

<https://helda.helsinki.fi>

---

## Early prediction of acute necrotizing pancreatitis by artificial intelligence : a prospective cohort-analysis of 2387 cases

Hungarian Pancreatic Study Group

2022-12-01

---

Hungarian Pancreatic Study Group , Kiss , S & Szentesi , A 2022 , ' Early prediction of acute necrotizing pancreatitis by artificial intelligence : a prospective cohort-analysis of 2387 cases ' , Scientific Reports , vol. 12 , no. 1 , 7827 . <https://doi.org/10.1038/s41598-022-11517-w>

---

<http://hdl.handle.net/10138/355011>

<https://doi.org/10.1038/s41598-022-11517-w>

---

cc\_by

publishedVersion

---

*Downloaded from Helda, University of Helsinki institutional repository.*

*This is an electronic reprint of the original article.*

*This reprint may differ from the original in pagination and typographic detail.*

*Please cite the original version.*



OPEN

## Early prediction of acute necrotizing pancreatitis by artificial intelligence: a prospective cohort-analysis of 2387 cases

Szabolcs Kiss<sup>1,2,3</sup>, József Pintér<sup>4</sup>, Roland Molontay<sup>4,5</sup>, Marcell Nagy<sup>4</sup>, Nelli Farkas<sup>2,6</sup>, Zoltán Sipos<sup>2</sup>, Péter Fehérvári<sup>2,7</sup>, László Pecze<sup>2</sup>, Mária Földi<sup>1,2,3</sup>, Áron Vincze<sup>8</sup>, Tamás Takács<sup>9</sup>, László Czakó<sup>9</sup>, Ferenc Izbéki<sup>10</sup>, Adrienn Halász<sup>1,10</sup>, Eszter Boros<sup>10</sup>, József Hamvas<sup>11</sup>, Márta Varga<sup>12</sup>, Artautas Mickevicius<sup>13</sup>, Nándor Faluhelyi<sup>14</sup>, Orsolya Farkas<sup>14</sup>, Szilárd Váncsa<sup>2,15</sup>, Rita Nagy<sup>2,3,15</sup>, Stefania Bunduc<sup>15,16</sup>, Péter Jenő Hegyi<sup>15,17</sup>, Katalin Márta<sup>15,17</sup>, Katalin Borka<sup>15,18</sup>, Attila Doros<sup>15,19</sup>, Nóra Hosszúfalusi<sup>15,20</sup>, László Zubeck<sup>15,21</sup>, Bálint Erőss<sup>15,17</sup>, Zsolt Molnár<sup>15,21,22</sup>, Andrea Párniczky<sup>2,3</sup>, Péter Hegyi<sup>15,17,33</sup>, Andrea Szentesi<sup>1,2,33</sup>✉ & Hungarian Pancreatic Study Group<sup>17\*</sup>

Pancreatic necrosis is a consistent prognostic factor in acute pancreatitis (AP). However, the clinical scores currently in use are either too complicated or require data that are unavailable on admission or lack sufficient predictive value. We therefore aimed to develop a tool to aid in necrosis prediction. The XGBoost machine learning algorithm processed data from 2387 patients with AP. The confidence of the model was estimated by a bootstrapping method and interpreted via the 10th and the 90th percentiles of the prediction scores. Shapley Additive exPlanations (SHAP) values were calculated to quantify the contribution of each variable provided. Finally, the model was implemented as an online application using the Streamlit Python-based framework. The XGBoost classifier provided an

<sup>1</sup>Doctoral School of Clinical Medicine, Faculty of Medicine, University of Szeged, Szeged 6720, Hungary. <sup>2</sup>Institute for Translational Medicine, Szentágotthai Research Centre, Medical School, University of Pécs, Szigeti út 12., II. Emelet, Pécs 7624, Hungary. <sup>3</sup>Heim Pál National Pediatric Institute, Üllői út 86, Budapest 1089, Hungary. <sup>4</sup>Human and Social Data Science Lab, Budapest University of Technology and Economics, Műgyetem rkp. 3, Budapest 1111, Hungary. <sup>5</sup>Stochastics Research Group, Hungarian Academy of Sciences, Budapest University of Technology and Economics, Egy József u. 1, Budapest 1111, Hungary. <sup>6</sup>Institute of Bioanalysis, Medical School, University of Pécs, Honvéd u. 1, Pécs 7624, Hungary. <sup>7</sup>Department of Biomathematics and Informatics, University of Veterinary Medicine, István u. 2, Budapest 1078, Hungary. <sup>8</sup>Division of Gastroenterology, First Department of Medicine, Medical School, University of Pécs, Ifjúság út 13, Pécs 7624, Hungary. <sup>9</sup>Department of Medicine, University of Szeged, Kálvária sgt. 57, Szeged 6725, Hungary. <sup>10</sup>Department of Internal Medicine, Szent György Teaching Hospital of County Fejér, Seregélyesi út 3, Székesfehérvár 8000, Hungary. <sup>11</sup>Bajcsy-Zsilinszky Hospital, Maglódi út 89-91, Budapest 1106, Hungary. <sup>12</sup>Department of Gastroenterology, BMKK Dr Rethy Pal Hospital, Gyulai út 18, Békéscsaba 5600, Hungary. <sup>13</sup>Vilnius University Hospital Santaros Clinics, Clinics of Abdominal Surgery, Nephrourology and Gastroenterology, Faculty of Medicine, Vilnius University, Santariškių g. 2, 08410 Vilnius, Lithuania. <sup>14</sup>Department of Medical Imaging, Medical School, University of Pécs, Ifjúság út 13, Pécs 7624, Hungary. <sup>15</sup>Centre for Translational Medicine, Semmelweis University, Üllői út 26, Budapest 1085, Hungary. <sup>16</sup>Doctoral School, Carol Davila University of Medicine and Pharmacy, Bulevardul Eroii Sanitari 8, 050474 Bucharest, Romania. <sup>17</sup>Division of Pancreatic Diseases, Heart and Vascular Center, Semmelweis University, Baross u. 23, Budapest 1082, Hungary. <sup>18</sup>2nd Department of Pathology, Semmelweis University, Üllői út 93, Budapest 1091, Hungary. <sup>19</sup>Department of Transplantation and Surgery, Semmelweis University, Baross u. 23, Budapest 1082, Hungary. <sup>20</sup>Department of Internal Medicine and Hematology, Semmelweis University, Szentkirályi u. 46, Budapest 1088, Hungary. <sup>21</sup>Department of Anaesthesiology and Intensive Therapy, Semmelweis University, Üllői út 78, Budapest 1082, Hungary. <sup>22</sup>Department of Anaesthesiology and Intensive Therapy, Poznan University of Medical Sciences, ul. św. Marii Magdaleny 14, 61861 Poznan, Wielkopolska, Poland. <sup>33</sup>These authors contributed equally: Péter Hegyi and Andrea Szentesi. \*A list of authors and their affiliations appears at the end of the paper. ✉email: szentesiai@gmail.com

**AUC value of 0.757. Glucose, C-reactive protein, alkaline phosphatase, gender and total white blood cell count have the most impact on prediction based on the SHAP values. The relationship between the size of the training dataset and model performance shows that prediction performance can be improved. This study combines necrosis prediction and artificial intelligence. The predictive potential of this model is comparable to the current clinical scoring systems and has several advantages over them.**

Acute pancreatitis (AP) affects about 34 per 100,000 people per year, and it is the most frequent gastrointestinal disease requiring acute hospitalization<sup>1,2</sup>. The overall mortality is around 3%<sup>3,4</sup>; however, in about 10–20% of AP cases, acute necrotizing pancreatitis (ANP) develops, thus further increasing the risk of morbidity and mortality<sup>5,6</sup>. The overall mortality of ANP is approximately 15–20%, of which there is a further twofold increase in a third of ANP cases where the necrotic tissue becomes infected<sup>7,8</sup>.

Early appraisal of severity and prognosis is crucial in AP, particularly on clinical admission, to identify patients at risk of developing life-threatening complications. In these cases, close monitoring and early intervention may prevent organ dysfunction and a fatal outcome<sup>9,10</sup>.

It has long been known that necrosis is a consistent prognostic factor in AP<sup>9</sup>. The diagnosis of this local complication strongly relies on contrast-enhanced computer tomography (CECT) because it has a much higher sensitivity to detect ANP than ultrasonography<sup>7</sup>. Despite being the gold standard method for diagnosing ANP, CECT has many disadvantages: (1) ANP usually becomes apparent only 72 h after the onset of symptoms; (2) early and inappropriate CECT may prolong hospitalization; and (3) it is not accessible in every case<sup>11</sup>. There is therefore a need for other methods to supplement ANP assessment.

As the underlying pathophysiology of AP becomes more and more familiar by the accumulation of scientific data, several potential therapeutic targets have been identified<sup>12,13</sup>. Since some of these specific therapies may be available soon, prompt initiation of treatment after early identification of ANP could be even more important.

Since ANP is associated with life-threatening complications and increased mortality and it is the principal determinant of the incidence of secondary infection in AP<sup>14</sup>, researchers have endeavored to find an accurate clinical scoring system or biomarker that can predict ANP, the severe disease course or mortality itself. As regards ANP, these systems are either too complicated or require data that is unavailable in the initial stage of hospitalization or lack sufficient sensitivity and specificity. They are therefore rarely used in everyday clinical practice.

As artificial intelligence (AI) can overcome the limitations provided by the complexity of the data and time-dependent variables, the number of AI tools is increasing in medicine<sup>15</sup>. AI applications in pancreatic diseases are also evolving quickly<sup>16</sup>. Four AI models aimed to predict the severity of AP on clinical admission, all of which seem to outperform the conventional prediction scores<sup>17–19</sup>. Despite their promising preliminary results, these AI tools are limited by the overlap between the patient group used for model preparation and internal validation and the relatively low patient number.

This study has two main goals: first, to overcome these limitations and build an AI model that provides an accurate prediction for ANP development; and, second, to create an online tool from the model that could aid physicians in the early prognosis of AP.

## Methods

This study was reported following the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) Statement<sup>20</sup>. Ethics approval was obtained from the Hungarian Medical Research Council's Scientific and Research Ethics Committee (22254-1/2012/EKU, 17787-8/2020/EÜIG). Written informed consent was obtained from all participants before enrolment. The study was conducted in accordance with the Helsinki Declaration.

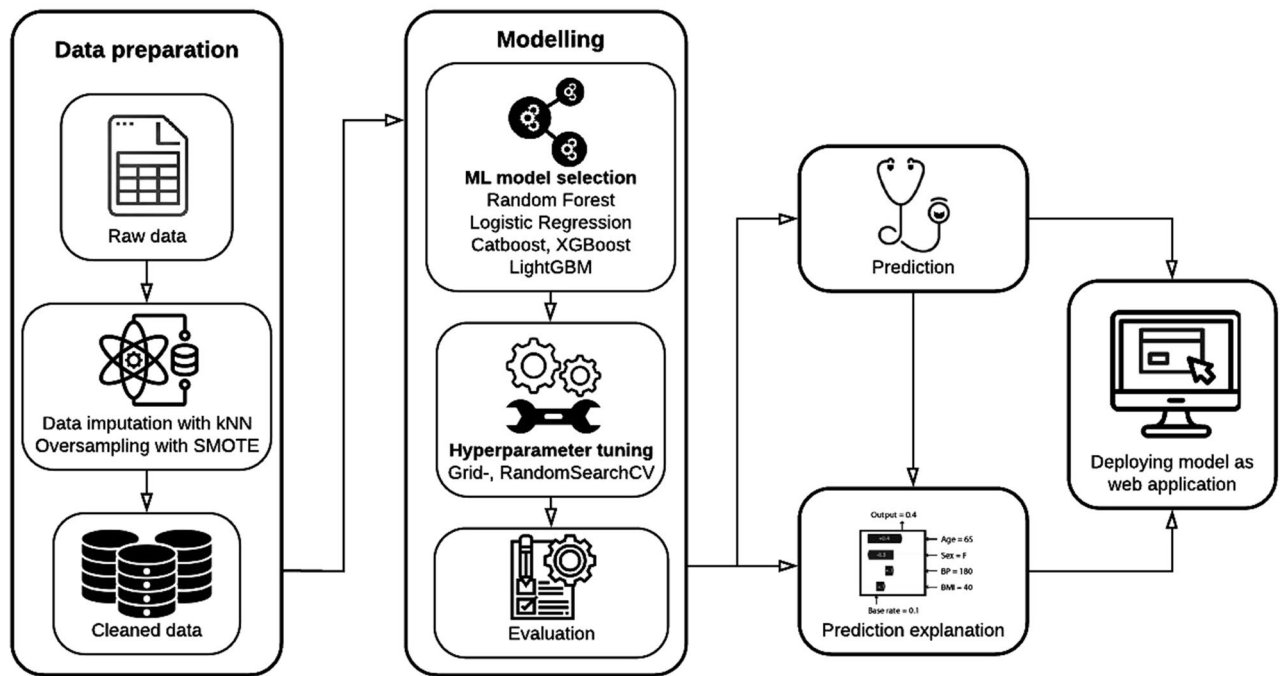
**Data source and eligibility criteria.** The analyzed dataset was collected by the Hungarian Pancreatic Study Group between 2012 and 2019. There were 2461 adult patients enrolled in the patient registry from 30 centers across 13 countries (Appendix A). All patients fulfilled two out of three AP diagnostic criteria based on the revised Atlanta classification<sup>21</sup>. Data were collected by physicians and trained clinical administrators on admission and each day during the whole hospital stay and were stored both on paper and electronically. Relevant clinical data underwent a four-level quality check system before analysis.

In all cases deemed eligible a CECT was performed during hospitalization to assess pancreatic necrosis formation. Exclusion criteria were as follows: (1) no pancreas imaging had been performed; and (2) the mere suspicion of necrosis formation by imaging, which was not confirmed later by CECT.

**Groups, outcomes, and predictors analyzed.** Eligible participants were divided into two groups: (1) pancreatic necrosis formation was confirmed by a radiologist by CECT during hospitalization; and (2) absence of necrosis development. The dataset was analyzed and compared accordingly.

ANP was defined as lack of parenchymal enhancement or findings of peripancreatic necrosis such as an acute necrotic collection on CECT<sup>22</sup>. Other local (acute peripancreatic fluid collection and pseudocyst) and systemic (new-onset diabetes, heart failure, renal failure, and respiratory failure) complications and disease severity were defined based on the revised Atlanta classification<sup>21</sup>. Data on in-hospital mortality, length of hospital stay, and etiology of AP were also collected.

The assessed predictors of ANP were gender, age, body mass index (BMI), and laboratory parameters measured in the first 24 h of clinical admission. The following were evaluated: alanine transaminase, albumin, amylase,



**Figure 1.** Flowchart representing the process of developing the model.

alkaline phosphatase (ALP), aspartate transaminase, blood urea nitrogen, calcium, C-reactive protein (CRP), creatinine, direct bilirubin, gamma-glutamyl transferase (GGT), glucose, estimated glomerular filtration rate (eGFR), glycosylated hemoglobin (HbA1c), hematocrit, hemoglobin, lactate dehydrogenase (LDH), lipase, potassium, procalcitonin, red blood cell count, sodium, thrombocyte, total bilirubin, total cholesterol, total protein, total white blood cell count (WBC), and triglyceride.

**Predictive modelling.** The process of predictive modelling is depicted in Fig. 1. Thirty-one variables have been used for modelling. Data quality is provided in Appendix A. Missing data were handled with a k-nearest-neighbor-based data imputer algorithm (KNNImputer)<sup>23</sup>. The SMOTE algorithm<sup>24</sup> was used to deal with the imbalance in class distribution (number of patients with and without ANP).

Random Forest, Logistic Regression, Catboost, XGBoost, and LightGBM were tested for modelling to identify the best performing machine learning algorithm<sup>25–28</sup>. The catboost, xgboost, lightgbm, and scikit-learn Python packages were applied. The optimal model was chosen based on the receiver operating characteristic (ROC) curve and area under the ROC curve (AUC) value after performing four-fold cross-validation. The confidence of the best performing model was estimated with a bootstrapping method, namely by re-sampling the training dataset and training a hundred independent copies of the model on these datasets. The confidence of the model prediction was interpreted with the aid of the 10th and the 90th percentiles of the prediction scores.

Shapley Additive exPlanations (SHAP) values were calculated<sup>29</sup> to locally explain the model prediction and to quantify the contribution of each variable provided. Finally, the model was deployed as an online application using the Streamlit Python-based framework.

**Other statistical analyses.** The presence of sampling bias was tested by assessing the representativeness between the cohort analyzed and the whole cohort (Appendix A). The prediction parameters were also compared between patients with and without ANP with the Kolmogorov–Smirnov test and the Chi-squared test. ANP was tested as a risk factor for mortality, severe AP, and local and systemic complications by calculating risk ratios (RR) with the corresponding 95% confidence interval (CI).

**Ethics approval.** Ethics approval was obtained from the Hungarian Medical Research Council's Scientific and Research Ethics Committee (22254-1/2012/EKU, 17787-8/2020/EÜIG). The study was conducted in accordance with the Helsinki Declaration.

**Consent to participate.** Written informed consent was obtained from all participants before enrolment.

**Consent for publication.** The corresponding author accepts responsibility for releasing this material on behalf of all co-authors.

Variable	Value (n = 2387)
Age in years, median (IQR)	57 (44–69)
Male, n (%)	1357 (56.85%)
BMI, median (IQR)	27.14 (23.88–31.25)
<b>Etiology, n (%)</b>	
Biliary	955 (40.01%)
Alcoholic	484 (20.28%)
Hypertriglyceridaemia	81 (3.39%)
Biliary and alcoholic	39 (1.63%)
Biliary and hypertriglyceridaemia	13 (0.54%)
Alcoholic and hypertriglyceridaemia	58 (2.43%)
Post-ERCP	67 (2.81%)
Idiopathic	432 (18.10%)
Other	258 (10.81%)
<b>Revised Atlanta classification</b>	
Mild, n (%)	1714 (71.81%)
Moderate, n (%)	551 (23.08%)
Severe, n (%)	122 (5.11%)
Mortality, n (%)	66 (2.76%)
Length of stay in days, median (IQR)	8 (6–12)
Patients with local complication, n (%)	623 (26.19%)
APFC, n (%)	510 (21.37%)
Pseudocyst, n (%)	179 (7.50%)
Acute necrotic collection, n (%)	233 (9.76%)
Patients with systemic complication, n (%)	202 (8.46%)
Respiratory failure, n (%)	136 (5.70%)
Heart failure, n (%)	52 (2.18%)
Renal failure, n (%)	83 (3.48%)
New-onset diabetes, n (%)	75 (3.14%)

**Table 1.** Characteristics of the analyzed study population. *APFC* acute peripancreatic fluid collection, *BMI* body mass index, *ERCP* endoscopic retrograde cholangiopancreatography, *IQR* interquartile range.

## Results

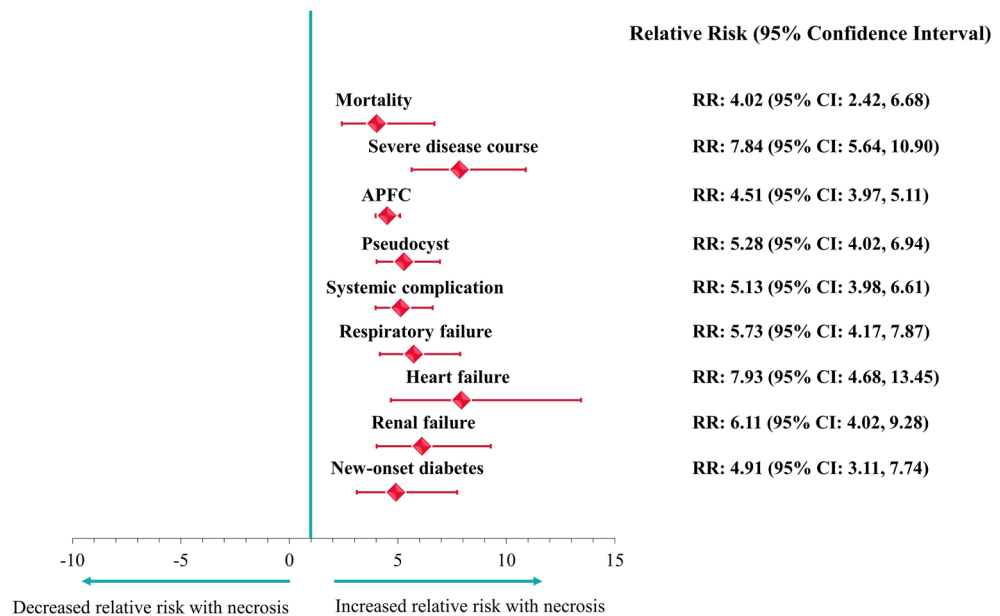
**Characteristics of the cohort analyzed.** 2387 of the 2461 patients with AP proved to be eligible for the analysis. Characteristics of this population are summarized in Table 1. In 9.76% of the cases, ANP was confirmed. There was a statistically significant difference between patients with and without ANP as regards age, gender, and BMI (Appendix B Supplementary Figs. 16–18). A detailed analysis of the results as regards other biomarkers can be found in Appendix B.

ANP was associated with a significantly higher risk for mortality, severe disease course, and all the investigated local and systemic complications (Fig. 2). ANP was also associated with longer hospitalization ( $9.13 \pm 6.21$  days vs.  $20.78 \pm 19.70$  days,  $p < 0.001$ ).

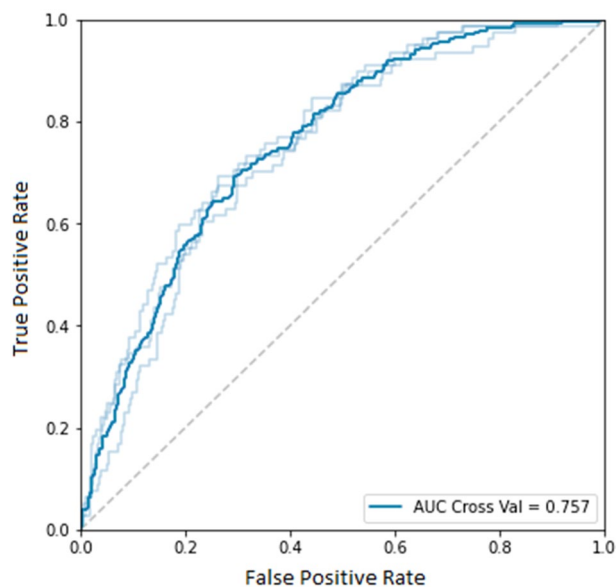
**Model selection and model performance.** After an evaluation of the machine learning algorithms, an XGBoost classifier was identified as the best performing model with an AUC value of 0.757 (standard deviation: 0.012) on cross-validation (Fig. 3). The relationship between the size of the data set and the model performance is depicted in Fig. 4. The steady increase of AUC values implies that our model has not yet reached its maximal prediction performance. Internal validation implies that our model has higher reliability near the endpoints of the prediction spectrum since the confidence intervals are narrower (Fig. 5).

The assessment of the impact on the model output showed that glucose, CRP, ALP, gender, and WBC have the five highest SHAP values. The most influential predictors are shown in Fig. 6 Panel A. Our assessment showed that the predictive potential depends on the number of biomarkers provided. The models built on the top  $k$  most influential predictors according to their SHAP values show an increasing performance as regards the predictive potential; however, the extent of this improvement decreases with the number of variables provided (Fig. 7).

**Application.** The current version of the model can be accessed at <http://necro-app.org/>. At least five of the available predictors must be provided to use the application. This limit was applied based on the relation between the size of the dataset and the desired accuracy<sup>30</sup>. The application is aided by a built-in BMI calculator and validations to filter out invalid values. The model offers a numerical probability value between 0 and 1. The higher the number, the higher the risk for ANP becomes. These numerical values are also supplied with a textual interpretation. For educational purposes, the effect of the biomarkers on prediction is also indicated (Fig. 6 Panel B).



**Figure 2.** Association between necrosis development and other complications in acute pancreatitis.



**Figure 3.** Receiver operating characteristic (ROC) curve for the XGBoost model.

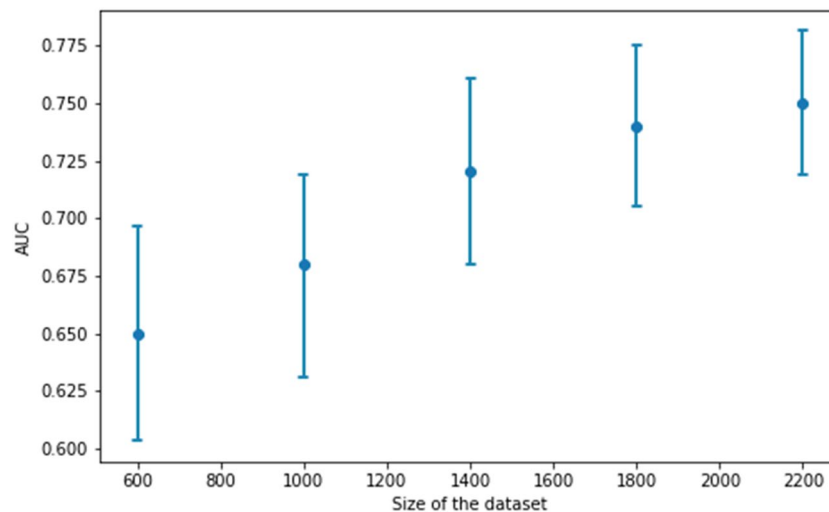
By checking an extra field, the application assigns a confidence interval in addition to the numerical value. This adds further clarification to the predicted necrosis probability; however, it takes extra time.

## Discussion

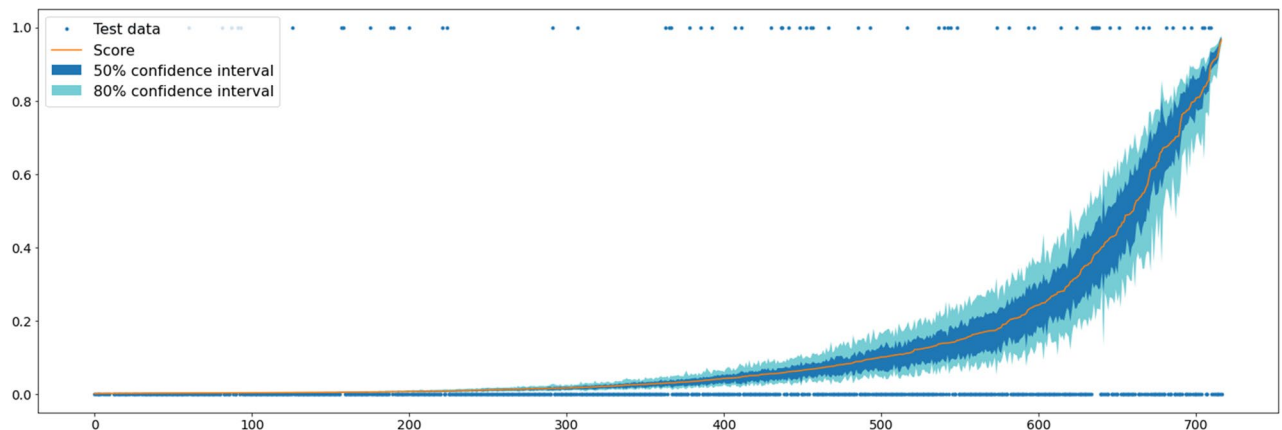
The current study describes the first AI model designed to predict ANP. In addition to creating this model, we also implemented it as an easily accessible online tool. In addition to these, ANP was extensively described in a large, prospective, multicenter, cohort study.

**Our cohort in the context of previous data.** With the occurrence of ANP in around one-tenth of patients, our results are comparable with previously reported data<sup>31,32</sup>. The importance of ANP in determining the disease course and outcome is well-known<sup>33,34</sup>. Schepers et al. found that 38% of the patients with ANP developed respiratory, cardiovascular or renal system failure<sup>35</sup>. In our cohort, necrosis was also associated with a four to eight-fold increased risk of local and systematic complications, severe disease course, and mortality. We also confirmed their observation regarding prolonged hospitalization indicating the impact of ANP on short-





**Figure 4.** The relationship between the size of the data set and the model performance. The blue dot represents the area under the ROC curve value and the vertical lines show the corresponding confidence intervals.

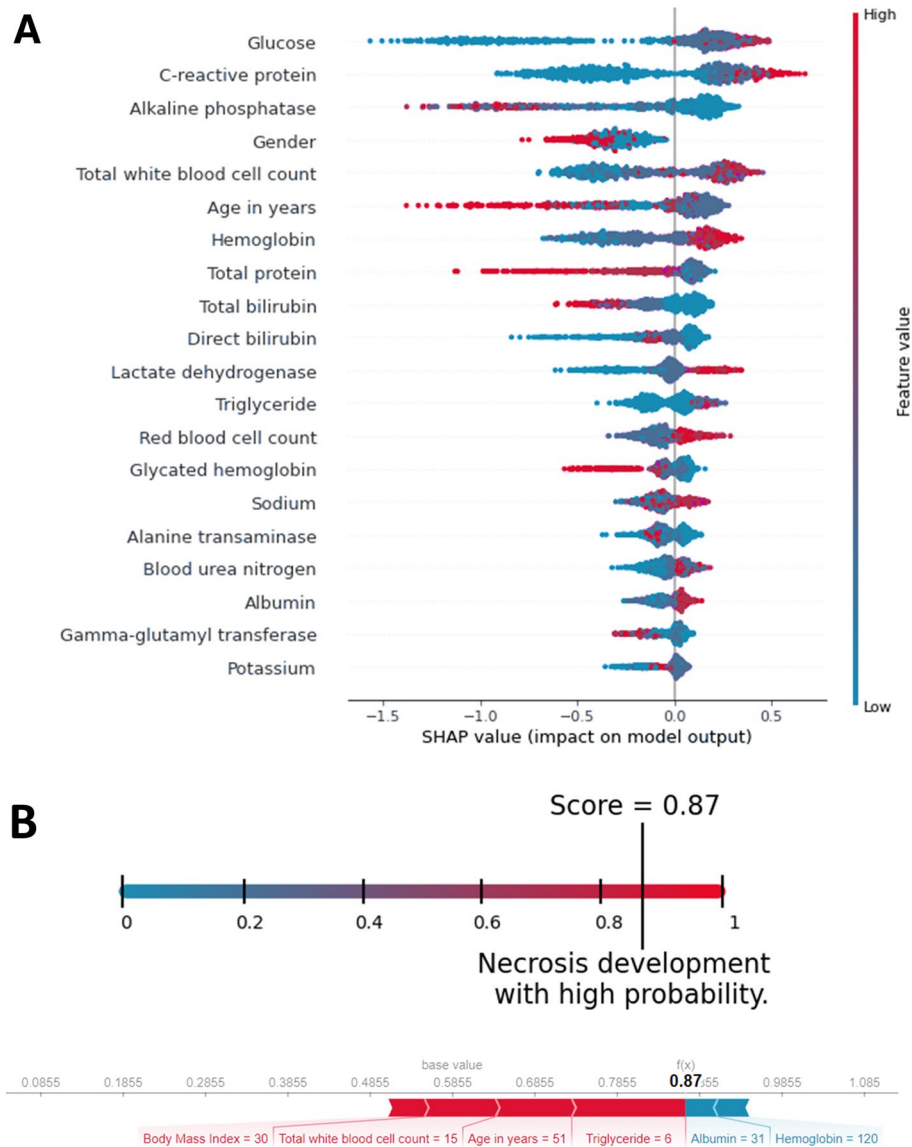


**Figure 5.** The predicted necrosis probabilities with the corresponding 50% (between the 25th and 50th percentiles) and 80% confidence (between the 10th and 90th percentiles).

term (i.e.: in-hospital) outcomes. However, the importance of pancreatic necrosis development also lies in the long-term complications.

Recent studies investigated this topic and shed light on long-term outcomes. A meta-analysis of long-term follow-up studies found that the pooled prevalence of exocrine pancreatic insufficiency (EPI) after ANP is between 41 and 58% depending on the extent of necrosis<sup>36</sup>. In a cohort study by Maatman et al., this ratio was only 19%<sup>22</sup>. The discrepancy in the frequency can be attributed to that. While the meta-analysis accounted for EPI during both the hospital stay and follow-up, the cohort assessed EPI after the resolution of AP. Furthermore, the retrospective nature of data has an inherent limitation, which can also explain this difference. In addition to the increased frequency of EPI, they found endocrine insufficiency in 35% of the patients with a median follow-up of 46 months. Despite the fact that our study covered the time of hospitalization, our results imply that necrosis formation increases the risk of new-onset diabetes.

**Currently existing clinical scores as predictors of necrosis.** Since ANP is a potent prognostic factor for the short-term severity of AP and could forecast long-term consequences, it would be ideal for identifying these patients as soon as possible. The prediction of ANP was attempted by numerous scoring systems and biomarkers<sup>37</sup>; however, each of them has its own limitations. The Balthazar Computer Tomography Severity Index (CTSI) possesses a higher positive predictive value for necrosis than most commonly used prediction methods<sup>38</sup>, e.g. the Ranson score and the APACHE II score, but it is limited by the availability of CECT. It must be noted that ANP usually becomes apparent after two to three days after disease onset, and that prevents on-admission prediction in certain cases. The application of other scoring systems without mandatory CECT is restricted by their complexity. The Ranson score has eleven factors, which have to be assessed on admission and



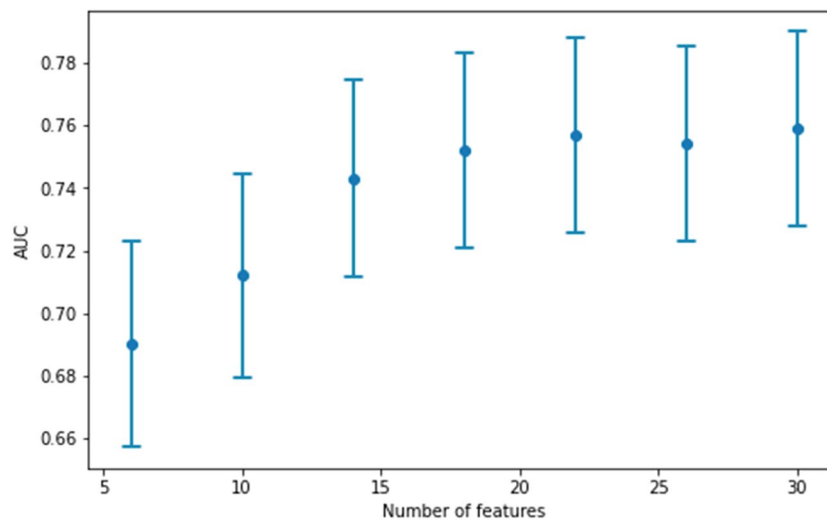
**Figure 6.** (A) The features with the highest impact on model output based on the SHAP values. The higher the predictor is on the list, the bigger the impact on model output. Each patient is represented by a dot. The x-axis represents the extent of the impact on prediction. The color of the dot shows the feature value (e.g. the red color implies higher values). (B) An example of prediction and its textual interpretation. The lower picture highlights the effect of individual predictors and the final necrosis probability provided by the model.

after 48 hours<sup>39</sup>. The APACHE II score is superior to the scores noted above in terms of flexibility and speed; however, its sensitivity and specificity are far lower<sup>40</sup>.

Two prospective studies compared CTSI, Ranson score, and APACHE II score in predicting necrosis development<sup>41,42</sup>. Despite limitation in terms of patient number and the slightly different AUC values for necrosis, they concluded that the positive predictive value decreases in the following order: CTSI, Ranson score, and APACHE II. It must be emphasized that these scoring systems are strongly limited by the conversion of continuous variables to binary ones and this topic should be investigated by more mathematical models with better accuracy<sup>42</sup>.

**Artificial intelligence in the prognosis of acute pancreatitis.** Artificial intelligence has appeared on the scene as a very intriguing modality of data-based decision support, and these models are extensively researched in numerous areas of medicine, including pancreaticobiliary diseases<sup>43</sup>. In the last decade, multiple AI algorithms have been developed in AP<sup>16</sup>. Most of these models were designed to predict the occurrence of a





**Figure 7.** The models build on the  $k$  predictors with the highest SHAP value.

specific complication or disease severity. The most commonly used score in critical care is the APACHE II score; however, three AP severity AI models have been reported to outperform this score<sup>17–19</sup>. The AI model developed by Keogan et al. was compared to the CTSI and Ranson scores, both of which were found inferior in terms of predicting the severity of AP<sup>44</sup>. It should be noted that this study assessed the disease severity with LOH and not with the revised Atlanta classification. Despite the positive results, these prediction systems, except for the artificial neural network by Mofidi et al.<sup>19</sup>, are limited by the overlap between the data used for model training and the validation. Furthermore, these models need another step after validation. Despite the tremendous efforts and scientific results, much of this knowledge has not been applied in everyday clinical practice<sup>45</sup>. In order to bring these complex models to the bedside, they need to be implemented as easy to use and broadly accessible tools<sup>46</sup>.

The current study was not designed to predict severity but to assess the probability of necrosis formation on clinical admission. Although we had a different outcome, we aimed to overcome the limitations of most previous models and to find a way to use our AI model. As suggested by Shung et al., AI-assisted tools have to overcome many challenges<sup>46</sup>. First of all, we must have high-quality data. This issue was addressed in our study with a four-level data quality check system. The second main challenge is ongoing data maintenance. Our model was constructed such that the new data could be incorporated after validation. Since the predictive potential of the model shows an increasing trend, this could contribute to better accuracy. Algorithmic understanding is also a key factor. The help of physicians, who will eventually use the AI model, is crucial to confirm the performance of such a tool. Furthermore, practitioners could help in differentiating between valid prediction with actual signals and distorted predictions masked by confounding variables<sup>46</sup>. Our web-based application shows the weighted impact of the individual biomarkers in each decision. This tool thus meets these expectations. Consequently, the next step will be screening for these confounding factors while continuously incorporating new data and monitoring the feasibility of the bedside application of this model.

**Strengths and limitations.** Our study has multiple strengths and some limitations. Although the predictive potential of this model is similar to that of currently available predictive scoring systems, it has multiple advantages over them. It provides risk assessment with any five of the predictors in our study, which are commonly assessed in daily practice. Therefore, this better reflects everyday clinical practice. To the best of our knowledge, this is the first AI model to strive to predict the development of ANP on clinical admission. We designed our model on a much larger population, as compared to the already existing prognostic AI models in AP, and there was no overlap between the original and validation population. Furthermore, we placed great emphasis on the interpretation of the model for physicians and its implementation by creating an online application. Nevertheless, in addition to predictive model development, ANP was extensively analyzed.

In addition to these strengths, the present study has several limitations. Firstly, as we move further from the endpoint of the prediction spectrum, the confidence of the model becomes wider, and prediction becomes less reliable. Secondly, the cross-validated AUC value of our XGBoost model is currently in the fair range<sup>47</sup>. Thirdly, data imputation can also introduce bias. Most of these limitations can be overcome. Based on our analyses, we could reach better predictive potential by increasing the training sample size and more data could provide more accurate imputation as well. Therefore, by using the application, making further predictions with more data, the model itself could improve.

It should be highlighted that AI models should not be considered as a substitute for human intelligence<sup>16</sup>. These tools, including our model, were designed to facilitate physicians' decision-making and every prediction should be interpreted in accordance with the clinical picture.

**Implication for practice and research.** Development of ANP is associated with several short- and long-term complications, e.g. endocrine insufficiency, but CECT is not performed solely and exclusively to confirm necrosis in AP. Therefore, by knowing the high risk for necrosis development, we can identify a group of patients who need closer follow-up. Nevertheless, this model can aid physicians when CECT is either contraindicated or not available. Also, as soon as new therapies emerge, early identification of ANP will become even more important. Further research is needed on other potential predictive factors, which could be incorporated in the current model to further improve predictions.

## Conclusion

This study is the first to combine prediction of necrosis development and artificial intelligence in AP. The predictive potential of this model is comparable to the already existing clinical scoring systems and the model is expected to further improve with use. The easy-to-use web application supported by the interpretation of the prediction facilitates early, on-admission prediction of necrosis and allows continuous data maintenance and algorithmic understanding.

## Data availability

The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 14 December 2021; Accepted: 7 April 2022

Published online: 12 May 2022

## References

- Boxhoorn, L. *et al.* Acute pancreatitis. *Lancet (London, England)* **396**, 726–734. [https://doi.org/10.1016/s0140-6736\(20\)31310-6](https://doi.org/10.1016/s0140-6736(20)31310-6) (2020).
- Xiao, A. Y. *et al.* Global incidence and mortality of pancreatic diseases: A systematic review, meta-analysis, and meta-regression of population-based cohort studies. *Lancet Gastroenterol. Hepatol.* **1**, 45–55. [https://doi.org/10.1016/s2468-1253\(16\)30004-8](https://doi.org/10.1016/s2468-1253(16)30004-8) (2016).
- Berger, Z. *et al.* Acute pancreatitis in Chile. A multicenter study on epidemiology, etiology and clinical outcome. Retrospective analysis of clinical files. *Pancreatology* **20**, 637–643. <https://doi.org/10.1016/j.pan.2020.04.016> (2020).
- Párniczky, A. *et al.* Prospective, multicentre, nationwide clinical data from 600 cases of acute pancreatitis. *PLoS ONE* **11**, e0165309. <https://doi.org/10.1371/journal.pone.0165309> (2016).
- Aranda-Narvaez, J. M., Gonzalez-Sanchez, A. J., Montiel-Casado, M. C., Titos-Garcia, A. & Santoyo-Santoyo, J. Acute necrotizing pancreatitis: Surgical indications and technical procedures. *World J. Clin. Cases* **2**, 840–845. <https://doi.org/10.12998/wjcc.v2.i12.840> (2014).
- Jha, A. K., Goenka, M. K., Kumar, R. & Suchismita, A. Endotherapy for pancreatic necrosis: An update. *JGH Open* **3**, 80–88. <https://doi.org/10.1002/jgh3.12109> (2019).
- Bugiantella, W. *et al.* Necrotizing pancreatitis: A review of the interventions. *Int. J. Surg.* **28**(Suppl 1), S163–171. <https://doi.org/10.1016/j.ijssu.2015.12.038> (2016).
- Colvin, S. D., Smith, E. N., Morgan, D. E. & Porter, K. K. Acute pancreatitis: An update on the revised Atlanta classification. *Abdominal Radiol. (New York)* **45**, 1222–1231. <https://doi.org/10.1007/s00261-019-02214-w> (2020).
- Balthazar, E. J., Robinson, D. L., Megibow, A. J. & Ranson, J. H. Acute pancreatitis: Value of CT in establishing prognosis. *Radiology* **174**, 331–336. <https://doi.org/10.1148/radiology.174.2.2296641> (1990).
- Ding, N., Guo, C., Li, C., Zhou, Y. & Chai, X. An artificial neural networks model for early predicting in-hospital mortality in acute pancreatitis in MIMIC-III. *Biomed. Res. Int.* **2021**, 6638919. <https://doi.org/10.1155/2021/6638919> (2021).
- IAP/APA evidence-based guidelines for the management of acute pancreatitis. *Pancreatology* **13**, e1–e15. <https://doi.org/10.1016/j.pan.2013.07.063> (2013).
- Li, J., Zhou, R., Zhang, J. & Li, Z. F. Calcium signaling of pancreatic acinar cells in the pathogenesis of pancreatitis. *World J. Gastroenterol.* **20**, 16146–16152. <https://doi.org/10.3748/wjg.v20.i43.16146> (2014).
- Petersen, O. H., Gerasimenko, J. V., Gerasimenko, O. V., Gryshchenko, O. & Peng, S. The roles of calcium and ATP in the physiology and pathology of the exocrine pancreas. *Phys. Rev.* **101**, 1691–1744. <https://doi.org/10.1152/physrev.00003.2021> (2021).
- Bradley, E. L. 3rd., Murphy, F. & Ferguson, C. Prediction of pancreatic necrosis by dynamic pancreatography. *Ann. Surg.* **210**, 495–504. <https://doi.org/10.1097/0000658-198910000-00010> (1989).
- Hu, H.-T. *et al.* Artificial intelligence assists identifying malignant versus benign liver lesions using contrast-enhanced ultrasound. *J. Gastroenterol. Hepatol.* <https://doi.org/10.1111/jgh.15522> (2021).
- Gorris, M., Hoogenboom, S. A., Wallace, M. B. & van Hooft, J. E. Artificial intelligence for the management of pancreatic diseases. *Digest. Endosc.* **33**, 231–241. <https://doi.org/10.1111/den.13875> (2021).
- Andersson, B., Andersson, R., Ohlsson, M. & Nilsson, J. Prediction of severe acute pancreatitis at admission to hospital using artificial neural networks. *Pancreatology* **11**, 328–335. <https://doi.org/10.1159/000327903> (2011).
- Pearce, C. B., Gunn, S. R., Ahmed, A. & Johnson, C. D. Machine learning can improve prediction of severity in acute pancreatitis using admission values of APACHE II score and C-reactive protein. *Pancreatology* **6**, 123–131. <https://doi.org/10.1159/000090032> (2006).
- Mofidi, R., Duff, M. D., Madhavan, K. K., Garden, O. J. & Parks, R. W. Identification of severe acute pancreatitis using an artificial neural network. *Surgery* **141**, 59–66. <https://doi.org/10.1016/j.surg.2006.07.022> (2007).
- Collins, G. S., Reitsma, J. B., Altman, D. G. & Moons, K. G. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD statement. *Ann. Intern. Med.* **162**, 55–63. <https://doi.org/10.7326/m14-0697> (2015).
- Banks, P. A. *et al.* Classification of acute pancreatitis—2012: Revision of the Atlanta classification and definitions by international consensus. *Gut* **62**, 102–111. <https://doi.org/10.1136/gutjnl-2012-302779> (2013).
- Maatman, T. K. *et al.* The continuum of complications in survivors of necrotizing pancreatitis. *Surgery* **168**, 1032–1040. <https://doi.org/10.1016/j.surg.2020.07.004> (2020).
- Troyanskaya, O. *et al.* Missing value estimation methods for DNA microarrays. *Bioinformatics* **17**, 520–525. <https://doi.org/10.1093/bioinformatics/17.6.520> (2001).
- Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **16**, 321–357 (2002).
- Chen, T. & Guestrin, C. in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 (Association for Computing Machinery, 2016).

26. Dorogush, A. V., Ershov, V. & Gulin, A. J. A. CatBoost: Gradient boosting with categorical features support. [abs/1810.11363](https://arxiv.org/abs/1810.11363) (2018).
27. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825 (2012).
28. Ke, G. *et al.* *Proceedings of the 31st International Conference on Neural Information Processing Systems*. 3149–3157. (Curran Associates Inc., 2017).
29. Lundberg, S. M. & Lee, S.-I. *NIPS*.
30. Althnani, A. *et al.* Impact of dataset size on classification performance: An empirical evaluation in the medical domain. *Appl. Sci.* **11**, 796 (2021).
31. Baron, T. H., DiMaio, C. J., Wang, A. Y. & Morgan, K. A. American gastroenterological association clinical practice update: Management of pancreatic necrosis. *Gastroenterology* **158**, 67–75.e61. <https://doi.org/10.1053/j.gastro.2019.07.064> (2020).
32. Boumitri, C., Brown, E. & Kahaleh, M. Necrotizing pancreatitis: Current management and therapies. *Clin. Endosc.* **50**, 357–365. <https://doi.org/10.5946/ce.2016.152> (2017).
33. Aparna, D., Kumar, S. & Kamalkumar, S. Mortality and morbidity in necrotizing pancreatitis managed on principles of step-up approach: 7 years experience from a single surgical unit. *World J. Gastrointest. Surg.* **9**, 200–208. <https://doi.org/10.4240/wjgs.v9.i10.200> (2017).
34. Harris, H. W., Barcia, A., Schell, M. T., Thoeni, R. F. & Schechter, W. P. Necrotizing pancreatitis: A surgical approach independent of documented infection. *HPB (Oxford)* **6**, 161–168. <https://doi.org/10.1080/136518204100033634> (2004).
35. Schepers, N. J. *et al.* Impact of characteristics of organ failure and infected necrosis on mortality in necrotising pancreatitis. *Gut* **68**, 1044–1051. <https://doi.org/10.1136/gutjnl-2017-314657> (2019).
36. Huang, W. *et al.* Exocrine pancreatic insufficiency following acute pancreatitis: Systematic review and meta-analysis. *Dig. Dis. Sci.* **64**, 1985–2005. <https://doi.org/10.1007/s10620-019-05568-9> (2019).
37. Leonard-Murali, S. *et al.* Necrotizing pancreatitis: A review for the acute care surgeon. *Am. J. Surg.* **221**, 927–934. <https://doi.org/10.1016/j.amjsurg.2020.08.027> (2021).
38. Leung, T. K. *et al.* Balthazar computed tomography severity index is superior to Ranson criteria and APACHE II scoring system in predicting acute pancreatitis outcome. *World J. Gastroenterol.* **11**, 6049–6052. <https://doi.org/10.3748/wjg.v11.i38.6049> (2005).
39. Abu-Eshy, S. A., Abolfotouh, M. A., Nawar, E. & AbuSabib, A. R. Ranson's criteria for acute pancreatitis in high altitude: Do they need to be modified?. *Saudi J. Gastroenterol.* **14**, 20–23. <https://doi.org/10.4103/1319-3767.37797> (2008).
40. Lankisch, P. G. *et al.* The APACHE II score is unreliable to diagnose necrotizing pancreatitis on admission to hospital. *Pancreas* **24**, 217–222. <https://doi.org/10.1097/00006676-200204000-00002> (2002).
41. Khanna, A. K. *et al.* 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. <https://doi.org/10.1155/2013/367581> (2013).
42. Papachristou, G. I. *et al.* Comparison of BISAP, Ranson's, APACHE-II, and CTSI scores in predicting organ failure, complications, and mortality in acute pancreatitis. *Am. J. Gastroenterol.* **105**, 435–441. <https://doi.org/10.1038/ajg.2009.622> (2010) ((**quiz 442**)).
43. Chen, P.-T. *et al.* Applications of artificial intelligence in pancreatic and biliary diseases. *J. Gastroenterol. Hepatol.* **36**, 286–294. <https://doi.org/10.1111/jgh.15380> (2021).
44. Keogan, M. T. *et al.* Outcome analysis of patients with acute pancreatitis by using an artificial neural network. *Acad. Radiol.* **9**, 410–419. [https://doi.org/10.1016/s1076-6332\(03\)80186-1](https://doi.org/10.1016/s1076-6332(03)80186-1) (2002).
45. Hegyi, P., Eröss, B., Izbéki, F., Párniczky, A. & Szentesi, A. Accelerating the translational medicine cycle: The Academia Europaea pilot. *Nat. Med.* **27**, 1317–1319. <https://doi.org/10.1038/s41591-021-01458-8> (2021).
46. Shung, D. L. & Sung, J. J. Y. Challenges of developing artificial intelligence-assisted tools for clinical medicine. *J. Gastroenterol. Hepatol.* **36**, 295–298. <https://doi.org/10.1111/jgh.15378> (2021).
47. El Khouli, R. H. *et al.* Relationship of temporal resolution to diagnostic performance for dynamic contrast enhanced MRI of the breast. *J. Magn. Reson. Imaging* **30**, 999–1004. <https://doi.org/10.1002/jmri.21947> (2009).

## Acknowledgements

The research was supported by project grants K131996 to PH, FK131864 to AM, K128222 to LC and University of Pécs Medical School Research Fund (300909) to AS. The work of JP, RM and MN was supported by the Ministry of Innovation and the National Research, Development and Innovation Office within the framework of the Artificial Intelligence National Laboratory Programme. The funders had no effect on the concept, data collection and analysis, or writing of the manuscript. We wish to thank all contributing members of the Hungarian Pancreatic Study Group, a full list of them can be found in the Supplementary Information, Appendix A.

## Author contributions

S.K.: conceptualisation, project administration, investigation, data curation, visualization, writing-original draft; J.P.: formal analysis, methodology, visualization, writing-original draft; R.M.: formal analysis, methodology, visualization, writing-original draft; M.N.: formal analysis, methodology, visualization, writing-original draft; N.F.: conceptualisation, formal analysis, methodology, visualization, writing-original draft; Z.S.: formal analysis, methodology, visualization, writing-original draft; P.F.: formal analysis, methodology, visualization, writing-review & editing; L.P.: formal analysis, methodology, visualization, writing-review & editing; M.F.: conceptualisation, project administration, methodology, investigation, data curation, visualization, writing-original draft; Á.V.: data curation, writing-review & editing; T.T.: data curation, writing-review & editing; L.C.: investigation, data curation, writing-review & editing; F.L.: data curation, writing-review & editing; A.H.: data curation, investigation, writing-review & editing; E.B.: data curation, investigation, writing-review & editing; J.H.: data curation, writing-review & editing; M.V.: data curation, writing-review & editing; A.M.: conceptualisation, data curation, writing-review & editing; N.F.: conceptualisation, data curation, writing-review & editing; O.F.: conceptualisation, data curation, writing-review & editing; S.V.: conceptualisation, investigation, data curation, writing-review & editing; R.N.: investigation, data curation, writing-review & editing; S.B.: data curation, investigation, writing-review & editing; P.J.H.: data curation, investigation, data curation, writing-review & editing; K.M.: investigation, data curation, writing-review & editing; K.B.: investigation, writing-review & editing; A.D.: investigation, writing-review & editing; N.H.: investigation, writing-review & editing; L.Z.: investigation, writing-review & editing; B.E.: investigation, data curation, writing-review & editing; Z.M.: investigation, writing-review & editing; A.P.: investigation, data curation, writing-review & editing; P.H.: conceptualisation; supervision; investigation, data curation, funding acquisition, writing-original draft; A.S.: conceptualisation; supervision; investigation, data curation, funding acquisition, writing-original draft. All authors certify that they have participated sufficiently

in the work to take public responsibility for the content, including participation in the concept, design, analysis, writing, or revision of the manuscript.

## Funding

Open access funding provided by University of Pécs. The research was supported by project grants K131996 to PH, FK131864 to AM, K128222 to LC and University of Pécs Medical School Research Fund to AS. The work of JP, RM and MN was supported by the Ministry of Innovation and the National Research, Development and Innovation Office within the framework of the Artificial Intelligence National Laboratory Programme. The funders had no effect on the concept, data collection and analysis, or writing of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-11517-w>.

**Correspondence** and requests for materials should be addressed to A.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

## Hungarian Pancreatic Study Group

Szabolcs Kiss<sup>1,2,3</sup>, Nelli Farkas<sup>2,6</sup>, Zoltán Sipos<sup>2</sup>, Péter Fehérvári<sup>2,7</sup>, László Pecze<sup>2</sup>, Mária Földi<sup>1,2,3</sup>, Áron Vincze<sup>8</sup>, Tamás Takács<sup>9</sup>, László Czako<sup>9</sup>, Ferenc Izbéki<sup>10</sup>, Adrienn Halász<sup>1,10</sup>, Eszter Boros<sup>10</sup>, József Hamvas<sup>11</sup>, Márta Varga<sup>12</sup>, Artautas Mickevicius<sup>13</sup>, Nándor Faluhelyi<sup>14</sup>, Orsolya Farkas<sup>14</sup>, Szilárd Váncsa<sup>2,15</sup>, Rita Nagy<sup>2,3,15</sup>, Stefania Bunduc<sup>15,16</sup>, Péter Jenő Hegyi<sup>15,17</sup>, Katalin Márta<sup>15,17</sup>, Katalin Borka<sup>15,18</sup>, Attila Doros<sup>15,19</sup>, Nóra Hosszúfalusi<sup>15,20</sup>, László Zubek<sup>15,21</sup>, Bálint Erőss<sup>15,17</sup>, Zsolt Molnár<sup>15,21,22</sup>, Andrea Párniczky<sup>2,3</sup>, Péter Hegyi<sup>2,15,17</sup>, Andrea Szentesi<sup>1,2</sup>, Judit Bajor<sup>8</sup>, Szilárd Gódi<sup>8</sup>, Patrícia Sarlós<sup>8</sup>, József Czimmer<sup>8</sup>, Imre Szabó<sup>8</sup>, Gabriella Pár<sup>8</sup>, Anita Illés<sup>8</sup>, Roland Hágendorn<sup>8</sup>, Balázs Csaba Németh<sup>9</sup>, Balázs Kúri<sup>9</sup>, Dóra Illés<sup>9</sup>, László Gajdán<sup>10</sup>, Veronika Dunás-Varga<sup>10</sup>, Roland Fejes<sup>10</sup>, Mária Papp<sup>23</sup>, Zsuzsanna Vitális<sup>23</sup>, János Novák<sup>24</sup>, Imola Török<sup>25</sup>, Melania Macarie<sup>25</sup>, Elena Ramírez-Maldonado<sup>26</sup>, Ville Sallinen<sup>27</sup>, Shamil Galeev<sup>28</sup>, Barnabás Bod<sup>29</sup>, Ali Tüzün Ince<sup>30</sup>, Dániel Pécsi<sup>2</sup>, Péter Varjú<sup>2</sup>, Márk Félix Juhász<sup>2,3</sup>, Klementina Ocskay<sup>2,15</sup>, Alexandra Mikó<sup>2,31</sup> & Zsolt Szakács<sup>2,32</sup>

<sup>1</sup>Institute for Translational Medicine, Szentágothai Research Centre, Medical School, University of Pécs, Szigeti út 12., II. Emelet, Pécs 7624, Hungary. <sup>3</sup>Heim Pál National Pediatric Institute, Üllői út 86, Budapest 1089, Hungary. <sup>15</sup>Centre for Translational Medicine, Semmelweis University, Üllői út 26, Budapest 1085, Hungary. <sup>17</sup>Division of Pancreatic Diseases, Heart and Vascular Center, Semmelweis University, Baross u. 23, Budapest 1082, Hungary. <sup>21</sup>Department of Anaesthesiology and Intensive Therapy, Semmelweis University, Üllői út 78, Budapest 1082, Hungary. <sup>22</sup>Department of Anaesthesiology and Intensive Therapy, Poznan University of Medical Sciences, ul. św. Marii Magdaleny 14, 61861 Poznan, Wielkopolska, Poland. <sup>23</sup>Division of Gastroenterology, Department of Internal Medicine, University of Debrecen, Debrecen, Hungary. <sup>24</sup>Pándy Kálmán Hospital of Békés County, Gyula, Hungary. <sup>25</sup>County Emergency Clinical Hospital of Târgu Mures-Gastroenterology Clinic, University of Medicine, Pharmacy, Sciences and Technology "George Emil Palade", Targu Mures, Romania. <sup>26</sup>General Surgery, Consorci Sanitari del Garraf, Sant Pere de Ribes, Barcelona, Spain. <sup>27</sup>Department of Transplantation and Liver Surgery, Helsinki University Hospital, University of Helsinki, Helsinki, Finland. <sup>28</sup>Saint Luke Clinical Hospital, St. Petersburg, Russia. <sup>29</sup>Dr. Bugyi István Hospital, Szentes, Hungary. <sup>30</sup>Hospital of Bezmialem Vakıf University, School of Medicine, Istanbul, Turkey. <sup>31</sup>Department of Medical Genetics, Medical School, University of Pécs, Pécs, Hungary. <sup>32</sup>First Department of Medicine, Medical School, University of Pécs, Pécs, Hungary.