡
Neutrophils***	8,394.00 (4,867.5-12,359.0)	5,678 (4,190.0-8,160.5)	<0.001‡
Lymphocytes*	2,074.89±965.93	2,395±997.19	0.008‡
Amylase***	65.00 (48.0-86.0)	56.00 (46.0-75.5)	0.026‡

** Values expressed as mean and standard deviation; ** values expressed in frequency and percentage; *** values expressed as median and interquartile range; † Mann-Whitney u-test; ǂ Pearson chi-square test; ǂǂ Student's t-test; [a] statistical significance after residuals analysis.*

Lower values for systolic and diastolic blood pressure, hematocrit, hemoglobin and lymphocytes were also statistically associated with the Case group. Basophil, eosinophil and monocyte levels did not show significant differences between groups, which is why they are deliberately omitted from table 1.

Regarding gender, evidence suggests that male patients are more likely to be in the Case group than female ones (p<0.001). We found no significant differences between groups regarding peripheral oxygen saturation values (p=0.450).

Partial or qualitative urine examination showed no significant difference between the groups regarding sedimentscopy. Research on abnormal urinary elements, on the other hand, suggests that the presence of ketone bodies (p=0.022) or bilirubin (p=0.009) is more associated with the Case group than to the Control group.

In the subset of 22 (6.7%) patients admitted for hospitalization without needed surgery, there were 22 different diagnoses, such as diverticulitis, hepatic laceration, cirrhosis, mesothelioma, and subileus. These admissions totaled 82 days of hospitalization, the longest being 13 days, with a median of four days.

The subgroup of 58 (17.6%) patients who were hospitalized and submitted to surgery accounted for 15 cholecystectomies, nine exploratory laparotomies, six appendectomies, six debridements or drainages of abscesses, five gastro or duodenorraphies, five colectomies or enterectomies, among less frequent others. This subgroup accumulated 478 days of hospitalization, with a median of 4, ranging from 1, in some cases of cholecystectomy and appendectomy, to 70 days of hospitalization, in one case of perforated gastric ulcer treated with gastrorraphy.

The 15 (4.5%) events that resulted in death together totaled about 100 days of hospitalization. The causes of death were cervical cancer (n=3), liver cancer (n=2), portal vein thrombosis (n=1), sepsis (n=1) and polytrauma (n=1). Bladder, rectum, lung, stomach and pancreatic cancers also had one occurrence each. Two deaths had no identified cause.

From the collected data, it was possible to calculate the OR of some selected variables in relation to the Case and Control groups. The chance of a patient with acute abdominal pain and age 50 years or older to suffer the outcome is 2.67 times greater than the chance of a younger patient (95%CI: 1.62-4.38).). Similarly, male gender is more likely to suffer the outcome, with an OR of 3.67 (95%CI: 2.23-6.05). Hematocrit below 35% (OR=4.25; 95%CI: 2.33-7.77) and mainly below 30% (OR=20.13; 95%CI: 4.47-90.51) as well.

Leukocytosis ≥10,000/mm^3 (OR=2.86; 95%CI: 1.62-4.43) and, more strongly, ≥16,000/mm^3 (OR=6.26; 95%CI: 3,11-12,62), is significantly more likely to allocate the patient to the Case group. Neutrophils above 7,700/mm^3 had an OR of 2.62 (95%CI: 1.60-4.28). Lymphopenia <2,000/mm^3 resulted in an OR of 1.70 (95%CI: 1.05-2.76) in favor of the Case group.

During the comprehensive analysis of the data collected in the present work, the authors envisioned the possibility of extrapolating the laboratory and clinical patterns found through a score, conceived based on values with statistically significant differences between the studied groups and subgroups, primarily based on their OR. We determined cutoff points for axillary temperature, leukocytes and neutrophils based on the literature[10,11], and of the other variables based on the analysis of the ROC curve of each test[8,9].

After assigning a score for each variable, was obtained a score ranging from 0 to 12 points. Hematocrit and amylase items had to be graded in order to improve the final properties of the score. Figure 2 shows the variables and risk groups according to the proposed acute abdomen score.

We established the cutoff points for the risk groups after statistical analysis of the mean scores according to the outcome. This analysis demonstrated successively higher averages for the most severe outcomes, as shown in figure 3. The OR applied to the final score shows that patients with a score of 3, ie, with a risk different from low, are 9.98 (95%CI: 5.52-18.05) times more likely to belong to the Case group rather than to the Control group.

DISCUSSION

The demographic characteristics of the study sample agree with other similar studies. The average age of patients seeking urgent or emergency care is widely variable and may differ further when analyzing specific diseases.

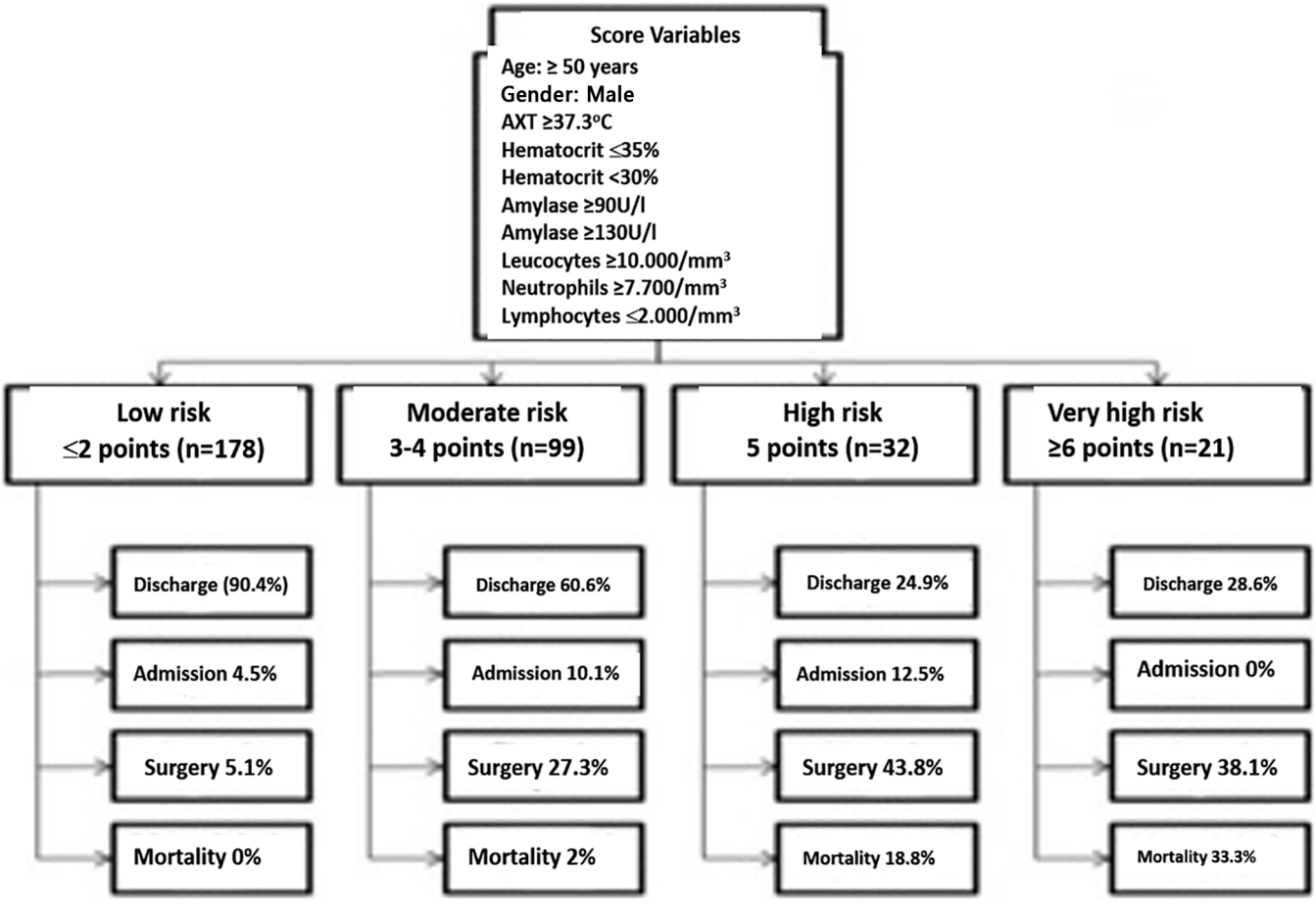

Figure 2. *Score for acute abdomen.*
In the presence of one of the variables shown in the figure, the patient should receive 1 point. In the case of amylase, he/she will receive 1 point if it is ≥90 U/l and 1 more point if it is ≥130 U/l, totaling a maximum of 2 points for the amylase item. For hematocrit, the patient will receive 1 point if ≤35% and 1 more point if ≤30%, totaling a maximum of 2 points for this item. Calculating the final score, the patient can be classified into one of four risk groups, as follows: low, with 0% mortality; moderate, with 2% mortality; high, with 18.8% mortality; and very high, with 33.3% mortality; AXT= axillary temperature.

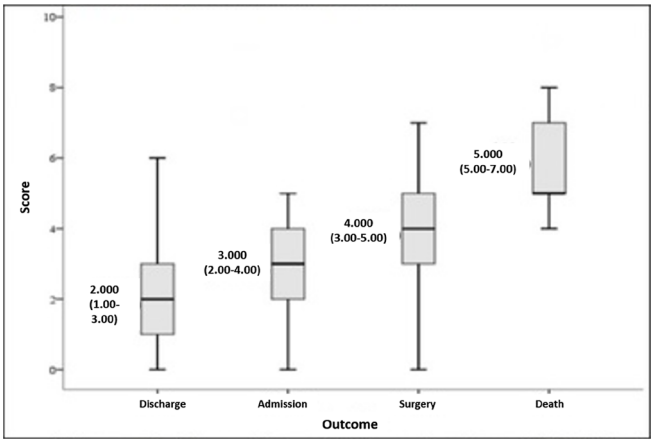

Figure 3. *Median range according to outcome. Median and interquartile range of the acute abdomen score according to outcome. The p-value <0,001 was obtained with the Krustal-Wallis h-test, showing statistical significance. The Dunn post-hoc test showed statistically significant differences in the comparisons Discharge-Surgery (p<0.001), Discharge-Death (p<0.001), Admission-Death (p=0.003), and non-significant differences in the comparisons Discharge-Admission (p=0.056), Admission-Surgery (p=0.406), and Surgery-Death (p=0.083).*

In the literature, the average seems to be 25.3 years for diseases such as acute appendicitis[10] and 51 years for diseases such as acute mesenteric ischemia[12].

In emergency room visits, the average for all causes can be around 46.7 years[12]. When analyzing the visits due to acute abdomen, the average age of 39.54 years found in this study is supported by literature with similar methodology, being 34 years in the study by Powers and Guertler[5] and 39 in the study by Acute Abdominal Pain (AAP) Study group[13].

Regarding gender, there are 58% of men in appendicitis studies[10] and 69.8% in acute mesenteric ischemia studies[14]. When observed the generic complaint of abdominal pain, the prevalence of women becomes noticeable, 54% to 68%[5,13]. These data validate those found in this series, with 64.8% women in the total sample and 56.8% men in the Case group.

The total hospitalization rate in the present study was 28.8%, which is 10% higher than the rate found by Powers and Guertler[5], but similar to the 27.4% rate of a study conducted in 1972 in the same institution[15]. Of the 76 surgical diagnoses identified in the Powers and Guertler[5] series, 36 (47.3%) were cholecystitis, which was also the most frequent diagnosis in our study, with 19% of occurrence. Moreover, 4.3% of 1,166 hospitalized patients died, which is proportionally similar to the 4.5% rate found in this study. Data from DATASUS[7] suggest 9.62% lethality among patients hospitalized with ICD10 R10.

The score suggested in this paper required the authors to arbitrate some cutoff points and to choose variables for its composition. The axillary temperature cutoff point that Alvarado[10] established in his homonymous score for the diagnosis of acute appendicitis was equal to or above 37.3ºC, with a sensitivity of 73% and a specificity of 50%. Another analysis[16] found a sensitivity of 70% and a specificity of 65%, using as a cutoff point the temperature above 37.7ºC.

Gans *et al.*[2] evaluated as low the sensitivity (66% to 78%) of leukocytosis above 10,000/mm^3 to discriminate urgent causes of acute abdomen, defined as those that need treatment within 24 hours to prevent complications. The specificity of the same test was determined to be 66%.

The study by Alvarado[10] retrospectively analyzed 277 medical records of hospitalized patients with suspected acute appendicitis. Of these, 89% had the diagnosis confirmed surgically. From these data, the author determined the statistical properties of clinical and laboratory findings and found variables that allowed the construction of a score that now bears his name, based on the accuracy of the tests.

Ozkan *et al.*[17] rated the statistical properties of the Alvarado score in a sample of 74 patients undergoing appendectomy, with an average age of 36 years and the prevalence of men (70.3%).

The score presented sensitivity of 54%, specificity of 73.3%, positive predictive value of 88.2%, negative predictive value of 29.7%, and accuracy of 58%. Comparatively, ultrasonography showed, respectively, the values 71.2%, 47%, 82.2%, 31.8% and 66%. It is important to note that the Alvarado and Ozkan samples were composed of hospitalized patients already diagnosed with appendicitis. It is expected that, in this population, the prevalence of such disease is quite high, which favors the performance of the score. In contrast, the present study evaluated care provided in the urgency and emergency sector, whose population is predominantly composed of patients without surgical pathology and even without a definite diagnosis. Even so, we found statistically significant differences between the subgroups and their outcomes, suggesting a good diagnostic capacity of the proposed score.

It is important to recognize that the records excluded from this study due to their incompleteness could possibly be part of the Control group. On the other hand, the inclusion criterion being only the generic diagnosis of abdominal pain, represented by the ICD10 R10, patients searched for the medical service and who were promptly given a specific diagnosis were beyond the reach of this study. Such patients would be part of the Case group. In addition, there is a possibility that some patients in the Control group may have returned to the hospital and suffered outcomes outside this study period. Certainly, such changes in the sample would impact the statistical properties of the proposed score.

A future study that uses this score in a prospective and blind manner, that is, without the emergency physician relying on its outcome to define conduct, could shed light on its usefulness and confirm its properties. The same study could evaluate the inclusion of elements of the physical examination, or of the anamnesis, and of variables capable of increasing the sensitivity of the score. There are other scores in the medical literature that aim to diagnose specific abdominal pathologies. However, we did not identify any whose scope was the risk classification of patients with abdominal pain syndrome, that is, regardless of the diagnosis. This fact is possibly the greatest virtue of the present study, since the use of the acute abdomen score could be considered an aid in deciding the need for hospitalization.

O'Brien[6] considers as a criterion for hospitalization the presence of signs or symptoms of severity in high-risk patients. Other authors[18] consider as severe pathology one in which fever coexisted with any of the following items: need for intravenous antibiotic therapy, intravenous vasoactive drugs, surgery, radiological drainage, intensive care unit admission, and white series, red series or platelet count alterations.

Using the score proposed in this study, the possibility of hospitalization could be considered in cases where the score was different from a low risk result (score 3) or, mainly, high or very high risk (score 5). Finally, the authors point out that even though they have ceaselessly sought to nullify all possible biases, either in the data collection or analysis stage, it is not possible to exclude the occurrence of primary measurement bias, since all data, mainly vital signs such as axillary temperature, were collected and recorded without methodological supervision.

Our study found variables statistically associated with the outcomes analyzed, which helps determine the profile of patients most prone to hospital admission, surgery or death. The authors developed a scoring system that demonstrates the ability to stratify patients with abdominal pain according to the chance of outcome, based solely on age, gender, axillary temperature, hematocrit, leukogram and serum amylase parameters. This tool can be further refined to the point of accurately and precisely distinguishing patients without disease from those at risk of death. For the time being, it was evident to the authors that the proposed score has a good ability to identify patients with a higher chance of death.

R E S U M O

Objetivo: *estabelecer a correlação entre dados objetivos coletados na avaliação inicial e os desfechos nos casos de dor abdominal aguda.* **Métodos:** *estudo retrospectivo, de caso controle, em que foram revisados prontuários de pacientes atendidos em um serviço de urgência e emergência no ano de 2016, diagnosticados com dor abdominal e pélvica.* **Resultados:** *de 2.352 prontuários avaliados, 330 foram considerados válidos para o estudo. Destes pacientes, 235 (71,2%) receberam alta e os 95 (28,8%) restantes foram internados, submetidos à cirurgia ou morreram. A análise estatística demonstrou que sexo masculino, idade ≥50 anos, temperatura axilar >37,3ºC, anemia, leucocitose >10.000/mm³, neutrocitose acima de 7.700/mm³, linfopenia <2.000/mm³ e hiperamilasemia são variáveis isoladamente associadas a pior desfecho.* **Conclusão:** *a presença de três ou mais das variáveis avaliadas aumenta fortemente a chance de um paciente sofrer os desfechos de cirurgia ou morte, sendo a chance de óbito tanto maior quanto maior o número de variáveis presentes.*

Descritores: *Dor Abdominal. Abdome Agudo. Pontuação de Propensão. Centro Cirúrgico Hospitalar.*

REFERENCES

1. Squires RA, Portier RG. Abdome agudo. In: Townsend Jr. CM, Beauchamp RD, Evers BM, Mattox KL, editores. Sabiston tratado de cirurgia: a base biológica da prática cirúrgica moderna. 19a ed. Rio de Janeiro: Elsevier; 2015. p. 1141-59.
2. Gans SL, Pols MA, Stoker J, Boermeester MA; expert steering group. Guideline for the diagnostic pathway in patients with acute abdominal pain. Dig Surg. 2015;32(1):23-31.
3. Feres O, Parra RS. Abdômen agudo. Medicina (Ribeirão Preto). 2008;41(4):430-6.
4. Meneghelli UG. Elementos para o diagnóstico do abdômen agudo. Medicina (Ribeirão Preto). 2003;36(2/4):283-93.
5. Powers RD, Guertler AT. Abdominal pain in the ED: stability and change over 20 years. Am J Emerg Med. 1995;13(3):301-3.
6. O'Brien MC. Acute abdominal pain. In: Tintinalli JE, Stapczynski JS, Ma OJ, Cline DM, Cydulka RK, Meckler GD; The American College of Emergency Physicians. Tintinalli's emergency medicine: a comprehensive study guide. 7th ed. New York: McGraw-Hill; 2016. p. 481-9.
7. Brasil. Ministério da Saúde. Sistema de Informações Hospitalares do SUS (SIH/SUS)/DATASUS. Morbidade Hospitalar do SUS: CID-10 R10, Período Jan/2018.
8. Oliveira Filho PF. Epidemiologia e bioestatística: fundamentos para a leitura crítica. 1a ed. Rio de Janeiro: Rubio; 2015.
9. Medronho RA, Bloch KV, Luiz RB, Werneck GL. Epidemiologia. 2a ed. São Paulo: Atheneu; 2009.
10. Alvarado A. A practical score for the early diagnosis of acute appendicitis. Ann Emerg Med. 1986;15(5):557-64.
11. Kabir SA, Kabir SI, Sun R, Jafferbhoy S, Karim A. How to diagnose an acutely inflamed appendix; a systematic review of the latest evidence. Int J Surg. 2017;40:155-62.
12. Barreto RF, Gomes CZL, da Silva RM, Signorelli AAF, Oliveira LF, Cavellani CL, et al. Avaliação de dor e do perfil epidemiológico, de pacientes atendidos no pronto-socorro de um hospital universitário. Rev Dor. 2012;13(3):213-9.
13. Acute Abdominal Pain (AAP) Study group. Diagnostic accuracy of surgeons and trainees in assessment of patients with acute abdominal pain. Br J Surg. 2016;103(10):1343-9.

14. Andersson RE, Hugander AP, Ghazi SH, Ravn H, Offenbartl SK, Nyström PO, et al. Diagnostic value of disease history, clinical presentation, and inflammatory parameters of appendicitis. World J Surg. 1999;23(2):133-40.
15. Brewer RJ, Golden GT, Hitch DC, Rudolf LE, Wangensteen SL. Abdominal pain. An analysis of 1,000 consecutive cases in a University Hospital emergency room. Am J Surg. 1976;131(2):219-23.
16. Wang Z, Chen JQ, Liu J, Tian L. A novel scoring system for diagnosing acute mesenteric ischemia in the emergency ward. World J Surg. 2017;41(8):1966-74.
17. Ozkan S, Duman A, Durukan P, Yildirim A, Ozbakan O. The accuracy rate of Alvarado score, ultrasonography, and computerized tomography scan in the diagnosis of acute appendicitis in our center. Niger J Clin Prac. 2014;17(4):413-8.
18. San Martín López JV, Casas Rojo JM, Mateo Rodríguez O, Frutos Pérez B, Farfán Sedano A, Cabello Clotet N, et al. Fiebre en urgencias: detección de patología grave. Rev Clin Esp. 2008;208(3):130-4.

Recebido em: 29/06/2019
Aceito para publicação em: 11/09/2019
Conflito de interesse: nenhum.
Fonte de financiamento: nenhuma.

Endereço para correspondência:
Felipe Antônio Cacciatori
E-mail: felipe_cacciatori@hotmail.com

CC BY