

A Novel 5-Cytokine Panel Outperforms Conventional Predictive Markers of Persistent Organ Failure in Acute Pancreatitis

Christopher Langmead, PhD¹, Peter J. Lee, MBChB², Pedram Paragomi, MD³, Phil Greer, PhD³, Kim Stello, RN³, Phil A. Hart, MD⁴, David C. Whitcomb, MD, PhD^{3,5} and Georgios I. Papachristou, MD, PhD^{3,4}

INTRODUCTION: Existing laboratory markers and clinical scoring systems have shown suboptimal accuracies for early prediction of persistent organ failure (POF) in acute pancreatitis (AP). We used information theory and machine learning to select the best-performing panel of circulating cytokines for predicting POF early in the disease course and performed verification of the cytokine panel's prognostic accuracy in an independent AP cohort.

METHODS: The derivation cohort included 60 subjects with AP with early serum samples collected between 2007 and 2010. Twenty-five cytokines associated with an acute inflammatory response were ranked by computing the mutual information between their levels and the outcome of POF; 5 high-ranking cytokines were selected. These cytokines were subsequently measured in early serum samples of an independent prospective verification cohort of 133 patients (2012–2016), and the results were trained in a Random Forest classifier. Cross-validated performance metrics were compared with the predictive accuracies of conventional laboratory tests and clinical scores.

RESULTS: Angiopoietin 2, hepatocyte growth factor, interleukin 8, resistin, and soluble tumor necrosis factor receptor 1A were the highest-ranking cytokines in the derivation cohort; each reflects a pathologic process relevant to POF. A Random Forest classifier trained the cytokine panel in the verification cohort and achieved a 10-fold cross-validated accuracy of 0.89 (area under the curve 0.91, positive predictive value 0.89, and negative predictive value 0.90), which outperformed individual cytokines, laboratory tests, and clinical scores (all $P \leq 0.006$).

DISCUSSION: We developed a 5-cytokine panel, which accurately predicts POF early in the disease process and significantly outperforms the prognostic accuracy of existing laboratory tests and clinical scores.

SUPPLEMENTARY MATERIAL accompanies this paper at <http://links.lww.com/CTG/A603>, <http://links.lww.com/CTG/A604>, <http://links.lww.com/CTG/A605>, <http://links.lww.com/CTG/A606>, <http://links.lww.com/CTG/A607>, and <http://links.lww.com/CTG/A608>

Clinical and Translational Gastroenterology 2021;12:e00351. <https://doi.org/10.14309/ctg.000000000000351>

INTRODUCTION

Acute pancreatitis (AP) is a complex clinical syndrome initiated by pancreatic injury from a variety of pathologic mechanisms (1). Its incidence has been increasing worldwide (2) and is among the most common gastrointestinal-related causes for hospitalization in the United States (3). Systemic inflammatory response develops in approximately half of all patients with AP (4) and may progress to organ failure, which can lead to death (5,6).

Important determinants of morbidity and mortality in AP are the development of pancreatic necrosis and persistent organ failure (POF) (6–8). Outcomes have recently improved in patients with pancreatic necrosis due to the development of effective and minimally invasive drainage/debridement techniques (9). On the other hand, the prediction of POF remains suboptimal despite extensive research (1,10). Predicting POF early in the disease course is important to accurately triage patients to proper

¹Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA; ²Division of Gastroenterology, Department of Medicine, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania, USA; ³Department of Medicine, Division of Gastroenterology, Hepatology and Nutrition, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA; ⁴Division of Gastroenterology, Hepatology, and Nutrition, Department of Internal Medicine, The Ohio State University Wexner Medical Center, Columbus, Ohio, USA; ⁵Departments of Physiology & Molecular Physiology, and Department of Human Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania, USA. **Correspondence:** Georgios I. Papachristou, MD, PhD. E-mail: georgios.papachristou@osumc.edu

Received December 16, 2020; accepted March 12, 2021; published online May 6, 2021

© 2021 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of The American College of Gastroenterology

hospitals and wards, prepare for timely organ support in those with impending organ failure, and risk stratify subjects for clinical trials.

Previous research efforts laid foundational groundwork in identifying and developing laboratory tests and scoring systems for prediction of POF in AP (10). Admission blood urea nitrogen (BUN), C-reactive protein, creatinine, and hematocrit represent conventional laboratory tests, whereas Acute Physiology and Chronic Health Evaluation (APACHE-II), Bedside Index of Severity of Acute Pancreatitis (BISAP), Ranson score, and systemic inflammatory response syndrome (SIRS) are scoring systems extensively evaluated in predicting POF (10). However, these have shown limited accuracies in early prediction of POF, and many of these tests and scores directly reflect the presence of organ failure, rather than serving as early prediction tools.

Local cytokine and chemokine levels derived from injured acinar cells and different inflammatory cell types increase early in AP in response to pancreatic injury. The magnitude of their circulating levels reflects the degree of tissue injury and innate immune response (11,12). To our knowledge, a machine learning approach using multiple cytokines has not been previously evaluated in the prognostication of AP severity. Therefore, we sought to discover and test a panel of cytokines that would accurately predict POF early in AP by applying information theory and a machine learning model and additionally compare its prognostic accuracy with conventional laboratory tests and scoring systems used in clinical practice.

METHODS

Study population

The Pancreatitis-associated Risk of Organ Failure (PROOF) is an observational, prospective, cohort study involving patients with AP enrolled at the beginning of their hospitalization. The PROOF protocol and cohort characteristics have been previously described (13,14). AP was defined according to established guidelines (15). Patients with chronic pancreatitis, pancreas cancer, others cancer requiring chemotherapy within the last 6 months, history of solid organ transplantation, pregnant women, and prisoners were excluded. Only direct admissions or patients transferred from an outside hospital within 12 hours from presentation were included. PROOF has been approved by the Institutional Review Board (IRB) of the University of Pittsburgh (Pro00000496) and registered at clinicaltrials.gov (NCT03075605).

Derivation phase (2007–2010)

Serum samples were obtained from 60 subjects on days 2 (enrollment), 3, and 4, relative to the onset of pain, and then stored at -80°C . They were batch analyzed using the Bio-Plex suspension array system, which included a fluorescent reader and Bio-Plex Manager analytical software (Bio-Rad Laboratories, Hercules, CA) at the University of Pittsburgh Cancer Institute Luminex Core (16). The Luminex platform was selected due to its ability to screen large number of markers in the same bio samples and based on available expertise at the University of Pittsburgh. Twenty-five cytokines associated with the development of an acute inflammatory response were studied using a proinflammatory cytokine panel, based on our hypothesis that early development of a proinflammatory milieu is associated with clinical decompensation, including organ failure (see Table A, Supplementary Digital Content, <http://links.lww.com/CTG/A603>). Their levels were measured in a total of 180 samples, i.e., at

3 different time points per subject (at days 2, 3, and 4). The cytokines were first discretized using an algorithm by Fayyad and Irani (17). Next, the Mutual Information between each cytokine (exposure variable) and POF (primary outcome) was computed (see Table B, Supplementary Digital Content, <http://links.lww.com/CTG/A604>) (18). Mutual Information scores range from 0 to 1 for a binary classification task, where 0 means the 2 variables are statistically independent and 1 means that they are perfectly correlated. Cytokines were deemed relevant to the outcome when the Mutual Information score was ≥ 0.2 for at least 2 of the 3 days calculated.

Verification phase (2012–2016)

Serum levels of the top 5 cytokines identified in the derivation phase were measured in the verification cohort before initiation of any interventions such as ERCP or plasmapheresis, using Meso Scale Discovery (MSD) MULTI-SPOT Assay System (MESO QuickPlex SQ 120 instrument) in the laboratory of investigators Whitcomb and Papachristou (19,20). The MSD platform was selected for the verification cohort due to its high reliability and sensitivity for measuring of levels at the lower detection threshold (21). The cytokines included angiopoietin 2 (Ang-2), hepatocyte growth factor (HGF), interleukin 8 (IL-8), resistin, and soluble tumor necrosis factor receptor superfamily member 1A (TNF-R1). They were measured on 3 separate assay kits: the V-Plex Human Proinflammatory for IL-8 and 2 custom human duplex kits combining Ang-2 and HGF on one and resistin and TNF-R1 on the other, purchased from MSD (Gaithersburg, MD). Using MSD Discovery Workbench analysis software v4.0 (Rockville, MD), standard curves were formed by fitting the electrochemiluminescence signal from calibrators to a 4-parameter logistic model with a $1/y^2$ weighting. R^2 values for the fitted curves were >0.95 for all cytokines.

Serum samples were run in duplicates with median intra-assay variability being 2.70%–9.56% and interassay variability 7.44–43.48 (see Table C, Supplementary Digital Content, <http://links.lww.com/CTG/A605>). At the time of statistical analysis, Ang-2 and TNF-R1 levels were found to be inaccurate in the first 36 serum samples measured. On further investigation, we discovered that the custom human duplex kits used for these samples were defective. These values were excluded from analysis and not repeated to avoid introduction of interbatch variability.

Statistical model

Descriptive statistics are presented as counts and proportions for categorical data and as median [interquartile range] for continuous data. The distributions of biomarker levels were assessed for normality using graphical methods and the Shapiro-Wilk test. Comparisons in discrete baseline characteristics between the derivation and verification AP cohorts and between AP subjects with and without POF were performed using the χ^2 test. Comparisons for continuous variables used the Mann-Whitney U test. The above statistical tests were performed using the R Project software (www.r-project.org). To adjust for multiple comparisons, the Hommel procedure was adopted for calculating adjusted P values. Two-tailed P values <0.05 were considered statistically significant.

Five cytokines with the highest mutual information gain were selected to build a novel cytokine panel. This panel was then used to train a Random Forest model using samples from the verification cohort of 133 subjects (22). Predictive performance metrics were computed for the primary end point (POF) (23). POF

was defined according to the Modified Marshall Score System as recommended by the Revised Atlanta Classification (7).

Ten-fold cross validation was performed to estimate how well the model would generalize to an independent data set. The accuracy of the Random Forest model was compared with each individual cytokine, as well as conventional laboratory tests, and clinical scoring systems, measured in the verification cohort at the time of study enrollment.

RESULTS

Demographic and clinical data of the derivation and verification cohorts are shown in Table 1.

Derivation phase

Eight biomarkers—Ang-2, HGF, IL-6, IL-8, IL-10, macrophage chemoattractant protein 1 (MCP-1), resistin, and TNF-RI—met the Mutual Information criteria (see Table B, Supplementary Digital Content, <http://links.lww.com/CTG/A604>). Of these, HGF was highly correlated with IL-10 ($R^2 = 0.77$), IL-8 with MCP-1 ($R^2 = 0.84$), and resistin with IL-6 ($R^2 = 0.79$). IL-10, MCP-1, and IL-6 were subsequently excluded from the selected cytokine panel because they

did not provide any independent information and showed lower information gain when compared to HGF, IL-8, and resistin on days 2 and 3. One potential approach could have been to exclude IL-8 because both Ang-2 and resistin were strongly correlated with IL-8. However, because IL-8 is a well-studied cytokine for the prediction of severity in AP, we felt that it was meaningful to retain in the final panel (24). Furthermore, the information gain of IL-8 was high (0.83 and 0.77 on days 2 and 3; see Table B, Supplementary Digital Content, <http://links.lww.com/CTG/A604>) supporting its inclusion. Mechanistically, IL-8 levels reflect the inflammatory pathway, whereas Ang-2 is implicated in the vascular leak and resistin is an adipokine associated with fat necrosis. Taken together, we decided to include IL-8 representing a unique pathway in the cytokine panel.

Thus, the final 5 cytokines selected were Ang-2, HGF, IL-8, resistin, and TNF-RI. Supplemental Table D, (see Supplementary Digital Content, <http://links.lww.com/CTG/A606>) lists the correlation coefficients between the selected cytokines with the median being relatively low ($R^2 = 0.52$), indicating that each cytokine may represent different aspects of the pathophysiologic mechanisms of disease. Comparisons of the 5 cytokine levels between patients without POF and with POF are depicted in boxplots (Figure 1).

Table 1. Demographic and clinical characteristics of the study population

Variable	Derivation cohort (n = 60)	Verification cohort (n = 133)	Overall cohort (n = 193)	P value ^b
Age				
Median [IQR], y	53 [40–67]	50 [36–67]	51 [37–67]	0.868
Sex				
Male (%)	34 (56.7)	63 (47.4)	97 (50.3)	0.868
Race				
White (%)	52 (86.7)	125 (94.0)	177 (91.7)	0.172
African American	8 (13.3)	5 (3.8)	13 (6.7)	
Other	0 (0.0)	3 (2.2)	3 (1.6)	
BMI				
Median [IQR], kg/m ²	29 [25–32]	30 [26–36]	30 [25–35]	0.246
Etiology				
Biliary (%)	17 (28.3)	60 (45.1)	77 (39.9)	0.003
Alcoholic	11 (18.3)	9 (6.8)	20 (10.4)	
Idiopathic	18 (30.0)	14 (10.5)	32 (16.6)	
Other	14 (23.3)	50 (37.6)	64 (33.2)	
Severity ^a				
Mild (%)	24 (40.0)	62 (46.6)	86 (44.6)	0.014
Moderately severe	5 (8.3)	34 (25.6)	39 (20.2)	
Severe	31 (51.7)	37 (27.8)	68 (35.2)	
Length of stay				
Median [IQR], days	12 [6.0–23.5]	6 [4.0–15.2]	7 [4.0–17.8]	0.029
In-hospital mortality				
Deaths (%)	6 (10.0)	9 (6.8)	15 (7.8)	0.868

Values for age and BMI are rounded to the nearest whole number.

Bold indicates statistically significant P value.

BMI, body mass index; IQR, interquartile range.

^aPer Revised Atlanta Criteria (7).

^bChi-square test for discrete values and Mann-Whitney U test for continuous variables.

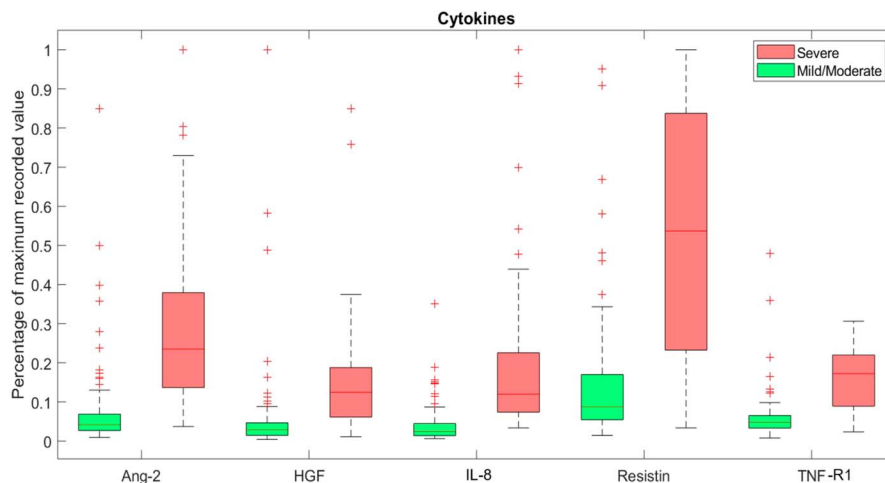


Figure 1. Comparisons of cytokine levels between the mild/moderately severe group (i.e., no POF; green) and the severe group (i.e., POF present; red) in the verification cohort. All *P* values were <0.001 when comparing cytokine levels between groups. Ang-2, angiopoietin 2; HGF, hepatocyte growth factor; IL-8, interleukin 8; POF, persistent organ failure; TNF-R1, tumor necrosis factor alpha receptor superfamily 1A.

Verification phase

Comparison of clinical characteristics and cytokine levels between subjects with and without POF in the verification cohort are shown in Table E (see Supplementary Digital Content, <http://links.lww.com/CTG/A607>). The cytokine panel was evaluated by 10-fold cross validation of a Random Forest classifier. The trained classifier was compared with a hand-picked combination of clinical variables (clinical panel), individual cytokines, existing laboratory parameters, and scoring systems. The hand-picked clinical panel included age, sex, body mass index, BUN, creatinine, and SIRS score at enrollment. The performance metrics of all the above for predicting severe pancreatitis are shown in Table F (see Supplementary Digital Content, <http://links.lww.com/CTG/A608>). The 5-cytokine panel showed the highest accuracy (0.89). The accuracy of the 5-cytokine panel was compared with the other study markers using a 1-sided Binomial test and a probability of success of 0.89 (the cross-validated accuracy of the 5-marker panel). The *P* values for the alternative models were determined using the empirical number of correct predictions for that model across the cohort, under a Binomial test. A Bonferroni correction factor of 18 was applied to account for testing the 18 alternative models.

In a sensitivity analysis including only subjects with complete data ($n = 97$), the cytokine panel performance improved slightly (overall accuracy 0.92, sensitivity 0.92, specificity 0.76, positive predictive value 0.92, and negative predictive values 0.93). In additional analyses, the cytokine panel's predictive performance did not improve when adding the 3 other high-performing cytokines (MCP-1, IL-10, and IL-6), clinical parameters, and/or laboratory tests to the modeling (data not shown).

The accuracies of the individual cytokines included in the panel, single laboratory tests, and scoring systems compared against the cytokine panel are shown in Table 2. The cytokine panel significantly outperformed all single cytokines, laboratory tests, and clinical scores assessed (all $P < 0.05$; Figure 2, Table 2).

DISCUSSION

The lack of an accurate tool to predict severe pancreatitis early in the disease course has hampered additional progress in clinical management and the execution of therapeutic trials. In this study, we developed a novel panel of 5 cytokines, which accurately predicts POF as early as within 24 hours from symptoms onset.

Importantly, the featured cytokines are biologically relevant in the pathophysiology of AP, which increases the likelihood for replication and future validation.

The 5-cytokine panel consists of Ang-2, HGF, IL-8, resistin, and TNF-R1, which have distinct roles in the pathophysiology of AP. Severe innate immune system response, lipolysis of peri-/intra-pancreatic fat, and vascular leak are recognized mechanisms that promote POF in AP (1). We theorize that the selected cytokines may reflect complementary pathways leading to POF (Figure 3). Independently, the selected cytokines outperformed most of the commonly used laboratory tests and clinical scoring systems for predicting POF with their performance being augmented when any of the 5 cytokine levels crossed the panel's threshold.

Previous studies on conventional laboratory parameters and clinical scores as predictors of severe AP laid the framework for this work; however, they have only reported modest accuracies (10,25). In a multicenter cohort study, the area under the curve (AUC) for admission BUN and hematocrit was 0.65 and 0.67 in predicting POF, respectively (26). Similarly, existing laboratory tests and scoring systems did not exceed an AUC of 0.75 in another large prospective study (10). In contrast, individual cytokines appear to perform slightly better. Malmstrom et al. reported the AUC of admission IL-6 and IL-8 levels to be 0.84 and 0.71, respectively, in predicting multiorgan failure (25). Ang-2 revealed AUCs of 0.81 and 0.85 for predicting POF in 2 independent studies (20,27). This can be partially explained by cytokine elevations likely occurring before any physiologic perturbations relevant to POF can be detected by laboratory parameters and clinical scores (Figure 3) (28). However, their prognostic performances varied widely between different studies, highlighting the disadvantage of using a single biomarker for AP severity prediction.

In our study, the accuracy of the 5-cytokine panel was significantly higher than each selected cytokine alone (Table 2). The contribution of different pathways mediating POF may vary between subjects across an AP cohort. Therefore, a panel of cytokines that reflects multiple nodes in the pathophysiology of POF has the potential to be highly accurate in predicting POF. To our knowledge, this is the first study to perform a head-to-head comparison on the prognostic accuracies of a cytokine panel versus individual cytokines, existing laboratory markers, and scoring systems in the same prospectively

Table 2. Comparative performance of conventional laboratory tests and existing clinical scoring systems against the cytokine panel

	Cutoff	Accuracy	P value ^a
Cytokine panel	NA	0.89	ref
Cytokines ^b			
Ang-2	23,690	0.81	0.005
HGF	3,211	0.81	0.005
IL-8	33	0.73	0.000
Res	5,482	0.74	0.000
TNF-RI	2,712	0.72	0.000
Laboratory tests ^c			
BUN	20.0	0.55	0.000
Cr	1.8	0.56	0.000
CRP	19.8	0.65	0.000
Hct	44.0	0.56	0.000
IL-6 ^b	19	0.78	0.000
MCP-1	978	0.74	0.000
Clinical scores			
APACHE-II	8	0.56	0.000
BISAP	2	0.63	0.000
Ranson	8	0.63	0.000
SIRS	2	0.54	0.000

Ang-2: angiotensin 2; APACHE-II: Acute Physiology and Chronic Health Evaluation; BISAP: Bedside Index of Severity of Acute Pancreatitis; BUN: blood urea nitrogen; Cr: creatinine; CRP: C-reactive protein; HCT: hematocrit; HGF: hepatocyte growth factor; IL-8: interleukin 8; MCP-1: macrophage chemoattractant protein 1; Res: resistin; SIRS: systemic inflammatory response syndrome; TNF-RI: tumor necrosis factor receptor superfamily IA.

^aOne-sided binomial test.
^bCytokine levels are given in pg/mL;
^cValues are given in mg/dL except in % for HCT.

enrolled AP cohort. The 5-cytokine panel was also compared and found to be more accurate to a hand-picked panel of selected clinically relevant variables. Our results suggest that early in the AP course, this novel cytokine panel performs with significantly higher accuracy than any available tools for predicting POF.

There are several strengths in this study. First, we prospectively enrolled patients with AP and collected blood early in the disease course. Timely enrollment is crucial to examine cytokine concentrations during the early phase of AP. By accomplishing this, the validity of the results was not only enhanced, but our findings could also contribute to the understanding of the inflammatory milieu in the early phase of AP. Second, detailed clinical data were prospectively collected for each enrolled subject with the disease severity being categorized according to the Revised Atlanta Classification, which represent the most updated severity categorization. Third, serum samples from 2 independent cohorts were analyzed with the 5-cytokine panel being developed in a derivation cohort and then verified in a separate cohort. The Luminex platform was selected for analyses in the derivation cohort due to its ability to rapidly screen many markers in an unbiased manner. The MSD platform was selected for the verification cohort due to its high reliability and improved sensitivity for measuring levels at the

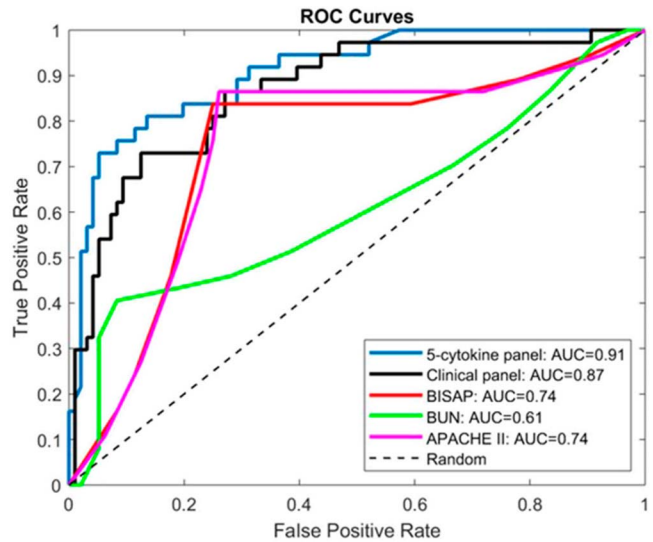


Figure 2. Receiver operating characteristic (ROC) curves comparing the cytokine panel (blue) to the clinical panel (black), BISAP (red), and BUN (green) in the verification cohort. AUC, area under the curve; BISAP, Bedside Index of Severity of Acute Pancreatitis; BUN, blood urea nitrogen.

lower detection threshold (29). The use of 2 different methods for cytokine measurements likely enhances the generalizability of our results. Application of information theory in the derivation cohort provided an unbiased statistical approach for identifying cytokines with the best predictive performance. The process of ranking all candidate cytokines according to their information gain contributed to the discovery of a novel set of 5 biomarkers that, when combined, was highly predictive of POF. The complexity of AP pathophysiology suggests a nonlinear relationship between circulating levels of cytokines (predictor variables) early in the course of AP and POF (outcome). We used Random Forests model, which is well suited to learning complex, nonlinear decision boundaries, because they are constructed from decision trees, which can learn arbitrary decision boundaries. In addition, mutual information filter and Random Forest training algorithm with bootstrap aggregation minimized the risk of overfitting (overestimation of a model's true accuracy) (30).

There are a few aspects of this study for consideration to accurately interpret the results. First, despite extensive efforts, our verification cohort was of a statistically modest sample size ($n = 133$) with only 9 deaths. Because of its low event rate, it was not feasible to build a prediction model for mortality with sufficient statistical power in our cohort. POF represents an independently meaningful clinical end point as it relates to length of stay and represents a surrogate end point for mortality. Similarly, the current sample size did not enable us to perform robust subgroup comparisons.

This illustrates the challenges of enrolling patients with AP and obtaining blood samples within the early hours from symptom onset. Second, a larger proportion of patients developed POF in the derivation cohort compared with the verification cohort. This could have resulted in an overestimation of the mutual information between selected cytokines and the primary end point. Nevertheless, the cytokine panel's high predictive performance in the verification cohort in which the prevalence of POF was lower indicates that our results were minimally affected by the above discrepancy. Third, there was a time period of 10 years from inception to completion of study. Thus, there was a potential risk of cytokine degradation at the time of measurement of their levels. In addition,

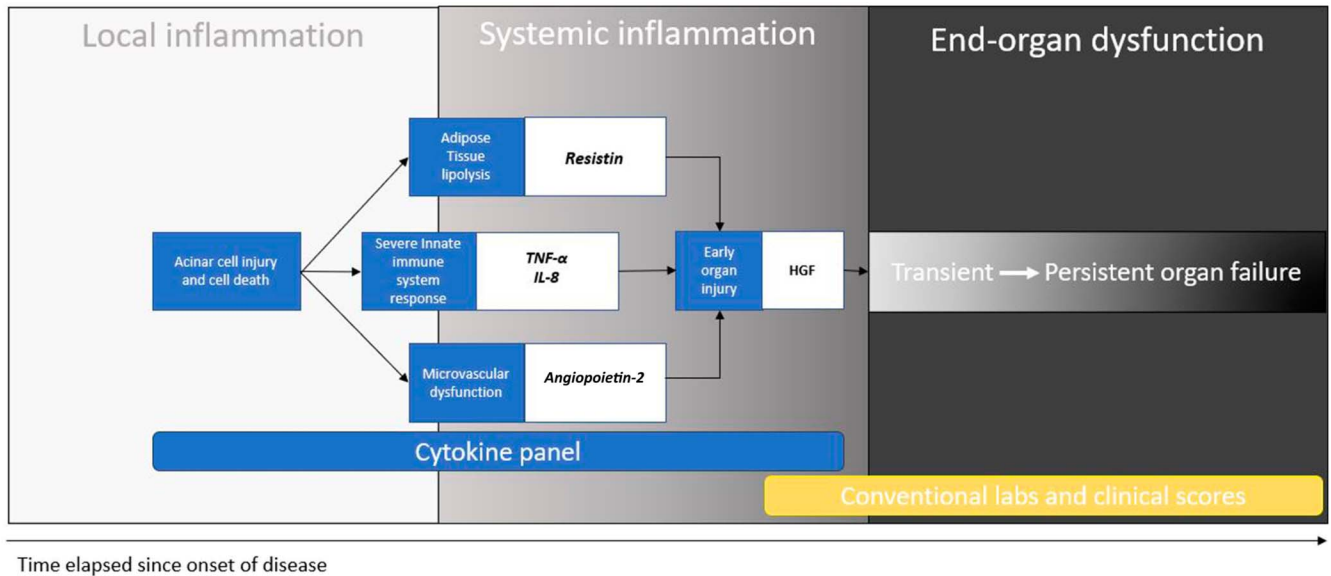


Figure 3. Proposed conceptual model for prediction of severe acute pancreatitis. The figure was constructed to aid discussion of the results and is not meant to exhaustively represent all cytokines and biomarkers associated with acute pancreatitis. IL, interleukin; HGF, hepatocyte growth factor; TNF- α , tumor necrosis factor alpha.

recently identified AP biomarkers were not included in this study. For example, disease-associated molecular patterns such as heat shock proteins, high mobility group box 1, histones and adenosine triphosphate have been associated with disease severity in AP (28). Endothelial leukocyte markers (e.g., P-, E- selectin) and prior targets of clinical trials (e.g., plasmin activating factor and trypsin) were also not explored. Inclusion of these markers in future studies may further enhance predictive performance of the cytokine panel. Finally, the cytokines included in our panel are not routinely measured in clinical laboratories (31). Therefore, the panel is not readily available for clinical decision making in its current form. Additional work is required to develop an assay that would provide fast cytokine measurements for clinical decision making. An example of this strategy has been the development of a point of care assay for IL-6 for the management of sepsis (32).

In conclusion, using innovative machine learning techniques, we developed a novel panel of biologically relevant cytokines that accurately predicts POF early in AP. This panel can potentially guide the early management of AP and inform the design of clinical trials.

CONFLICTS OF INTEREST

Guarantor of the article: Georgios I. Papachristou, MD, PhD.

Specific author contributions: Christopher Langmead, PhD, and Peter J. Lee, MBChB are first co-authors. Christopher Langmead, PhD, and Peter J. Lee, MBChB, contributed equally to this work. David C. Whitcomb, MD, PhD, and Georgios I. Papachristou, MD, PhD, codirected this project. C.L.: planning statistical methodology, statistical analysis, interpretation of data, and manuscript review. P.J.L.: interpretation of data and manuscript drafting. P.P.: data collection and manuscript review. P.G.: execution of the experiments, data collection, interpretation of data, and manuscript review. K.S.: execution of the experiments and data collection. P.A.H.: interpretation of data and manuscript review. D.C.W. conception of the study, interpretation of data, and manuscript review. G.I.P.: conception of the study, interpretation of data, manuscript drafting, and direct supervision of the manuscript.

Financial support: This research was funded by the Department of Veterans Affairs Merit Review I01CX000272 (G.I.P.), the U.S. Department of Defense Congressionally Directed Medical Research Programs (CDMRP) Awards W81XWH-14-1-0376 (D.C.W.) and W81XWH-17-1-0502 (D.C.W.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Potential competing interests: The authors report no conflicts of interest directly related to this research.

Study Highlights

WHAT IS KNOWN

- ✓ Many laboratory tests and clinical scores are used to predict persistent organ failure in acute pancreatitis but are only modestly accurate.
- ✓ Lack of an accurate tool to predict persistent organ failure in acute pancreatitis early in the disease course has significantly hampered progress in clinical management and execution of therapeutic trials.

WHAT IS NEW HERE

- ✓ A highly accurate panel of 5 cytokines was developed to predict persistent organ failure in acute pancreatitis using machine learning techniques, which significantly outperformed existing laboratory tests and clinical scores.

TRANSLATIONAL IMPACT

- ✓ Once developed for clinical use, this panel may accelerate identification of patients at risk of persistent organ failure and promptly assign appropriate level of care in acute pancreatitis.
- ✓ This panel may be useful for the design of future drug trials in acute pancreatitis.

REFERENCES

- Lee PJ, Papachristou GI. New insights into acute pancreatitis. *Nat Rev Gastroenterol Hepatol* 2019;16:479–96.
- Petrov MS, Yadav D. Global epidemiology and holistic prevention of pancreatitis. *Nat Rev Gastroenterol Hepatol* 2019;16:175–84.
- Peery AF, Crockett SD, Barritt AS, et al. Burden of Gastrointestinal, liver, and pancreatic diseases in the United States. *Gastroenterology* 2015;149:1731–41.e3.
- Singh VK, Wu BU, Bollen TL, et al. Early systemic inflammatory response syndrome is associated with severe acute pancreatitis. *Clin Gastroenterol Hepatol* 2009;7:1247–51.
- Zubia-Olaskoaga F, Maravi-Poma E, Urreta-Barallobre I, et al. Comparison between revised Atlanta classification and determinant-based classification for acute pancreatitis in intensive care medicine. Why do not use a modified determinant-based classification? *Crit Care Med* 2016;44:910–7.
- Dellinger EP, Forsmark CE, Layer P, et al. Determinant-based classification of acute pancreatitis severity: An international multidisciplinary consultation. *Ann Surg* 2012;256:875–80.
- Banks PA, Bollen TL, Dervenis C, et al. Classification of acute pancreatitis—2012: Revision of the Atlanta classification and definitions by international consensus. *Gut* 2013;62:102–11.
- Schepers NJ, Bakker OJ, Besselink MG, et al. Impact of characteristics of organ failure and infected necrosis on mortality in necrotising pancreatitis. *Gut* 2019;68:1044–51.
- van Brunschot S, van Grinsven J, van Santvoort HC, et al. Endoscopic or surgical step-up approach for infected necrotising pancreatitis: A multicentre randomised trial. *Lancet* 2018;391:51–8.
- 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;142:1476.
- Kany S, Vollrath JT, Relja B. Cytokines in inflammatory disease. *Int J Mol Sci* 2019;20:6008.
- Gong T, Liu L, Jiang W, et al. DAMP-sensing receptors in sterile inflammation and inflammatory diseases. *Nat Rev Immunol* 2020;20:95–112.
- Vipperla K, Papachristou GI, Easler J, et al. Risk of and factors associated with readmission after a sentinel attack of acute pancreatitis. *Clin Gastroenterol Hepatol* 2014;12:1911–9.
- Papachristou GI, Muddana V, Yadav D, 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* 2010;105:435–41; quiz 442.
- Tenner S, Baillie J, DeWitt J, et al. American College of Gastroenterology guideline: Management of acute pancreatitis. *Am J Gastroenterol* 2013;108:1400–15; 1416.
- Papachristou GI, Sass DA, Avula H, et al. Is the monocyte chemoattractant protein-1 -2518 G allele a risk factor for severe acute pancreatitis? *Clin Gastroenterol Hepatol* 2005;3:475–81.
- Fayyad UM, Irani KB. Multi-interval discretization of continuous-valued attributes for classification learning. *JCAI* 1993:1022–7.
- MacKay DJC. *Information Theory, Inference & Learning Algorithms*. Cambridge University Press: New York, NY, 2002.
- MESO QuickPlex SQ 120 Features and Specifications. (https://www.mesoscale.com/~media/files/product_highlights/quickplex_features_specifications.pdf). Accessed February 3, 2020.
- Whitcomb DC, Muddana V, Langmead CJ, et al. Angiotensin-2, a regulator of vascular permeability in inflammation, is associated with persistent organ failure in patients with acute pancreatitis from the United States and Germany. *Am J Gastroenterol* 2010;105:2287–92.
- Keustermans GCE, Hoeks SBE, Meering JM, et al. Cytokine assays: An assessment of the preparation and treatment of blood and tissue samples. *Methods* 2013;61:10–7.
- Breiman L. Random Forests. *Mach Learn* 2001;45:5–32.
- Matthews BW. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim Biophys Acta* 1975;405:442–51.
- Aoun E, Chen J, Reighard D, et al. Diagnostic accuracy of interleukin-6 and interleukin-8 in predicting severe acute pancreatitis: A meta-analysis. *Pancreatol* 2009;9:777–85.
- Malmstrom ML, Hansen MB, Andersen AM, et al. Cytokines and organ failure in acute pancreatitis: Inflammatory response in acute pancreatitis. *Pancreas* 2012;41:271–7.
- Koutrompakis E, Wu BU, Bakker OJ, et al. Admission hematocrit and rise in blood urea nitrogen at 24 h outperform other laboratory markers in predicting persistent organ failure and pancreatic necrosis in acute pancreatitis: A post hoc analysis of three large prospective databases. *Am J Gastroenterol* 2015;110:1707–16.
- Buddingh KT, Koudstaal LG, van Santvoort HC, et al. Early angiotensin-2 levels after onset predict the advent of severe pancreatitis, multiple organ failure, and infectious complications in patients with acute pancreatitis. *J Am Coll Surg* 2014;218:26–32.
- Garg PK, Singh VP. Organ failure due to systemic injury in acute pancreatitis. *Gastroenterology* 2019;156:2008–23.
- Dabitao D, Margolick JB, Lopez J, et al. Multiplex measurement of proinflammatory cytokines in human serum: Comparison of the Meso Scale discovery electrochemiluminescence assay and the cytometric bead array. *J Immunol Methods* 2011;372:71–7.
- Burnham KP, Anderson DR. *Model Selection and Multimodel Inference*. Chapter 1. Springer-Verlag New York, 2002, pp 49–97.
- Panch T, Mattie H, Celi LA. The “inconvenient truth” about AI in healthcare. *NPJ Digit Med* 2019;2:77.
- Fischer SK, Williams K, Wang L, et al. Development of an IL-6 point-of-care assay: Utility for real-time monitoring and management of cytokine release syndrome and sepsis. *Bioanalysis* 2019;11:1777–85.

Open Access This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.