World Journal of Emergency Surgery

STUDY PROTOCOL

Open Access

Machine learning for the rElapse risk eValuation in acute biliary pancreatitis: The deep learning MINERVA study protocol

Mauro Podda[1*], Adolfo Pisanu[1], Gianluca Pellino[2,3], Adriano De Simone[4], Lucio Selvaggi[2], Valentina Murzi[1], Eleonora Locci[1], Matteo Rottoli[5], Giacomo Calini[5], Stefano Cardelli[5], Fausto Catena[6], Carlo Vallicelli[6], Raffaele Bova[6], Gabriele Vigutto[6], Fabrizio D'Acapito[7], Giorgio Ercolani[7], Leonardo Solaini[7], Alan Biloslavo[8], Paola Germani[8], Camilla Colutta[8], Savino Occhionorelli[9], Domenico Lacavalla[9], Maria Grazia Sibilla[9], Stefano Olmi[10,11], Matteo Uccelli[11], Alberto Oldani[11], Alessio Giordano[12], Tommaso Guagni[12], Davina Perini[12], Francesco Pata[13,14], Bruno Nardo[13,14], Daniele Paglione[14], Giusi Franco[14], Matteo Donadon[15], Marcello Di Martino[15], Dario Bruzzese[4] and Daniela Pacella[4]

Abstract

Background Mild acute biliary pancreatitis (MABP) presents significant clinical and economic challenges due to its potential for relapse. Current guidelines advocate for early cholecystectomy (EC) during the same hospital admission to prevent recurrent acute pancreatitis (RAP). Despite these recommendations, implementation in clinical practice varies, highlighting the need for reliable and accessible predictive tools. The MINERVA study aims to develop and validate a machine learning (ML) model to predict the risk of RAP (at 30, 60, 90 days, and at 1-year) in MABP patients, enhancing decision-making processes.

Methods The MINERVA study will be conducted across multiple academic and community hospitals in Italy. Adult patients with a clinical diagnosis of MABP, in accordance with the revised Atlanta Criteria, who have not undergone EC during index admission will be included. Exclusion criteria encompass non-biliary aetiology, severe pancreatitis, and the inability to provide informed consent. The study involves both retrospective data from the MANCTRA-1 study and prospective data collection. Data will be captured using REDCap. The ML model will utilise convolutional neural networks (CNN) for feature extraction and risk prediction. The model includes the following steps: the spatial transformation of variables using kernel Principal Component Analysis (kPCA), the creation of 2D images from transformed data, the application of convolutional filters, max-pooling, flattening, and final risk prediction via a fully connected layer. Performance metrics such as accuracy, precision, recall, and area under the ROC curve (AUC) will be used to evaluate the model.

Discussion The MINERVA study aims to address the specific gap in predicting RAP risk in MABP patients by leveraging advanced ML techniques. By incorporating a wide range of clinical and demographic variables, the

*Correspondence:
Mauro Podda
mauro.podda@unica.it

Full list of author information is available at the end of the article

MINERVA score aims to provide a reliable, cost-effective, and accessible tool for healthcare professionals. The project emphasises the practical application of AI in clinical settings, potentially reducing the incidence of RAP and associated healthcare costs.

Trial registration ClinicalTrials.gov ID: NCT06124989.

Keywords Acute biliary pancreatitis, Recurrence, Hospital readmission, Cholecystectomy, Machine learning, Artificial intelligence

Introduction

Biliary acute pancreatitis (BAP) represents a significant healthcare burden globally, with a high incidence rate and associated mortality [1–3]. Current guidelines from major organisations such as the American College of Gastroenterology [4], American Gastroenterological Association [5], International Association of Pancreatology [6], World Society of Emergency Surgery [2], and British Society of Gastroenterology [7] recommend performing early cholecystectomy (EC) during the same hospital admission for patients with mild acute biliary pancreatitis (MABP) according to the revised Atlanta classification [8]. This recommendation is based on evidence that EC significantly lowers the risk of recurrent acute pancreatitis (RAP) and other biliary events without increasing surgical risks [9, 10].

Recent meta-analyses [11] showed that the recurrence rate among patients with MABP managed conservatively is significantly higher than in patients submitted to EC (35% vs. 11%). In Stevens et al. [12], patients who had cholecystectomy had a recurrence rate of 20% versus 43% of those managed conservatively. Several randomised controlled trials comparing EC versus delayed cholecystectomy (DC) for gallstone disease reported recurrence of symptoms/complications in the waiting period before surgery [13–18]. Some authors state that EC also shortens the total hospital stay in people with MABP under the condition that appropriate facilities and expertise are available [19–21].

Recently, the MANCTRA-1 study highlighted discordant gaps between daily clinical practice and recommendations from BAP guidelines, especially regarding the implementation of EC strategies [22, 23]. Beyond the mere lack of penetration of EC guidelines in MABP, potential knowledge-to-action gaps in EC implementation persist widely due to logistical obstacles, historical assumptions, and poor awareness of evidence.

The risk of RAP represents a fundamental outcome that the medical staff needs to carefully assess for every patient with MABP before deciding on a course of action. This assessment should be fast, reliable, accurate, and free of costs and extra procedures. However, to date, no agreed-upon, convenient, or economic methods have been developed to predict RAP. A recent review [24] on machine learning models (ML) for acute pancreatitis-related outcomes reported only two studies out of a total of 24 which tried to predict RAP, showing that it was the least explored outcome. All studies that tried to use statistical or artificial intelligence (AI) models for RAP prediction had several limitations: they focused on differential diagnosis [25] without providing a risk prediction, and required the extraction of radiomics features [26, 27]. Moreover, all of them were conducted on a small cohort of patients and/or had a retrospective design; thus, they lacked a model and score validation and a standardised data collection procedure. Finally, none used an ML model for feature extraction, but they all used a separate statistical model, such as LASSO regression [27].

Since MBAP patients represent a very varied and heterogeneous population, every medical information that can be recorded at index admission represents a useful vector of evidence to assess the risk of RAP. Within this context, the MINERVA (Machine learnINg for the rElapse Risk eValuation in Acute biliary pancreatitis) study stems from the need in the clinical practice of taking an operational decision in patients that are admitted to the hospital with a diagnosis of MABP.

Objectives

The MINERVA study aims to advise clinicians wisely about the right decision to take by providing a validated and standardised score of RAP risk that considers each patient's personal history, demographic data, and laboratory characteristics. The MINERVA score will allow to assess the risk of hospital readmission due to RAP for patients diagnosed with MABP who did not undergo EC during the index hospital admission using ML and AI. Beyond the construction of the model, the MINERVA study aims to reach the validation of the MINERVA score on an extensive, multicentric, prospective cohort and allow national and international clinicians, medical staff, researchers and the general audience to freely and easily access the MINERVA score computation and use it in their daily clinical practice.

Methods

This study protocol has been produced in accordance with the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) [28] and SPIRIT-AI (SPIRIT–Artificial Intelligence) guidelines [29].

Study setting

The MINERVA project will be conducted across a mix of academic and community hospitals in Italy. The participating centres are strategically chosen to ensure a comprehensive and representative sample of patients, facilitating the development and validation of a robust predictive model for assessing the risk of RAP in MABP patients (Supplementary Table 1).

Inclusion criteria at the patient's level

All consecutive adult patients (≥ 18 years old) of both sexes admitted to the participating centres (surgical departments and/or gastroenterology departments and/or internal medicine departments) with a clinical diagnosis of MABP (according to the Revised Atlanta Classification) [8], and not submitted to cholecystectomy or ERCP/ES (Endoscopic Retrograde CholangioPancreatography/Endoscopic Sphincterotomy) with the aim of definitive therapy during the same hospital admission will be included in the study.

Exclusion criteria at the patient's level

Patients with acute pancreatitis of aetiology other than gallstones, moderately-severe and severe acute pancreatitis (according to the Revised Atlanta Classification), presence of pancreatic necrosis, pregnant patients, patients not able to sign the informed consent to participate in the study will be excluded [Table 1].

Eligibility criteria for study centres

Participant centres must be located in Italy and have the necessary facilities to diagnose and treat patients with BAP. Centres have been selected based on the number of patients (higher first) recruited by each hospital for the "coMpliAnce with evideNce-based cliniCal guidelines in the managemenT of acute biliaRy pancreAtitis" (MANCTRA-1) study [22]. Centres must agree to follow the study protocol, including the recruitment, data collection, and follow-up procedures and obtain local Institutional Review Board (IRB) approval for participation in the study.

Table 1 Exclusion criteria based on the revised Atlanta classification of acute pancreatitis severity

Severity of Acute Pancreatitis	Revised Atlanta Classification	Exclusion Criteria in MINERVA Study
Mild	No organ failure, no local/systemic complications	Included in the study
Moderately Severe	Transient organ failure (< 48 h) and/or local/systemic complications without persistent organ failure	Excluded from the study
Severe	Persistent organ failure (> 48 h), single or multiple	Excluded from the study

*Notes: Organ failure is assessed based on the Modified Marshall Scoring System. Local complications include peripancreatic fluid collections, pancreatic necrosis, and abscesses. Systemic complications refer to exacerbation of pre-existing co-morbidities due to the acute pancreatitis episode

Eligibility criteria for investigators performing interventions

Surgeons, gastroenterologists, and other healthcare professionals involved in the MINERVA study must be fully qualified and licensed to practice in Italy. Healthcare professionals must have experience in diagnosing and treating BAP. Individuals must follow the study protocol and ensure accurate and timely data collection and reporting.

Primary outcome

The primary outcome of the MINERVA study is predicting the risk of RAP in patients after a first episode of MABP treated conservatively. Recurrence will be evaluated at 30, 60, 90 days, and at 1-year follow-up to ensure a comprehensive assessment of the primary outcome. This outcome will be reached by developing and validating a novel risk score. The MINERVA score will be grounded upon an ML model that considers patients' demographic and laboratory variables and data that can be easily collected and recorded at index patient admission.

Secondary outcomes

The secondary outcomes of the MINERVA study are to compare the accuracy of the MINERVA ML model to other traditional ML models previously adopted in literature (such as ANN and SVM) and with statistical models (such as multiple regression), and apply the MINERVA ML model to the prediction of MABP complications.

Study phases

In the development phase, the model will initially be developed using retrospective data from patients previously diagnosed with MABP in the MANCTRA-1 study [22]. Demographics, clinical history, and laboratory data will be collected to train the ML model. In the validation phase (study start date: January 1, 2024; primary completion date: December 31, 2024; study completion date: December 31, 2025), the model will be validated on a new, prospective cohort of patients diagnosed with MABP (Fig. 1). The accuracy, sensitivity, specificity, and overall performance of the predictive score will be compared to existing models.

Supplementary Table 2 reports the data points recorded for a patient's data to be included in the study.

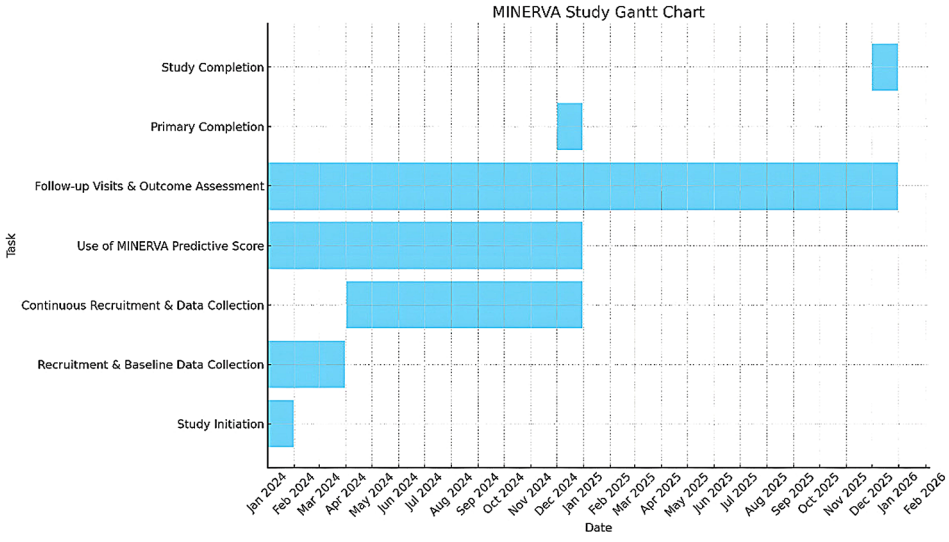

Fig. 1 MINERVA study gantt chart

Data will be excluded from the study if any criteria are met: incomplete demographic or clinical data, incomplete laboratory data, or incorrect or inconsistent data. The procedure for assessing and handling poor quality or unavailable input data in the MINERVA project will involve a combination of automated and manual data checks, data correction, imputation, and exclusion processes. Continuous data quality monitoring, proactive measures, and quality assurance audits will be implemented to maintain high data integrity standards.

Sample size

There is no strict or agreed scientific method to establish the minimum sample size required to train a Convolutional Neural Network (CNN) effectively. While deep models benefit in terms of accuracy the larger the sample size is, establishing the minimum required is still an open question. One common rule of thumb [30], which is the most widely adopted up-to-date, is the "10 times" rule. The rule proposes having at least 10 dataset examples per each neural network weight. To calculate the number of weights in a convolutional layer, it is sufficient to multiply the number of filters (n) times the height (h) and width (w) of the filter times the number of channels (c) and finally sum the number of biases that are equal to the number of filters, obtaining: n * h * w * c + n. The CNN model presented in the MINERVA study has 4 filters (equal to the 4 types of variables) of size 3×3 and 3 channels (RGB). Thus, the minimum sample size required by the rule equals 112*10 = 1120. A minimum of 692 (retrospective, from the MANCTRA-1 study) and 430 (prospective) patients (for a total of 1122) will be recruited (Fig. 2).

Statistics

The MINERVA study model will consist of the following steps (Fig. 3):

DeepInsight Spatial transformation All model variables are processed with kernel Principal Component Analysis (kPCA); the Convex Hull of the scatterplot of the main components is computed and the smallest rectangle is extracted; the rectangle is transformed into a 2D image with a fixed resolution using feature averaging and normalisation.

Convolutional neural network structure convolution layer, in which 4 filters are applied to the input; a max-pooling layer, which further reduces the dimensionality of the image; flattening operation of the obtained matrix, which combines the information extracted from the filters into a single one-dimensional vector; a fully connected layer; output layer, in which the risk prediction made by the model is computed.

Performance To prevent overfitting, the dataset will be split into a training set, a test set, and a validation set. Additionally, k-fold cross-validation will be used. The performance of the MINERVA model will be evaluated using the most adopted accuracy measures, such as precision, recall, and AUC (Area Under the ROC Curve). Additionally, its performance will be compared with that achieved using traditional machine learning methods (SVM, ANN).

Data collection and management methods

The MINERVA study will use REDCap (Research Electronic Data Capture) for data entry, coding, security, and storage. At hospital admission, healthcare providers will enter patient demographic, clinical history, and

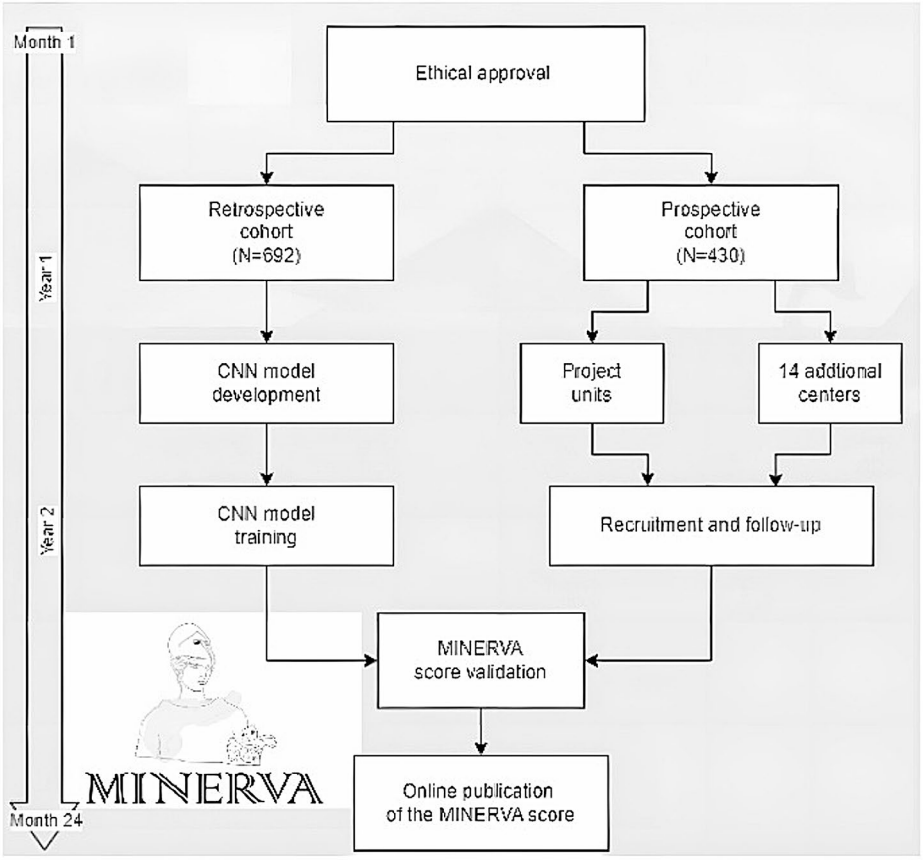

Fig. 2 MINERVA study flow chart

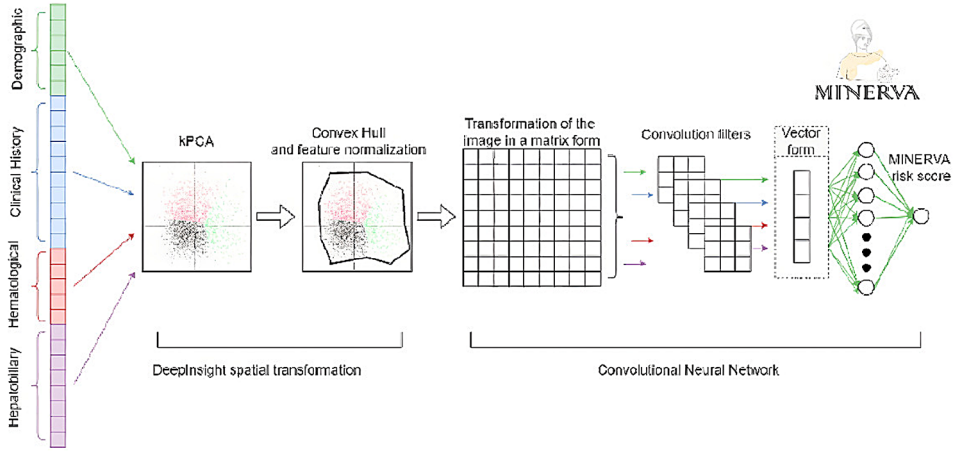

Fig. 3 MINERVA study ML-AI model (deepinsight spatial transformation, convolutional neural network, performance)

laboratory data into the MINERVA system. Standardised electronic data collection forms will be implemented within REDCap to ensure uniformity in data entry across all centres. Each participating unit includes two to three trained researchers responsible for data collection, all of whom have been trained in the use of REDCap and the study protocol to ensure consistency and accuracy. Rigorous data quality checks and double data entry will be implemented for critical data points to minimise errors. This involves two independent entries of the same data, followed by a comparison to identify and resolve discrepancies. Data will be entered in real-time or as soon as possible after patient interactions to ensure timely and accurate data capture. Implementing range checks within REDCap will automate data quality checks, validate data entries, and flag values outside predefined acceptable

ranges. The University of Cagliari will continuously perform data collection and entry quality checks during the study to minimise the effects of missing and inaccurate data.

Missing data will be managed using standard imputation methods, such as mean or median imputation for continuous variables and mode imputation for categorical variables, depending on the nature of the data. Advanced imputation techniques, such as multiple imputation or k-nearest neighbors (k-NN), will be employed if deemed appropriate during data analysis.

Interim analyses for the MINERVA project will be conducted every three months by the University of Cagliari and the University of Naples Federico II. The primary purposes of these analyses are to monitor participant safety, ensure data quality, and assess the preliminary efficacy of the AI predictive model.

Data safety information

Personal information about potential and enrolled patients will be collected at the time of hospital admission and during follow-up visits. Participants will be informed about the type of personal information collected and the purpose of its collection during the informed consent process. Written informed consent for the MINERVA study will be obtained before any personal information is recorded by trained study coordinators and research nurses under the oversight of the local coordinator at each participating centre. The local coordinator will provide the patient with a copy of the consent form. In cases where participants cannot consent, authorised surrogates can provide consent following the same ethical procedures.

Patients will be informed about how their data will be collected, shared, and maintained during consent. They will also be informed about their rights to access, correct, or withdraw personal data. In the unlikely event of a data breach, participants will be promptly informed about the breach, the potential impact, and the measures taken to mitigate any risks. Comprehensive training, documentation, and ethical compliance measures will be in place to ensure the integrity of the consent process.

The confidentiality of the online REDCap database will be ensured by appropriate standard operating procedures, including the use of passwords available only to the project staff involved. Personal identifiers will be replaced with unique participant codes to ensure de-identification. A master list linking patient codes to personal identifiers will be maintained separately and securely by the study coordinators at each centre. Only authorised personnel involved in the study (Principal Investigator, study coordinators, data managers) will have access to personal information, and role-based access control within the REDCap system will limit access to sensitive data based on the user's role and responsibilities.

When the validation and editing process is concluded, the formal 'locking' of the database will be documented. All electronic data will be stored in the REDCap system, hosted on secure servers with regular backups, while physical documents will be stored in locked cabinets in secure areas with restricted access. Personal data will be retained only as long as necessary to fulfil the purposes of the study and comply with regulatory requirements. However, for safety reasons and audit and inspection after the project completion, documents and any database records will be retained for at least five years following the end of the project.

At the end of the study, only aggregated and anonymised data will be shared with third parties (regulatory bodies, scientific publications).

Access to data

The local coordinator at each participating centre will have access to the final trial dataset. The DMC (Data Monitoring Committee) members will have access to the final trial dataset to perform their oversight duties, including safety monitoring and data quality assessments. Qualified statistical analysts involved in the data analysis phase will have access to the final dataset to perform statistical evaluations and generate study results. All individuals with access to the final dataset will be required to sign confidentiality agreements to ensure that data is handled in accordance with ethical standards and data protection regulations. These agreements will specify the responsibilities of each individual in maintaining data confidentiality and integrity.

Conflict of interest and independence

Investigators will retain full access to the final trial dataset without restrictions imposed by contractual agreements, ensuring their independence in conducting analyses and reporting results. Any potential conflicts of interest will be disclosed and managed in accordance with institutional policies.

Risk management plan

The MINERVA study will implement continuous monitoring and error identification mechanisms to ensure the AI model's reliability and safety. Real-time data monitoring will track the model's performance through a web-based platform, focusing on key metrics like sensitivity, specificity, positive predictive value, and negative predictive value. Automated logging will be developed to capture instances where the model's predictions deviate significantly from clinical outcomes. Additionally, healthcare providers will be encouraged to report any observed performance errors or unexpected outcomes manually.

Table 2 Risk management plan for procedural and technical risks

Risk	Proposed solution
Procedural risk: risks related to patient recruitment.	If the number of patients is not met by the 6th month of the project, the following centers will be contacted for the enrollment of additional patients: 1. General and Oncologic Surgery Unit, Santa Croce and Carle Hospital, Cuneo, Italy 2. Chirurgia Generale 2, ASST Spedali Civili di Brescia, Brescia, Italy 3. Policlinico Umberto I Sapienza University of Rome, Rome, Italy
Technical Risk: Difficulties and potential delays in the implementation of the algorithm.	Machine learning experts will be sought, who have already collaborated with the PI and local leads of the project such as: • Institute of Cognitive Sciences and Technologies, National Research Council, CNR, Italy; • Smarted srl, Italy. A start-up with years of expertise in national and international research projects in the field of ML; • Prof. Barbara Webb, from the Institute for Perception, Action and Behaviour, School of Informatics, University of Edinburgh, United Kingdom • ML experts Prof. Davide Marocco and Dr. Onofrio Gigliotta (Federico II University, Naples, Italy)

When identified, performance errors will be categorised based on their nature, such as data input errors, algorithmic errors, or implementation issues. The severity of each error will be assessed based on its potential impact on patient care and outcomes. For significant errors, a thorough root cause analysis will be conducted, involving a detailed review of data inputs, algorithm processes, and system logs. Clinical and data science experts will participate in these analyses to ensure a comprehensive understanding of the errors and their implications.

Insights from the error analysis will refine and retrain the AI model, enhancing its accuracy and reliability. This may involve adjusting algorithm parameters and thresholds based on observed error patterns. Additional data validation checks will be implemented to prevent future errors, and the dataset will be enriched with relevant features that could improve model performance. System enhancements, such as updating the user interface to be more intuitive and integrating decision support tools, will also be made to help healthcare providers use the AI model more effectively and identify potential errors.

Performance reviews will be conducted to evaluate the AI model's effectiveness and identify any recurring issues. Feedback from healthcare providers using the AI model will be regularly collected to identify practical challenges and areas for improvement. Monitoring patient outcomes will ensure that the AI model positively contributes to clinical decision-making and patient care.

Potential risks associated with the AI model, such as incorrect predictions, data breaches, and implementation challenges, will be identified and assessed for their likelihood and impact. Preventive measures will be implemented to minimise the occurrence of these risks. Contingency plans will be developed to address identified risks promptly and effectively. Detailed error reports will be maintained, documenting each identified performance error's nature, cause, and resolution. Regular updates on the AI model's performance and reliability will be provided to all stakeholders, including healthcare providers, researchers, and regulatory bodies.

Risk management plans for procedural and technical risks are reported in Table 2.

Dissemination and exploitation policy

Dissemination activities and exploitation of the results of the MINERVA study will be carried out continuously throughout the project. Dissemination outputs will target clinicians in the fields of general surgery and gastroenterology, academic communities, potential stakeholders in biostatistics and clinical research and the general public. The dissemination activities will aim to reach out to healthcare professionals and clinicians in order to disseminate the scope of the MINERVA study and teach in actual practice the use and the interpretation of the MINERVA score and its online dashboard, disseminate the MINERVA score model development (and validation) in the scientific community of ML for healthcare so to collect national and international experts' feedback and inform the general audience at the national and international as well as other medical workers about the MINERVA score development and use.

To achieve these objectives, a website (https://www.minervaproject.org/) has been created to disseminate pub lic deliverables, updates, and news on the project, publish scientific articles, and hold advertising events where the project is presented. The prototype of the implemented ML system, accessible free of charge and in full in a GitHub repository dedicated to the project, will be available online. The project's methodological framework will be presented at specific national and international conferences.

Communication plan for trial results

Enrolled patients will receive a summary report of the study results in plain language that is easy to understand. This report will include key findings and any relevant implications for their health. The patients will have the opportunity to discuss the results with their healthcare providers during follow-up visits. This ensures they receive personalised information and can ask questions about their participation and the study outcomes.

Authorship and acknowledgment

The MINERVA study follows the International Committee of Medical Journal Editors (ICMJE) guidelines

for authorship to ensure that all contributors are appropriately recognised. All funding sources and contributions from participating centres and individuals will be acknowledged in all publications and presentations. The study will adopt a collaborative group authorship approach, with individual contributions detailed in an appendix or supplementary material.

Data sharing

The study results will be published in open-access journals to maximise accessibility, and any data-sharing arrangements will comply with ethical standards and data protection regulations, including the General Data Protection Regulation (GDPR).

Expected impact

The expected impact of the MINERVA project will be on three different levels:

1. *At the patient's level.* The patients, who are the main focus of the MINERVA project, will finally benefit the most from applying the MINERVA score. Patients with similar conditions will receive the same treatment thanks to the MINERVA score. The best treatment choice will not be just on the specific clinicians. However, it will become justified by a validated, reliable, and free methodology that can be computed immediately with demographic and laboratory data without further elaboration. Although the MINERVA score will be validated only on the Italian population, the future of the MINERVA application will be to expand the model dataset to international data, allowing the use of the MINERVA score for other populations.
2. *At the medical staff level.* The aim of the MINERVA score is to deliver a tangible application, easy to reach from any health structure and/or mobile device, easy to interpret for specialists and non-specialists and easy to compute with data that can be easily collected at the patients' index admission, without requiring further or longer elaboration. Clinicians and health professionals will also be guided to the use and interpretation of the MINERVA score since the website that will host the MINERVA score dashboard and computation will include documentation that will explain not only how to use the dashboard and how to obtain the MINERVA score from the patient's data, but also how the score is computed and what is its accuracy.
3. *At the healthcare structure and economic level.* The economic impact of the MINERVA study will be two-fold. On the one hand, it will represent the first-ever ML-based assessment methodology developed to compute the risk of relapse using only index demographic and haematologic patients' data. The MINERVA score will also be completely free to access and compute for all healthcare professionals, researchers, and any interested public. On the other hand, reliably predicting the risk of RAP allows the clinician to make an informed and standardised decision on the patients' disease treatment course that prevents a second unexpected hospital admission, complications or a chronic evolution of the disease.

Monitoring impact

Several qualitative and quantitative indicators have been identified to measure the impact of the MINERVA study. Quantitative measures include the number of patients and of clinicians that use the MINERVA score, visits to the MINERVA website, people reached by the dissemination activities in each local recruitment unit and project unit, spontaneous users of the technology (not connected with our recruitment centres), relevant beneficiaries and stakeholders that attend dissemination events, citations of project publications. Quantitative measures include self-assessment questionnaires, evaluation surveys, interviews, feedback and reports received from health structures, stakeholders and the general public about the experience with MINERVA.

The Steering Committee will monitor the impact of the communication efforts through metrics such as citations, media coverage, and feedback from healthcare professionals and participants. Moreover, it will establish a feedback mechanism to gather input from all stakeholders, including participants, healthcare professionals, and collaborators.

Limitations

The MINERVA study protocol has certain limitations. The reliance on retrospective data from the MANCTRA-1 study in the development phase may introduce biases related to data quality and completeness, potentially impacting model accuracy. While the sample size adheres to established guidelines for CNN training, it may limit the model's ability to capture rare or complex patterns. The study's focus on Italian healthcare settings could restrict the external validity of findings to other international contexts. Furthermore, the complexity of CNN-based models may pose challenges in interpretability, potentially limiting their usability among clinicians without technical expertise.

Ethics

The study has undergone Institutional Review Board Ethical approval (ID 2.7 09/01/2024, University of Cagliari, Italy). Before the study is conducted, participants and caregivers, if applicable, will be fully informed about the

methods, procedures, handling of data, potential risks, and right to withdraw. Participants will be fully debriefed about the study aims and hypotheses in verbal and written form. The debrief will also include information about participants' right to withdraw, data handling, and contact information of the researchers and ethics committee.

Participants have the right to withdraw from the study at any point (before, during, or after). This will be clearly communicated in the information sheet and debrief, which also contain the contact details of the project principal investigator and local coordinator.

Sponsor

The study is sponsored by the University of Cagliari (Italy), Naples Federico II (Italy) and Naples Università della Campania "Luigi Vanvitelli" (Italy), and funded by the Fondo per il Programma Nazionale di Ricerca e Progetti di Rilevante Interesse Nazionale (PRIN), under the grant number 202273A4YP. The funder is not involved in the study design, collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication. It will also not influence the decision to submit the results for publication.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13017-025-00594-7.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

The authors thank Dr. Cesare Lofaro from the University of Cagliari for his invaluable assistance in using REDCap for data management in this study.

Author contributions

MP, AP, GP, ADS, LS, VM, EL, MR, GC, SC, FC, CV, RB, GV, FDA, GE, LS, AB, PG, CC, SO, DL, MGS, SO, MU, AO, AG, TG, DP, FP, BN, DP, GF, MD, MDM, DB and DP.

Funding

The study is sponsored by the University of Cagliari (Italy), Naples Federico II (Italy), and Naples Università della Campania "Luigi Vanvitelli" (Italy) and funded by the Fondo per il Programma Nazionale di Ricerca e Progetti di Rilevante Interesse Nazionale (PRIN) under grant number 202273A4YP.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Consent for publication

All authors of this study protocol provide their full consent for the publication of this work.

Competing interests

The authors declare no competing interests.

Ethics approval

The study has undergone Institutional Review Board Ethical approval (ID 2.7 09/01/2024, University of Cagliari, Italy).

Author details

[1]Department of Surgical Science, Emergency Surgery Unit, University of Cagliari, Cagliari, Italy
[2]Department of Advanced Medical and Surgical Sciences, Università Degli Studi Della Campania "Luigi Vanvitelli", Naples, Italy
[3]Colorectal Surgery, Vall d'Hebron University Hospital, Universitat Autonoma de Barcelona UAB, Barcelona, Spain
[4]Department of Public Health, University of Naples Federico II, Naples, Italy
[5]Department of Medical and Surgical Sciences, Alma Mater Studiorum University of Bologna, Bologna, Italy
[6]Department of General and Emergency Surgery, Bufalini Hospital-Level 1 Trauma Center, Cesena, Italy
[7]Department of Medical and Surgical Sciences (DIMEC), University of Bologna, Forlì, Italy
[8]General Surgery Department, Trieste University Hospital, Trieste, Italy
[9]Emergency Surgery Department, Ferrara University Hospital, Ferrara, Italy
[10]Università Vita-Salute San Raffaele, Milan, Italy
[11]Department of General and Oncological Surgery, Minimally Invasive Surgery Center, San Marco Hospital GSD, Bergamo, Zingonia, Italy
[12]Unit of Emergency Surgery, Careggi University Hospital, Florence, Italy
[13]Department of Pharmacy, Health and Nutritional Sciences, University of Calabria, Rende 87036, Italy
[14]Azienda Ospedaliera Annunziata, Cosenza, Italy
[15]Department of Health Sciences, University of Piemonte Orientale, Novara, Italy

Received: 15 July 2024 / Accepted: 22 February 2025
Published online: 03 March 2025

References

1. Roberts SE, Morrison-Rees S, John A, Williams JG, Brown TH, Samuel DG. The incidence and aetiology of acute pancreatitis across Europe. Pancreatology 2017 Mar-Apr;17(2):155–65. https://doi.org/10.1016/j.pan.2017.01.005. Epub 2017 Jan 19. PMID: 28159463.
2. Leppäniemi A, Tolonen M, Tarasconi A, Segovia-Lohse H, Gamberini E, Kirkpatrick AW, Ball CG, Parry N, Sartelli M, Wolbrink D, van Goor H, Baiocchi G, Ansaloni L, Biffl W, Coccolini F, Di Saverio S, Kluger Y, Moore E, Catena F. 2019 WSES guidelines for the management of severe acute pancreatitis. World J Emerg Surg. 2019;14:27. https://doi.org/10.1186/s13017-019-0247-0. PMID: 31210778; PMCID: PMC6567462.
3. Umans DS, Hallensleben ND, Verdonk RC, Bouwense SAW, Fockens P, van Santvoort HC, Voermans RP, Besselink MG, Bruno MJ, van Hooft JE, Dutch Pancreatitis Study Group. Recurrence of idiopathic acute pancreatitis after cholecystectomy: systematic review and meta-analysis. Br J Surg. 2020;107(3):191–9. Epub 2019 Dec 25. PMID: 31875953; PMCID: PMC7003758.
4. Tenner S, Vege SS, Sheth SG, Sauer B, Yang A, Conwell DL, Yadlapati RH, Gardner TB. American college of gastroenterology guidelines: management of acute pancreatitis. Am J Gastroenterol. 2024;119(3):419–37. https://doi.org/10.14309/ajg.0000000000002645. Epub 2023 Nov 7. PMID: 38857482.
5. Crockett SD, Wani S, Gardner TB, Falck-Ytter Y, Barkun AN, American Gastroenterological Association Institute Clinical Guidelines Committee. American Gastroenterological Association Institute guideline on initial management of acute pancreatitis. Gastroenterology. 2018;154(4):1096–1101. https://doi.org/10.1053/j.gastro.2018.01.032. Epub 2018 Feb 3. PMID: 29409760.
6. Working Group IAP/APA Acute Pancreatitis Guidelines. IAP/APA evidence-based guidelines for the management of acute pancreatitis. Pancreatology. 2013;13(4 Suppl 2):e1–15. https://doi.org/10.1016/j.pan.2013.07.063. PMID: 24054878.
7. Working Party of the British Society of Gastroenterology; Association of Surgeons of Great Britain and Ireland. UK guidelines for the management of acute pancreatitis. Gut. 2005;54(Suppl 3):iii1–9. https://doi.org/10.1136/gut.2004.057026. Pancreatic Society of Great Britain and Ireland; Association of Upper GI Surgeons of Great Britain and Ireland. PMID: 15831893; PMCID: PMC1867800.
8. Banks PA, Bollen TL, Dervenis C, Gooszen HG, Johnson CD, Sarr MG, Tsiotos GG, Vege SS, Acute Pancreatitis Classification Working Group. Classification of acute pancreatitis–2012: revision of the Atlanta classification and definitions by international consensus. Gut. 2013;62(1):102–11. https://doi.org/10.1136/gutjnl-2012-302779. Epub 2012 Oct 25. PMID: 23100216.

9. Podda M, Di Martino M, Ielpo B, Catena F, Coccolini F, Pata F, Marchegiani G, De Simone B, Damaskos D, Mole D, Leppaniemi A, Sartelli M, Yang B, Ansaloni L, Biffl W, Kluger Y, Moore EE, Pellino G, Di Saverio S, Pisanu A. The 2023 MANCTRA acute biliary pancreatitis care bundle: A joint effort between human knowledge and artificial intelligence (ChatGPT) to optimize the care of patients with acute biliary pancreatitis in Western countries. Ann Surg. 2024;279(2):203–12. Epub 2023 Jul 17. PMID: 37450700; PMCID: PMC10782931.
10. Di Martino M, Ielpo B, Podda M. Optimal Timing of Cholecystectomy for Severe Acute Pancreatitis and Understanding Complications and Comorbidities-Reply. JAMA Surg. 2024;159(4):466–467. https://doi.org/10.1001/jamasurg.2023.6932. PMID: 38198179.
11. da Costa DW, Bouwense SA, Schepers NJ, Besselink MG, van Santvoort HC, van Brunschot S, Bakker OJ, Bollen TL, Dejong CH, van Goor H, Boermeester MA, Bruno MJ, van Eijck CH, Timmer R, Weusten BL, Consten EC, Brink MA, Spanier BWM, Bilgen EJS, Nieuwenhuijs VB, Hofker HS, Rosman C, Voorburg AM, Bosscha K, van Duijvendijk P, Gerritsen JJ, Heisterkamp J, de Hingh IH, Witteman BJ, Kruyt PM, Scheepers JJ, Molenaar IQ, Schaapherder AF, Manusama ER, van der Waaij LA, van Unen J, Dijkgraaf MG, van Ramshorst B, Gooszen HG, Boerma D, Dutch Pancreatitis Study Group. Same-admission versus interval cholecystectomy for mild gallstone pancreatitis (PONCHO): a multicentre randomised controlled trial. Lancet. 2015;386(10000):1261–1268. https://doi.org/10.1016/S0140-6736(15)00274-3. PMID: 26460661.
12. Stevens CL, Abbas SM, Watters DA. How Does Cholecystectomy Influence Recurrence of Idiopathic Acute Pancreatitis? J Gastrointest Surg. 2016;20(12):1997–2001. doi: 10.1007/s11605-016-3269-x. Epub 2016 Sep 23. PMID: 27663692; PMCID: PMC5104786.
13. Nebiker CA, Frey DM, Hamel CT, Oertli D, Kettelhack C. Early versus delayed cholecystectomy in patients with biliary acute pancreatitis. Surgery. 2009;145(3):260–4. https://doi.org/10.1016/j.surg.2008.10.012. Epub 2009 Feb 1. PMID: 19231577.
14. Schmidt M, Søndenaa K, Vetrhus M, Berhane T, Eide GE. A randomized controlled study of uncomplicated gallstone disease with a 14-year follow-up showed that operation was the preferred treatment. Dig Surg. 2011;28(4):270–6. https://doi.org/10.1159/000329464. Epub 2011 Jul 9. PMID: 21757915.
15. Mador BD, Panton ON, Hameed SM. Early versus delayed cholecystectomy following endoscopic sphincterotomy for mild biliary pancreatitis. Surg Endosc. 2014;28(12):3337–42. https://doi.org/10.1007/s00464-014-3621-8. Epub 2014 Jun 25. PMID: 24962855.
16. Riquelme F, Marinkovic B, Salazar M, Martínez W, Catan F, Uribe-Echevarría S, Puelma F, Muñoz J, Canals A, Astudillo C, Uribe M. Early laparoscopic cholecystectomy reduces hospital stay in mild gallstone pancreatitis. A randomized controlled trial. HPB (Oxford). 2020;22(1):26–33. Epub 2019 Jun 22. PMID: 31235428.
17. Loozen CS, Oor JE, van Ramshorst B, van Santvoort HC, Boerma D. Conservative treatment of acute cholecystitis: a systematic review and pooled analysis. Surg Endosc. 2017;31(2):504–15. Epub 2016 Jun 17. PMID: 27317033.
18. Bagepally BS, Haridoss M, Sasidharan A, Jagadeesh KV, Oswal NK. Systematic review and meta-analysis of gallstone disease treatment outcomes in early cholecystectomy versus Conservative management/delayed cholecystectomy. BMJ Open Gastroenterol. 2021;8(1):e000675. https://doi.org/10.1136/bmjgast-2021-000675. PMID: 34261757; PMCID: PMC8280848.
19. Gurusamy KS, Nagendran M, Davidson BR. Early versus delayed laparoscopic cholecystectomy for acute gallstone pancreatitis. Cochrane Database Syst Rev. 2013;(9):CD010326. https://doi.org/10.1002/14651858.CD010326.pub2. PMID: 23996398.
20. Yuan X, Xu B, Wong M, Chen Y, Tang Y, Deng L, Tang D. The safety, feasibility, and cost-effectiveness of early laparoscopic cholecystectomy for patients with mild acute biliary pancreatitis: A meta-analysis. Surgeon. 2021;19(5):287–96. Epub 2020 Jul 22. PMID: 32709425.
21. Sharma A, Vans E, Shigemizu D, Boroevich KA, Tsunoda T. DeepInsight: a methodology to transform a non-image data to an image for convolution neural network architecture. Sci Rep. 2019;9(1):11399. https://doi.org/10.1038/s41598-019-47765-6. PMID: 31388036; PMCID: PMC6684600.
22. Podda M, Pacella D, Pellino G, Coccolini F, Giordano A, Di Saverio S, Pata F, Ielpo B, Virdis F, Damaskos D, De Simone B, Agresta F, Sartelli M, Leppaniemi A, Riboni C, Agnoletti V, Mole D, Kluger Y, Catena F, Pisanu A, MANCTRA-1 Collaborative Group. CoMpliAnce with evideNce-based clinical guidelines in the management of acute biliary pancreAtitis): The MANCTRA-1 international audit. Pancreatology. 2022;22(7):902–16. Epub 2022 Jul 16. PMID: 35963665.
23. Di Martino M, Podda M, Raptis D, Machairas N, Ielpo B, Pata F, Pellino G, Di Saverio S, Catena F, Pisanu A, Donadon M, Pawlik TM, MANCTRA-1 Collaborative Group. The influence of socioeconomic inequity and guidelines compliance on clinical outcomes of patients with acute biliary pancreatitis. An international multicentric cohort study. HPB (Oxford). 2024 May 13:S1365–182X(24)01684-8. https://doi.org/10.1016/j.hpb.2024.05.001. Epub ahead of print. PMID: 38796347.
24. Zhou Y, Ge YT, Shi XL, Wu KY, Chen WW, Ding YB, Xiao WM, Wang D, Lu GT, Hu LH. Machine learning predictive models for acute pancreatitis: A systematic review. Int J Med Inf. 2022;157:104641. https://doi.org/10.1016/j.ijmedinf.2021.104641. Epub 2021 Nov 10. PMID: 34785488.
25. Mashayekhi R, Parekh VS, Faghih M, Singh VK, Jacobs MA, Zaheer A. Radiomic features of the pancreas on CT imaging accurately differentiate functional abdominal pain, recurrent acute pancreatitis, and chronic pancreatitis. Eur J Radiol. 2020;123:108778. https://doi.org/10.1016/j.ejrad.2019.108778. Epub 2019 Dec 11. PMID: 31846864; PMCID: PMC7968044.
26. Chen Y, Chen TW, Wu CQ, Lin Q, Hu R, Xie CL, Zuo HD, Wu JL, Mu QW, Fu QS, Yang GQ, Zhang XM. Radiomics model of contrast-enhanced computed tomography for predicting the recurrence of acute pancreatitis. Eur Radiol. 2019;29(8):4408–17. https://doi.org/10.1007/s00330-018-5824-1. Epub 2018 Nov 9. PMID: 30413966.
27. Hu X, Yang B, Li J, Bai X, Li S, Liu H, Zhang H, Zeng F. Individualized Prediction of Acute Pancreatitis Recurrence Using a Nomogram. Pancreas. 2021;50(6):873–878. https://doi.org/10.1097/MPA.0000000000001839. PMID: 34347724.
28. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, Hróbjartsson A, Mann H, Dickersin K, Berlin JA, Doré CJ, Parulekar WR, Summerskill WS, Groves T, Schulz KF, Sox HC, Rockhold FW, Rennie D, Moher D. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med. 2013;158(3):200–7. https://doi.org/10.7326/0003-4819-158-3-201302050-00583. PMID: 23295957; PMCID: PMC5114123.
29. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ, SPIRIT-AI, CONSORT-AI Working Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Lancet Digit Health. 2020;2(10):e549–60. https://doi.org/10.1016/S2589-7500(20)30219-3. Epub 2020 Sep 9. PMID: 33328049; PMCID: PMC8212701.
30. Alwosheel A, van Cranenburgh S, Chorus CG. Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis. J Choice Modelling. 2018;28:167–82.

Publisher's note