Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology) in press
www.jzus.zju.edu.cn; www.springer.com/journal/11585
E-mail: jzus_b@zju.edu.cn

Correspondence

https://doi.org/10.1631/jzus.B2300539

A novel clinical data management platform for acute pancreatitis

Shiyin CHEN[1*], Cheng ZHANG[1*], Zhien WANG[2], Jian ZHANG[1], Wenqiao YU[3], Yanshuai WANG[1], Weiwei SI[1], Tingbo LIANG[1✉], Yun ZHANG[1✉]

[1]*Department of Hepatobiliary and Pancreatic Surgery, Zhejiang Provincial Key Laboratory of Pancreatic Disease, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China*
[2]*Department of Rehabilitation Medicine, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China*
[3]*Department of Comprehensive Intensive Care Unit, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China*

This study presents a multi-center clinical data management platform that facilitates unified and structured management of real-world data and serves as an ideal tool to enhance the quality and progress of clinical research related to severe acute pancreatitis (SAP). The use of the platform enables clinical teams to obtain safe, accurate, unified structure, feasible traceability, clear scene, and fully functional real-world medical data in the diagnosis, treatment, and research of acute pancreatitis (AP).

SAP is a life-threatening disease characterized by rapid progression, multiple organ damage, treatment challenges, and high mortality. Consequently, there is an urgent need for clinical studies to enhance existing treatment strategies and improve clinical outcomes (L.Boxhoorn et al., 2020). Given the extensive data generated from SAP and the heterogeneity of clinical treatment approaches to it (S.Finazzi et al., 2020), it has become imperative to establish a collaborative major data management model across multiple medical centers, with a tailored data structure designed for SAP-related requirements (A.Leppäniemi et al., 2019). In this study, we developed a novel multicenter data-management platform for AP, primarily focusing on SAP cases. The key components of this platform are a standardized data format based on pre-defined structures (Figs. 1a and S1, Table S1), visualization of various chart-based data scenarios (Figs. 1b, 2, and S2), storage and protection of transfer data in the cloud server (Fig. 2), homogeneous case data tracked with markers (Table S1, Fig. 3), and a research process simplified by using randomized controlled trial (RCT) modules (Fig. 3).

The "Full data acceptance + post-structured application" data-management mode (Table S1), exemplified by the MIMIC-IV database (A.Johnson et al., 2020) is not optimal for reproducing clinical scenes accurately. To effectively manage the high heterogeneity of data in multi-center studies, experts from various countries have made numerous attempts to collect and structure the unstructured data. However, these attempts have not been completely effective for several reasons. We compared our platform with other clinical data platforms for AP, such as Medbit (K.Lu, et al., 2022), REDCap (A.B.Docherty et al., 2021), and eICU (T.J. Pallard et al., 2018) (Table 1). It is important to note that these structured databases were established under project guidance rather than representing comprehensive real-world databases. Once the related project is completed, such databases

✉ Yun ZHANG, bigzyun1977@zju.edu.cn
Tingbo LIANG, liangtingbo@zju.edu.cn
[*]The two authors contributed equally to this work
Yun ZHANG, https://orcid.org/0000-0002-0339-4223
Tingbo LIANG, https://orcid.org/0000-0003-0143-3353
Shiyin CHEN, https://orcid.org/0009-0005-0313-9045
Cheng ZHANG, https://orcid.org/0000-0001-8631-4411

Received July 30, 2023; Revision accepted Mar. 6, 2024;
Crosschecked xxx. xx, 20xx; Published online xxx. xx, 20xx

© Zhejiang University Press 2024

Unedited

become invalid for future use due to their inability to restore contextual details. On the other hand, the "pre-structured setting + valuable data application" novel data-management mode (Table S1) involves the definition of disease-specific clinical data requirements in advance by experienced specialists to ensure comprehensiveness and accuracy. This approach aims to reflect life-like medical scenarios and incorporate objective and subjective data experiences. The characteristics of each case are described in a highly generalized manner with uniform feature markers within this real-scene reproduction data platform, facilitating the identification of homogeneous cases that meet research requirements before the implementation of clinical studies. Consequently, this solution significantly enhances efficiency and quality in conducting complex disease-related clinical research.

The data-entry interface consists of log data and image data and serves as the core technology for medical data management on this platform. It requires doctors to professionally customize the structure of the SAP-related data. The data-acquisition methods included the following. ①Check input: all data that cannot be objectively measured, such as diagnosis and treatment, were designed as optional. The check was performed by experienced clinicians based on the actual situation and subjective judgment. These data that cannot be objectively measured are the most valuable part of the data platform (Figs. S1c and S1e). ②Fill-in-the-blank: all measurable digital data, including physical, treatment results (e.g.: drainage volume), and laboratory indicators were entered by filling in the blank. Due to the labor-intensive aspect of this input mode, part of such data could be replaced by automatically importing data from the hospital information system (HIS) or bedside equipment (Figs. S1b, S1c, and S1e). ③Major-event button: when any major treatment event (eg: mechanical ventilation) starts or stops, a start or stop button can be pressed. The platform can automatically draw a Gantt chart (if prompted by the user) to reproduce the scenario of major events (Figs. S1d, S1f, and S2). ④Inheritance entry: for all subjective and objective data associated with major treatment events (eg: mechanical ventilation mode and parameters), the platform can automatically inherit the settings of the previous day until they are manually changed, thus reducing the workload (Fig. S1d). ⑤Free entry: some data that surpass the pre-set data structure were filled in freely based on their importance and particularity for future open retrieval (Fig. S1a). ⑥ Automatic import from HIS: the platform can open the interface to each subsystem of the hospital and automatically collect the required data from the corresponding systems, such as the laboratory information system (LIS), picture archiving and communication systems (PACS), critical care system, and medical order system. ⑦ Case report form (CRF) data from RCT research can be imported sequentially from log data. Different modes of data export include quantification tables (Excel), trend charts, and Gantt charts, all of which contribute to achieving the reproduction of real diagnosis and treatment scenarios.

To meet the requirements of data standardization, the platform presents the data category in a unified description that fulfills the diagnosis and treatment of AP. We classified all clinical data into three types: quantifiable objective data, non-quantifiable judgment data, and major events (Table S2, Fig. 1a). Objective data refers to the data obtained through direct bedside observation, laboratory tests, and examinations, that can be directly entered by filling in the required information. On the other hand, judgment data is diagnostic information that has been analyzed, evaluated, and categorized by physicians and often requires selecting options on the platform. Major-event data refers to significant therapeutic occurrences throughout the entire diagnosis and treatment process of SAP (Table S2, Fig. S1), such as antibiotic usage, mechanical ventilation, kidney-replacement therapy, fluid resuscitation, vasoactive drug administration enteral nutrition support drainage procedures debridement surgery glucocorticoid utilization artificial liver intervention; all of which were pre-set with selectable options. Finally, all SAP-related data was divided into nine systems based on clinical characteristics: inflow and outflow dynamics, respiratory function, infection, gastrointestinal tract and nutritional status, organ function and test results, drainage procedures, surgical interventions, complications, and special scores.

Once entered into the platform, data can accurately reproduce a specific diagnosis and treatment scenario with the following tools: ①The log-type data-entry interface of the nine major systems is itself a reproduction

Unedited

of the real diagnosis and treatment scenario, and the integration of the relevant observation, judgment, and intervention data of each specific system (circulation, respiration, nutrition, surgery, and so on) every day is the daily routine of the doctor; this is in contrast to the lack of scene reproduction afforded by the simple and mechanical storage of data categories in the past database. ②The dynamic data show the spatiotemporal changes in patients' treatment, including: ⑴ Gantt chart: The platform can automatically plot a Gantt chart of any period that can visually show the cause, content, and start and end time of all major treatment events that are occurring or have ended, and can record 26 major events in 7 categories (Figs. 2 and S2, Table S2). ⑵ Trend chart: This is a dynamic presentation of all objective data. The platform can select one or more interrelated objective data points to draw a dynamic trend line following the needs of disease interpretation in a certain period that can incorporate different period data points depending on research and clinical diagnosis and treatment needs, to show the change rule and relationship of relevant data within a specific time. It can also record 142 kinds of dynamic monitoring data in 7 categories (Table S2, Fig. 2). ③A three-dimensional model of Excel data export (Fig. 1b): All objective digital data can be exported to an Excel table simultaneously, providing the final tool for statistical analysis. As a result, such data were all marked as three dimensions of variables such as different patients, time points, or objective digital data. The system can export any two of these variables.

The key point of multi-center research is standardized and unified data which is preserved completely, along with tools to assist in clinical research (N.J.Schepers et al., 2019). The data-management mode of pre-structural design lays a solid foundation for a series of functions such as scene presentation, homogeneous case tracking, and RCT research. Researchers and consultants can grasp the diagnosis and treatment process of complex cases quickly due to the dynamic presentation of objective quantitative data and the application of the Gantt chart of major treatment events. When patients from different hospitals in a data-sharing network transfer, "online and offline" parallel data and patient transfer can be carried out. Before transfer, all data can be automatically shared with the second hospital, and thus accurate and efficient handover of patient information, data maintenance for transferred patients, and complete information for doctors can be achieved (Fig. 2). Homogeneous patient retrieval, based on feature marks and a three-dimensional model of Excel data export, enables users to rapidly lock and export the cases needed for research, and thus perform efficient analysis of all data. Multi-center users in the network can initiate and participate in RCT research through the main accounts of hospitals at all levels. The benefits of this platform include the networking of the entire RCT research process, automation and standardization of data import, facilitation of the implementation of effective multi-task and multi-center RCT, and ultimately improvement in the quality and efficiency of research.

Of note, our team has obtained the following significant achievements based on this data platform: I. We applied for two software copyrights: a computer software copyright for "Severe Pancreatitis Clinical Data Collection and Analysis Software" (registration number: 2009SR054112, issued by the National Copyright Administration, November 23, 2009), and a copyright registration certificate for "Chengqian Severe Acute Pancreatitis Data Platform Software" (registration number: 2018SR1030454, issued by the National Copyright Administration, December 18, 2018). II. Seven articles have been published (Y.Zhang et al., 2014; S.Y.Zhang et al., 2014; Y.Zhang et al., 2015; W.Q.Yu et al., 2016; Y.Zhang et al., 2020; Y.Zhang et al., 2021; F.F.Huang et al., 2020). III. More than 40 medical centers have participated in platform use and data sharing. IV. Eight online RCT studies are being conducted.

Our prospects for the platform are as follows. In order to enhance the utilization rate of the platform, we will look for leading regional experts to modify the data types and preset options based on the diagnosis and treatment habits of doctors and equipment characteristics in each regional center. The goal is to build a data platform with characteristics of multiple regional centers, thereby creating a data-sharing ecology with multi-center homogeneity and personalization. After the two-step approach for assessing the severity of APhas been recognized by experts (Y. Zhang et al., 2021), we plan to further expand the number of cases (30,000 cases) by 100 times based on multi-center joint application of the pancreatitis data platform, to further verify the accuracy and universality of the two-step approach. New data platforms for clinical research, such as REDCap and Castor EDC, have already been designed by clinicians to obtain data from electronic medical records, but they are all

Unedited

designed for a single clinical study and still belong to the post-structured data-management mode (M.Botta et al., 2021). Meanwhile, there is no pre-structured model for a single disease in hospital electronic medical record systems (N.S.Abul-Husn et al., 2019). We have successfully developed a new function for the automatic import of objective data from hospital information systems to the platform, effectively reducing the input workload for doctors. Based on the synchronous connection between hospital information systems, equipment interfaces, and single-disease data platforms, our design of a single-disease electronic medical record of AP should serve as a model. IV. To expand this type of single-disease data platform to other diseases, we have established several critical-disease-based data platforms for scenarios such as sepsis, acute renal failure, and intensive-care-unit enteral nutrition work.

The data-management platform does have some limitations. First, it is not an open-access web platform, which potentially limits its general application and practical value. Second, insufficient promotion of the application means that broader coverage of medical centers is needed to facilitate more extensive data sharing. Third, the follow-up design is inadequate. We only designed the data of the acute phase and paid insufficient attention to the data of the long-term prognosis of patients after discharge. The follow-up section will be further designed in the future, including the follow-up of three months, six months, or even more than one year after discharge, focusing on the internal and external secretion function and quality of life of patients.

Data availability statement

The entire dataset of the platform is available for these medical centers that have joined and contributed to the data entry.

Acknowledgments

We are very grateful to Hangzhou Chengqian Yishu Technology Co., Ltd., and Dr. Chuliang Hu from the Institute of Computing Technology of the Chinese Academy of Sciences for their help in the development of the platform. This work was supported by the National Key Research and Development Project (grant number 2019YFC1316000) and National Natural Science Foundation of China (grant numbers 81830089, 81801640, 81802359, 81871545, 81802334).

Author contributions

Yun ZHANG, Shiyin CHEN, and Zhien WANG wrote and corrected the manuscript; Shiyin CHEN and Cheng ZHANG revised the manuscript; Cheng ZHANG, Jian ZHANG, Yanshuai WANG, and Weiwei SI contributed to data entry and daily maintenance; Yun ZHANG, Tingbo LIANG, and Wenqiao YU designed and constructed the data platform and daily maintenance. All authors have read and approved the final manuscript, and therefore, have full access to all the data in the study and took responsibility for the integrity and security of the data.

Compliance with ethics guidelines

Shiyin CHEN, Zhien WANG, Cheng ZHANG, Jian ZHANG, Wenqiao YU, Yanshuai WANG, Weiwei SI, Tingbo LIANG, and Yun ZHANG declare that they have no conflict of interest.

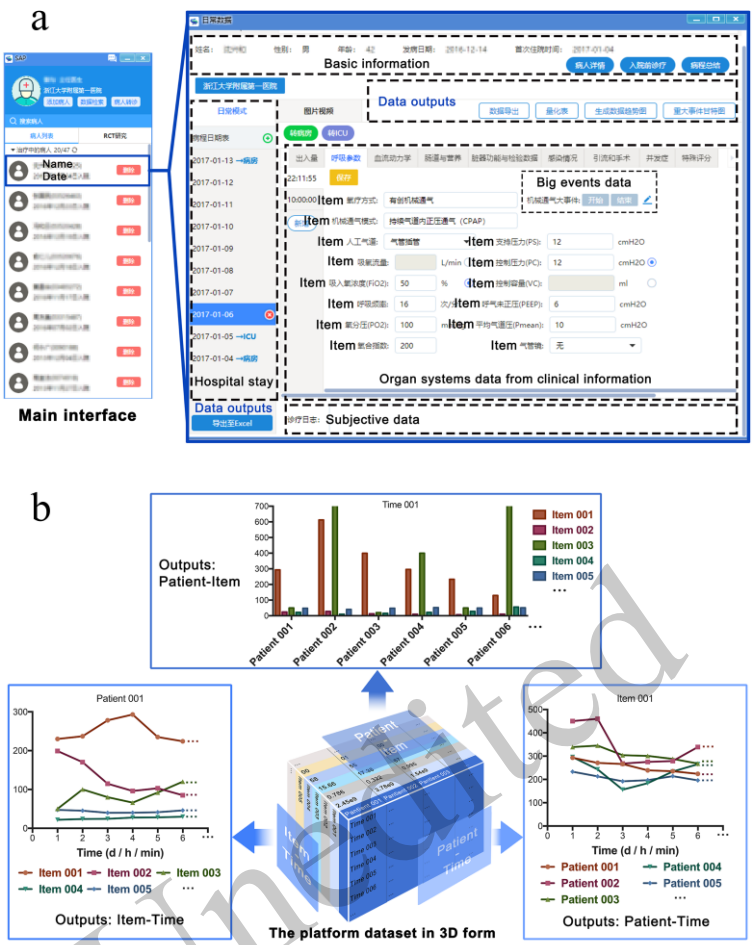
This article does not contain any studies with human or animal subjects performed by any of the authors.

Unedited

Table 1 The data platform compared with others

	Ours	Medbit	REDCap	eICU
Nature	Database+clinical research system	Clinical research system	Clinical research system	Database
Data characteristics	“Pre-structured data” encompasses all the features related to a specific disease	Structured data pertaining to individual trial characteristics	Structured data pertaining to individual trial characteristics	"Post-structured data" solely focuses on the characteristics of the data
Feasibility of conducting homogeneous case retrieval	The ability to efficiently and precisely retrieve information based on feature labels	Not feasible	Not feasible	Not feasible
Study type	Prospective study, retrospective study, cross-sectional study, and real-world study	Prospective study, retrospective study, and cross-sectional study	Prospective study, retrospective study, and cross-sectional study	Retrospective study and cross-sectional study
Trial conduct convenience	Convenient	Convenient	Moderately convenient	Inconvenient
Whether suitable for multi-center study	Suitable	Suitable	Moderately suitable	Not suitable
Whether suitable for single-disease research	Suitable for single disease of acute pancreatitis	Suitable for specific research of single disease	Not suitable	Not suitable
Conduct multi-task and multi-achievement scientific research mode	The feature marker search enables quick determination of multiple research directions and corresponding cases, while the RCT loading module allows for simultaneous execution of multiple test tasks.	Able to perform multi-task and multi-achievement research mode	Unable to perform multi-task and multi-achievement research mode	Unable to perform multi-task and multi-achievement research mode

Unedited



Unedited

Fig. 1

Data input and export of the clinical data management platform. (a) Platform user interface. The user clicks on "case" from the main interface to enter the detailed interface. The detailed case information includes basic patient information, hospital-stay information, major-event data, organ-system data, subjective input data, and data outputs. (b) The three-dimensional data platform comprises the patient (case), item, and time. Data can be output in three forms: patient-item, item-time, and patient-time.

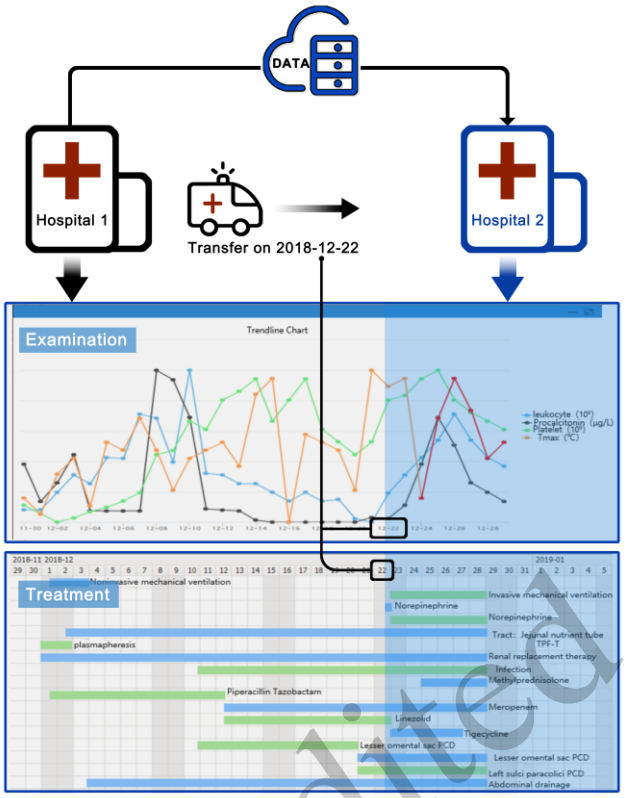


Fig. 2

Data transmission among transferred patients. When a patient is transferred from hospital 1 to hospital 2, all relevant clinical data on the patient, including treatment-related major-event data presented in the Gantt chart and examination-related measurement data presented in the trend chart, automatically transfers to the next hospital via the data platform cloud to ensure the continuity and integrity of the case data.

Unedited

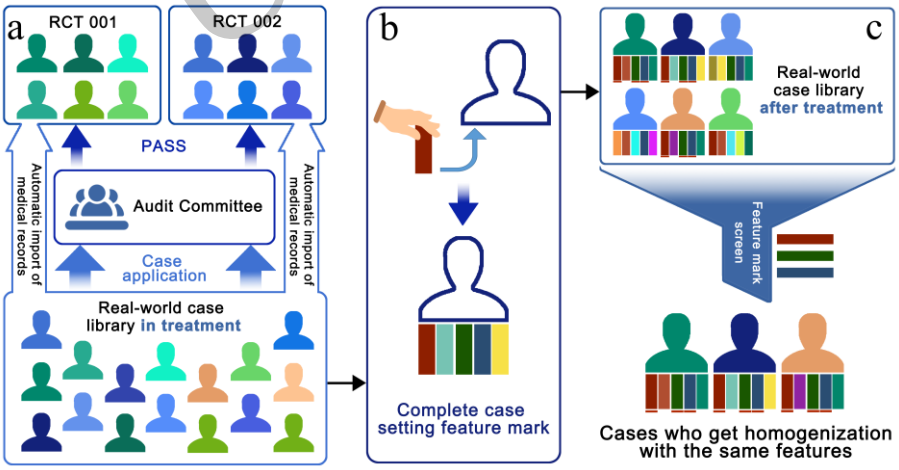


Fig. 3

Research mode of randomized controlled trial (RCT) and establishment and retrieval of homogeneous patients based on feature marks. (a) Real-world cases with heterogeneity in the treatment process will be submitted to the audit committee for RCT inclusion evaluation by the doctor in charge. If the case fulfills the enrollment criteria and is approved by the audit committee, it will be entered into the corresponding RCT study. Meanwhile, the relevant clinical data of these enrolled cases before enrollment will be automatically imported into the RCT study through the clinical data platform, and the RCT study data will then be recorded on the platform. (b) In the clinical treatment process, all cases are labeled in the clinical data platform with feature marks which include the disease-classification standard, organ-dysfunction

classification, imaging features, infection, and major treatment events , and these labels highly condense the clinical characteristics of cases. (c) After treatment, cases are already labeled with feature marks. Researchers can screen for specific feature marks based on their research needs, to obtain a series of homogeneous cases.

References

Boxhoorn L, Voermans RP, Bouwense SA, et al., 2020. Acute pancreatitis. *Lancet*, 396(10252):726-734. https://doi.org/10.1016/S0140-6736(20)31310-6

Finazzi S, Paci G, Antiga L, et al., 2020. PROSAFE: a European endeavor to improve quality of critical care medicine in seven countries. *Minerva Anestesiol*, 86(12):1305-1320. https://doi.org/10.23736/S0375-9393.20.14112-9

Leppäniemi A, Tolonen M, Tarasconi A, et al., 2019. WSES guidelines for the management of severe acute pancreatitis. *World J Emerg Surg*. https://doi.org/10.1186/s13017-019-0247-0

Johnson A, Bulgarelli L, Pollard T, et al., 2020. MIMIC-IV (version 0.4). *PhysioNet*. https://doi.org/10.13026/a3wn-hq05

Ke L, Zhou J, Mao W, et al., 2022. Immune enhancement in patients with predicted severe acute necrotising pancreatitis: a multicentre double-blind randomised controlled trial. *Intensive care medicine*, 48(7), 899–909. https://doi.org/10.1007/s00134-022-06745-7

Pollard TJ, Johnson AEW, Raffa JD, et al., 2018. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Scientific data*, 5, 180178. https://doi.org/10.1038/sdata.2018.178

Docherty AB, Mulholland RH, Lone NI, et al., 2021. Changes in in-hospital mortality in the first wave of COVID-19: a multicentre prospective observational cohort study using the WHO Clinical Characterisation Protocol UK. *The Lancet.* Respiratory medicine, 9(7), 773–785. https://doi.org/10.1016/S2213-2600(21)00175-2

Rajaram A, Hickey Z, Patel N, et al., 2020. Training medical students and residents in the use of electronic health records: a systematic review of the literature. *J Am Med Inform Assoc*, 27 (1) 175-180. https://doi.org/10.1093/jamia/ocz178

Zhang SY, Liang ZY, Yu WQ, et al., 2014. Early enteral nutrition with polymeric feeds was associated with chylous ascites in patients with severe acute pancreatitis. *Pancreas*, 43(4):553-558. https://doi.org/10.1097/MPA.0000000000000067

Zhang Y, Gao SL, Zhang SY, et al., 2014. Six cases of severe acute pancreatitis complicated with vancomycin-resistant enterococcus enteritis. *Shock*, 42(5):400-406. https://doi.org/10.1097/SHK.0000000000000237

Zhang Y, Zhang SY, Gao SL, et al., 2015. Successful Resolution of Gastric Outlet Obstruction Caused by Pancreatic Pseudocyst or Walled-Off Necrosis After Acute Pancreatitis: The Role of Percutaneous Catheter Drainage. *Pancreas*, 44(8):1290-1295. https://doi.org/10.1097/MPA.0000000000000429

Yu WQ, Zhang Y, Zhang SY, et al., 2016. Impact of misplaced subclavian vein catheter into jugular vein on transpulmonary thermodilution measurement variables. *J Zhejiang Univ Sci B*,17(1):60-66. https://doi.org/10.1631/jzus.B1500167

Zhang Y, Yu WQ, Zhang J, et al., 2020. Efficacy of Early Percutaneous Catheter Drainage in Acute Pancreatitis of Varying Severity Associated With Sterile Acute Inflammatory Pancreatic Fluid Collection. *Pancreas*, 49(9):1246-1254. https://doi.org/10.1097/MPA.0000000000001666

Huang F, Meng H, Sun Y, et al., 2020. Factors Causing Abnormal Heart Failure After Successful Fluid Resuscitation on Acute Pancreatitis Patients. *Pancreas*, 49(4):e33-e35. https://doi.org/10.1097/MPA.0000000000001532

Zhang Y, Zhang C, Yu WQ, et al., 2021. The "Two-Step" approach for classifying the severity of acute pancreatitis: A retrospective study. *J Hepatobiliary Pancreat Sci*, 28(10):902-912. https://doi.org/10.1002/jhbp.943

Schepers NJ, Bakker OJ, Besselink MG, et al., 2019. Impact of characteristics of organ failure and infected necrosis on mortality in necrotising pancreatitis. *Gut*, 68(6):1044-1051. https://doi.org/10.1136/gutjnl-2017-314657

Botta M, Tsonas AM, Pillay J, et al., 2021. Ventilation management and clinical outcomes in invasively ventilated patients with COVID-19 (PRoVENT-COVID): a national, multicentre, observational cohort study. *Lancet Respir Med*, 9(2):139-148. https://doi.org/10.1016/S2213-2600(20)30459-8

Abul-Husn NS, Kenny EE, 2019. Personalized Medicine and the Power of Electronic Health Records. *Cell*, 177(1):58-69. https://doi.org/10.1016/j.cell.2019.02.039

Supplementary information:

Tables S1 and S2; Figs. S1 and S2

Unedited