Artificial intelligence, clinical decision support and machine learning

P0155

DEVELOPMENT OF A MACHINE LEARNING BASED MORTALITY PREDICTION MODEL USING CLOT WAVEFORM ANALYSIS

V. Iyilikci [1], B. Isbilen Basok [3], A. Colak [4], A. Saritas [5], D.I. Topcu [2]

[1]*Department of Medical Biochemistry, Gaziantep City Hospital, Gaziantep, Turkey*

[2]*Department of Medical Biochemistry, Izmir City Hospital, Izmir, Turkey*

[3]*Department of Medical Biochemistry, UHS Izmir Faculty of Medicine, Dr. Behçet Uz Child Disease and Pediatric Surgery Training and Research Hospital, Izmir, Turkey*

[4]*Department of Medical Biochemistry, UHS Izmir Faculty of Medicine, Tepecik Training and Research Hospital, Izmir, Turkey*

[5]*Department of Reanimation and Anesthesiology, UHS Izmir Faculty of Medicine, Tepecik Training and Research Hospital, Izmir, Turkey*

BACKGROUND-AIM

Clot waveform analysis (CWA) monitors changes in the transparency of a plasma sample during clotting tests such as activated partial thromboplastin time and prothrombin time. CWA is an economically convenient and simple test that can aid clinical stratification and management by implementing an algorithm into the software of an automated coagulometer. Studies suggest that, in addition to abnormal waveforms, the peak times, and heights observed in the derivative curves of CWA are valuable for assessing hemostatic abnormalities in clinical conditions. In our study, we aimed to develop a machine learning (ML) mortality prediction model based on laboratory data plus CWA parameters.

METHODS

Data from 400 intensive care unit patients whose hospitalization exceeded 72 hours was used. Age, gender, and laboratory parameters at the 24th-72nd hour of ICU admission were used as input variables; and survival status according to the 7th-28th day endpoint was used as output variable. Data analysis was performed using Python within the Jupyter Notebook interface. Various ML models including random forest(RF) and Light GBM(LGBM) were developed by the k=10-fold cross-validation method. The performance metrics accuracy, precision, recall, and F1 scores were calculated. ROC analysis was conducted to determine the models' discrimination power. Parameter significance levels were calculated using Shapley additive explanations (SHAP) and feature-importance methods for model explainability analyses.

RESULTS

In the final performance evaluation, 0.87/0.88 accuracy, 0.89/0.88 AUC, 0.76/0.79 F1 score values were obtained for RF and LGBM at the 7th-day outcome, and 0.83/0.83 accuracy, 0.88/0.90 AUC, 0.83/0.83 F1 score values were obtained at the 28th-day outcome. For both outcomes, CWA parameters were ranked in the top 5 in the importance ranking of the models.

CONCLUSIONS

When conventional and ML-based mortality prediction studies in the literature are analyzed, our study is the first study using CWA parameters. While our RF and LGBM models successfully predicted mortality, CWA played an important role in their estimations. Our study reveals the potential of ML in clinical chemistry and that CWA will be included in clinical use in the future with its simple, cheap, and fast prognosis prediction.

Artificial intelligence, clinical decision support and machine learning

P0156

ARTIFICIAL INTELLIGENCE (AI) AND SAFETY IN THE WORKPLACE: POSSIBLE APPLICATIONS IN THE LABORATORY? PRELIMINARY EVALUATION OF THE ITALIAN SOCIETY OF CLINICAL PATHOLOGY AND LABORATORY MEDICINE (SIPMEL) HEALTH AND SAFETY STUDY GROUP

R. Lamanna [5], E. Rimini [7], L.F. Simula [2], S. Casani [8], A. Pinna [7], T. Bettini [1], M.R. Fanello [3], M. Pradella [2], M. Moffa [2], V. Izzo [6], N. Camusso [4], L. Rossi [5]
[1]*Brescia Hospital, Italy*
[2]*Italian Society of Clinical Pathology and Laboratory Medicine (SIPMEL), Castelfranco Veneto (TV), Italy*
[3]*Local Health Centre, Savona, Italy*
[4]*Oristano Hospital, Italy*
[5]*Pisa University Hospital, Italy*
[6]*San Giovanni di Dio e Ruggi D'Aragona Salerno Hospital, Italy*
[7]*Territorial Analysis Laboratory, Sassari, Italy*
[8]*Udine Hospital, Italy*

BACKGROUND-AIM

In recent years, artificial intelligence (AI) has made great strides in occupational safety. AI-based technologies are transforming injury prevention, offering precise and efficient tools to protect workers. In health and safety, it is crucial that the opportunities offered by the proper use of AI are not lost, but expanded on a large scale.

METHODS

Some areas of application of AI to Laboratory Medicine were analyzed: Risk Assessment; Prevention and Protection Measures; Predictive Analysis and Preventive Actions; Use of Personal Protective Equipment (PPE).

RESULTS

AI improves risk assessment, for instance by analyzing data collected from environmental sensors and surveillance cameras to identify situations that could lead to accidents, helping to develop and implement more effective preventive measures. It is used to customize healthcare training, adapting it to the specific needs and risks of the individual worker, improving its effectiveness by developing machine learning algorithms that analyze historical accident data, identifying trends and patterns; it is related to PPE, applied in different ways, from simply verifying the actual use of mandatory PPE and the absence of damage, in support of those responsible for controlling and supervising the implementation of safety.

CONCLUSIONS

AI represents a breakthrough in accident prevention. The ability to analyze amounts of data in real time, identify patterns and trends, and provide predictions and recommendations to prevent accidents makes it an invaluable resource for healthcare safety. AI solutions not only protect workers but also improve the efficiency and sustainability of organizations, reducing injuries and illnesses, and avoiding additional costs related to holidays, absences. The use of AI systems such as 'Copilot' help in the compilation of procedures to be shared with laboratory staff. The success of AI depends on effective integration with other safety practices and procedures; companies must maintain a holistic approach to safety, combining AI with appropriate training, careful supervision and a safety culture. IA offers enormous benefits in accident prevention, but its ethical and social implications must be considered to ensure that workers rights and dignity are respected through responsible and transparent use.

Artificial intelligence, clinical decision support and machine learning

P0157

EVALUATION OF LARGE LANGUAGE MODELS IN CLINICAL QUESTIONS FOR AUTOIMMUNE DISEASES: A COMPARATIVE ANALYSIS BASED ON CHATGPT 4O, CLAUDE 3.5 SONNET, AND GEMINI 1.5 PRO

M. Ning [3], Y. Chen [2], Y. Tao [1]

[1]*Department of Laboratory Medicine, Nanji Drum Tower Hospital Clinical College of Nanjing University of Chinese Medicine, Nanjing, Jiangsu, China*

[2]*Department of Laboratory Medicine, Nanjing Drum Tower Hospital Clinical College of Nanjing University of Chinese Medicine, Nanjing, Jiangsu, China*

[3]*Department of Laboratory Medicine, NanjingDrum Tower Hospital Clinical College of Nanjing University of Chinese Medicine, Nanjing, Jiangsu, China*

BACKGROUND-AIM

Large language models (LLMs) have established a presence in providing medical services to patients and supporting clinical practice for doctors.

METHODS

To explore the ability of LLMs in answering clinical questions related to autoimmune diseases, this study was designed with 65 questions related to autoimmune diseases, covering five domains: concepts, report interpretation, diagnosis, prevention and treatment, and prognosis. These questions were answered by three LLMs: ChatGPT 4o, Claude 3.5 Sonnet, and Gemini 1.5 Pro. The responses were then evaluated by 8 clinicians based on criteria including relevance, completeness, accuracy, safety, readability, and conciseness.

RESULTS

The results showed that the performance of the three LLMs in the evaluation of autoimmune diseases significantly surpassed that of both junior and senior doctors. Notably, Claude 3.5 Sonnet excelled in providing comprehensive and accurate responses to clinical questions on autoimmune diseases, demonstrating the great potential of LLMs in assisting doctors with the diagnosis, treatment, and management of autoimmune diseases.

CONCLUSIONS

LLMs are able to provide answers to AIDs-related questions with specificity and safety profiles. Comparative analysis reveals that the performance of the three large language models (LLMs) significantly outperforms both junior and senior doctors. Our findings highlight that Claude 3.5 Sonnet excels in delivering comprehensive, accurate, and well-structured responses to clinical questions related to autoimmune diseases. Its ability to interpret and analyze complex clinical issues in this field is particularly outstanding, even surpassing the expertise of both junior and senior doctors. This demonstrates that LLMs, especially Claude 3.5 Sonnet, have the potential to play a crucial role in assisting healthcare professionals with the diagnosis, treatment, and management of autoimmune diseases, providing valuable support in clinical practice.

Artificial intelligence, clinical decision support and machine learning

P0158

EVALUATION OF AN AUTOMATED ALGORITHM FOR THE DETECTION OF HYPERANDROGENISM IN THE LABORATORY

M.I. Pastor García [1], J.A. Delgado Rodríguez [1], M.I. Vázquez Tavero [1], S. Jimeno Pulido [1]
[1]*Hospital Universitario Son Espases*

BACKGROUND-AIM

Multitude of patient requests from primary care are received in the laboratory for the study of menstrual ataxia or infertility. Polycystic ovary syndrome (PCOS) is one of the most common cause. However, in many instances, the tests requested are insufficient or inadequate for diagnostic guidance.
Luteinizing hormone (LH) and follicle stimulating hormone (FSH) are the basic tests for the study of menstrual ataxia. Although the LH/FSH ratio has not been considered to have diagnostic value for PCOS, it can serve as a trigger for the adition of tests that are part of the differential diagnosis and have not been requested.
Our aim was to establish an automated algorithm for laboratory scale-up of tests for the detection of hyperandrogenic states and to evaluate its cost.

METHODS

Retrospective study between January-September 2024. We considered requests from Primary Care for the study of menstrual irregularity, infertility or clinical hyperandrogenism in women of childbearing age who had been asked for FSH and LH. The calculation of the LH/FSH ratio was automated and the determination of progesterone was added in those patients who had not started the menopausal transition (FSH <13mIUI/ML) and with an LH/FSH ratio>1.6, to verify collection in follicular phase. If the progesterone concentration<0.8ng/mL, testosterone determination was added, if it had not been requested. The LH/FSH ratio of 1.6 was taken from a previous study performed in our laboratory to detect PCOS. That point was selected as it reported a specificity of 80%.

RESULTS

7604 requests were considered. 1051 requests met the criteria of FSH<13 and Ratio >1.6. Progesterone was added to 68 patients (6.5%) and testosterone to 304 (28.9 %). This represents a monthly cost of approximately 68 euros for the laboratory, considering only the price of the reagent .The algorithm made possible the detection of 92 women with biochemical hyperandrogenism (30% of the additions).

CONCLUSIONS

The establishment of an automated algorithm for the detection of hyperandrogenism makes it possible to better identify the cause of the menstrual irregularity. This avoids a delay in diagnosis, future requests for laboratory tests or unnecessary referrals to specialists at a low cost for the laboratory

Artificial intelligence, clinical decision support and machine learning

P0159

AN INNOVATIVE EVIDENCE-BASED LABORATORY MEDICINE (EBLM) TEST TO HELP DOCTORS IN THE ASSESSMENT OF THE PARATHYROID FUNCTION AND BONE METABOLISM

A. Roca Vidalmargó [2], S.J. Calleja Freixes [2], J.D. Santotoribio Camacho [1]
[1]*Hospital Universitario Puerto Real*
[2]*Kience Inc., Wilmington, Delaware, US*

BACKGROUND-AIM

To develop a novel non-invasive, evidence-based laboratory medicine (EBLM) test to assist doctors in assessing the parathyroid function and bone metabolism and to evaluate its accuracy in detecting the main parathyroid diseases, such as hypoparathyroidism and hyperparathyroidism, and their types (primary or secondary), as well as bone metabolism diseases, such as osteopenia and osteoporosis.

METHODS

This study is part of a previous one already published at the European Society for Medical Oncology (ESMO) Congress 2024, which is focused on the accuracy evaluation of a novel non-invasive test for Multi-Cancer Early Detection (MCED). To develop the algorithm, several combinations of analytes were analyzed to identify the most significant groupings related to the parathyroid function and bone metabolism. The algorithm's efficiency was then enhanced using serial and parallel approximation techniques. Its performance was trained with a dataset of 4,746 patients. The validation of the algorithmic test was performed through a randomized controlled trial (RCT) with a sample size of 152 patients. Their blood samples were tested by Laboratorio Echevarne (Spain), using their biochemistry techniques.

RESULTS

For the RCT, the sensitivity achieved was 1.00 and the specificity was 1.00. Additionally, the area under the receiver operating characteristic (AUROC) curve, the positive predictive value (PPV), and the negative predictive value (NPV), were 1.00, 1.00, and 1.00, respectively. This indicates a strong correlation between the algorithm outcomes and the high likelihood of having parathyroid and bone metabolism diseases, as well as calcium and vitamin D metabolism disorders.

CONCLUSIONS

This innovative non-invasive blood-based biomarker algorithm holds promise in helping doctors in providing timely and accurate assessment of parathyroid and bone metabolism diseases —even in early stages—, as well as reduce medical errors or misdiagnoses. These results advocate further exploration, prompting our intention to conduct a new RCT involving 26,000 participants.

Artificial intelligence, clinical decision support and machine learning

P0160

CAUSAL ASSOCIATION BETWEEN 39 LABORATORY PARAMETERS AND TYPE 1 DIABETES: MACHINE LEARNING COMBINED WITH MENDELIAN RANDOMIZATION

L. Ze [1], X. Jiancheng [1]
[1]*The First Hospital of Jilin University*

BACKGROUND-AIM

The exact pathogenic mechanisms of type 1 diabetes are not yet fully understood. This study aims to identify Type 1 diabetes risk factors by combining machine learning and Mendelian randomization for more accurate analysis in Asian and European populations.

METHODS

210 Type 1 diabetes patients and 210 healthy controls from the First Hospital of Jilin University were selected. LassoCV with 5-fold cross-validation screened 39 laboratory parameters. Nine machine learning algorithms were compared, and the top 10 parameters were selected. Mendelian randomization analyzed the causal relationship between the selected parameters and Type 1 diabetes using SNPs from the UK Biobank.

RESULTS

XgBoost performed best (AUC: 1.000). The top 10 parameters were retinol binding protein, total iron binding capacity, total protein, albuminumin, unsaturated iron binding capacity, creatine kinase, hemoglobin, ferritin, indirect bilirubin, and aspartate aminotransferase. Mendelian randomization found a significant causal relationship between albumin and Type 1 diabetes in Europeans ($P < 0.05$). 186 SNPs significantly affected this relationship. The direction of causality from albumin to Type 1 diabetes was verified. Other parameters did not show a significant causal relationship.

CONCLUSIONS

The study identified a causal link between albumin and Type 1 diabetes in Europeans, suggesting that abnormal albumin levels can promote Type 1 diabetes development. This method offers a new approach for finding disease risk factors, improving research efficiency and benefiting clinical patients.

Artificial intelligence, clinical decision support and machine learning

P0161

LLM AND MACHINE LEARNING METHODS USAGE FOR STANDARDIZING CLINICAL LABORATORY PRODUCTION TESTS CATALOGS

A. Amosova [1], A. Chekunov [1], V. Aidarova [1], R. Kuvshinov [1], A. Andreeva [1], P. Gontard [1]
[1]*Gontard & Cie Group, 157 Kantonsstrasse, FREIENBACH 8807*

BACKGROUND-AIM

Standardization in laboratory medicine historically aims to laboratory tests as a part of medical information. However, for operative tasks it is necessary to consider a single test as a production unit. To address this task, we developed an approach for defining tests and created a standardized catalog based on software, Lab-Data. We define a production test as a combination of triplets: a particular analyte, the type of biological specimen utilized, and the analytical instrument employed. Manual standardization in a particular laboratory requires specialized staff and it is costly. This study explores using large language models (LLMs) and Machine Learning (ML) to efficiently align laboratory tests with Lab-Data catalog.

METHODS

The main challenge lies in the diverse local test designations (languages, codes, etc.) and the need to match triplets of test elements rather than single text strings. To address these challenges, LLMs were utilized for naming standardization, and the traditional semantic text-matching approach was adapted by developing a tree-based structure, performing candidate selection via k-nearest neighbours, and using tree distillation to refine shortlist candidates. Reranking was handled by the YetiRank algorithm, utilizing diverse numerical extracted and engineered features.

RESULTS

The automated approach achieved moderate average accuracy of ~80-85%, despite challenges posed by test designation clarity. The study highlighted the necessity of evaluating model quality beyond mere matching accuracy by analyzing errors through domain expertise. ~10-15% of fundamentally unmatchable items caused by inadequate and/or ambiguous test naming suggests that the solution's true matching potential exceeds what accuracy metrics alone reveal.

CONCLUSIONS

Laboratory community often lacks consensus on standardization of production elements, leading to unstructured data and complicating lab benchmarking. Standardized approach to defining tests as a production unit and naming conventions can address these issues. LLMs have shown potential in converting laboratory catalogs into standardized format of Lab-Data. By advocating for the standardization of production tests as a core component of laboratory practice, we contribute to the broader objective of enhancing the quality and coherence IVD services.

Artificial intelligence, clinical decision support and machine learning

P0162

THE EVOLUTION AND TRANSFORMATIVE IMPACT OF ARTIFICIAL INTELLIGENCE IN HEALTHCARE AND LABORATORY MEDICINE

K. Bhat [2], N. Wadhwa [2], P. Sharma [1]
[1]*Ex All India Institute of Medical Sciences, Jodhpur*
[2]*Himalayan Institute of Medical Sciences, Swami Ram University*

BACKGROUND-AIM

Artificial Intelligence (AI) has evolved from being a theoretical idea to becoming a transformative force in modern technology, especially in healthcare and laboratory medicine. Its foundation was laid by pioneers like Alan Turing, who introduced the concept of machine intelligence in his seminal work “Computer Machinery and Intelligence” in 1950, and John McCarthy, who coined the term “Artificial Intelligence” in 1955. Over the decades, AI has progressed through various phases, including periods of rapid innovation and challenges like the “AI Winter.” Despite these hurdles, it has emerged as an indispensable tool for improving diagnostics, optimizing laboratory workflows, and enhancing patient care.

METHODS

A comprehensive review of AI's evolution was conducted, highlighting key developments from foundational concepts by Alan Turing and John McCarthy to recent innovations in deep learning, natural language processing, and robotics. The study also examined AI applications in diagnostics, digital pathology, laboratory information systems (LIS), and picture archiving and communication systems (PACS).

RESULTS

AI advancements such as expert systems, wireless health technologies, and laboratory-on-a-chip (LOC) devices have improved diagnostic accuracy, optimized workflows, and enhanced research capabilities. Digital pathology and PACS have streamlined image acquisition, storage, and retrieval, leading to faster turnaround times and improved collaboration. LIS and PACS integration have further enhanced operational efficiency, supporting quality control, research, and education.

CONCLUSIONS

The integration of AI into healthcare and laboratory medicine has ushered in a new era of efficiency, accuracy, and innovation. By automating complex processes, enhancing diagnostic precision, and supporting groundbreaking research, AI has transformed the way healthcare is delivered and managed. As technologies continue to evolve, AI promises to unlock new possibilities, from advancing precision medicine to democratizing access to high-quality diagnostics globally. The future of AI in healthcare is not just about technology but about improving lives through smarter, faster, and more accessible medical care.

Artificial intelligence, clinical decision support and machine learning

P0163

AN INNOVATIVE EVIDENCE-BASED LABORATORY MEDICINE (EBLM) TEST TO HELP DOCTORS IN THE ASSESSMENT OF THE IRON TRANSPORT AND STORAGE FUNCTION

S.J. Calleja Freixes [2], A. Roca Vidalmargó [2], J.D. Santotoribio Camacho [1]
[1]*Hospital Universitario Puerto Real, Cádiz, Spain*
[2]*Kience Inc., Wilmington, Delaware, US*

BACKGROUND-AIM

To develop a novel non-invasive, evidence-based laboratory medicine (EBLM) test to assist doctors in assessing the iron transport and storage function and to evaluate its accuracy in detecting the main anemia types (normocytic, microcytic, macrocytic), their main causes (liver disease, vitamin B12 deficiency, bleeding or chronic kidney disease), and many other hematologic diseases.

METHODS

This study is part of a previous one already published at the European Society for Medical Oncology (ESMO) Congress 2024, which focused on the accuracy evaluation of a novel non-invasive test for Multi-Cancer Early Detection (MCED). To develop the algorithm, several combinations of analytes were analyzed to identify the most significant groupings related to iron transport and storage function. The algorithm's efficiency was enhanced using serial and parallel approximations. Its performance was trained with a dataset of 2,160 patients. The validation of the algorithmic test was performed through a randomized controlled trial (RCT) with a sample size of 152 patients. Blood samples were tested by Laboratorio Echevarne (Spain), using their hematology and biochemistry techniques.

RESULTS

For the RCT, the sensitivity achieved was 1.00 and the specificity was 1.00. The area under the receiver operating characteristic (AUROC) curve, the positive predictive value (PPV), and the negative predictive value (NPV), were 1.00, 1.00, and 1.00, respectively. This indicates a strong correlation between the algorithm outcomes and the high likelihood of having hematologic disease.

CONCLUSIONS

This innovative non-invasive blood-based biomarker algorithm holds promise in helping doctors in providing timely and accurate basic assessment of hematologic diseases —even in early stages, as well as reducing medical errors or misdiagnoses. These results advocate further exploration, prompting our intention to conduct a clinical study involving 26,000 participants to enhance our findings and inform clinical practice.

Artificial intelligence, clinical decision support and machine learning

P0164

AN INNOVATIVE EVIDENCE-BASED LABORATORY MEDICINE (EBLM) TEST TO HELP DOCTORS IN THE ASSESSMENT OF THE LIVER AND BILE-PANCREATIC EXOCRINE FUNCTION

S.J. Calleja Freixes [2], A. Roca Vidalmargó [2], J.D. Santotoribio Camacho [1]
[1]*Hospital Universitario Puerto Real, Cádiz, Spain*
[2]*Kience Inc., Wilmington, Delaware, US*

BACKGROUND-AIM

To develop a novel non-invasive, evidence-based laboratory medicine (EBLM) test to assist doctors in assessing the liver and bile-pancreatic exocrine function and to evaluate its accuracy in detecting a wide range of hepatocellular and cholestatic diseases.

METHODS

This study is part of a previous one already published at the European Society for Medical Oncology (ESMO) Congress 2024, which is focused on the accuracy evaluation of a novel non-invasive test for Multi-Cancer Early Detection (MCED). To develop the algorithm, several combinations of analytes were analyzed to identify the most significant groupings related to the liver and bile-pancreatic exocrine function. The algorithm's efficiency was then enhanced using serial and parallel approximation techniques. Its performance was trained with a dataset of 99,631 patients. The validation of the algorithmic test was performed through a randomized controlled trial (RCT) with a sample size of 152 patients. Their blood samples were tested by Laboratorio Echevarne (Spain), using their biochemistry and immunoassay techniques.

RESULTS

For the RCT, the sensitivity achieved was 1.00 and the specificity was 1.00. Additionally, the area under the receiver operating characteristic (AUROC) curve, the positive predictive value (PPV), and the negative predictive value (NPV), were 1.00, 1.00, and 1.00, respectively. This indicates a strong correlation between the algorithm outcomes and the high likelihood of having hepatocellular or cholestatic diseases, something very important with the global obesity epidemic, a primary risk factor for them.

CONCLUSIONS

This innovative non-invasive blood-based biomarker algorithm holds promise in helping doctors in providing timely and accurate assessment of hepatocellular and cholestatic diseases —even in early stages—, as well as reduce medical errors or misdiagnoses. These results advocate further exploration, prompting our intention to conduct a clinical study involving 26,000 participants to enhance our findings and inform clinical practice.

Artificial intelligence, clinical decision support and machine learning

P0165

IDENTIFICATION OF ROUTINE LABORATORY BLOOD TESTS USING MACHINE LEARNING FOR PREDICTING NON-ST-ELEVATION MYOCARDIAL INFARCTION

Y.M. Liou [3], W.J. Luo [1], C.H. Hsiao [1], C.S. Hung [4], C.H. Huang [2], K.Y. Su [1]

[1]*Department of Clinical Laboratory Sciences and Medical Biotechnology, College of Medicine, National Taiwan University, Taipei, Taiwan.*

[2]*Department of Emergency Medicine, National Taiwan University Medical College Hospital, Taipei, Taiwan.*

[3]*Department of Laboratory Medicine, National Taiwan University Hospital, Taipei, Taiwan.*

[4]*Division of Cardiology, Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan*

BACKGROUND-AIM

Routine laboratory tests provide essential insights into physiological conditions and underpin clinical decision-making. Several hematological and biochemical biomarkers have demonstrated associations with mortality in non-ST-elevation myocardial infarction (NSTEMI), offering diagnostic or prognostic value beyond cardiac troponin (cTn). Since no single test can reliably identify infarcted patients, leveraging a combination of laboratory indices via machine learning offers a promising alternative. This study aims to identify routine laboratory blood tests critical for predicting NSTEMI using machine learning models.

METHODS

Data were retrospectively collected from 44,000 patients who underwent at least one cTn measurement in the emergency department of National Taiwan University Hospital from 1 May, 2016 to 31 Dec, 2021. A total of 45 lab tests, including cTn, were evaluated upon admission. We tested different missing data thresholds or chi-square p-values above which tests were excluded. The XGBoost model was used to predict NSTEMI, with performance assessed using the area under receiver-operating-curves (AUROC) and average precision of precision-recall curves (APPRC) from five-fold cross-validation. Feature importance was ranked using SHAP (SHapley Additive exPlanations) values.

RESULTS

Following the exclusion of 157 patients diagnosed with STEMI, 1,256 without CBC data, 730 without available cTn within 12 hours before diagnosis, and 26,761 whose diagnosis did not rely on second cTn, 15,096 patients remained for analysis. This cohort included 690 NSTEMI cases and 14,406 controls. Among tested thresholds, a missing value threshold of 30% and a p-value threshold of 0.05 excluded rarely ordered tests without compromising model's performance. 23 lab tests were identified and incorporated with demographic features to construct the model showing an AUROC of 0.919 and an APPRC of 0.648. SHAP analysis highlighted the top 10 features: hs-cTnT, CK-MB, PT-INR, CK, age, PT, eosinophils, basophils, lymphocytes, and RDW-CV.

CONCLUSIONS

We identified 23 routinely ordered laboratory tests that, when integrated using machine learning, enhanced the prediction of NSTEMI. This approach offers a data-driven foundation for improving diagnostic accuracy in clinical settings.

Artificial intelligence, clinical decision support and machine learning

P0166

A COMPARATIVE ANALYSIS OF LARGE LANGUAGE MODELS ON CLINICAL QUESTIONS FOR AUTOIMMUNE DISEASES

M. Ning [1], Y. Tao [1], Y. Chen [1]
[1]*Nanjing Drum Tower Hospital*

BACKGROUND-AIM

The development of artificial intelligence (AI) has made great strides in recent years, especially after the launch of large language models (LLM). Our study evaluated the performance in delivering clinical questions related to autoimmune diseases (AIDs)

METHODS

46 questions on six aspects of AIDs were compiled and enter into ChatGPT 3.5, ChatGPT 4.0, and Gemini from April 1st, 2024 to May 1st, 2024. The replies of those three chatbots were collected and sent to three laboratory specialists for scoring according to relevance, correctness, completeness, helpfulness, and safety. Scores for three chatbots in five quality dimensions and the scores of the replies to the questions under each quality dimension were analyzed.

RESULTS

ChatGPT 4.0 showed superior performance than ChatGPT 3.5 and Gemini in all five quality dimensions. ChatGPT 4.0 scored higher than ChatGPT 3.5 on completeness, correctness, and safety in answering questions related to prevention and treatment. ChatGPT 4.0 scored higher than ChatGPT 3.5 on relevance in answering questions related to prognosis. ChatGPT 4.0 and Gemini scored higher than ChatGPT 3.5 on completeness in answering questions related to diagnosis. Concerning helpfulness, ChatGPT 4.0 scored higher than ChatGPT 3.5 and Gemini in answering questions related to clinical feathers and report interpretation. ChatGPT 4.0 scored higher than ChatGPT 3.5 in answering all areas of questions, except for diagnostic questions. Gemini scored higher than ChatGPT 3.5 on helpfulness in answering questions related to clinical features. Furthermore, The length of ChatGPT 4.0's replies was the longest, followed by ChatGPT 3.5, and the length of Gemini's replies was the shortest.

CONCLUSIONS

Our findings highlight the potential of LLMs, particularly ChatGPT 4.0 is superior to delivering comprehensive and accurate responses to AIDs-related clinical questions.

Artificial intelligence, clinical decision support and machine learning

P0167

"EVALUATION OF THE EFFICACY OF AN ARTIFICIAL INTELLIGENCE ALGORITHM FOR PREDICTING THE ESTIMATED VALUE OF PLATELET CONCENTRATE DURING APHERESIS"

J. Raraz-Vidal [1]
[1]*Villa el Salvador Emergency Hospital - Peru*

BACKGROUND-AIM

The aim of this study was to develop and evaluate the effectiveness of an AI algorithm for predicting the estimated value of platelet concentrate (PC) obtained by apheresis.

METHODS

This was an applied study using data from 30 patients, including age, BMI, gender, complete blood count (hematocrit, blood type, platelets), number of cycles, apheresis time, processed blood volume, and platelet concentrate value. Linear regression in Python was used for platelet concentrate prediction. The Mann-Whitney U test with a significance level of $p < 0.05$ was used to compare the mean difference between the PC obtained by the HAEMONETICS MCS®+ 9000 device and the estimated value by the algorithm.

RESULTS

The mean age was 36.43 ± 10.93 years, the average weight was 78.56 ± 10.54 kg, and the mean height was 1.69 ± 0.05 meters. From the hemogram, the mean hematocrit was 44.48 ± 3.78%, and the mean platelet count was 289,833.3 ± 62.2 platelets per microliter. During the apheresis procedure with the HAEMONETICS MCS®+ 9000 equipment, the average duration was 83.4 ± 7.20 minutes, the mean number of cycles was 7.8 ± 0.48, the average platelet concentrate of the equipment was 5.30 ± 1.01 × 10^11 platelets, and the obtained platelet volume was 394.2 ± 70.67. The AI algorithm identified weight, height, hematocrit, platelet count, and cycles as predictive variables for estimating the platelet concentrate in apheresis ($p < 0.05$). The analysis using the Mann-Whitney U test revealed no significant difference between the means of the platelet concentrate obtained by the HAEMONETICS MCS®+ 9000 equipment and the value estimated by the AI algorithm ($p > 0.05$).

CONCLUSIONS

The artificial intelligence algorithm used to estimate the platelet concentrate has shown similar results to the apheresis equipment. This indicates that the algorithm can be a useful tool in the selection of platelet donors.

Artificial intelligence, clinical decision support and machine learning

P0168

COMPARATIVE ANALYSIS OF THE PERFORMANCE OF AUTOMATED DIGITAL CELL MORPHOLOGY ANALYZERS FOR LEUKOCYTE DIFFERENTIATION IN HEMATOLOGIC MALIGNANCIES: MINDRAY MC-80 VERSUS WEST MEDICAL HEMA VISION

S. Sacchetti [1], M. Bellia [3], M. Vidali [2], L. Giacomini [1], V. Zanotti [1], A. Patriarca [3], U. Dianzani [1], G. Gaidano [3], R. Rolla [1]

[1]*Clinical Biochemistry Laboratory, "Maggiore della Carità" University Hospital, Department of Health Sciences, University of Eastern Piedmont, Novara, 28100, Italy*

[2]*Clinical Pathology Unit, Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico, 20122 Milan, Italy*

[3]*Division of Hematology, "Maggiore della Carità" University Hospital, Department of Translational Medicine, University of Eastern Piedmont, Novara, 28100, Italy*

BACKGROUND-AIM

The hematology laboratory has enhanced its diagnostic capabilities by using advanced artificial intelligence tools to analyze digital images of peripheral blood cells. The Mindray MC-80 (MC80) has performed excellently in various independent studies. This study aims to compare the leukocyte differential performance of the MC80 with that of HemaVision (HV) and the gold standard, manual microscopy.

METHODS

75 patients (M: F 53:47%; median (min-max) age 63 ys (1-90)), with hematological malignancies (ALL= 4, B-CLL=20, AML=20, CML=5, lymphoma= 20, infection=6) were analyzed. Their smears were compared using the MC-80, HV, and manual microscopy. According to REF, the agreement between microscopy (reference method, REF), HV, and, MC80, was expressed as the median (IQR) of a given cell population/feature, with REF-HV and REF-MC80 differences expressed as bias and 95% limits of agreement.

RESULTS

Concordance was calculated for all complete blood count parameters, but only the following are reported: Neu% [REF: 23.5% (6.5-36.7); REF-HV: 0.09 (-0.35 to 0.54); REF-MC80: 0.21 (-1.16 to 1.57)]; Ly% [REF: 45% (12.5-77.8); REF-HV: -2.56 (-6.72 to 1.60); REF-MC80: 23.03 (16.99 to 29.08)]; Mo% [REF: 2.00% (0.50-4.9); REF-HV: -2.15 (-3.57 to -0.73); REF-MC80: -1.47 (-2.42 to - 0.51)]; Eo% [REF: 1.0% (0.0-2.0); REF-HV: -0.44 (-0.77 to -0.11); REF-MC80: 0.08 (-0.25 to 0.40)]; Baso% [REF: 0.0% (0.0-0.5); REF-HV: -0.76 (-1.73 to 0.21); REF-MC80: -2.22 (-3.17 to -1.28)]; band cells [REF: 0.5% (0.0-1.5); REF-HV: -0.01 (-0.19 to 0.17); REF-MC80: -1.87 (-2.52 to -1.23)]; myelocytes [REF: 0.00% (0.00-0.5); REF-HV: 0.18 (-0.16 to 0.51); REF-MC80: -4.10 (-5.81 to -2.40)]; metamyelocytes [REF: 0.00% (0.00-0.4); REF-HV: 0.33 (0.04 to 0.63); REF-MC80: -0.56 (-0.98 to - 0.14)]; blasts, all samples [REF: 0.0% (0.0-34.6); REF-HV: 10.07 (5.17 to 14.97), REF-MC80: -2.05 (-7.06 to 2.96)]; blasts, in acute leukemia [REF: 61.2% (31.5-91.5, 2.0-98.0); REF-HV: 32.55 (21.55 to 43.56), REF-MC80: 17.70 (10.18 to 25.23)]; smudge cells in CLL [REF: 64.8% (42.9-100.3); REF-HV: 0.67 (-2.03 to 3.36), REF-MC80: -48.43 (-70.86 to -26.00)]

CONCLUSIONS

The study shows that MC80 has a higher sensitivity in identifying blasts than HV. However, HV shows better agreement with microscopy than MC80.

Artificial intelligence, clinical decision support and machine learning

P0169

IMPROVING PERSONALIZED ANTIBIOTIC STEWARDSHIP THROUGH MACHINE LEARNING-DRIVEN PROCALCITONIN PREDICTION USING LAB DATA

T. Sahin [2], E. Ispir [2], D.I. Topcu [2], S. Calik [1], I. Karakoyun [2]
[1] *Department of Infectious Disease, Izmir City Hospital, Izmir, Türkiye*
[2] *Core Laboratory, Department of Medical Biochemistry, Izmir City Hospital, Izmir, Türkiye*

BACKGROUND-AIM

Antimicrobial resistance is a major global health issue, worsened by the overuse of antibiotics. Procalcitonin (PCT), a biomarker for bacterial infections, helps guide antibiotic use, reduce unnecessary prescriptions, and improve decision-making. Clinical decision support (CDS) tools that incorporate predictive modeling for PCT binary classification can significantly enhance antibiotic stewardship efforts and prevent the overuse of antibiotics. This study aimed to develop machine learning (ML) models to predict PCT levels (≥0.5 ng/mL) using common laboratory data like complete blood count, C-reactive protein (CRP), and creatinine (Cr), thereby improving resource use and antibiotic management.

METHODS

This study used routine laboratory data, including CBC, CRP, and creatinine, from 7,047 patient results at İzmir City Hospital to develop and validate ML models predicting PCT positivity (≥0.5 ng/mL). Four ML models were tested: Random Forest, XGBoost, Support Vector Machine (SVM), and Logistic Regression. Performance metrics included accuracy, area under the curve (AUC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Feature importances and SHAP (SHapley Additive exPlanations) were analyzed and composite variables like P-LCR×CRP and CRP×Cr were created to boost accuracy. External validation on 1,265 independent patient results assessed model.

RESULTS

Random Forest and XGBoost outperformed other models, achieving 85% accuracy and AUCs of 0.93 and 0.92, respectively. Logistic Regression (72% accuracy, AUC: 0.78) and SVM (70% accuracy, AUC: 0.75) showed lower performance, particularly for PCT-positive cases. SHAP analysis identified composite features (e.g., P-LCR×CRP, CRP×Cr) as key predictors. External validation confirmed the robustness of Random Forest and XGBoost, both achieving 86% accuracy, ≥92% specificity, PPV of 0.80, and NPV of 0.88. XGBoost slightly outperformed Random Forest with a higher kappa value (0.66) and AUC (0.912).

CONCLUSIONS

XGBoost and Random Forest models, integrated into clinical workflows, can optimize PCT testing, reduce costs, and improve antibiotic stewardship, addressing the global challenge of antimicrobial resistance.

Artificial intelligence, clinical decision support and machine learning

P0170

MULTIPLEX PCR BASED DETECTION OF ANTIMICROBIAL RESISTANCE GENES AND STEWARDSHIP WITH COMPLEX ALGORITHMS

R. Shrestha [3], S. Beqaj [2], P. Neopane [1]
[1]*Department of Microbiology, Indiana University Health, Indianapolis, IN, USA*
[2]*Medical Database, Inc, Irvine, CA, USA*
[3]*Mountain View Medical Laboratory, Orange, CA, USA*

BACKGROUND-AIM

Standard diagnostic practices using routine microbiological procedures take 2–3 days for pathogen identification and antibiotic susceptibility testing leading to delayed diagnosis and an increase in healthcare expenditures. Molecular-based detection has the advantage of identifying pathogens directly from samples within hours. An increasing number of laboratories are using antimicrobial resistance (ABR) genes as an alternative to traditional antimicrobial susceptibility tests (AST). However, such approach reports ABR genes specific to drug classes rather than individual drugs making laboratory reports difficult to interpret by clinicians. Therefore, we aimed to develop software to analyze such complex molecular testing data into easy-to-read reports and suggest antimicrobial treatment utilizing a guidelines-based algorithm.

METHODS

Software named Laboratory Decision System Rx that can translate raw text files from PCR amplification data into laboratory reports with interpretation of results of pathogens and ABR gene detection was developed. Complex algorithm was built that rules out all antibiotics based on ABR gene results and scores the potential drugs based on guidelines. 20 urine samples from suspected UTI patients were analyzed with conventional microbiological procedures and multiplex PCR-based assay for pathogens and ABR genes.

RESULTS

Significant bacteriuria was identified in 14 of 20 samples by molecular method while only 9 samples were culture positive of which 3 were reported as mixed flora and unable to perform the conventional AST. Individual pathogens were identified in such samples with molecular methods and are also positive for multiple ABR genes including ESBL, Macrolide, Fosfomycin, Sulfonamide, and Tetracycline. Suggested drugs by LDSRx were also susceptible in AST in 3 of 6 samples.

CONCLUSIONS

LDSRx can be a potential platform that simplifies complex infectious disease molecular assay results into easy-to-read reports. The platform incorporates guideline-based algorithm to score potentially useful drugs based on the ABR gene results.

Artificial intelligence, clinical decision support and machine learning

P0171

AN INNOVATIVE EVIDENCE-BASED LABORATORY MEDICINE (EBLM) TEST TO HELP DOCTORS IN THE ASSESSMENT OF THE PANCREATIC ENDOCRINE FUNCTION

S.J. Calleja Freixes [2], A. Roca Vidalmargó [2], J.D. Santotoribio Camacho [1]
[1]*Hospital Universitario Puerto Real*
[2]*Kience Inc., Wilmington, Delaware, US*

BACKGROUND-AIM

To develop a novel non-invasive, evidence-based laboratory medicine (EBLM) test to assist doctors in assessing the pancreatic endocrine function and to evaluate its accuracy in detecting the main pancreatic endocrine diseases, such as insulin resistance (IR), prediabtes, and type 2 diabetes mellitus (DM2), as well as the 7.5-year risk of developing DM2.

METHODS

This study is part of a previous one already published at the European Society for Medical Oncology (ESMO) Congress 2024, which is focused on the accuracy evaluation of a novel non-invasive test for Multi-Cancer Early Detection (MCED). To develop the algorithm, several combinations of analytes were analyzed to identify the most significant groupings related to the pancreatic endocrine function. The algorithm's efficiency was then enhanced using serial and parallel approximation techniques. Its performance was trained with a dataset of 9,391 patients. The validation of the algorithmic test was performed through a randomized controlled trial (RCT) with a sample size of 152 patients. Their blood and urine samples were tested by Laboratorio Echevarne (Spain), using their biochemistry techniques.

RESULTS

For the RCT, the sensitivity achieved was 1.00 and the specificity was 1.00. Additionally, the area under the receiver operating characteristic (AUROC) curve, the positive predictive value (PPV), and the negative predictive value (NPV), were 1.00, 1.00, and 1.00, respectively. This indicates a strong correlation between the algorithm outcomes and the high likelihood of having pancreatic endocrine diseases, something very important with the global obesity epidemic, a primary risk factor for them.

CONCLUSIONS

This innovative non-invasive blood and urine-based biomarker algorithm holds promise in helping doctors in providing timely and accurate assessment of pancreatic endocrine diseases —even in early stages, as well as reduce medical errors or misdiagnoses. These results advocate further exploration, prompting our intention to conduct a new RCT involving 26,000 participants.

Artificial intelligence, clinical decision support and machine learning

P0172

AN INNOVATIVE EVIDENCE-BASED LABORATORY MEDICINE (EBLM) TEST TO HELP DOCTORS IN THE BASIC ASSESSMENT OF MAIN NON-MALIGNANT, HIGHLY PREVALENT, MORBID, CANCER-PRECURSOR, AND DEADLY DISEASES

S.J. Calleja Freixes [2], A. Roca Vidalmargó [2], J.D. Santotoribio Camacho [1]
[1]*Hospital Universitario Puerto Real, Cádiz, Spain*
[2]*Kience Inc., Wilmington, Delaware, US*

BACKGROUND-AIM

To develop a novel non-invasive, evidence-based laboratory medicine (EBLM) test to assist doctors in assessing the main body functions, systems and metabolisms, and to evaluate its accuracy in detecting the main non-malignant, highly prevalent, morbid, cancer-precursor, and deadly diseases.

METHODS

This study is part of a previous one already published at the European Society for Medical Oncology (ESMO) Congress 2024, which is focused on the accuracy evaluation of a novel non-invasive test for Multi-Cancer Early Detection (MCED). To develop the algorithm, several combinations of analytes were analyzed to identify the most significant groupings related to the main body functions, systems and metabolisms. The algorithm's efficiency was then enhanced using serial and parallel approximation techniques. Its performance was trained with a dataset of 185,882 patients. The validation of the algorithmic test was performed through a randomized controlled trial (RCT) with a sample size of 152 patients. Their blood samples were tested by Laboratorio Echevarne (Spain), using their hematology and biochemistry techniques.

RESULTS

For the RCT, the sensitivity achieved was 1.00 and the specificity was 1.00. Additionally, the area under the receiver operating characteristic (AUROC) curve, the positive predictive value (PPV), and the negative predictive value (NPV), were 1.00, 1.00, and 1.00, respectively. This indicates a strong correlation between the algorithm outcomes and the high likelihood of having a non-malignant, highly prevalent, morbid, cancer-precursor, and deadly disease.

CONCLUSIONS

This innovative non-invasive blood-based biomarker algorithm holds promise in helping doctors in providing timely and accurate basic assessment of the main non-malignant, highly prevalent, morbid, cancer-precursor, and deadly diseases —even in early stages—, as well as reduce medical errors or misdiagnoses. These results advocate further exploration, prompting our intention to conduct a clinical study involving 26,000 participants to enhance our findings and inform clinical practice.

Artificial intelligence, clinical decision support and machine learning

P0173

AI OUTPERFORMING HUMANS: ANALYSING CHATGPT'S RESULTS IN THE 2023 BIR EXAM

A. Calvera [1], G. Verdú [1], A.B. Fabregat-Bolufer [2]

[1]*Biochemistry and Molecular Genetics Department, Biomedical Diagnostic Center (CDB), Hospital Clínic of Barcelona, Barcelona, Spain.*

[2]*CORE Laboratory, Biochemistry and Molecular Genetics Department, Biomedical Diagnostic Center (CDB), Hospital Clínic of Barcelona, Barcelona, Spain.*

BACKGROUND-AIM

The rapid advancement of generative AI, particularly large language models like OpenAI's ChatGPT, is revolutionizing diverse fields, including healthcare. These constantly evolving models have emerged as powerful tools across scientific, educational, and professional domains. This study evaluates the performance of three advanced ChatGPT models (GPT-4o mini, GPT-4o, and GPT-o1) in the Spanish 2023 BIR exam—a specialized test in biology-related disciplines that serves as a gateway to health training programs in Spain—focusing on their accuracy, reproducibility, and implications for healthcare education.

METHODS

The 200 multiple-choice questions from the 2023 BIR exam were presented individually to GPT-4o mini, GPT-4o and GPT-o1 in the same format as the original exam, including question text and answer options. Accuracy rates, exam scores and hypothetical rank predictions compared to human candidates were calculated for each model. Reproducibility was tested by re-evaluating 35 selected questions across the three models. Statistical analyses, including consistency metrics and error correction rates, were performed using RStudio (Posit PBC).

RESULTS

GPT-4o mini achieved a 93% accuracy rate (186/200), ranking second among human candidates. GPT-4o excelled with a 97.5% accuracy (195/200), securing the top position, while GPT-o1 followed closely at 97% (194/200), also ranking first. Performance across disciplines was consistent, although minor weaknesses were identified in cellular biology and physiology. In terms of reproducibility, all models demonstrated high consistency (96.7%) for initially correct responses. GPT-o1 exhibited superior error correction ability, revising 80% of its incorrect answers, compared to 60% for GPT-4o and 20% for GPT-4o mini.

CONCLUSIONS

ChatGPT models showcased exceptional performance, surpassing most human candidates in the 2023 BIR exam. Their high accuracy, consistent reproducibility and ability to correct errors -notably in GPT-o1- underscore their potential as transformative tools in healthcare education. These results highlight the evolving capabilities of AI in specialized exams, paving the way for innovative applications in medical and scientific education.

Artificial intelligence, clinical decision support and machine learning

P0174

INTEGRATING AI IN BIOCHEMISTRY: OPPORTUNITIES AND CHALLENGES

G. Demirel [1]
[1]*Selcuk University*

BACKGROUND-AIM

Industry 4.0 has revolutionized biochemistry and biotechnology, utilizing vast datasets from metabolomics, proteomics, and pharmacogenomics. These innovations, including AI and machine learning, enhance precision in analyzing metabolic pathways and molecular interactions. However, challenges such as data quality, interpretation complexity, and ethical concerns persist. This study explores the capabilities and limitations of AI in solving biochemical problems, emphasizing the need for accurate data curation to optimize research and clinical outcomes.

METHODS

A systematic review of AI applications in biochemistry was conducted, analyzing peer-reviewed studies from the last five years. PubMed was used to identify research on AI advancements in metabolomics, enzyme prediction, pharmacogenomics, and drug discovery.

RESULTS

AI has advanced drug discovery by speeding up the identification of bioactive molecules and enhancing toxicity and biochemical interaction predictions. It integrates omics data from genomics, proteomics, and metabolomics, offering a deeper understanding of metabolic pathways and disease mechanisms. In pharmacogenomics, AI identifies genetic markers linked to enzymatic and metabolic profiles, enabling personalized treatments. However, challenges remain, such as poor-quality datasets affecting AI predictions in enzyme kinetics and metabolic flux analysis. Interpreting dynamic biochemical systems requires expert validation to ensure reliability. Ethical and regulatory issues, including data ownership, ethical use, and compliance with standards, also persist.

CONCLUSIONS

AI and ML are revolutionizing biochemistry by driving innovation in drug discovery, enzyme analysis, and metabolic pathway research. These technologies improve efficiency and accelerate timelines while uncovering new insights. However, their success depends on addressing challenges like data accuracy, fostering interdisciplinary collaboration, and adhering to ethical and regulatory standards. Overcoming these obstacles will ensure AI continues to advance biochemistry with precision, reliability, and fairness.

Artificial intelligence, clinical decision support and machine learning

P0175

INNOVATING AN INTELLIGENT COGNITIVE RATIO SYSTEM FOR PROACTIVE EARLY DETECTION OF PATHOGENS

M. Halak [1]
[1]*My Reseaerch*

BACKGROUND-AIM

-Preventing cases of medical misdiagnosis that pose a significant risk and harm to human health-Reducing the spread of diseases and the incidence of them through early detection of their causes, treatment, and elimination with solutions based on analysis with the highest accuracy, comprehensiveness, and depth, and with integration and coordination of all relevant specialties. -Saving billions of dollars annually on medical operations, medications, treatment sessions, and first aid cases

METHODS

smart path that relies on the Internet of Things, artificial intelligence techniques, deep learning algorithms, big data science and analysis, and cloud computing for a comprehensive base for collecting and analyzing data for all tables of measurements of the proportions of important rare elements in the body, heavy metals harmful to it, vitamins in the body, body mass, and fat volume. Stored in the body to be processed, classified, and sorted into all fields using artificial intelligence algorithms, which are obtained from the outputs of smart measuring devices that operate with high-scanning technology. Frequency, as well as data from analysis centers' measurement tables that measure the levels of harmful minerals, important rare elements, vitamins, and blood, urine, stool, kidney, and heart analyses and Endocrine and liver tests and ...With genetic and immunological medicine analysis tables and microbiome science to build a huge database that increases

RESULTS

The ratio system for early detection of pathogens and providing innovative solutions to get rid of them has the advantage of:Reviewing the medical history of the patient who lost consciousness and the option to show all his medical tests through his fingerprint .This feature gives paramedics and the emergency department ,a great ability to take the correct measures to rescue unconscious patients While the full details of their health condition appear By reading their fingerprint

CONCLUSIONS

The first path will be awareness-raising and preventive guidance, in person and virtually, And the second path:A website (platform) and two applications for Android and iPhone On an interactive basis,

Artificial intelligence, clinical decision support and machine learning

P0176

EVALUATION OF ARTIFICIAL INTELLIGENCE PREDICTION OF ACUTE LEUCEMIA (AI-PAL) MODEL THROUGH CASES FROM PRACTICE

N. Klapuh-Bukvić [1], Z. Kurtanović [3], D. Šeper [4], A. Herić [2], M. Skurić-Tomić [2]

[1]*1. Clinical Center University of Sarajevo, Bolnicka 25, Sarajevo, B&H 2. University of Sarajevo, Faculty of Health Studies, Stjepan Tomic 1, B&H*

[2]*1.Clinical Center University of Sarajevo, Bolnicka 25, Sarajevo, B&H*

[3]*3.University of Sarajevo, Faculty of Pharmacy, Zmaja od Bosne 8, Sarajevo, B&H 4.Public Institution Health Center of Sarajevo Canton, Vrazova 11, Sarajevo, B&H*

[4]*3.University of Sarajevo, Faculty of Pharmacy, Zmaja od Bosne 8, Sarajevo, B&H 5.Public Institution General Hospital "Prim.dr. Abdulah Nakas" Sarajevo, Silvija Strahimira Kranjcevica 12, B&H*

BACKGROUND-AIM

Acute leukemias are hematological cancers that represent life-threatening conditions. They are are characterized by infiltration of transformed immature hematopoietic cells into the blood and bone marrow, rapid disease progression and high mortality rate. Establishing a diagnosis is a complex multiphase process that begins with a clinical examination and anamnesis, followed by routine biochemical-hematological analyses, cytological-histological diagnostics, immunophenotyping, cytogenetic analyzes and molecular diagnostics. These stages require highly qualified staff, sophisticated equipment and time. The speed of establishing an accurate diagnosis in acute leukemias is one of the key factors affecting the success of treatment. For this reason, a model (AL-PAL) based on artificial intelligence algorithms was developed in France, which uses the results of 10 routine parameters determined during patient admission which can be used to predict the type of acute leukemia in suspected cases. Objective is to evaluate the capabilities of the AI-PAL model by applying it to practical cases.

METHODS

The study analized 23 patients with a definitive diagnosis of acute myeloid leukemia (AML) or acute lymphocytic leukemia (ALL). All patients are over 18 years old, treated at the Hematology Clinic of the Clinical Center of the University of Sarajevo. The results of routine analyzes performed during the admission of patients to the clinic were used to evaluate the AL-PAI model.

RESULTS

23 patients were analyzed, aged 20 to 83 years, with a gender structure: 39% women and 61% men. In 21 patients, the final diagnosis established by standard procedures confirmed the presence of AML, while in 2 patients the diagnosis was ALL. Using the results of routine analyzes performed during the admission of patients and the application of the AI-PAL model, out of 23 tested patients, the same diagnosis was obtained in 22 cases, demonstrating the high prognostic accuracy of this model.

CONCLUSIONS

The analysis shows highly positive results regarding the prognostic accuracy of this model, supporting its primary purpose. However, a more detailed evaluation should be done on a larger number of samples.

Artificial intelligence, clinical decision support and machine learning

P0177

ELECTRONIC CLINICAL DECISION SUPPORT TOOL FOR CRITICAL HYPERBILIRUBINEMIA IN PRETERM NEONATES

Y. Xiao [2], M. Palmucci [1], L. Carlin [1], C. Lee [2], M. O'Gorman [2], S. Nair [3], L. Yieh [3], E. Leung [2]
[1]*Clinical Information Services, Children's Hospital Los Angeles, Los Angeles, CA, USA*
[2]*Department of Pathology and Laboratory Medicine, Children's Hospital Los Angeles, Los Angeles, CA, USA*
[3]*Department of Pediatrics, Division of Neonatology, Fetal and Neonatal Institute, Children's Hospital Los Angeles, Los Angeles, CA, USA*

BACKGROUND-AIM

Critical hyperbilirubinemia in preterm neonates is treated with phototherapy or exchange transfusion when bilirubin results exceed gestational age and age-specific medical decision levels (MDLs) to prevent bilirubin-induced neurological damage. The conventional evaluation at Children's Hospital Los Angeles involves multiple manual steps and is poised to inconsistencies and delays in treatment.

METHODS

We designed and implemented an electronic clinical decision support (CDS) tool in our electronic medical record system to identify and alert Neonatal Intensive Care Unit (NICU) clinicians of critical hyperbilirubinemia with a SmartZone alert. We evaluated the performance of our manual evaluation workflow, the accuracy of the electronic CDS tool, and the outcome of the electronic CDS tool to reduce the time to place orders for interventions.

RESULTS

Among the 22 patients who met the criteria to have phototherapy ordered before implementing the electronic CDS tool, 20 (90%) had phototherapy ordered. 14 (70%) phototherapy orders were placed <24 hours, four were placed 24 - 72 hours, and two were placed >72 hours after bilirubin results exceeded the corresponding MDLs. Among the 15 patients who met the criteria to have phototherapy ordered after implementing the electronic CDS tool, all (100%) received phototherapy orders, with 14 (93%) placed <24 hours and one order placed <48 hours. The electronic CDS tool identified all eligible patients correctly. The proportion of phototherapy ordered <24 hours increased from 70% to 93% after implementing the electronic CDS tool.

CONCLUSIONS

The electronic CDS tool promoted more appropriate and timely intervention orders to manage critical hyperbilirubinemia in preterm neonates.

Artificial intelligence, clinical decision support and machine learning

P0178

UNIVERSAL NOMOGRAM FOR PREDICTING REFERABLE DIABETIC RETINOPATHY: A VALIDATED MODEL FOR COMMUNITY AND OPHTHALMIC OUTPATIENT POPULATIONS USING EASILY ACCESSIBLE INDICATORS

D. Niu [1], H. Liu [1], T. Lei [1], Y. Zhang [2]

[1]*Clinical Laboratory Center, Xi'an People's Hospital (Xi'an Fourth Hospital), Affiliated People's Hospital of Northwest University, Xi'an, Shaanxi, China*

[2]*Shaanxi Eye Hospital, Xi'an People's Hospital (Xi'an Fourth Hospital), Affiliated People's Hospital of Northwest University, Xi'an, Shaanxi, China*

BACKGROUND-AIM

This study aimed to develop and validate a universal nomogram for predicting referable diabetic retinopathy (RDR) in type 2 diabetes mellitus (T2DM) patients, using easily accessible clinical indicators for both community and ophthalmic outpatient populations.

METHODS

A cross-sectional study was conducted with 1,830 T2DM patients from 14 communities in Xi'an, Shaanxi, China. Participants completed questionnaires, underwent physical exams, and ophthalmic assessments. Univariate analysis and least absolute shrinkage and selection operator (LASSO) regression identified key predictors for RDR. A nomogram was developed using multivariable logistic regression. Model performance was evaluated through area under the curve (AUC), accuracy, precision, recall, F1 score, Youden index, calibration curves, and decision curve analysis (DCA). The dataset was split into training (80%) and test (20%) sets, with external validation using 123 T2DM outpatients from Shaanxi Eye Hospital.

RESULTS

Seven key predictors were identified: serum creatinine, urea nitrogen, urine glucose, HbA1c, urinary microalbumin, diabetes duration, and systolic blood pressure. The nomogram exhibited moderate predictive accuracy, with AUCs of 0.730 (95% CI: 0.691–0.759), 0.767 (95% CI: 0.704–0.831), and 0.723 (95% CI: 0.610–0.835) for the training, test, and external validation sets, respectively. DCA indicated that the threshold probability for identifying RDR in diabetic patients ranged from 8% to 72%, substantiating its clinical applicability.

CONCLUSIONS

This nomogram, based on readily available clinical indicators, provides a reliable and scalable tool for predicting RDR risk in both community and ophthalmic settings. It offers a practical solution for early detection and personalized management of RDR, with broad applicability and clinical potential.

Artificial intelligence, clinical decision support and machine learning

P0179

PATHOLOGICAL CLASSIFICATION OF PRIMARY NEPHROTIC SYNDROME BASED ON ARTIFICIAL INTELLIGENCE OF CLINICAL LABORATORY DATA

L. Pang [1], H. Li [1]
[1]*Department of Clinical Laboratory, Peking University First Hospital*

BACKGROUND-AIM

Primary nephrotic syndrome (PNS) is one of the most common diseases of kidney. Its etiology is complex, pathological types are varied, clinical diagnosis and treatment is difficult. At present, renal biopsy, as the gold standard of renal pathological diagnosis, is a traumatic and invasive examination with many contraindications and complications. This study aimed at developing a machine learning model to predict pathological types of PNS based on clinical laboratory data.

METHODS

A total of 564 patients with PNS confirmed by renal biopsy at Peking University First Hospital from 2015 to 2019 were used for modeling. Clinical laboratory data from patients were randomly split into a training set (70%) and an internal validation set (30%) for model building. In addition, the data of PNS patients from 2020 to 2021 were used as an external validation set to further evaluation. Five machine learning models were used to build the prediction model, and the performance was evaluated by receiver operating characteristic (ROC) curve.

RESULTS

In scheme one, the XGBoost model has the highest predictive value, with an average AUC of 0.8090 in the internal validation set and 0.8003 in the external validation. It has high predictive value for MN and MPGN, but low predictive value for MsPGN. Among the 12 feature variables, ALB has the largest weight in the XGBoost model. In scheme two, the Gradient Boosting model has the highest predictive value, with an average AUC of 0.8550 in the internal validation set and 0.8410 in the external validation. It has the highest predictive value for MN and a low predictive value for MsPGN. Among the 13 feature variables, anti-PLA2R has the largest weight.

CONCLUSIONS

Among the five machine learning models, XGBoost and Gradient Boosting algorithms show the best prediction performance and can be used for fast prediction of pathological types of PNS.

Artificial intelligence, clinical decision support and machine learning

P0180

CAPABILITIES AND LIMITATIONS OF NINE LARGE LANGUAGE MODELS IN SOLVING SPECIALIZED CLINICAL QUIZZES

H. Park [1], W.Y. Heo [1]
[1]*Department of Laboratory Medicine and Genetics, Samsung Medical Center, Seoul, Korea*

BACKGROUND-AIM

Large language models (LLMs) have shown potential in specialized fields such as medicine, but their performance in niche domains like clinical chemistry and laboratory management is not fully understood. This study evaluated the capabilities of nine LLMs in solving clinical quizzes from the Laboratory Medicine Online journal (eISSN 2093-6338). The goal was to assess their ability to handle domain-specific tasks without prior fine-tuning.

METHODS

A dataset of 109 quizzes, including single-select, open-ended, and multi-select questions, was used. The models included GPT-4o (OpenAI, San Francisco, CA, USA), Claude 3 Opus (Anthropic, San Francisco, CA, USA), and Gemini 1.5 Pro (Google, Mountain View, CA, USA), along with their earlier or smaller versions. The quizzes were translated from Korean to English, and zero-shot prompting was applied, assigning roles like "clinical chemist" to the models. Statistical scoring was performed by experts, and translation bias was analyzed. Contextual prompting with original article data was also tested to improve accuracy for complex questions.

RESULTS

The GPT-4o model achieved the highest accuracy (81.7%), followed by GPT-4 Turbo (76.1%), Claude 3 Opus (74.3%), and Gemini 1.5 Pro (69.7%). Single-choice and table-based questions yielded the best results, with GPT-4o correctly answering 67 out of 84 single-select questions. However, performance declined for figure-based and calculation-heavy quizzes, where no model consistently excelled. Contextual prompting improved accuracy significantly on difficult questions, with GPT-4o correctly answering 7 out of 8 previously unanswered questions, compared to 4 for Claude 3 Opus and 5 for Gemini 1.5 Pro. Translation bias analysis revealed minimal impact, with consistent results across Korean and English datasets. However, eight quizzes remained unsolved by all models, highlighting gaps in domain-specific knowledge.

CONCLUSIONS

This study demonstrates that LLMs like GPT-4o are capable of addressing many domain-specific tasks in laboratory medicine, especially when supported with contextual information. Despite promising results, performance on niche topics and ethical considerations like reliability and bias require further exploration.

Artificial intelligence, clinical decision support and machine learning

P0181

AN ELECTRONIC ALERT SYSTEM FOR IDENTIFICATION OF THROMBOTIC MICROANGIOPATHIES

K. Reimand [1], K. Vaagen [1], M. Rosenberg [1]
[1]*Tartu University Hospital*

BACKGROUND-AIM

Though rare, thrombotic microangiopathies (TMAs) are medical emergent diseases that carry very high mortality if not recognized and treated promptly. Only a few cases of various TMAs (atypical hemolytic uremic syndrome (aHUS); thrombotic thrombocytopenic purpura (TTP); postinfectious, mainly Shiga-toxin associated hemolytic uremic syndrome (HUS)) have been diagnosed in Tartu University Hospital (TUH) in recent years. Considering disease incidence and size of the hospital catchment area, as well as the hospital profile, there may be some undetected cases of TMA. The aim of our project was to create and validate an electronic alert (e-alert) system for rapid and accurate identification of TMAs in TUH.

METHODS

An e-alert system was created and incorporated into our Laboratory Information System (LIS) in February 2024. A computerized algorithm uses changes in serum creatinine (>130 µmol/L), hemoglobin concentration (<90 g/L) and platelet count (<90 x 109/L). Alert is triggered when all abovementioned changes occur simultaneously within 24 hours. A corresponding notification about the possibility of TMA is displayed in LIS/HIS. During the validation period notifications are visible only to laboratory personnel. All generated e-alerts from February 9th to September 11th were reviewed by two laboratory physicians.

RESULTS

During validation period 506 alerts for 145 patients (56% male, 44% female, mean age 67.7 years) were generated. The most common pathologies were infections (often with underlying hematological disease), various oncological diseases, liver cirrhosis, patients on extracorporeal membrane oxygenation treatment. aHUS was confirmed in one patient, a three-year-old girl (received Eculizumab treatment, very good response). TTP was confirmed in one 59-year-old woman (received plasma exchange therapy and Rituximab). On retrospective data analysis we have found some (~5) suspicious cases for TMA (e-alert, elevated LDH levels), but data necessary for differential diagnosis (ADAMTS-13 activity, SC5b-9 level) were not available.

CONCLUSIONS

An e-alert system is a valuable tool for improving the diagnosis of TMAs in TUH and partner hospitals that share the same LIS.

Artificial intelligence, clinical decision support and machine learning

P0182

AN INNOVATIVE EVIDENCE-BASED LABORATORY MEDICINE (EBLM) TEST TO HELP DOCTORS IN THE ASSESSMENT OF THE HYDROELECTROLYTIC METABOLISM

A. Roca Vidalmargó [2], S.J. Calleja Freixes [2], J.D. Santotoribio Camacho [1]
[1]*Hospital Universitario Puerto Real, Cádiz, Spain*
[2]*Kience Inc., Wilmington, Delaware, US*

BACKGROUND-AIM

To develop a novel non-invasive, evidence-based laboratory medicine (EBLM) test to assist doctors in assessing the hydroelectrolytic metabolism and to evaluate its accuracy in detecting the main electrolytes imbalance-related diseases, such as hyponatremia, hyperkalemia, or hypomagnesemia, among others.

METHODS

This study is part of a previous one already published at the European Society for Medical Oncology (ESMO) Congress 2024, which is focused on the accuracy evaluation of a novel non-invasive test for Multi-Cancer Early Detection (MCED). To develop the algorithm, several combinations of analytes were analyzed to identify the most significant groupings related to the hydroelectrolytic metabolism. The algorithm's efficiency was then enhanced using serial and parallel approximation techniques. Its performance was trained with a dataset of 2,626 patients. The validation of the algorithmic test was performed through a randomized controlled trial (RCT) with a sample size of 152 patients. Their blood samples were tested by Laboratorio Echevarne (Spain), using their biochemistry techniques.

RESULTS

For the RCT, the sensitivity achieved was 1.00 and the specificity was 1.00. Additionally, the area under the receiver operating characteristic (AUROC) curve, the positive predictive value (PPV), and the negative predictive value (NPV), were 1.00, 1.00, and 1.00, respectively. This indicates a strong correlation between the algorithm outcomes and the high likelihood of having an electrolytes imbalance-related disease.

CONCLUSIONS

This innovative non-invasive blood-based biomarker algorithm holds promise in helping doctors in providing timely and accurate assessment of electrolytes imbalance-related diseases —even in early stages—, as well as reduce medical errors or misdiagnoses. These results advocate further exploration, prompting our intention to conduct a clinical study involving 26,000 participants to enhance our findings and inform clinical practice.

Artificial intelligence, clinical decision support and machine learning

P0183

AN INNOVATIVE EVIDENCE-BASED LABORATORY MEDICINE (EBLM) TEST TO HELP DOCTORS IN THE BASIC ASSESSMENT OF THE THYROID FUNCTION

A. Roca Vidalmargó [2], S.J. Calleja Freixes [2], J.D. Santotoribio Camacho [1]
[1]*Hospital Universitario Puerto Real*
[2]*Kience Inc., Wilmington, Delaware, US*

BACKGROUND-AIM

To develop a novel non-invasive, evidence-based laboratory medicine (EBLM) test to assist doctors in assessing the thyroid function and to evaluate its accuracy in detecting the main thyroid diseases, such as hypothyroidism and hyperthyroidism, as well as their origin (primary or secondary).

METHODS

This study is part of a previous one already published at the European Society for Medical Oncology (ESMO) Congress 2024, which is focused on the accuracy evaluation of a novel non-invasive test for Multi-Cancer Early Detection (MCED). To develop the algorithm, several combinations of analytes were analyzed to identify the most significant groupings related to the thyroid function. The algorithm's efficiency was then enhanced using serial and parallel approximation techniques. Its performance was trained with a dataset of 6,516 patients. The validation of the algorithmic test was performed through a randomized controlled trial (RCT) with a sample size of 152 patients. Their blood samples were tested by Laboratorio Echevarne (Spain), using their biochemistry techniques.

RESULTS

For the RCT, the sensitivity achieved was 1.00 and the specificity was 1.00. Additionally, the area under the receiver operating characteristic (AUROC) curve, the positive predictive value (PPV), and the negative predictive value (NPV), were 1.00, 1.00, and 1.00, respectively. This indicates a strong correlation between the algorithm outcomes and the high likelihood of having thyroid diseases.

CONCLUSIONS

This innovative non-invasive blood-based biomarker algorithm holds promise in helping doctors in providing timely and accurate basic assessment of thyroid diseases —even in early stages, as well as reduce medical errors or misdiagnoses. These results advocate further exploration, prompting our intention to conduct a new RCT involving 26,000 participants.

Artificial intelligence, clinical decision support and machine learning

P0184

¿CAN ARTIFICIAL INTELLIGENCE REPLACE THE PERFORMANCE OF CERTAIN CLINICAL LABORATORY TESTS? ANALYZING A LATENT RISK

J.M. Sanchez Gonzalez [6], M.C. Moran Moguel [3], A.E. Rivera Cisneros [2], J.H. Portillo Gallo [5], M. Lopez Rodriguez [4], L. Figueroa Montes [1]

[1]*Clinical Pathology Service, Hospital III Suárez Angamos, EsSalud, Former President of the Latin American Association of Clinical Pathology and Laboratory Medicine (ALAPAC/ML), Lima, Peru.*

[2]*Dean, University of Football and Sports Sciences, Pachuca, Hidalgo, Mexico.*

[3]*Department of Philosophical, Methodological, and Instrumental Disciplines. University Center for Health Sciences. University of Guadalajara. Guadalajara, Jalisco, Mexico.*

[4]*President, Latin American Association of Clinical Pathology and Laboratory Medicine, A.C., Guadalajara, Jalisco, Mexico.*

[5]*Protreasurer, Latin American Association of Clinical Pathology and Laboratory Medicine, A.C., Chihuahua, Chihuahua, Mexico.*

[6]*Vicepresident, Mexican Federation of Clinical Pathology, A.C. (FEMPAC), Treasurer, Latin American Association of Clinical Pathology and Laboratory Medicine (ALAPAC/ML), Director, National Institute of Learning, Skills and Research in Sciences, S.C (INAHIC*

BACKGROUND-AIM

The analysis of comprehensive clinical parameters powered by Artificial Intelligence (AI) is driving a transformation in Clinical Laboratory (CL) testing, enhancing the prediction, prevention, diagnosis, and prognosis of errors and even diseases. Experts attribute AI the ability to efficiently analyze and process large datasets, facilitating the development of new diagnostic and predictive models.

METHODS

We posed the specific question, “Will AI replace the performance of certain tests in the CL?” as part of a broader survey on AI among CL professionals in Mexico. The question was framed using a Likert scale: 1) Strongly agree; 2) Agree; 3) Neither agree nor disagree; 4) Disagree; and 5) Strongly disagree. The responses were analyzed using descriptive statistical methods and cross-referenced with variables such as gender, age, educational level, and CL size.

RESULTS

214 professionals participated (128 women and 86 men) of which 95 work in public CLs and 119 in private ones. Participants’ ages ranged from 20-81 years. 17 were technicians, 93 held bachelor's degrees, 33 had specializations, and 71 had postgraduate degrees. 30.84% agreed or strongly agreed that AI could replace the performance of certain CL tests, while 47.66% disagreed or strongly disagreed. Meanwhile, 21.5% neither agreed nor disagreed. Among those who agreed, 39.4% worked in micro CLs (>10 employees); 34.84% in small CLs (10–50); 16.66% in medium CLs (50–100); and 9.09% in large CLs (>100). Conversely, among those who disagreed, the percentages were 30.39%, 42.16%, 19.6%, and 7.84%, respectively, for micro, small, medium, and large CLs

CONCLUSIONS

Although AI shows enormous potential to transform clinical laboratories, most respondents believe it will not fully replace the performance of laboratory tests but rather complement human expertise. This balance between technology and human expertise defines the collaborative future of AI in CL. The results revealed a divided perception in Mexico highlighting a preference for using AI as a complementary tool rather than a substitute. This acceptance of AI appears to be influenced by the size of the laboratory, with greater acceptance observed in smaller labs where AI could address operational limitations. Importantly, its use must be framed by regulations, responsibility, and ethics.

Artificial intelligence, clinical decision support and machine learning

P0185

LEVERAGING ARTIFICIAL INTELLIGENCE TO UNCOVER THE GENETIC LANDSCAPE OF CARDIOMYOPATHY IN OMANI PATIENTS

A. Al Amri [1], A. Al Balushi [1], N. Al Hashmi [1], M. Al Araimi [1], M. Al Rawahi [2], T. Al Lawatia [2], N. Al Mahrami [3], M. Al Balushi [1], B. Al Shamsi [1]
[1]*National Genetics Centre of the Royal Hospital*
[2]*National Heart Centre of the Royal Hospital*
[3]*Oman College of Health Sciences*

BACKGROUND-AIM

Cardiomyopathy, a heterogeneous group of myocardial disorders, remains a significant health burden worldwide, with genetic factors playing a pivotal role in its pathogenesis. In Oman, the genetic underpinnings of cardiomyopathy are not well-characterized. This study aims to elucidate the genetic causes of cardiomyopathy in the Omani population using Whole Exome Sequencing (WES) and advanced Artificial Intelligence (AI) methodologies.

METHODS

A cohort of 100 Omani patients diagnosed with various forms of cardiomyopathy was recruited. Genomic DNA was extracted, and WES was performed to identify pathogenic and likely pathogenic variants associated with the disease. Comprehensive bioinformatic analysis was conducted to interpret the identified variants. Additionally, an AI model was developed to prioritize variants and predict their pathogenicity, integrating clinical phenotypes and genetic data for enhanced diagnostic accuracy.

RESULTS

The study successfully identified several novel and previously reported variants linked to cardiomyopathy. The AI model demonstrated high predictive accuracy, significantly reducing the time required for variant interpretation and enabling the identification of key genetic contributors with clinical relevance. This approach not only deepens our understanding of the genetic basis of cardiomyopathy in Oman but also provides a framework for implementing AI in genomic medicine, enhancing diagnostic precision and personalized care.

CONCLUSIONS

Our findings emphasize the importance of integrating next-generation sequencing and AI in the investigation of complex genetic disorders. This study sets a precedent for future research in precision medicine and highlights the potential of AI-driven approaches in tackling the challenges of genetic variant interpretation.

Artificial intelligence, clinical decision support and machine learning

P0186

AN INNOVATIVE EVIDENCE-BASED LABORATORY MEDICINE (EBLM) TEST TO HELP DOCTORS IN THE ASSESSMENT OF THE CARDIOVASCULAR FUNCTION

S.J. Calleja Freixes [2], A. Roca Vidalmargó [2], J.D. Santotoribio Camacho [1]
[1]*Hospital Universitario Puerto Real*
[2]*Kience Inc, Wilmington, Delaware*

BACKGROUND-AIM

To develop a novel non-invasive, evidence-based laboratory medicine (EBLM) test to assist doctors in assessing the cardiovascular function and to evaluate its accuracy in detecting the main cardiovascular diseases (CVD), as well as the 10-year risk of developing them.

METHODS

This study is part of a previous one already published at the European Society for Medical Oncology (ESMO) Congress 2024, which is focused on the accuracy evaluation of a novel non-invasive test for Multi-Cancer Early Detection (MCED). To develop the algorithm, several combinations of analytes were analyzed to identify the most significant groupings related to the cardiovascular function. The algorithm's efficiency was then enhanced using serial and parallel approximation techniques. Its performance was trained with a dataset of 15,309 patients. The validation of the algorithmic test was performed through a randomized controlled trial (RCT) with a sample size of 152 patients. Their blood samples were tested by Laboratorio Echevarne (Spain), using their biochemistry techniques.

RESULTS

For the RCT, the sensitivity achieved was 1.00 and the specificity was 1.00. Additionally, the area under the receiver operating characteristic (AUROC) curve, the positive predictive value (PPV), and the negative predictive value (NPV), were 1.00, 1.00, and 1.00, respectively. This indicates a strong correlation between the algorithm outcomes and the high likelihood of having CVD, something very important with the global obesity epidemic, a primary risk factor for them.

CONCLUSIONS

This innovative non-invasive blood and urine-based biomarker algorithm holds promise in helping doctors in providing timely and accurate assessment of CVD —even in early stages, as well as reduce medical errors or misdiagnoses. These results advocate further exploration, prompting our intention to conduct a clinical study involving 26,000 participants to enhance our findings and inform clinical practice.

Artificial intelligence, clinical decision support and machine learning

P0187

IDENTIFICATION OF ANTIFUNGAL RESISTANCE IN CANDIDA TROPICALIS BY USING MALDI-TOF MS AND ARTIFICIAL INTELLIGENCE APPROACHES

S. Chen [1], Y. Huang [1]
[1]*Department of Laboratory Medicine, National Taiwan University Hospital, Taipei, Taiwan.*

BACKGROUND-AIM

Invasive candidiasis and candidemia are nosocomial severe infections normally found in patients with weak immune systems, postoperative situations, or central venous catheters. Azole antifungal drugs are widely used in clinical treatment to reduce candida infection. Clinical trials have indicated that the proportion of candida resistance to azole drugs has increased annually. Among the non-albicans Candida species, Candida tropocalis is the most drug-resistant in Taiwan. Traditional methods of identifying antifungal drug resistance take a longer time and require more complicated steps. Hence, this study aims to build an artificial intelligence platform by the combination of mass spectrums based on Matrix-Assisted Laser Desorption/ Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS) technology and drug susceptibility test.

METHODS

In this study, 160 clinical isolates were collected from 2021 to 2024 and the spectra were performed using the MALDI-TOF MS Biotyper system (Bruker, MALDI Biotyper® sirius). The mass spectra database is established with the peak characteristics and distribution and underwent preprocessing steps. In the meanwhile, the susceptibility tests are performed by using commercial products (Thermo Scientific™, Sensititre™ YeastOne™). The machine learning models were trained using preprocessed structured data and corresponding labels.

RESULTS

Among the susceptibility to fluconazole, 53 (33.1%) were resistant, 16 (10.0%) were susceptible-dose dependent, and 91 (56.9%) were susceptible. Data augmentation was achieved through sample repetition to enhance the models' performance. Due to the microbial spectra and individual variability, the molecular features were represented with the average MALDI-TOF signal and the heatmap of all isolates. Several algorithms were applied to determine the most appropriate approach for detecting fluconazole resistance. Then, the accuracies of all pipelines were compared, and the pipelines associated with the highest accuracies were selected.

CONCLUSIONS

With better generalization abilities, our artificial neural network model can serve as a reliable screening tool for C. tropicalis susceptibility to fluconazole. It could provide not only quality improvement but also better efficiency of clinical treatment.

Artificial intelligence, clinical decision support and machine learning

P0188

APPLICATION OF MACHINE LEARNING METHODS IN PREDICTING COMPLICATIONS IN PATIENTS WITH END STAGE KIDNEY DISEASE ON ERYTHROPOIETIN STIMULATING AGENTS THERAPY

D. Ćujić [3], M. Baralić [1], N. Dmitrašinović [4], G. Dmitrašinović [2]

[1]*Department of Nephrology, University Clinical Center of Serbia, 11000 Belgrade, Serbia; School of Medicine, University of Belgrade, 11000 Belgrade, Serbia*

[2]*Faculty of Pharmacy, University of Belgrade, 11221 Belgrade, Serbia*

[3]*Institute for the Application of Nuclear Energy, University of Belgrade, 11080 Belgrade, Serbia*

[4]*Singidunum University, 11010 Belgrade, Serbia*

BACKGROUND-AIM

Erythropoietin stimulating agents (ESA) are common in the management of anemia in patients with end stage kidney disease (ESKD). The aim of this study was to investigate if, application of machine learning (ML) methods could be useful in prediction of development of ESA therapy complications (like thrombosis, myocardial infarction (MI) or cerebrovascular insult (CVI) in ESKD patients.

METHODS

Excel database contained retrospective data of 440 patients, given in 440 rows and 28 columns. Firstly, proper data formatting and selection of columns of interest for medical evaluation were done. ESA therapy was initiated after establishment of primary diagnose (DG1). In case it caused any of the three side effects, these effects were found in columns after secondary (DG2), tertiary (DG3) or after each subsequent examination (DG4). For each examination, "zero" column (DG20, DG30, DG40) was added and effects of interests were marked numerically: thrombosis (1), CVI (2) and MI (3). Label Encoder is used for conversion of textual to numeric data. "Zero" columns were created to check if ML can help distinguish these three diseases from the others. Random forest algorithm was used with application of RandomForestClassifier with several estimators and GradientBoostingClassifier for fine optimization.

RESULTS

For column DG2 RandomForestClassifier with 150 estimators gave the best results of 0.49367, GradientBoostingClassifier gave the result of 0.37975, while with cross-validation the result was 0.48101. The side effects examined were not sufficiently present in DG20 for machine learning model to be applied. RandomForestRegressor was not sufficiently precise because given values were mostly negative. Duration of hemodialysis given as number of months was shown as the most dominant characteristic with the biggest influence on result (0,16332).

CONCLUSIONS

The amount of data in patient database was the biggest limitation factor in confirming the primary hypothesis. However, these preliminary results are the basis for further investigations, with the extended scale database and standardized diagnostic coding system, that would facilitate the application of ML methods in disease management of ESKD patients.

Artificial intelligence, clinical decision support and machine learning

P0189

DEVELOPMENT AND VALIDATION OF AN INTELLIGENT AUTOMATED WORKFLOW FOR ACCURATE PLATELET COUNTING IN HEMATOLOGY LABORATORY

A. Giovannelli [2], F.G. Viola [3], E. Nicolai [5], F. Tomassetti [2], M. Pelagalli [2], M.A. Irno Consalvo [4], S. Gargiulo [4], F. Buccisano [1], M. Pieri [2], S. Bernardini [2]

[1]*Department of Biomedicine and Prevention, University of Rome Tor Vergata*
[2]*Department of Experimental Medicine, University of Tor Vergata, Rome, Italy*
[3]*Department of Laboratory Medicine, Tor Vergata University Hospital, Rome, Italy*
[4]*U.O.S.D. Advanced Diagnostics of Oncohematology, PTV Foundation, Tor Vergata Polyclinic*
[5]*UniCamillus-Saint Camillus International University of Health and Medical Sciences, Rome, Italy*

BACKGROUND-AIM

Platelet (PLT) count is crucial for assessing hemostasis and coagulopathies, with variations in platelet levels often linked to various disorders. Current methods, including microscopy, impedance, optical, fluorescent, and immunological techniques, face challenges that may lead to resource and cost inefficiencies. This study aims to validate an intelligent workflow designed to automatically provide accurate PLT counts, emphasizing its accuracy and efficiency.

METHODS

A total of 1208 samples, including 909 routine samples (randomly selected) and 299 abnormal samples with suspected platelet interference, were collected at Tor Vergata University Hospital. Analyses were performed using the Mindray BC-6800plus hematology analyzer (PLT-I and PLT-O techniques), with blood smears analyzed via the SC-120 slide maker and MC-80 morphology analyzer (PLT-M). Abnormal samples were further evaluated using CD41/CD61 immune-PLT methods to assess the workflow's accuracy.

RESULTS

Among the 299 abnormal samples, 187 samples were reported with PLT-I, showing strong correlation with immune-PLT ($R2 = 0.9334$, mean bias = -18.18). For the remaining samples, 15 were reported with PLT-O, yielding a correlation coefficient of 0.9827. The rest (PLT-O < 100) required morphology analysis with PLT-M to validate the accuracy of PLT-O. Of these, 93 samples passed validation and could report PLT values based on PLT-O, correlating well with the reference method ($R2 = 0.9746$). Only 4 samples required manual verification, as PLT-M exceeded the tolerance range of PLT-O, necessitating manual judgment of the sample status. Regarding the overall efficiency of the intelligent workflow, 1,203 (99.59%) of the samples passed the automated intelligent PLT workflow and reported PLT values. 234 (19.4%) required PLT-O retest, 161 (13.3%) required microscopy review, and 5 (0.41%) required manual review.

CONCLUSIONS

The intelligent PLT counting workflow achieves over 99% accuracy, minimizes fluorescent counting and smear reviews, and enhances efficiency, making it ideal for large hospitals by reducing manual effort, reagent costs, and advancing laboratory automation.

Artificial intelligence, clinical decision support and machine learning

P0190

MACHINE LEARNING FOR DETECTING IRON DEFICIENCY THROUGH COMPREHENSIVE BLOOD ANALYSIS

P. Hsueh [1], Y. Chang [1], C. Chen [1], C. Hsiao [1], H. Shih [1], D. Cho [1]
[1]*China Medical University Hospital, China Medical University, Taichung, Taiwan*

BACKGROUND-AIM

Iron deficiency (ID) is a prevalent global health issue with a major impact on health and well-being. Early detection of ID is crucial but challenging due to its nonspecific symptoms and the limitations of traditional diagnostic tests such as a ferritin blood test; thus, these tests are not feasible for large-scale screening.

METHODS

This study proposes a machine learning approach using complete blood count (CBC) data and cell population data (CPD) for detecting ID in the general population. CBC data and CPD were obtained using the Beckman Coulter 900 (Beckman Coulter, Miami, Florida, USA). We retrospectively collected patient data from three hospitals to train and validate five machine learning models including CBC, CPD, and demographic data. After selecting the highest-performing model, we analyzed the impact of various feature sets on model performance, followed by feature selection and hyperparameter tuning to optimize the accuracy of the model. Additionally, we evaluated the model performance in different subgroups to ensure its robustness in diverse populations.

RESULTS

We retrospectively enrolled 9,608 patients from three hospitals. The model achieved area under the receiver operating characteristic curve values of >0.94 and precision–recall curve values of >0.85, demonstrating the high reliability and robustness. Although models incorporating fewer features did not have significantly lower performance, the model with all features provided improved predictions for some cohorts. Furthermore, Subgroup analysis showed the model performed lower in male and nonanemic populations but maintained acceptable accuracy and effectiveness.

CONCLUSIONS

In conclusion, our study highlights the effectiveness of a machine learning model integrating CPD with CBC parameters for screening ID in the general population.

Artificial intelligence, clinical decision support and machine learning

P0191

AUTOVERIFICATION OF COAGULATION TESTS AND THE TRANSFORMATIVE ROLE OF ALGORITHMS AND DIGITALIZATION IN POST-ANALYTICS

H. Lame [2], E. Refatllari [2], N. Heta [2], I. Korita [2], A. Coraj [2], V. Tole [2], M. Lika [2], A. Bulo [1]

[1]*Laboratory Department, Faculty of Medicine, University of Medicine, Tirane*

[2]*Laboratory Department, Faculty of Medicine, University of Medicine, Tirane; Laboratory Networks, Mother Theresa University Hospital Centre*

BACKGROUND-AIM

Digital transformation of medical laboratories involves the use of digital technology to collect data and automate processes, including post-analytics. Autoverification (AV) allows rapid and accurate verification of laboratory tests. This study's aim was designing, implementation and evaluation of an AV model for coagulation tests (PT, INR, APTT, Fibrinogen, D-Dimer) in Laboratory Networks, Laboratory Department, Mother Theresa University Hospital Centre, Albania.

METHODS

The working group created to design an AV model for STAT samples, chose the rules for Coagulation tests AV following the recommendations of CLSI AUTO 10A:2006, CLSI AUTO15-ED1:2019, ISO 15189:2022, CLSI EP33:2023, CLSI H21-A5:2008 and CAP GEN.43875. AV range limits were set based on CLSI recommendations and assays' linearity. Delta checks were established after RCV calculations and clinical experience adjustment. Logical rules, sample quality indicators, Quality Control results and instrument error messages were also integrated in the algorithm. After implementation on the lab middleware system (AMS), AV was validated for 3 months. During regular operation, model's performance was evaluated for 1000 consecutive STAT samples of coagulation. Data were processed and analyzed using Excel and SPSS.

RESULTS

STAT coagulation samples were autoverified only if every test of the sample fulfilled AV criteria, otherwise they were kept for traditional verification by the laboratory doctors. Delta check rules were prioritized. The performance analysis showed an AV rate of 69.1%. 691 samples containing 2137 single tests were autoverified. There was a statistically significant difference of the AV rate according to the number of tests ordered in each sample. AV failure causes were range limits violation (66.1%), Delta Check failure 20.6%, simultaneous range and delta limits failure (11.3%) and instrument errors (1.9%).

CONCLUSIONS

Laboratory systems digitalization and algorithms' integration transform post-analytics and test results verification process in particular. AV helps in the detection of mismatched samples and acute deterioration of patients' clinical condition. The AV model increases lab performance improving the healthcare service and reducing repetitive tasks for the medical staff, sparing time for more creative activities.

Artificial intelligence, clinical decision support and machine learning

P0192

APPLICATIONS OF AI IN MEDICAL EQA – DATA ACQUISITION, ANALYSIS AND PROGNOSIS OF FUTURE LABORATORY PERFORMANCE

K. Nichani [1], K. Frost [1], B. Colson [1], K. Hettwer [1], K. Simon [1], S. Uhlig [1]
[1]*QuoData GmbH*

BACKGROUND-AIM

The application of artificial intelligence (AI) in external quality assurance (EQA) for medical laboratories represents a transformative approach, focusing on three key areas: data acquisition, analysis, and performance prognosis.

METHODS

In the data acquisition phase, large language models (LLMs) coupled with retrieval augmented generation (RAG) systems effectively extract and consolidate method-related information from both structured and unstructured data sources, significantly improving the breadth and depth of laboratory documentation. Combining meta data and measurement data opens possibilities to not only improve workflows but also provide many more actionable KPIs to laboratories. The analysis component employs robust AI algorithms to process historical EQA data, providing deeper insights into performance patterns than traditional statistical approaches. A novel AI-driven prognostic model evaluates the probability of failure in future EQA exercises by taking into account several factors such as historical trends, data variability and correlations between consecutive test results.

RESULTS

This approach thus overcomes an important limitation of being able to collect unstructured information on procedures and methods and large measurement datasets (e.g. spectra and sequences) in an EQA. The proposed approach demonstrates capability in effectively mapping the data and using it to detect subtle patterns. The methodology's effectiveness was validated using simulated data from a number of laboratories. from a number of laboratories, showing improved accuracy in predicting failure compared to traditional statistical methods.

CONCLUSIONS

The integration of AI technologies in EQA will translate to a more robust and predictive approach to performance assessment while providing actionable insights for proactive quality management to thus laboratories and streamlining quality control procedures.

Artificial intelligence, clinical decision support and machine learning

P0193

THE USE OF GAAD MEDICAL ALGORITHM IN ROUTINE CLINICAL PRACTICE: A CASE SERIES

Đ. Kralj [1], D. Obradović [2], S. Stanković [3]
[1]*Clinical Hospital Centre Zvezdara*
[2]*Roche Diagnostics Poland, Faculty of Pharmacy*
[3]*University Clinical Centre of Serbia*

BACKGROUND-AIM

Hepatocellular carcinoma (HCC) is the most prevalent primary liver malignancy. Timely diagnosis and risk stratification are crucial for improving HCC prognosis. The GAAD (Gender, AFP, Age and DCP) medical algorithm combines clinical biomarkers and parameters such AFP, PIVKA-II, gender and age to provide an HCC risk assessment. GAAD enables prompt evaluation aiding in diagnostic accuracy, faster treatment and follow-up planning. This study evaluates implementation of GAAD in a real-world setting through a series of complex patient cases.

METHODS

This analysis included five patients treated at a tertiary medical center. Data were extracted from clinical records presenting features, diagnostic workups, GAAD medical algorithm scores (risk score system with cut off value of 2.57 - higher risk number scored leads more likely to HCC is developed) and therapeutic decisions.

RESULTS

Case 1: A 55-year-old woman with NASH-related HCC underwent resection with an GAAD score of 1.87 after resection. Despite initial R0 resection, recurrence within two years emphasized the importance of dynamic GAAD assessment in surveillance protocols.
Case 2: A 69-year-old woman with advanced HCC disease presented with a GAAD score of 8.03, prompting lesion biopsy and systemic therapy.
Case 3: A 61-year-old man with cirrhosis scored a low GAAD score (0.5), corresponding to non-malignant imaging findings.
Case 4: A male patient with advanced HCC and HCV. GAAD initial score was 9.29. In combination with other clinical parameters management strategy was shaped.
Case 5: Multifocal HCC in a patient with significantly elevated PIVKA-II level reflected a poor prognosis. The GAAD score (10) supported the decision to prioritize palliative care.
The concordance on the HCC presence, in this case GAAD and other diagnostics procedures (US, hepatogram, others) was 100%.

CONCLUSIONS

The GAAD medical algorithm score is a highly valuable diagnostic-prognostic arsenal for HCC disease, improving patient stratification and personalized care within a multidisciplinary framework. Its integration into routine practice underscores the importance of combining biomarker-based assessments with clinical judgment. This series confirms the potential of GAAD to refine clinical decisions and enhance outcomes in diverse patient populations.

Artificial intelligence, clinical decision support and machine learning

P0194

EVALUATING HISTOPATHOLOGY EQA SCHEME RESULTS: HUMAN EXPERTISE VERSUS AI ALGORITHMS

J. Pelanti [4], P. Eloranta [4], J. Juhila [2], A. Wester [1], T. Tolonen [3], M. Laurila [3], H. Berghäll [4]
[1]*Aiforia Technologies, Helsinki, Finland*
[2]*Aiforia Technologies, Helsinki, Finland (at the time of this research project)*
[3]*Department of Pathology, Fimlab Laboratories, Tampere, Finland*
[4]*Labquality EQAS by Aurevia, Aurevia, Helsinki, Finland*

BACKGROUND-AIM

Accurate prostate cancer identification depends on Gleason score evaluation, combining the most common and aggressive grades to determine a Grade Group (GG) from 1 to 5, with 5 being the most aggressive. The developments in digital pathology and artificial intelligence (AI) allow the use of whole slide images (WSI) alongside expert evaluations. Labquality EQAS by Aurevia conducts a virtual histopathology external quality assessment (EQA) scheme twice a year. In round 2-2023, participants analyzed 7 whole specimen scanned slides on prostate cancer. We compared the visual image GG analysis of the scheme participants with two different AI (AI1 and AI2) standalone results. AI is intended to be used as a supportive tool for the medical professional and this needs to be taken into consideration.

METHODS

Clients received relevant clinical histories for 7 scanned virtual prostate biopsies. The formalin-fixed samples were stained with hematoxylin and eosin. 149 pathologists graded the samples, which were also analyzed by two AI models detecting tumor epithelium and Gleason patterns from WSIs.

RESULTS

The AI results aligned with the participants' Gleason scoring in 3 of 7 cases (cases 1, 2, and 6). For GG, there was consensus in cases 2 and 6. In case 1, 28% of participants and AI1 graded it as GG3, while AI2 graded it as GG4. In case 3, 49% of participants graded it as GG2, and 28% as GG3, which matched AI1; AI2 could not read this case. For case 4, 55% of participants and AI2 graded it as GG3, while AI1 graded it as GG5 with a minority vote for Gleason pattern GG5. Case 5 showed the most discrepancy, with participants grading from GG2-5; 19% graded it as GG2, matching both AIs. In case 7, 61% of participants graded it as GG1, while AI graded it as GG2, agreeing with 32% of participants.

CONCLUSIONS

AI models can analyze WSIs, helping to reduce the workload of medical professionals. In this study, the grading of samples differed slightly between participants and AI models. However, variability in Gleason scoring and GG among participants highlights the challenges in diagnosis. In all cases, both participants and AI models graded the clinical outcomes similarly, ensuring patients could receive comparable treatment.

Artificial intelligence, clinical decision support and machine learning

P0195

BUILDING AN EXPLAINABLE MACHINE LEARNING MODEL FOR PREDICTING BONE MARROW METASTASIS OF MALIGNANT SOLID TUMORS USING HAEMATOLOGICAL PARAMETERS

H. Li [1], Z. Du [1], W. Xu* [1], L. Qu* [1]
[1]*Department of Laboratory Medicine, the First Hospital of Jilin University*

BACKGROUND-AIM

To develop and validate a preferable explainable machine learning model for predicting bone marrow metastasis (BMM) in malignant solid tumors using haematological parameters.

METHODS

Methods: Patients with malignant solid tumors (744 cases) admitted to the First Hospital of Jilin University from January 2018 to December 2024 were collected and split into derivation cohort (496 cases for model building and internal validation) and validation cohort (248 cases for external validation). Basic clinical information was queried from the hospital electronic medical record system, including 45 laboratory indicators, such as blood routine, coagulation routine, tumour markers, liver function, and ion, which were taken as candidate feature parameters, and bone marrow metastasis as the ending event. Nine machine learning (ML) algorithms were used to build the predicting model. The maximum area under the curve (AUC), sensitivity, specificity, and F1 score were calculated to evaluate the optimal model and the combination of feature parameters. The Shapley Additive exPlanation method was used to rank the importance of features and to explain the final model.

RESULTS

Among 9 ML models, the BMM-RF-4 model from random forest with 4 parameters CA153, NSE, PLT and RBC showed the highest AUC both in derivation cohort and in validation cohort, 0.906 and 0.910, respectively, and has been transformed into a convenient tool to facilitate its application in clinical settings.

CONCLUSIONS

We built an explainable ML predicting model for bone marrow metastasis of malignant solid tumors, BMM-RF-4, which was developed based on haematological parameters to offer a laboratory basis for BMM predicting.

Artificial intelligence, clinical decision support and machine learning

P0196

IDENTIFICATION OF KOCH BACILI THROUGH MACHINE LEARNING

J. Raraz-Vidal [1]
[1]*Villa el Salvador Emergency Hospital - Peru*

BACKGROUND-AIM

the aim of the study was to develop and evaluate the clinical efficacy of the automated method based on artificial intelligence to identify koch bacilli in Ziehl-Neelsen (ZN) stained sheets.

METHODS

Application study. A pilot study was carried out. An automated method (based on AI) for the identification of mycobacteria was developed. We prepared a training data set with 85 positive and 85 negative slides with the same size and color, from ZN-stained slides scanned and published on the internet. Which were confirmed by 2 clinical pathologists confirming positive and negative lamina. A neural network model based on machine learning algorithms was created to identify Koch's bacillus through its characteristics, in addition to training the neural network to improve identification. There was a sample of 44 slides (22 positives). Sensitivity, specificity, PPV, NPV, LR were estimated.

RESULTS

The pilot study was performed with 44 images (22 images with Koch's bacilli and 22 images without Koch's bacilli). We compared the pathologists' results obtained by separately evaluating the images and the results obtained through the neural network. The test obtained a sensitivity of 90% and a specificity of 80% by the AI-assisted method, for the detection of AFB. The PPV 82%, NPV 89%, FP 20%, FN 10%. The positive Likelihood ratio (LR) obtained 4.5 and the negative LR 0.12.

CONCLUSIONS

The artificial intelligence program presented good sensitivity and specificity to identify koch bacilli

Artificial intelligence, clinical decision support and machine learning

P0197

AN INNOVATIVE EVIDENCE-BASED LABORATORY MEDICINE (EBLM) TEST TO HELP DOCTORS IN THE BASIC ASSESSMENT OF THE PROSTATE FUNCTION

A. Roca Vidalmargó [2], S.J. Calleja Freixes [2], J.D. Santotoribio Camacho [1]
[1]*Hospital Universitario Puerto Real, Cádiz, Spain*
[2]*Kience Inc., Wilmington, Delaware, US*

BACKGROUND-AIM

To develop a novel non-invasive, evidence-based laboratory medicine (EBLM) test to assist doctors in assessing the prostate function and to evaluate its accuracy in detecting the main prostate diseases, such as benign prostate hyperplasia (BPH), prostatitis and prostate cancer (PCa).

METHODS

This study is part of a previous one already published at the European Society for Medical Oncology (ESMO) Congress 2024, which is focused on the accuracy evaluation of a novel non-invasive test for Multi-Cancer Early Detection (MCED). To develop the algorithm, several combinations of analytes were analyzed to identify the most significant groupings related to the prostate function. The algorithm's efficiency was then enhanced using serial and parallel approximation techniques. Its performance was trained with a dataset of 2,160 patients. The validation of the algorithmic test was performed through a randomized controlled trial (RCT) with a sample size of 152 patients. Their blood samples were tested by Laboratorio Echevarne (Spain), using their immunoassay techniques.

RESULTS

For the RCT, the sensitivity achieved was 1.00 and the specificity was 1.00. Additionally, the area under the receiver operating characteristic (AUROC) curve, the positive predictive value (PPV), and the negative predictive value (NPV), were 1.00, 1.00, and 1.00, respectively. This indicates a strong correlation between the algorithm outcomes and the high likelihood of having a prostate disease, including PCa.

CONCLUSIONS

This innovative non-invasive blood-based biomarker algorithm holds promise in helping doctors in providing timely and accurate basic assessment of prostate disease —even in early stages— and PCa, as well as reduce medical errors or misdiagnoses. These results advocate further exploration, prompting our intention to conduct a clinical study involving 26,000 participants to enhance our findings and inform clinical practice.

Artificial intelligence, clinical decision support and machine learning

P0198

ASSESSMENT OF KNOWLEDGE AND ATTITUDE OF LABORATORY STAFF ON USE OF ARTIFICIAL INTELLIGENCE USE IN LABORATORY; AUDIT DONE IN LOWER MIDDLE-INCOME COUNTRY: SRI LANKA.

S. Pavithra [1], D. Jayasekara [1], P. Gamage [1], T. Nallaiah [1], G. Katulanda [1]
[1]*National Hospital Sri Lanka*

BACKGROUND-AIM

This study aimed to assess the knowledge and attitude of laboratory staff regarding the use of Artificial Intelligence (AI) in laboratory settings in Sri Lanka, a lower-middle-income country.

METHODS

A cross-sectional audit was conducted among laboratory professionals in National Hospital Sri Lanka to evaluate their familiarity with AI technologies and their perspectives on the integration of AI in laboratory operations. The study utilized a structured questionnaire, which was distributed to laboratory technicians, medical technologists, laboratory doctors and postgraduate trainees and other relevant staff members.

RESULTS

The results revealed a wide variation in knowledge levels, with many participants having limited awareness of AI's capabilities and the definition at the laboratory level. While some staff members were positive about AI's potential to enhance diagnostic accuracy and improve laboratory efficiency, concerns about the cost of implementation, the need for extensive training, and the risk of job displacement were commonly expressed. It was common opinion that AI might affect the human creativity and there should be human AI bridging to make the service better.

CONCLUSIONS

The findings emphasize the necessity for targeted educational initiatives to improve laboratory staff's understanding of AI and its practical use in resource-constrained settings. The study also highlights the challenges associated with adopting AI in low-income countries, including cultural and infrastructural barriers.

Artificial intelligence, clinical decision support and machine learning

P0199

A COMPUTATIONAL METHODOLOGY FOR IDENTIFYING POTENTIAL THERAPEUTIC AGENTS FOR LUNG ADENOCARCINOMA

A. Tuncal [1], R. Kalkan [2]

[1]*Department of Medical Biochemistry, Faculty of Medicine, Cyprus Health and Social Sciences University, Mersin 10, Morphou 99750, Northern Cyprus*

[2]*Faculty of Medicine, European University of Lefke, Mersin 10, Lefke 99728, Northern Cyprus*

BACKGROUND-AIM

Lung adenocarcinoma (LUAD) is a variety of non-small cell lung cancer (NSCLC), and it is the most prevalent type of lung cancer. The prognosis of LUAD is significantly influenced by the stage at which the diagnosis is made. Patients diagnosed with early-stage disease typically exhibit a more favorable prognosis when surgical intervention is performed; conversely, those with advanced-stage disease frequently experience a less favorable outcome. However, it is noteworthy that introducing targeted therapies and immunotherapies has led to an enhancement in survival rates for certain patients.

The genetic components of LUAD play a major role in its initiation, progression, and response to therapy. In recent decades, significant advances have been made in uncovering the genetic mutations and alterations that drive LUAD, leading to the development of targeted therapies. These therapies are specifically tailored to an individual's unique tumor genetic makeup to improve outcomes in select patients. TP53 is the most frequently mutated gene in LUAD, with somatic mutations have been identified in approximately 70% of patient samples. The present study aims to identify candidate compounds that could target the TP53 gene and assess their potential anti-tumor effects using in-silico approaches.

METHODS

Candidate compounds that target TP53 based on their biological functions were identified from the Gene2drug database. Drug set enrichment analysis (DSEA) was used to assess the pathway-based similarity of these drugs. Candidate drug antitumor activities were extracted from DepMap subsequent to the PRISM viability assay on the lung adenocarcinoma cell lines and ranked by Gene2drug output ($P < 1E{-}5$).

RESULTS

A total of 1,117 compounds were analyzed using the Drug Sensitivity (PRISM Repurposing Public 24Q2) 24Q2 Tool. TP53 damaging mutation and Drug Sensitivity PRISM Repurposing Public 24Q2 was used to sort lung adenocarcinoma cell lines for sensitivity to candidate drugs. The sensivity of atropine was extracted from DepMap (6.83E-5).

CONCLUSIONS

This in-silico study shows sensitivity to atropine and suggests its potential in tumor-targeted strategies. Further investigation into efficacy and mechanisms of action is advised.

Artificial intelligence, clinical decision support and machine learning

P0200

REFERENCE INTERVAL (RI) INDIRECT ESTIMATION: VITAMIN B12 IN ADULTS IN BRAZIL, LAB RI TOOL IN MINING OF DATALAKE

D.G.R. Canali [1], M.d.A.V. Feistauer, [1], K.T. Avi [1], L.M.F. Vieira [2]
[1]*Dasa*
[2]*LabRI - Dasa*

BACKGROUND-AIM

Vitamin B12, known as cobalamin, is a water-insoluble micronutrient found in foods of animal origin and is involved in vital metabolic processes in humans. The objective was to evaluate Reference Intervals (RI) described by the manufacturer Roche® in comparison with data from Brazilian clinical laboratory with operations localized in two states (Rio de Janeiro and Santa Catarina), using the LabRI® statistical tool with a combination of parametric, non-parametric and robust statistical algorithms and techniques to exclude latent abnormal values, outliers and estimate reference limits and their respective confidence intervals.

METHODS

A total of 6.224 vitamin B12 results were obtained within the linear range from routine laboratory tests performed on outpatients of both sexes, aged 19 to 60 years, and with exclusion criteria predefined by the medical staff. All results were analyzed using the electrochemiluminescence method on Roche® equipment, Cobas 6000 and 8000 models. Sample processing followed standardized pre-analytical and analytical criteria controlled by good laboratory practices and current legislation, in addition to being covered by the National Accreditation Standard of the Clinical Laboratory Accreditation Program (PALC).

RESULTS

After statistical processing of the data, 66 outliers (1,1%) were detected and excluded, maintaining the mean and median results with symmetrical distribution. The RI proposed by the manufacturer was 197 to 771 pg/mL for both genders. In the estimation of new IRs, using a 90% confidence interval, the range of 245 to 985 pg/mL was found.

CONCLUSIONS

After utilizing a robust statistical tool and multidisciplinary analysis of the results found, the RI provided by the manufacturer was replaced by RI specific to the adult population served.

Artificial intelligence, clinical decision support and machine learning

P0201

UNSUPERVISED DEEP LEARNING FOR DETECTING PATTERNS OF METABOLIC DISEASES IN LABORATORY MEDICINE

S. Amusat [1]
[1]*Chemical Pathology Department, The Directorate of Laboratory Medicine, Beaumont Hospital, Dublin, Ireland*

BACKGROUND-AIM

Unsupervised learning is a machine learning approach that identifies patterns in unlabelled data without predefined outputs. Unlike supervised learning, it autonomously categorises data into meaningful groups, such as normal versus abnormal, based on intrinsic attributes. This study applied artificial neural networks (ANNs) and anomaly detection to analyse metabolic profile data, aiming to flag potential metabolic diseases risks.

METHODS

Using data from 1,308 subsamples in the 2015-2016 NHANES survey, the study population consisted of males and females aged 18–59 years. An autoencoder, a specialised ANN was implemented using H2O's deep learning framework in R. The model featured a 5-layer neural network with Tanh activation functions for non-linear transformations, and regularisation techniques (L1 and L2) to prevent overfitting. Trained with a quadratic loss function and Gaussian distribution, the autoencoder minimised reconstruction errors and flagged cases exceeding a threshold (0.0775, 95% CI). A low training mean squared error (MSE) of 0.028 demonstrated the model accuracy.

RESULTS

The analysis revealed 17 abnormal cases among 328 test samples, characterised by significantly higher mean reconstruction errors compared to normal samples of MSE of 0.114 versus 0.023, respectively. Abnormal cases showed pronounced deviations in variables such as body composition parameters, glucose, lipid profile, and inflammatory markers, suggesting potential metabolic or systemic dysregulation. Density plots highlighted a distinct abnormal subgroup with shared anomalies across multiple parameters, though with overlapping intervals, validating the model's ability to discern clinically relevant patterns

CONCLUSIONS

This study demonstrates the utility of unsupervised deep learning techniques for anomaly detection in unlabelled clinical datasets. The autoencoder successfully identified anomalies linked to clinical conditions like type-2 diabetes, metabolic syndrome, dyslipidemia, and obesity. These findings highlight the potential of unsupervised deep learning in uncovering hidden patterns, offering powerful tools for exploratory analysis in clinical research.

Artificial intelligence, clinical decision support and machine learning

P0202

TRAINING AND VALIDATING A MACHINE LEARNING ALGORITHM FOR SEPSIS PREDICTION IN EMERGENCY DEPARTMENT USING A LARGE LIST OF ADMISSION BIOCHEMICAL TESTS

C.L. Brasen [3], E.S. Andersen [3], J.B. Madsen [2], J. Hastrup [2], H. Christensen [1], D.P. Andersen [4], P.M. Lind [2], N. Mogensen [3], P.H. Madsen [5], A.F. Christensen [5], J.S. Madsen [1], E. Ejlersen [6], I. Brandslund [2]

[1]*Department of Biochemistry and Immunology, Lillebaelt Hospital, University Hospital of 9 Southern Denmark, Beriderbakken 4, 7100 Vejle, Denmark.*

[2]*Department of Biochemistry and Immunology, Lillebaelt Hospital, University Hospital of Southern Denmark, Beriderbakken 4, 7100 Vejle, Denmark.*

[3]*Department of Biochemistry and Immunology, Lillebaelt Hospital, University Hospital of Southern Denmark, Beriderbakken 4, 7100 Vejle, Denmark. Faculty of Health Sciences, Department of Regional Health Research, University of Southern Denmark, Campusvej*

[4]*Department of emergency, Kolding Hospital, Lillebaelt Hospital, University Hospital of Southern Denmark, Sygehusvej 24, 6000 Kolding, Denmark*

[5]*Department of Medicine, Kolding Hospital, Lillebaelt Hospital, University Hospital of Southern Denmark, Sygehusvej 24, 6000 Kolding, Denmark.*

[6]*Department of Medicine, Vejle Hospital, Lillebaelt Hospital, University Hospital of Southern Denmark, Beriderbakken 4, 7100 Vejle, Denmark.*

BACKGROUND-AIM

The patients in the medical emergency department are difficult to diagnose due to a wide range of symptoms and diseases. The emergency department often has a high number of patients and this is expected to increase in many countries due to an ageing population. In this study we have investigated the feasibility of developing machine learning algorithms to assist physicians handling the complex situation of diagnosing sepsis in the medical emergency departments. Using algorithms like this is expected to reduce diagnostic errors and improve patient outcome.

METHODS

We included a total of 9,190 consecutive patients contacting one of two medical emergency departments for diagnosis and treatment in this cohort study. All patients had a large biochemical workup performed at admission including blood and urine analyses on clinical decision totaling 260 biomarkers. After adding nurse-registered data we trained machine learning algorithms on a random 80% sample of the patients and validated the results on the remaining 20%. Patients who were admitted or developed sepsis during their hospital stay were used as the target for supervised model training using 10-fold cross validation.

RESULTS

A total of 199 patients either had sepsis at admission or developed it during the hospital stay. Area under the ROC curve reached 90.1% in the training cohort and 90.0% in the holdout cohort.

CONCLUSIONS

We have shown that it is possible to develop a machine-learning algorithm with high AUC for use in medical emergency departments using laboratory results performed at admission combined with data from nurse triage.

Artificial intelligence, clinical decision support and machine learning

P0203

PREDICTIVE MODEL OF TYPE 2 DIABETES MELLITUS BASED ON CLINICAL, BIOCHEMICAL AND GUT MICROBIOTA PROFILES: DEEP LEARNING DEVELOPMENT.

S. Castañeda Nieto [2], L. Valentín Aragón [2], M.J. Cano Llergo [2], F. Rodríguez Cantalejo [2], A. Camargo García [1], J. Caballero Villarraso [3]
[1]*Maimónides Biomedical Research Institute of Córdoba (IMIBIC)*
[2]*Reina Sofia University Hospital*
[3]*Reina Sofia University Hospital/Maimónides Biomedical Research Institute of Córdoba (IMIBIC)*

BACKGROUND-AIM

Type 2 diabetes mellitus (T2DM) is a chronic disorder of carbohydrate metabolism associated with risk factors -such as obesity, dyslipidemia and hypertension-, and unhealthy lifestyles that lead to the development of severe macro- and microvascular complications. The microbiota, especially the intestinal microbiota, plays a key role in metabolic homeostasis given its participation in different metabolic processes, as well as in the development of DM-2 and its complications. With the advance of artificial intelligence (AI) it is possible to develop novel models based on machine learning (ML) to predict the risk of developing T2DM and facilitate its diagnosis. Our aim was to develop a predictive model of the risk of developing T2DM based on clinical, biochemical and gut microbiota parameters, which estimates the time margin for developing this disease.

METHODS

A Deep Learning Multilayer Perceptron (MLP) algorithm was developed and trained with real patient data from a epidemiological study. Data were normalized and augmented to increase their diversity and avoid over-fitting the model. Subsequently, the model was trained using Google Colab, a free online environment that allows Phyton code to be run and, once trained, the neural network was constructed. The developed network was normalized and optimized, and the best hyperparameters were chosen for model construction using Bayesian optimization. To increase the available data in order to better train the model, a common ML technique of aggregating 'noise' was employed. This allows an increase in the diversity of the data, avoiding overfitting, but keeping the model within reasonable limits (using a normal distribution on the real data) and the original shape (size).

RESULTS

The predictor model works as a computer application in a web environment that returns a numerical result corresponding to the number of months it will take for a particular individual to develop T2DM (based on their clinical, biochemical and microbiota parameters) with an accuracy of 95.2%.

CONCLUSIONS

Accessible computer development has been achieved, with a friendly and intuitive interface and with clinical utility for application in preventive medicine. It could be implemented as a primary prevention tool to minimize the risk of developing T2DM.

Artificial intelligence, clinical decision support and machine learning

P0204

ASSOCIATION RULE MINING FOR ANALYZING CALCIUM-VITAMIN D METABOLISM AND ANTI-EPILEPTIC MEDICATIONS: A BEHAVIORAL APPROACH

P.K. Dabla [2], K. Upreti [1], S. Gupta [3]
[1]*Department of Computer Science, CHRIST (Deemed to be University), Delhi NCR, Ghaziabad, India*
[2]*Department of 'Biochemistry' G.B. Pant Institute of Postgraduate Medical Education & Research (GIPMER), New Delhi, India*
[3]*Department of 'Neurology' G.B. Pant Institute of Postgraduate Medical Education & Research (GIPMER), New Delhi, India*

BACKGROUND-AIM

Epilepsy is a prevalent neurological disorder that often requires Anti-Epileptic Drugs (AEDs) as the primary treatment. However, there is limited research on the impact of AEDs on calcium metabolism, particularly using association rule mining (ARM) techniques. ARM is a data analysis method used to identify frequent patterns, correlations, and associations within patient behavior datasets.

METHODS

In this study, association rule mining was applied to examine the relationship between calcium-vitamin D metabolism and AED use. Participants were epileptic patients, both male and female, attending neurology outpatient and inpatient clinics. The patients were divided into three groups: Group 1 (one AED), Group 2 (two AEDs), and Group 3 (more than two AEDs). The analysis focused on key biomarkers, including total calcium, phosphorus, alkaline phosphatase, ionized calcium, and vitamin D levels.

RESULTS

A total of 150 patients were studied, with 50 patients in each group. The cohort was 60% male, and 86 patients had generalized epilepsy, while 64 presented with partial seizures. Notably, 42% of the patients had been on AEDs for more than five years. The analysis revealed that polytherapy (the use of multiple AEDs) was associated with lower calcium and vitamin D levels when compared to monotherapy and dual therapy. Additionally, polytherapy was linked to elevated alkaline phosphatase and phosphorus levels.

CONCLUSIONS

The application of association rule mining provided valuable insights into the impact of AED use on calcium-vitamin D metabolism. Key factors such as age, gender, and polytherapy were found to influence parathyroid hormone levels in patients taking AEDs. These findings underscore the importance of considering bone health in epilepsy management, with potential interventions, such as prophylactic vitamin D supplementation, recommended for patients undergoing long-term AED treatment.

Artificial intelligence, clinical decision support and machine learning

P0205

DIAGNOSTIC SIGNIFICANCE OF SLC2A9 GENE POLYMORPHISMS AND SERUM BIOMARKERS IN GOUT AND HYPERURICEMIA

W. Feng [2], L. Cui [2], X. Zhao [1]
[1]*Beijing Friendship Hospital, Capital Medical University*
[2]*Peking University Third Hospital*

BACKGROUND-AIM

Uric acid, the final product of purine metabolism, is primarily excreted through the kidneys and intestines. Dysregulation in uric acid production or excretion can result in hyperuricemia and gout. GLUT9, a key protein for uric acid excretion, has garnered significant attention, particularly for two SNPs (rs3733591 and rs1014290) on its encoding gene SLC2A9. However, their relationship with gout and hyperuricemia in the Han Chinese population has not been researched.

METHODS

This study investigated 498 individuals, including 300 patients with hyperuricemia or gout and 198 healthy controls. The genotypes of rs3733591 and rs1014290 were determined using the multicolor melting curve analysis (MMCA) method. Additionally, clinical and laboratory data from 433 participants were used to construct diagnostic models.

RESULTS

The results showed that the proportion of the wild-type (C) allele at rs3733591 and the mutant (A) allele at rs1014290 were significantly higher in patients with hyperuricemia and gout compared to healthy controls. Additionally, individuals with the homozygous mutant genotype at rs3733591 had significantly lower uric acid levels compared to those with the wild-type and heterozygous genotypes, whereas homozygous mutants of rs1014290 exhibited higher uric acid levels. Among the various models, the Logistic Regression model based on eight factors—gender, age, eGFR, WBC, HDL, MCHC, rs3733591, and rs1014290—demonstrated the best diagnostic performance for gout and hyperuricemia, achieving an AUC of 0.8737.

CONCLUSIONS

The results revealed the SLC2A9 gene plays a pivotal role in regulating serum uric acid levels and contributes to the development of gout and hyperuricemia in the Han Chinese population. And polymorphisms of SLC2A9 can be used in combination with a variety of other indicators to diagnose gout and hyperuricemia.

Artificial intelligence, clinical decision support and machine learning

P0206

MACHINE LEARNING ALGORITHMS FOR THE PREDICTION OF URINARY TRACT INFECTION USING DEMOGRAPHICS AND DIPSTICK REFLECTANCE RESULTS

J. Favresse [1], J. Cabo [1], M. Bosse [1], B. Lardinois [1], J. Cadrobbi [1], K. Laffineur [1], M. Elsen [1], J. Douxfils [4], L. Roelandts [3], S. De Bruyne [2]
[1]*Clinique Saint-Luc Bouge*
[2]*Ghent*
[3]*QUALIblood*
[4]*UNamur*

BACKGROUND-AIM

Urinary tract infections (UTIs) are among the most common infections encountered in both primary and secondary healthcare settings. Current diagnostic practices for UTIs often require 24–48 hours due to the time needed for culture results. Given that 70–80% of urine cultures return negative, there is significant interest in rapidly identifying negative samples to reduce unnecessary antibiotic use. This study aimed to evaluate the performance of six machine learning models in predicting UTIs.

METHODS

Urine samples from 22,961 patients collected between September 28, 2023, and June 29, 2024, were analyzed. Six machine learning models were assessed for their ability to predict UTIs based on five definitions incorporating pyuria and culture outcomes. The dataset was randomly divided into a training set (70%, n=16,072) and an independent test set (30%, n=6,889). Seventeen predictive parameters, including dipstick reflectance results (cobas u 601, Roche Diagnostics) and demographic variables (age, sex, and hospital visit type), were evaluated.

RESULTS

The CatBoost Classifier emerged as the best-performing model, achieving an area under the curve (AUC) between 92.5% and 94.5% depending on the UTI definition, with a negative predictive value (NPV) consistently exceeding 95%. In comparison, the predictive performance of nitrite (NIT; AUC = 72.5%) or leukocyte esterase (LEU; AUC = 73.6%) alone was significantly lower than any machine learning model ($p<0.0001$).

CONCLUSIONS

Machine learning models, particularly the CatBoost Classifier, demonstrate high accuracy and offer a promising tool to aid clinicians in UTI diagnosis. Unlike traditional culture methods, these models deliver results within an hour. Further external validation with an independent dataset and prospective studies assessing the impact on antibiotic prescribing practices are recommended.

Artificial intelligence, clinical decision support and machine learning

P0207

HARNESSING DATA AUGMENTATION IN A DEEP LEARNING MODEL FOR THE INTERPRETATION OF LIPOPROTEIN GEL ELECTROPHORESIS IN HYPERLIPIDAEMIA

S.L. Kakako [1], R. Punchoo [4], I. Fabris-Rotelli [3], T.S. Pillay [2]
[1]*Department of Chemical Pathology, University of Pretoria, South Africa*
[2]*Department of Chemical Pathology, University of Pretoria/National Health Laboratory Service (NHLS), South Africa*
[3]*Department of Statistics, University of Pretoria, South Africa*
[4]*Division of Chemical Pathology, University of Cape Town/National Health Laboratory Service (NHLS), South Africa*

BACKGROUND-AIM

As the morbidity and mortality rate of atherosclerotic cardiovascular disease (ASCVD) continues to surge, all efforts must be harnessed to reduce it. The situation is particularly acute in low- and middle-income countries, where four out of five cardiovascular disease deaths occur from ASCVD thereby necessitating the need for accurate, timely, and accessible diagnosis. Lipid gel electrophoresis is pivotal in the diagnosis of hyperlipidaemia and the interpretation is mostly based on Fredrickson's classification. The correct interpretation of gel results requires human expertise and experience. However, the interpretation may be subjective, and time-consuming, and a generalized standard is not well established. Considering the success of artificial intelligence in repetitive human tasks, we adopted an augmented deep-learning model to interpret lipoprotein electrophoretic gel.

METHODS

A total of 836 labelled gels were analysed. The dataset was split into 70% and 30% training and test sets, respectively. For improved robustness and generalization, training dataset augmentation was transformed, leaving each class with 300 gels. A convolutional neural network (CNN) model was trained with the augmented dataset, while the raw 30% dataset was used for the testing.

RESULTS

All the Fredrickson types and normal classes had uneven gel distribution, but type I was not represented in the cohort. The CNN performs best in the type IIb hyperlipidaemia class while overall accuracy was 73%. Each class precision, recall and F1-score were as follows: Normal (72%, 87%, and 79%), type IIa (76%, 56%, and 64%), type IIb (84%, 94%, and 89%), type III (72%, 67%, and 70%), type IV (58%, 79%, and 67%), and type V (87%, 57%, and 69%).

CONCLUSIONS

An augmented deep learning model promises to be valuable in classifying lipoprotein gels in hyperlipidaemia, especially with a reasonable number of real-life datasets. This will enhance the resilience of the model to variation and improve the interobserver bias traditionally observed in the human expert interpretation of lipoprotein gel.

Artificial intelligence, clinical decision support and machine learning

P0208

NEW LABEL-FREE SERUM EXOSOMES DETECTION METHOD BASED ON HIERARCHICAL SERS SUBSTRATE FOR DIAGNOSIS OF PANCREATIC CANCER USING AI

Y. Li [2], L. Du [1], J. Li [2], P. Li [2]
[1]*Department of Clinical Laboratory, Qilu Hospital of Shandong University, Jinan*
[2]*Department of Clinical Laboratory, The Second Hospital of Shandong University, Jinan*

BACKGROUND-AIM

Early diagnosis significantly enhances the 5-year survival rate of pancreatic cancer (PaC) patients. Obtaining information on molecular phenotypic changes in exosomes provides prospects for early non-invasive diagnosis of PaC. Unfortunately, current detection modes for exosomes are time-consuming and still not sensitive enough, so methods that can directly obtain exosome information in complex biological fluids are urgently needed.

METHODS

In this study, we developed a new method for early diagnosis of PaC by obtaining a spectral set of serum exosomes on a hierarchical surface-enhanced Raman scattering (SERS) substrate and analyzing them use artificial intelligence (AI).

RESULTS

We designed a micro-lens array/silver nanowires/silver nanoparticles hierarchical SERS substrate (MLA/AgNWs/AgNPs H-SERS substrate) that exhibited a minimum detection concentration of 10-9 M and a minimum relative standard deviation of 7.68%. The superior performance of the substrate increased the strength and stability of exosome biological information acquisition. Furthermore, through the spectral analysis of exosome from 149 serum samples using AI, we performed PaCs diagnosis easily with an area under the receiver operating curve (AUROC) of 0.96 and successfully classified 24 cases of early PaCs. Moreover, the maximum diagnostic positive rate of 161 cases of non-pancreatic cancer was 4.44%, supporting the fact that the model was specific.

CONCLUSIONS

This study provides a new sensitive diagnostic method for early diagnosis of PaC.This label-free Raman spectral analysis can potentially be extended to identify multiple cancers, offering a non-invasive diagnostic approach for clinic.

Artificial intelligence, clinical decision support and machine learning

P0209

CLASSIFICATION OF CERVICAL LESIONS BASED ON HPV GENOTYPING USING MACHINE LEARNING ALGORITHMS

R. Liu [1], S. Wang [1], W. Miao [1], P. Yin [1], L. Dai [1]
[1]*Sansure Biotech Incoporation*

BACKGROUND-AIM

Persistent infection with high-risk human papillomavirus (HPV) is a major risk factor for cervical cancer. Timely detection and treatment of cervical precancerous lesions can reduce the occurrence of cancer. It aims to improve the efficiency of early diagnosis by establishing machine learning models to predict the risk of cervical lesions based on HPV genotyping and patient information.

METHODS

A retrospective analysis was conducted on HPV genotyping from 158,565 women, of which 19,707 had ThinPrep cytologic testing (TCT) and 7,539 had colposcopy and 4,762 had biopsy. A variety of machine learning algorithms were evaluated on the results of TCT, colposcopy and biopsy as target values to evaluate the importance of HPV genotypes and patient information, and to construct models to predict the risk of different grades of lesions.

RESULTS

(1) The overall prevalence of HPV infection was 17.89%, and the top types was 52 (4.44%), 58 (2.10%), 53 (1.96%), 81 (1.85%), 42 (1.75%), 16 (1.44%). (2) Lasso regression was used for feature evaluation, in which the patient age, HPV co-infection, and HPV16/18/33/39/52/58 had relatively higher weights. (3) The models of low-grade lesions had the performance with an area under the curve (AUC) of 0.665 in TCT, 0.652 in colposcopy, and 0.561 in biopsy. (4) The models of high-grade lesions had the performance with an AUC of 0.706 in TCT, 0.700 in colposcopy, and 0.663 in biopsy. (5) The model of malignant tumors had the performance with an AUC of 0.850 in biopsy.

CONCLUSIONS

The patient age, HPV co-infection, and HPV16/18/33/39/52/58 contributed significantly more to the risk of cervical lesions. Algorithms such as neural network and logistic regression had better performance in predicting the risk of various cervical lesions based on HPV genotyping and patient age. The models of high-grade lesions had higher AUC than that of the models of low-grade lesions. The AUC of logistic regression of malignant tumor was up to 0.85. It is recommended that patients with high risk in high-grade lesions or malignant tumors should undergo cervical biopsy in time to confirm the diagnosis and be treated by doctors, so as to reduce the risk of cervical cancer.

Artificial intelligence, clinical decision support and machine learning

P0210

PATIENT-BASED REAL-TIME QUALITY CONTROL INTEGRATING NEURAL NETWORKS AND JOINT PROBABILITY ANALYSIS

Y. Xia [2], W. Zheng [1], H. Xue [2], M. Feng [3], Q. Zhang [2], B. Li [2], X. Li [4], H. Qi [5], Y. Liu [5], T. Badrick [6], L. Zheng [4], L. Ji [3]

[1] *Shenzhen Mindray Bio-Medical Electronics Co, Ltd, Shenzhen, China*

[2]*Department of Laboratory Medicine, Peking University Shenzhen Hospital, Shenzhen, Guandong, China*

[3]*Department of Laboratory Medicine, Peking University Shenzhen Hospital, Shenzhen, Guandong, China#The First School of Clinical Medicine, Guangdong Medical University, Zhanjiang, Guandong, China*

[4]*Department of Laboratory Medicine, Southern Medical University Nanfang Hospital, Guangzhou, Guandong, China*

[5]*Shenzhen Mindray Bio-Medical Electronics Co, Ltd, Shenzhen, China*

[6]*The Royal College of Pathologists of Australasia*

BACKGROUND-AIM

Patient-based real-time quality control (PBRTQC) continuously monitors laboratory test quality using patient results but faces limitations, including high false alarm rates (FAR) and reduced sensitivity due to biological variations in diverse patient populations. These challenges restrict its ability to distinguish sample fluctuations from instrument errors, limiting real-life applications. This study introduced a neural network-based PBRTQC (NN-PBRTQC) system integrating neural networks (NN) and joint probability analysis. The model aims to enhance error detection by addressing patient demographic variability and reducing FAR, offering a more accurate and practical solution for laboratory quality control.

METHODS

Test data for various analytes were collected from Peking University Shenzhen Hospital and Nanfang Hospital Southern Medical University. A neural network model was trained to predict test results by incorporating patient demographics. Residuals between predicted and actual test results served as inputs for statistical process control algorithms to detect system errors. To reduce FAR, an intelligent alarm system based on joint probability analysis was implemented. The performance of NN-PBRTQC was evaluated using the number of patients required for error detection (NPeds) under different desired FAR (DFAR) levels, comparing it to traditional PBRTQC.

RESULTS

NN-PBRTQC significantly enhanced the clinical performance of PBRTQC. Data transformation using the neural network significantly reduced data fluctuations by removing the influence of patient demographics. At a DFAR of 0.1%, NN-PBRTQC required 63.9% fewer samples to detect errors compared to traditional PBRTQC. Even at a 100-fold lower DFAR (0.001%), NN-PBRTQC needed 30.2% fewer samples for error detection than traditional PBRTQC operating at a DFAR of 0.1%.

CONCLUSIONS

NN-PBRTQC offers a robust improvement to PBRTQC by addressing sample variability and minimizing false alarms. By lowering the false alarms and NPeds at equivalent DFAR levels, it enhances the effectiveness of PBRTQC, accelerating its implementation in clinical laboratories.

Artificial intelligence, clinical decision support and machine learning

P0211

IMPROVEMENT IN TIME TO FIRST MEDICAL ACTION RELATED TO MYOCARDIAL INJURY AND REDUCTION IN MORTALITY AFTER IMPLEMENTATION OF A TROPONIN RESULTS NOTIFICATION SYSTEM VIA SMS

C. Camarasa [2], M. Lopez-Garrigós [1], A. Blasco [1], R. Torreblanca [1], I. Gutierrez [1], R. Carbonell [1], E. Flores [1]
[1]*Clinical Laboratory Department, Hospital Universiari de Sant Joan, San Juan de Alicante*
[2]*Universidad Miguel Hernández*

BACKGROUND-AIM

Troponin measurement in the diagnosis of acute coronary syndrome (ACS) is crucial and highly sensitive for detecting myocardial injury.
The overcrowding of emergency departments (ED) is an increasingly significant problem. Reducing waiting times in ED represents an ongoing challenge. It is essential to implement strategies that efficiently manage patient flow in ED through more sophisticated triage tools or improved coordination between departments.
One contributing factor to prolonged stays is the inability to detect the availability of laboratory test results in real time.
The aim was to evaluate the time to the first medical action related to ACS treatment and 28-day mortality after implementing automatic and immediate notification of elevated troponin levels via mobile messaging using a Clinical Decision Support System (CDSS) (Abbott®) and compare this with a retrospective cohort.

METHODS

A quasi-experimental study was designed. Patients under 65 years old who presented to ED with an initial troponin result above 100 ng/μL were included. Patients with a prior diagnosis, from other institutions, transferred by ambulance or with a classical ACS presentation were excluded.
Through the CDSS, an intervention was implemented to send an SMS to the physicians responsible for patients meeting the inclusion criteria. The message contained demographic information, the patient's location in ED, and the troponin result.
The time to the first medical action related to ACS was calculated. The 28-day mortality incidence was also determined by reviewing medical records and compared to the pre-intervention period for patients with the same characteristics.

RESULTS

A total of 66 patients were included in both periods. The populations in both periods were homogeneous, with no differences in age or sex.
The time to the first medical action (median, P90) in the pre-intervention period was 178, 388 minutes, compared to 156, 285 minutes in the post-intervention period. The 28-day mortality rate in the pre-intervention period was 4.5%, compared to 1.5% in the post-intervention period.

CONCLUSIONS

The notification of troponin results via SMS improves the time to the first medical action related to myocardial injury and reduces 28-day mortality in these patients.

Artificial intelligence, clinical decision support and machine learning

P0212

MACHINE LEARNING MODELS ACCELERATE RISK STRATIFICATION FOR NON-ST-ELEVATION MYOCARDIAL INFARCTION

W.J. Luo [1], Y.M. Liou [3], C.H. Hsiao [1], C.S. Hung [5], H.Y. Pan [4], C.H. Huang [2], K.Y. Su [1], P.C. Yang [6]

[1]*Department of Clinical Laboratory Sciences and Medical Biotechnology, College of Medicine, National Taiwan University, Taipei, Taiwan*

[2]*Department of Emergency Medicine, National Taiwan University Hospital, Taipei, Taiwan*

[3]*Department of Laboratory Medicine, National Taiwan University Hospital, Taipei, Taiwan*

[4]*Division of Cardiology, Department of Internal Medicine, National Taiwan University Hospital Hsin-Chu Branch, Hsin-Chu, Taiwan*

[5]*Division of Cardiology, Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan*

[6]*Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan*

BACKGROUND-AIM

Diagnosing non-ST-elevation myocardial infarction (NSTEMI) often involves serial cardiac troponin (cTn) measurements, leading to diagnostic delays and overcrowded emergency departments. This study proposes a machine learning-based approach for early risk stratification and rapid triage to provide accurate and safe diagnoses within an hour for patients with suspected NSTEMI.

METHODS

This retrospective diagnostic study included 54,636 patients who underwent at least one cTn measurement in the emergency departments of two independent medical centers in Taiwan between May 1, 2016, and Dec 31, 2021. STEMI cases and patients lacking hematological or serial cTn measurements were excluded. Cases were defined as patients diagnosed with acute myocardial infarction by experienced physicians, while the remaining patients served as controls. Machine learning models incorporated demographic data and 23 NSTEMI-related routine laboratory tests from initial blood draws to compute risk scores. Diagnostic performance was assessed using the positive predictive value (PPV) and negative predictive value (NPV), alongside comparisons with the European Society of Cardiology (ESC) 0 h/1 h algorithm and cTn alone.

RESULTS

The machine learning model achieved a superior area under the receiver-operating-characteristic curve (0.921 [0.907–0.936], $p<0.0001$) compared to high-sensitivity cTn alone (0.776 [0.751–0.800]). Independent validation showed that risk scores <1.8 and ≥38.5 identified 48.3% of patients as low risk (NPV 98.8% [98.5–99.1%]) and 2.6% as high risk (PPV 78.1% [73.2–82.4%]). When concatenated with the ESC 0 h/1 h algorithm, the machine learning model stratified 68.4% of patients with PPV 83.3% (77.3–87.9%) and NPV 100% (99.5–100%) at the 0-hour mark, and stratified 85.3% of patients with a PPV of 84.9% (79.5–87.7%) and NPV of 100% (99.6–100%) at the 1-hour mark, outperforming the standalone ESC 0 h/1 h algorithm.

CONCLUSIONS

Incorporating machine learning with routine laboratory test results from the first blood draw enhances risk stratification for NSTEMI. This approach allows for early and precise rule-in of high-risk patients and safe rule-out of low-risk individuals, improving diagnostic efficiency when integrated into current diagnostic pathways.

Artificial intelligence, clinical decision support and machine learning

P0213

COMPUTATIONAL ANALYSIS OF WHOLE SLIDE IMAGES OF HISTOLOGICAL SECTIONS FROM ISCHEMIC STROKE THROMBI

V. Martínez [2], E. López [2], M. Negrotto [1], V. Escobar [1], R. Medeiros [2]
[1]*Facultad de Medicina, Universidad de la República, Hospital de Clínicas*
[2]*Facultad de Química, Universidad de la República, Hospital Maciel*

BACKGROUND-AIM

Ischemic stroke is a leading cause of morbidity and mortality worldwide. Technological advancements in ischemic stroke treatment, specifically mechanical thrombectomy, have enabled the retrieval of thrombus samples for study. High-resolution imaging of histological slides and subsequent analysis could provide additional insights into the etiology of the event, which remains largely unknown in a significant proportion of cases.

METHODS

Thrombi were extracted from patients undergoing thrombectomy, and their etiology was classified using the TOAST (Trial of Org 10172 in Acute Stroke Treatment) score. Following sectioning and hematoxylin and eosin staining, clot components (red blood cells [RBCs], fibrin-platelet aggregates [FPs], and white blood cells [WBCs]) were segmented in whole-slide images. For this, tools as Orbit Image Analysis, PyRadiomics and ImageScope were used. Histomic features were developed to capture the structural distribution of RBC/FP regions, encompassing radiomics, radial composition, and object features of RBC/FP. Textural features from nuclear and extranuclear regions of WBCs were computed to define classes, summarized by class frequency distributions.

RESULTS

We established a computational workflow for high-resolution histological image analysis, enabling characterization of RBC-FP content and WBC complexity in ischemic stroke thrombi of varying etiologies. RBC-FP features include FP 70th, first order statistic 10th and first order statistic mean absolute deviation. WBC features include several class frequency distributions. Some of these features showed differences between cardioembolic and atherothrombotic thrombi.

CONCLUSIONS

The proposed histomic workflow provides insights into thrombus composition and serves as a valuable tool for analyzing thrombi, aiding in etiological determination.

Artificial intelligence, clinical decision support and machine learning

P0214

VISION TRANSFORMER-BASED PREDICTIVE MODEL FOR THE AUTOMATIC IDENTIFICATION OF BOTRYOID NEUTROPHILS

A. Merino [1], K. Barrera [2], G. Verdú [1], P. San-José [3], M. Vallvé [3], A. Molina [1], J. Rodellar [2]

[1]*CORE Laboratory. Biochemistry and Molecular Genetics Department. Biomedical Diagnostic Center. Hospital Clinic. Barcelona. Spain*

[2]*Department of Mathematics. Technical University of Catalonia. Barcelona. Spain*

[3]*Hematology Department. Hospital de la Santa Creu i Sant Pau. Barcelona. Spain*

BACKGROUND-AIM

With the increasing frequency of heat waves due to climate change, heat-related illnesses are becoming more common. Heatstroke is a life-threatening condition characterized by hyperthermia and multiple organ failure. A rare morphological feature, botryoid nuclei showing hypersegmented nucleus organized around a network of interchromatin filament bridges, has been identified in the peripheral blood (PB) of patients with hyperthermia. Cobalamin and folate deficiencies, myelodysplastic syndromes and specific treatments are also associated with the appearance of hypersegmented nuclei in neutrophils. This work presents an automated classification model to distinguish between normal, botryoid and hypersegmented neutrophils.

METHODS

A total of 4626 images of normal neutrophils, 64 botryoid and 147 hypersegmented were collected from PB using the CellaVision DM96, DM9600, DI60 and Mindray MC80 automatic blood cell analyzers and May Grünwald-Giemsa staining. Botryoid neutrophils were acquired from six patients with hyperthermia. To address class imbalance, only 2.5% of the total number of images corresponding to normal neutrophils were considered for training, A lightweight vision transformer-based architecture was used for the model development. The basic concept is that input images are divided into 4x4-pixel patches and processed using 7x7 sliding windows, requiring low computational resources and time during execution.

RESULTS

The model was validated using a test set of 4689 images. Botryoid neutrophil detection showed accuracy of 96.9%, sensitivity of 96.88%, f1-score of 96.9%, Jaccard index of 93.94% and a Matthews correlation coefficient MCC of 96.85%. Hypersegmented neutrophils achieved accuracy values of 98.6%, sensitivity of 99.3%, f1 of 96.96%, Jaccard index of 91.8%, and MCC of 95.61%. Normal neutrophils achieved an accuracy of 99.73%, sensitivity of 99.73%, f1 of 100%, a Jaccard index of 99.67%, and MCC of 95.78%. Overall, the model achieved a global accuracy of 98.4%.

CONCLUSIONS

The classification system proposed in this work proved to be effective in differentiating between normal, hypersegmented and botryoid neutrophils. This model provides an accessible and accurate tool for laboratory specialists in heatstroke cases detection.

Artificial intelligence, clinical decision support and machine learning

P0215

FEASIBILITY OF A LARGE LANGUAGE MODEL-BASED VIRTUAL ASSISTANT FOR LABORATORY DECISION SUPPORT: A PILOT STUDY

A. Musayeva [1]
[1]*Gakh District Central Hospital (Azerbaijan)*

BACKGROUND-AIM

Large language models (LLMs) are increasingly explored as cost-effective tools for decision support in medical fields; however, practical data on their performance in routine laboratory scenarios remain scarce. This study aims to investigate the feasibility of using an off-the-shelf LLM as a virtual assistant for interpreting common laboratory test profiles and suggesting follow-up diagnostics in real-world settings.

METHODS

A pilot framework was designed to evaluate the LLM's interpretative accuracy and clinical relevance. A set of anonymized test result profiles, covering frequent clinical scenarios (e.g., suspected anemia, metabolic disturbances), was prepared. Each result profile—together with brief patient background—was converted into a textual prompt. The LLM's responses were structured to address three areas: (1) identification of key abnormalities, (2) possible differential diagnoses, and (3) recommended next tests or clinical steps. Two experienced laboratory specialists independently assessed these responses for both clinical appropriateness and clarity.

RESULTS

Preliminary observations revealed that the LLM could consistently identify major test abnormalities and propose plausible follow-up measures for a variety of routine cases. Early feedback from the laboratory specialists indicated that the system was particularly effective at flagging results that fell substantially outside normal ranges. However, in more nuanced or borderline scenarios, the LLM's suggestions sometimes required manual review or modification. Further data collection is ongoing, and detailed quantitative findings will be analyzed and presented upon completion of the study.

CONCLUSIONS

These initial findings suggest that an LLM-based virtual assistant may serve as a practical, low-resource decision support tool in laboratory medicine. Although human expertise remains crucial—especially for complex cases—this approach has the potential to streamline the interpretation of standard test profiles and expedite recommendations for follow-up testing. Future steps will include expanding the dataset, refining prompts, and examining the model's integration with laboratory information systems to enhance its reliability and utility.

Artificial intelligence, clinical decision support and machine learning

P0216

REAL-TIME DIAGNOSIS OF SEPSIS USING MACHINE LEARNING: A MINIMALISTIC APPROACH WITH LABORATORY VARIABLES

A. Rojo Gómez [3], Y. Abril [2], J.M. Salinas [2], R. Torreblanca [1], I. Gutiérrez [1], M. López-Garrigós [1], Á. Blasco [1], R. Carbonell [1], E. Flores Pardo [1]

[1]*Clinical Analysis Service, Hospital Universitario San Juan de Alicante*
[2]*IT and Telecommunications Service, Hospital Universitario San Juan de Alicante*
[3]*University of Alicante*

BACKGROUND-AIM

Sepsis requires immediate diagnosis and treatment to prevent severe complications. Studies have shown that each hour of delay in initiating appropriate treatment increases mortality rates by 7–8%. Current diagnostic methods often involve complex and time-consuming laboratory tests, delaying intervention. This study proposes a machine learning (ML) model capable of diagnosing sepsis in real time using a minimal set of laboratory variables to improve diagnostic speed and patient outcomes.

METHODS

A retrospective analysis was conducted using a dataset comprising 123,317 cases after preprocessing, including sepsis diagnoses and controls. The dataset contained 62,535 control cases and 60,782 sepsis cases. The model employs a stacking approach with three base classifiers—Light Gradient Boosting Machine (LightGBM), Extreme Gradient Boosting (XGBoost), and Random Forest (RF)—combined using a logistic regression metamodel. The selected variables, derived from the complete blood count (CBC), included age, sex, hemoglobin, red blood cells, white blood cells, mean corpuscular volume, platelets, and procalcitonin.

RESULTS

The stacking model achieved an area under the receiver operating characteristic curve (AUC) of 0.8439 (95% confidence interval [CI]: 0.8390–0.8481), demonstrating excellent discriminatory power. It showed strong overall performance, with a precision of 0.7573, indicating that 75.73% of predicted sepsis cases were correct, and a recall of 0.7834, reflecting its ability to detect 78.34% of actual sepsis cases. The F1-score of 0.7701 highlights the balance between precision and recall, and the accuracy was 76.21%.
Among individual classifiers, RF achieved an AUC of 0.8176 (95% CI: 0.8124–0.8225), XGBoost obtained 0.8328 (95% CI: 0.8278–0.8373), and LightGBM performed best with 0.8402 (95% CI: 0.8354–0.8447). The stacking model combined the strengths of these classifiers, achieving superior performance overall.

CONCLUSIONS

The proposed machine learning model offers a precise and rapid method for diagnosing sepsis in real time using a minimal set of laboratory variables. By reducing reliance on extensive tests, this approach enhances diagnostic efficiency and supports timely interventions that can improve patient outcomes.

Artificial intelligence, clinical decision support and machine learning

P0217

DEVELOPMENT AND PERFORMANCE EVALUATION OF CLINICAL DECISION SUPPORT SYSTEM(CDSS) ALGORITHMS FOR THE CLINICAL BIOCHEMISTRY LABORATORY

Z. Deniz [2], S. Uysal [1]
[1]*Department of Medical Biochemistry, Balıkesir University Medical School, Balıkesir, Türkiye*
[2]*Medical Biochemistry Laboratory, Yalova State Hospital, Yalova, Türkiye*

BACKGROUND-AIM

Manual validation of test results in clinical chemistry laboratories is time-consuming, labor-intensive, and subject to variability among laboratory personnel. The aim of this study was to establish a set of autoverification (AV) rules for complete blood count, clinical chemistry, and immunochemistry analyses and to evaluate the performance of these rules.

METHODS

Algorithms were developed for complete blood count (DXH 800), clinical chemistry (AU 680), and immunochemistry (DXI 600/800) panels (Beckman Coulter Inc., CA, USA). During the implementation process, the REMISOL Advance middleware program (Beckman Coulter Inc., USA) was utilized. Initially, pre-analytical and analytical process rules were established. For post-analytical processes, reference change values (RCV) were used to determine delta check thresholds. Algorithms for glucose and creatinine were adjusted based on sigma-metric performance, and for these two tests, the AV range was extended to the assay's measurement range when sigma values exceeded 5. The AV behavior of the system was compared with that of expert users. Following the activation of the clinical decision support system (CDSS), changes in turnaround time (TAT) and AV rates were evaluated.

RESULTS

With the established algorithms, test-specific AV rates of 79.5% for clinical chemistry and 74.9% for immunochemistry were achieved, while sample-specific AV rates were 24.5% and 57%, respectively. For complete blood count subparameters, sample-specific validation was applied, limiting the AV rate to 17.3%. When the AV behavior of the system was compared with two expert users, no autoverified results were observed among those not directly approved by the experts. After full activation of the system, despite an approximately 10% increase in average laboratory test volume, a 10% reduction in TAT was observed.

CONCLUSIONS

The contribution of CDSS to laboratory workflows is undeniable. Integrating CDSS with sigma-metric data and moving average elements is considered essential for reliably autoverifying a greater proportion of results.

Artificial intelligence, clinical decision support and machine learning

P0218

ASSESSING GENERATIVE AIS PERFORMANCE IN 2023 SPAIN'S PHARMACY RESIDENCY EXAM (FIR): ARE HUMAN EXPERTS STILL ON TOP?

G. Verdú [1], A. Calvera [1], A.B. Fabregat-Bolufer [2]

[1]*Biochemistry and Molecular Genetics Department, Biomedical Diagnostic Center (CDB), Hospital Clínic, Barcelona, Spain.*
[2]*CORE Laboratory, Biochemistry and Molecular Genetics Department, Biomedical Diagnostic Center (CDB), Hospital Clínic, Barcelona, Spain*

BACKGROUND-AIM

Generative artificial intelligence (AGIs) and large language models (LLMs) are transforming medical education by simulating test scenarios, providing instant feedback, and benchmarking question difficulty. AI tools like ChatGPT have already shown strong performance in high-stakes licensing exams (e.g., Japanese Medical Licensing Exam, Korean Pharmacist Licensing Exam). This study evaluates the accuracy and reproducibility of three ChatGPT versions in Spain's 2023 FIR (Pharmaceutical Internal Resident) exam, assessing their potential for both academic preparation and clinical decision support in clinical laboratories.

METHODS

Three ChatGPT versions (4o mini, 4o, and o1 preview) were each prompted with all 200 official 2023 FIR questions, exactly as in the original exam, and tested in duplicate. We assessed accuracy (overall score and predicted rank) and reproducibility (consistency of correct responses), comparing AI performance to human candidates.

RESULTS

All tested models achieved scores sufficient for FIR residency positions. ChatGPT 4o mini attained 80.5% accuracy (14th place), struggling especially with complex, image-based, or clinical-case questions. ChatGPT 4 improved to 87% (3rd place), reducing errors in analytical techniques and pharmaceutical chemistry. GPT o1 preview performed best, reaching 92.46% accuracy (1st place), although it left one image-based question unanswered. In terms of reproducibility for initially correct answers, ChatGPT 4o and o1 each displayed 96.67% identical repeats, slightly exceeding 4o mini's 90%.

CONCLUSIONS

Rising model complexity and computational power correlate with improved FIR exam performance among these AI systems. Image-based questions and intricate clinical cases remain more challenging, though GPT o1 demonstrated the highest overall accuracy and acknowledged limitations in visual data interpretation. These findings underscore AGI's expanding role in reshaping healthcare professional training and assessment.

Artificial intelligence, clinical decision support and machine learning

P0219

OPTIMIZING COLONY DETECTION WITH DEEP LEARNING: ENHANCED IMAGE ACQUISITION AND INTEGRATED GUI DEVELOPMENT

J. Du [1], M. Sun [1], Z. Li [1], Q. Sun [1]
[1]*The First Bethune hospital of Jilin University*

BACKGROUND-AIM

Background: With the continuous emergence of novel pathogens, rendering traditional methods reliant on manual operation inadequate to meet the demands of modern standards. Deep learning techniques, with their capability for in-depth analysis of large-scale data and autonomous feature learning, have demonstrated significant potential in the field of automated colony detection. To further expand the application of deep learning, it is crucial to enhance model performance and reduce the barriers to technical implementation.

METHODS

Methods: This study optimized the image acquisition process and developed an integrated graphical user interface system. Seven common clinical pathogenic bacteria were selected, with 30 high-resolution images captured for each species using the MicroChronos™ device. The device, equipped with a high-performance camera, captured images at a vertical upward angle of the inverted plate in a sealed environment, ensuring uniform illumination. Exposure time was adjusted based on the optical properties of the culture dish background to maintain stable image brightness and quality. Image augmentation expanded the dataset, followed by training the colony detection model with YOLOv8. A GUI system was developed using PySide6 and OpenCV.

RESULTS

Result: A total of 210 high-resolution images were collected, covering seven clinically relevant pathogenic bacteria, with 7,274 colony targets annotated. After image enhancement, the number of colonies increased to 17,788. The YOLOv8 training results indicate that the model achieves an overall precision of 97.7%. The developed graphical user interface system enables real-time detection, classification, and counting of colony images.

CONCLUSIONS

Conclusion: The improved image acquisition technique and precise control of exposure time in this study significantly enhanced the quality of the dataset. In addition, the development of the integrated graphical user interface has lowered the application barrier of deep learning techniques and streamlined the automated colony detection process.

Artificial intelligence, clinical decision support and machine learning

P0220

A DIAGNOSTIC MODEL FOR PRETEST PROBABILITY OF LUNG CANCER IN PATIENTS WITH SOLITARY PULMONARY NODULES: INTERPRETATION BASED ON SHAP-ENHANCED XGBOOST MACHINE LEARNING MODEL

q. Zhang [1], m. Feng [2], h. Zhang [1], y. Zhang [1], s. Jia [1], z. Wang [2], y. Ma [1]
[1]*Shaanxi University of Chinese Medicine*
[2]*The Affiliated Hospital of Shaanxi University of Chinese Medicine*

BACKGROUND-AIM

Isolated pulmonary nodules (SPN) are common incidental findings on CT scans, with prevalence ranging from 8% to 51% in lung cancer screening trials. The estimation of the clinical probability of malignancy in patients with SPN can facilitate subsequent diagnosis and improve the 5-year survival rate. Therefore, early screening plays a pivotal role in improving treatment outcomes and enhancing survival rates in lung cancer.

METHODS

we identified miRNA-941 and miRNA-574-3p as differentially expressed between patients with benign and malignant lung nodules through bioinformatics analysis. Subsequently, we collected clinical data and laboratory test indicators from 82 patients with malignant lung nodules and 85 patients with benign lung nodules. Primers were designed to detect the expression levels of miRNA-941 and miRNA-574-3p in these 167 patients. In this study, logistic regression was employed for feature selection, while the XGBoost algorithm was used to construct the machine learning model. The model's performance was comprehensively evaluated through receiver operating characteristic (ROC) curve analysis, calibration curve analysis, and clinical decision curve analysis, alongside key metrics including sensitivity, specificity, accuracy, and F1-score.

RESULTS

The XGBoost model demonstrated superior performance, achieving a higher area under the AUC in predicting the occurrence of lung cancer (training set: 0.989, 95% CI: 0.975–0.996; test set: 0.857, 95% CI: 0.757–0.943).

CONCLUSIONS

Our constructed XGBoost machine learning model exhibited high performance in predicting lung cancer. Furthermore, the SHAP (SHapley Additive exPlanations) method was utilized to provide interpretability for the model's predictions.The results indicated that the most significant predictive features, ranked by importance, were: 'nodule diameter,' 'miRNA-941,' 'CEA,' 'miRNA-574-3p,' and 'presence of spiculated signs'.

Artificial intelligence, clinical decision support and machine learning

P0221

A MACHINE LEARNING ALGORITHM BASED ON A 15-AUTOANTIBODY PROFILE BY A NOVEL FULLY AUTOMATED MULTIPLEXED MICROARRAY IMMUNOASSAY FOR THE DIAGNOSIS OF AUTOIMMUNE CONNECTIVE TISSUE DISEASES

G. Gomez [3], Y. Cheng [2], K. Nita [2], M. Hausmann [1], C. Fischer [3], Y. Ataman-Önal [1]
[1]*Research & Development, AliveDx, Switzerland.*
[2]*Research & Development, AliveDx, UK.*
[3]*Scientific & Medical Affairs, AliveDx, Switzerland.*

BACKGROUND-AIM

Detection of relevant autoantibodies is key in the identification of autoimmune connective tissue diseases (CTD). The evaluation of multiple autoantibodies for a more comprehensive serological profiling may improve the diagnosis of these conditions. We evaluated the diagnostic utility, in patients with CTD and controls, of machine learning classifiers based on the 15-autoantibody profile performed by a novel, single-use, multiplexed microarray immunoassay, used with its fully automated high-throughput proprietary system for the detection of IgG autoantibodies directed to dsDNA, SS-A 60, TRIM21 (SS-A 52), SS-B, Sm, Sm/RNP, U1RNP, Jo-1, Scl-70, Centromere B, Chromatin, Ribosomal P, DFS70, RNAP III and CCP2.

METHODS

Sera from 475 patients diagnosed with CTD in accordance with current guidelines [127 patients with systemic lupus erythematosus (SLE), 74 with systemic sclerosis, 76 with Sjögren's syndrome (SjS), 71 with idiopathic inflammatory myopathies, 54 with mixed CTD, 73 with rheumatoid arthritis] and 652 patients with other disorders, who served as disease controls were analyzed using the investigational MosaiQ AiPlex® CTDplus (AliveDx, CH) assay. Classification models were developed using all 15 autoantibodies or a selected subset, employing the RandomForest algorithm. Diagnostic performance was evaluated by receiver operating characteristic curve analysis.

RESULTS

A RandomForest classifier incorporating all 15 autoantibodies demonstrated robust performance in predicting SLE, achieving an area under the curve (AUC) of 0.92. In comparison, the individual SLE-specific markers dsDNA and Sm yielded lower AUCs of 0.68 and 0.60, respectively. For SjS, the 15-plex RandomForest classifier achieved an AUC of 0.83, outperforming a 3-plex RandomForest classifier based on SS-A 60, TRIM21, and SS-B autoantibodies, which had an AUC of 0.62. The individual AUCs for these markers were 0.63, 0.59, and 0.58, respectively. Similarly, for other CTDs, 15-plex classifiers consistently outperformed the individual disease-specific markers.

CONCLUSIONS

Multiplex autoantibody testing combined with machine learning algorithms has the potential to improve the diagnosis of autoimmune CTD.

Artificial intelligence, clinical decision support and machine learning

P0222

ROLE OF PANCREATIC K-ATP CHANNEL GENETIC POLYMORPHISM IN DIAGNOSIS OF T1DM AND SULPHONYLUREA TREATMENT ELIGIBILITY.

P. Chauhan [1], P. Jana [1], S. Ahamad [3], D. Gupta [3], A. Seth [2], R. Singh [1]
[1]*Dept of Biochemistry, Lady Hardinge Medical College & Associated Hospitals, University of Delhi, New Delhi*
[2]*Dept of Pediatrics, Lady Hardinge Medical College & Associated Hospitals, University of Delhi, New Delhi*
[3]*Translational Bioinformatics Group, International Center for Genetic Engineering & Biotechnology, New Delhi*

BACKGROUND-AIM

Type 1 Diabetes Mellitus (T1DM) is a multifactorial disorder with a strong genetic component. The 10th International Diabetes Federation Atlas 2021 states that 2,294,000 Indians below 19 years had T1DM. Disruption of potassium-sensitive adenosine tri-phosphate sensitive channel potential and consequent dysregulated insulin secretion, are observed due to mutations in two genes on chromosome 11p15.1, the potassium inwardly-rectifying channel subfamily J member 11(KCNJ11) and ATP binding cassette subfamily C member 8 (ABCC8) genes. Responsible for encoding channel subunits potassium inward rectifier (Kir6.2) and sulfonylurea receptor (SUR1) respectively, studies have linked KCNJ11-rs5219 and ABCC8-rs1799854 single nucleotide polymorphisms (SNPs) with pathological hyperglycemia. Our study examined the association of KCNJ11-rs5219 and ABCC8-rs1799854 with Kir6.2/SUR1 expression in T1DM.

METHODS

This cross-sectional study enrolled 100 T1DM pediatric cases and 100 age-sex-matched healthy controls, excluding known cases of other autoimmune disorders. EDTA blood samples were collected for DNA extraction, PCR amplification, RFLP, and detection by agarose gel electrophoresis. Kir 6.2 and SUR1 levels were estimated by sandwich ELISA. Bioinformatic tools were used in molecular dynamic simulation (MDS) for Kir6.2/SUR1-Sulphonylurea interaction.

RESULTS

KCNJ11-rs5219 and ABCC8-rs1799854 were significantly associated with T1DM as polymorphic alleles of both genes showed significantly high frequency among cases compared to controls ($p<0.001$). Kir 6.2 and SUR1 showed significantly higher expression in cases than controls and were also significantly associated with the dominant models of their respective encoding polymorphic genes. Diagnostic accuracy of Kir6.2 (75.5%) and SUR1 (69%) was estimated by ROC curve analysis. Three machine learning modules (KNN, Naïve Bayes, and Clustered SVM) were devised with KCNJ11-rs5219 having an importance score of 19.41, which diagnosed T1DM with 94.7% accuracy. MDS showed significant interaction between altered Kir6.2/SUR1 with Sulphonylurea.

CONCLUSIONS

Early detection of KCNJ11-rs5219 and ABCC8-rs1799854 coupled with Kir 6.2/SUR1 expression analysis can serve as valuable diagnostic markers of T1DM and help indicate eligibility for Sulphonylurea treatment.

Artificial intelligence, clinical decision support and machine learning

P0223

ENSEMBLE LEARNING APPROACH TO CLASSIFICATION OF HYPERLIPIDAEMIA IN LIPOPROTEIN ELECTROPHORESIS GELS

S.L. Kakako [1], R. Punchoo [4], I. Fabris-Rotelli [3], T.S. Pillay [2]
[1]*Department of Chemical Pathology, University of Pretoria, South Africa*
[2]*Department of Chemical Pathology, University of Pretoria/National Health Laboratory Service (NHLS), South Africa*
[3]*Department of Statistics, University of Pretoria, South Africa*
[4]*Division of Chemical Pathology, University of Cape Town/National Health Laboratory Service (NHLS), South Africa*

BACKGROUND-AIM

Hyperlipidaemia significantly increases the risk of atherosclerotic cardiovascular disease (ASCVD), a leading cause of death globally. One method to assess hyperlipidaemia is lipoprotein electrophoresis and classification is important clinically and is based on Fredrickson's phenotypes. Lipid electrophoresis produces a gel interpreted by experts by visual inspection that is subjective, rigorous, and not standardized. The need for accurate and reproducible determination of lipoprotein phenotypes is key to reducing the burden of hyperlipidaemia, which is the aim of this work. In this study, ensemble classifiers of machine learning models were presented to interpret lipoprotein gels.

METHODS

Using 836 annotated lipoprotein gels, feature extraction was carried out. Data from the extraction and the corresponding lipid profile were presented to the ensemble machine-learning models of random forest (RF) and extra gradient boost (XGBoost). With Visual Studio Code in Python, the models were trained and tested with 80% and 20% of the data, respectively. The model's evaluation was on type Ila, IIb, and IV owing to their higher support in the sample distribution.

RESULTS

The XGBoost performance was superior with an accuracy of 84% compared to 79% of RF. The XGBoost achieved the following precision, recall, F1-score, and AUROC: for type Ila (0.94, 0.93, 0.93, and 0.94), for type IIb (0.84, 0.92, 0.93, and 0.93), and type IV (0.78, 0.76, 0.77, and 0.84) respectively. Conversely, RF precision, recall, F1-score, and AUROC for type Ila (0.89, 0.96, 0.92, and 0.94), type IIb (0.62, 0.87, 0.74, and 0.86), and type IV (0.82, 0.72, 0.74, and 0.86) respectively.

CONCLUSIONS

The XGBoost model showed better performance and would run with less variability compared to the visual inspection of the human expert currently used for interpretation. This study highlights that the model can be used for gel interpretation without human intervention.

Artificial intelligence, clinical decision support and machine learning

P0224

USE OF A MACHINE LEARNING MODEL FOR THE DIAGNOSIS OF DROWNING BY STRONTIUM AND RUBIDIUM ASSAYS IN BONE MARROW

B. Lelievre [3], S. Malbranque [2], D. Zabet [2], F. Genevieve [1], N. Jousset [2], J. Gaulier [4]
[1]*Hematology laboratory, University hospital, Angers*
[2]*Legal medicine, University hospital, Angers*
[3]*Pharmacology-toxicology laboratory, University hospital, Angers*
[4]*Toxicology laboratory, University hospital, Angers*

BACKGROUND-AIM

Machine learning has many applications, including diagnostic classifications. Drowning is defined on the basis of contextual data, autopsy observations and biomarkers (diatoms, etc.). However, in cases of freshwater drowning, conclusions can be difficult to draw, especially in the case of putrefied corpses. We aimed to evaluate biological biomarkers in bone marrow samples and machine learning models for drowning classification.

METHODS

Our study focused on 68 cases of cadaver discovery in the Anjou region and 24 alive controls. Strontium (Sr) and rubidium (Rb) concentrations were determined using ICP-MS (7800, Agilent). Database included anthropometric data, Sr and Rb concentrations in bone marrow, Sr/Rb ratio, degree of putrefaction (0 to 4, assessed by two coroners) and presumed immersion delay. Four groups were defined: wet drowning, false drowning, alive control, deceased negative control. Groups were assigned in advance on the basis of standard criteria. After statistical analyses and a principal component analysis (PCA), an unsupervised (Kmeans for clustering) and five supervised machine learning models were tested. An 80/20 split was applied to divide data in into training and validation (internal and external) groups. Models were optimized and evaluated using standard metrics and confusion matrix. A further analysis was performed by removing data with a degree of putrefaction greater than 3. Finally, an oversampling (to limit the imbalance in sample sizes) and the testing of the best-performing models were carried out and evaluated.

RESULTS

Statistical analyses highlighted that putrefaction influenced Sr concentration in bone marrow, suggesting difficulties in interpreting concentrations at putrefaction levels above 3. PCA showed that the Sr/Rb ratio was the “most discriminating” variable for group assignment. The best-performing models were AdaBoost and Random Forest (accuracy: 89%, sensibility: 100% for both), but they were less satisfactory for the false drowning group (the smallest group). The elimination of samples with high degree of putrefaction improved the performances (accuracy: 100%) of models.

CONCLUSIONS

Oversampling produced better results (accuracy:95% for Random Forest), and a larger-scale evaluation is planned to confirm these initial results.

Artificial intelligence, clinical decision support and machine learning

P0225

DEVELOPMENT OF A NOVEL MACHINE LEARNING-BASED SCREENING MODEL FOR INFECTIOUS MONONUCLEOSIS IN FEBRILE PEDIATRIC PATIENTS USING HEMATOLOGY ANALYZER SCATTERGRAMS

Q. Cai [2], B. Ye [1], W. Zheng [1], S. Zhang [5], J. Zhang [1], Y. Shen [3], D. Yao [1], H. Zhang [4], Z. Huang [2], J. Hu [1], Y. Liu [1], Y. Ma [1], J. Wang [2], Y. Wang [3]

[1]*Clinical Department (IVD), Shenzhen Mindray Bio-Medical Electronics Co., Ltd, Shenzhen 518057, China;*
[2]*Department of Clinical Laboratory, Ruijin Hospital, Shanghai Jiaotong University School of Medicine, Shanghai 200025, China;*
[3]*Department of Clinical Laboratory, The First Affiliated Hospital of Soochow University, Suzhou 215006, China;*
[4]*Department of Clinical Laboratory, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450008, China*
[5]*Department of Clinical Laboratory, The First Affiliated Hospital Sun Yat-sen University, Guangzhou 510062, China;*

BACKGROUND-AIM

Infectious mononucleosis (IM), primarily caused by Epstein-Barr virus (EBV), is an acute disease with diverse clinical presentations. Due to the lack of specific symptoms, IM is often misdiagnosed in febrile children, leading to inappropriate antibiotic use or chronic EBV infection. This study aims to develop a novel screening model based on hematology analyzer scattergrams to improve IM diagnosis in febrile pediatric patients.

METHODS

This retrospective study included 5511 patients (518 with EBV-induced IM) under 14 years old with fever, visiting the Children's Hospital Zhejiang University School of Medicine from January to December 2023. All subjects underwent complete blood count (CBC) using a Mindray BC-7500 CRP hematology analyzer and other tests, with hematology scattergrams and 41 parameters analyzed. Scattergrams of IM patients were analyzed using probability density distribution to select relevant features. Recursive feature elimination (RFE) was used for feature selection, and four machine learning algorithms were evaluated. The random forest classifier (RFC) model was selected as the best-performing model. The models were evaluated using AUC, sensitivity, specificity, confusion matrices, and SHapley Additive exPlanations (SHAP).

RESULTS

The Novel IM screening model (Novel IMs) achieved optimal performance with 11 key features, including scatter plot parameters like D-Lym-Peak-Valley-R and D-Lym-SFL-CV. The model showed accuracy ≥91.92%, specificity ≥91.8%, and sensitivity ≥89.71% across training, validation, and test datasets. In the test set, sensitivity reached 93.33%, effectively distinguishing IM patients. The Novel IMs model outperformed traditional clinical parameters (e.g., Lym% and NLR) with an AUC of 0.96 and fewer false positives/negatives. SHAP analysis revealed the important contribution of novel scattergram features. Adjusting the model threshold improved specificity (95.99%) and accuracy (95.19%) for acute IM, while reducing sensitivity for nonacute cases, highlighting its adaptability for clinical priorities.

CONCLUSIONS

This is the first study to develop a machine learning model for screening IM in febrile children using parameters from blood cell scattergrams. This low-cost, convenient technique could enhance IM screening, improving diagnosis and reducing misdiagnosis.

Artificial intelligence, clinical decision support and machine learning

P0226

DEVELOPMENT AND VALIDATION OF THE SMART-ROMA MACHINE LEARNING MODEL FOR OVARIAN CANCER DIAGNOSIS: A MULTICENTER STUDY

W. Luo [2], Y. Xiong [1], Y. Lin [2]
[1]*Department of Laboratory Medicine, Nanfang Hospital, Southern Medical University, Guangzhou, China*
[2]*Mindray Medical International Ltd, Shenzhen,Guangdong*

BACKGROUND-AIM

Ovarian cancer, the most lethal gynecological malignancy, lacks reliable biomarkers for early diagnosis, and the diagnostic efficacy of single laboratory markers is limited. The Risk of Ovarian Malignancy Algorithm (ROMA), a widely used tool for ovarian cancer risk assessment, calculates risk in pre- and post-menopausal women based on CA125 and HE4 levels. However, ROMA has limitations, including low sensitivity for early-stage diagnosis and challenges in determining menopausal status, which requires considering menstrual history, age, symptoms, and laboratory tests. This restricts its clinical applicability. This study aims to develop a machine learning model, Smart-ROMA, to provide a more effective, accessible, and user-friendly tool for ovarian cancer diagnosis.

METHODS

In this multicenter cohort study, clinical and laboratory data were collected from women with and without ovarian cancer admitted to three hospitals in China between January 1, 2017, and April 30, 2023. Three key clinical features (HE4, CA125, and age) were selected based on their relevance to ovarian cancer diagnosis and cost-effectiveness for the model. The diagnostic model, Smart-ROMA, was developed using a Linear Support Vector Classification (LinearSVC) machine learning algorithm. The model was trained with 382 patients and evaluated using an internal validation set of 743 individuals, along with two external validation sets totaling 1,274 individuals.

RESULTS

Smart-ROMA achieved an AUC of 0.924 (95% CI 0.902 - 0.946) in the internal validation set, and AUCs of 0.973 (95% CI 0.954 - 0.992) and 0.960 (95% CI 0.940 - 0.980) in the two external validation sets. These results show that Smart-ROMA provides diagnostic performance comparable to the ROMA model (AUC = 0.919), using only three key indicators. By replacing menopausal status with age, the model enhances usability and accessibility in clinical practice.

CONCLUSIONS

Smart-ROMA demonstrated reliable performance in diagnosing ovarian cancer, highlighting its potential as a promising tool for improving ovarian cancer diagnosis and management. This updated model is particularly valuable in settings where determining menopausal status is difficult, offering a more convenient, cost-effective, and efficient diagnostic solution.

Artificial intelligence, clinical decision support and machine learning

P0227

ARTIFICIAL NEURAL NETWORK IN THE DIAGNOSIS OF URINARY TRACT INFECTIONS.

E. Nicolai [3], C. Ialongo [1], M. Ciotti [4], A. Giovannelli [2], F. Tomassetti [2], M. Pelagalli [2], M. Pieri [2], S. Bernardini [2]
[1]*Department of Experimental Medicine, Policlinico Umberto I, 'Sapienza' University, Rome.*
[2]*Department of Experimental Medicine, University of Rome Tor Vergata, Rome.*
[3]*Departmental Faculty of Medicine, UniCamillus-Saint Camillus International University of Health and Medical Sciences, Rome*
[4]*Laboratory of Microbiology and Virology, Policlinico Tor Vergata University Hospital, Rome.*

BACKGROUND-AIM

Urine microbial analysis is a frequently requested test, often associated with contamination during collection or storage, leading to false positive diagnosis and reporting delay.
In the era of digitalization, laboratory tests could be improved by implementing screening algorithms directly within laboratory middleware or even in the instruments themselves.
This study aims to evaluate the feasibility of using automated urinalysis counts to screen samples for which culture testing is requested. It leverages the capabilities of a simple artificial neural network (ANN) to pre-identify negative and contaminated (false positive) samples.

METHODS

The dataset for training the ANN includes 8,181 individual samples with available results from urinary cytology, dipstick tests, and culture testing. The dataset was randomly split 65%/25% for training and applying a multilayer perceptron (MLP). The MLP was refined by excluding inputs with normalized importance <0.2 to create a final model as simple as possible.

RESULTS

Physical and chemical parameters were not included in the final model. Among cytological parameters, only microbial count and leukocyte count were retained, while epithelial cells and red blood cells were excluded. The PPV and NPV for yeast presence were 0% and 100%, respectively; PPV and NPV for bacteria were 88.5% and 97.4%, respectively.
The model's performance demonstrates that the MLP can be used as a reliable screening test for negative samples, with NPV values of 97.4% and 100% for bacteria and yeast, respectively. Conversely, the MLP is completely unreliable for predicting positivity for both bacteria and yeast.
The MLP also fails to correctly identify contaminated samples, classifying them as either positive or negative for both bacteria and yeast. This can be explained by the fact that contamination is characterized not only by count but also by species present. Since urinalysis does not discriminate between species, this information is unavailable for ANN training, highlighting the insufficiency of count information alone in suspension.

CONCLUSIONS

In conclusion, the ANN applications can serve as an effective system for screening negative samples. Its NPV could be further improved if morphological information were available for training the MLP, reducing the risk of false negatives.

Artificial intelligence, clinical decision support and machine learning

P0228

COMPREHENSIVE BLOOD GLUCOSE LEVEL PREDICTION FROM HBA1C LEVELS USING MACHINE LEARNING MODELS ACROSS THE BIOLOGICAL RANGE

T.M. Okyay [2], I. Yilmaz [1], M. Koldas [2]
[1]*Health Science University Istanbul Haseki Training and Research Hospital*
[2]*UNIVERSITY OF HEALTH SCIENCES, ISTANBUL*

BACKGROUND-AIM

Accurately predicting blood glucose levels from HbA1c measurements has significant implications for personalized health monitoring and diabetes management. While previous studies have explored this relationship, existing models often fail to account for data imbalances and hyperglycemic outliers, limiting their predictive accuracy. This study introduces a novel machine-learning-based framework to address these challenges and improve blood glucose predictions.

METHODS

This study investigates a dataset of 197,180 patient samples, focusing on key features such as age, glucose, and HbA1c levels. The performance of 42 machine learning models was evaluated, with various transformations applied through Kernel Density Estimates (KDE) analysis. Targeted oversampling was implemented to enhance model robustness across these features.

RESULTS

Results show that the MLP Regressor achieved a moderate R^2 with the raw data. Logarithmic transformation effectively reduced RMSE. Gaussian KDE identified a low-density region around 550 mg/dL in hyperglycemia, prompting targeted oversampling. This, combined with log transformation, resulted in an R^2 of 0.93 and the lowest RMSE with the LGBM model, indicating strong predictive robustness for blood glucose levels.

CONCLUSIONS

In summary, the combination of logarithmic transformation, oversampling, and the LGBM Regressor presents a promising approach for accurately predicting blood glucose levels from HbA1c levels. The proposed approach offers a significant advancement over traditional methods by enhancing predictive accuracy, particularly for hyperglycemic outliers, and sets a new benchmark for leveraging HbA1c measurements in both healthy and diabetic management.

Artificial intelligence, clinical decision support and machine learning

P0229

DEVELOPING A MACHINE LEARNING MODEL FOR PREDICTING AUTOIMMUNE DISEASES USING HEMATOLOGICAL PARAMETERS

W. Liu [2], M. Zhang [2], L. He [1], L. Qu [2]
[1] *The First Hospital of Jilin University*
[2]*Department of Laboratory Medicine, the First Hospital of Jilin University*

BACKGROUND-AIM

To explore a screening model for autoimmune diseases based on hematological parameters results using machine learning tools in order to improve the accuracy of early diagnosis and provide convenient auxiliary diagnostic tools for clinical practice.

METHODS

Patients with AID positive (275 cases) and AID negative (116 cases) admitted to the First Hospital of Jilin University from January to March 2023 were collected and split into derivation cohort (275cases for model building and internal validation) and validation cohort (116 cases for external validation). Basic clinical information was queried from the hospital electronic medical record system, including 49 laboratory indicators, such as patient information, coagulation routine, blood routine, autoantibody test, clinical diagnosis and audit results. Nine machine learning (ML) algorithms were used to build the predicting model. The maximum area under the curve (AUC), sensitivity, specificity, and F1 score were calculated to evaluate the optimal model and the combination of feature parameters. The Shapley Additive exPlanation method was used to rank the importance of features and to explain the final model.

RESULTS

Among 9 ML models, the AID screening model from GBDT with 8 parameters LA screening Ratio, CRP and C3 showed the highest AUC both in derivation cohort and in validation cohort, 0.957 and 0.793, respectively, and has been transformed into a convenient tool to facilitate its application in clinical settings.

CONCLUSIONS

In this study, we developed a screening model for autoimmune diseases based on hematology parameters using a GBDT algorithm and explored the evaluation and selection of the model by integrating laboratory data. This model helps clinicians to identify high-risk patients in early screening, improve diagnostic efficiency, and optimize medical resource allocation.

Artificial intelligence, clinical decision support and machine learning

P0230

BUILDING A MACHINE LEARNING BASED PREDICTIVE MODEL FOR ACUTE KIDNEY INJURY USING URINE ANALYSIS RESULTS

C. Zhao [1], L. Qu [1], X. Zhao [2], W. Xu* [1]
[1]*Department of Laboratory Medicine, the First Hospital of Jilin University*
[2]*the First Hospital of Jilin University*

BACKGROUND-AIM

To build a predictive model for acute kidney injury (AKI) in hospitalized patients based on machine learning (ML) tools using urine analysis results.

METHODS

A total of 525 hospitalized patients were recruited from the First Hospital of Jilin University from August 2023 to June 2024. The basic information of the patients and the results of urine dry chemistry and urine formed fraction were obtained through the clinical information center of the hospital. The diagnostic criteria of AKI were used to categorize the patients into two groups: the AKI group, which included 94 cases, and the non-AKI group, which comprised 431 cases. The laboratory data of the 525 patients were randomly divided into training and validation sets. The ratio of the training and validation sets was set at 7:3, with AKI designated as the outcome variable, basic information of the patients and 23 laboratory tests as the characterization parameter.Logistic regression (Logistic), Adaptive Boosting (Ada Boosting), Gaussian NB (GNB), Support Vector Machine (SVM), and k-nearest neighbor classification (KNN) five machine learning (ML) models were implemented. The Receiver Operating Characteristic (ROC) curve was plotted, and the area under the curve (AUC) of the five models was compared to select the best model. The contribution of the 25 feature parameters in the best model was ranked using SHAP value analysis.The final model and the parameters in the model were selected and determined in decreasing order.

RESULTS

Among the five ML models, the logistic model integrating seven feature parameters had the highest AUC of 0.929, an accuracy of 0.862, a sensitivity of 0.83, a specificity of 0.871, an NPV of 0.609, a PPV of 0.959, and an F1 score of 0.688 in the validation set.

CONCLUSIONS

We built a Logistic modeling based on machine learning (ML) which facilitated the predicting of hospitalized acute kidney injury (AKI) patients, thereby serving as an promising foundation for the early diagnosis of AKI.

Artificial intelligence, clinical decision support and machine learning

P0231

GENERATION OF AN ARTIFICIAL INTELLIGENCE MODEL FOR SEPSIS DIAGNOSIS USING CLASSIC AND INNOVATIVE HEMOGRAM PARAMETERS

F.d.B. Téllez Hernández [2], I. Osuna Benavente [2], I. Almazo Guerrero [2], S. Martin Perez [2], R. Mondejar Garcia [1], A. Leon Justel [2]

[1]*Clinical Biochemistry Department, Nuestra Señora del Valme University Hospital, Sevilla*
[2]*Clinical Biochemistry Department, Virgen Macarena Univerisity Hospital, Sevilla*

BACKGROUND-AIM

The diagnosis of sepsis is mainly of a clinical nature, being SOFA and quick SOFA scales the main tools employed. However, some laboratory parameters can serve to support this diagnosis.
The aim of this study is to generate a diagnostic model for sepsis using hemogram parameters. This approach would be faster than biochemistry (since it does not need centrifugation) and involves no additional cost.

METHODS

A total of 23,011 emergency patients with C-reactive protein(CRP) and hemogram results from August to October of 2024 were analysed. Hematological, oncological and rheumatological patients were excluded to avoid confounding variables. The hemogram was analysed using Sysmex XN-10, and biochemistry was performed with Abbott Alinity. The hemogram analyser includes two new parameters: NEUT-RI (neutrophil reactivity) and NEUT-GI (neutrophil granularity).
Patients with CRP over 200mg/l and suspicion of sepsis were designated as septic, as well as those exhibiting procalcitonine levels over 5ng/ml. The algorithm included suspected sepsis, a process which consisted of examination of the clinical summaries provided by physicians to ascertain the presence of any of the ten lexical items pertaining to sepsis, including septic, fever or infection. Patients with CRP and procalcitonine levels below 5mg/l and 0.5ng/ml were considered as non-septic.
The utilisation of an in-house artificial intelligence programme generated a "Random Forest" model: an algorithm that, following a series of trials and errors, undertakes a specific path in bifurcations based on the value of certain parameters to reach a probabilistic diagnosis.

RESULTS

The generated algorithm employed solely the lymphocyte count, NEUT-RI, and the lymphocyte/neutrophil ratio.
The algorithm obtained a sensitivity of 0.71, a specificity of 0.92 and an accuracy of 0.9.

CONCLUSIONS

Taking into account that the reference diagnosis of sepsis was based on biochemistry and a clinical summary, it is worthwhile to emphasise the favourable results obtained with hemogram data. This methodology can be further developed using higher-quality data that includes full clinical information. It is also imperative to acknowledge that external validation remains an essential component of the process.

Artificial intelligence, clinical decision support and machine learning

P0232

DEVELOPMENT AND VALIDATION OF AN AUTOVERIFICATION SYSTEM (AVS) FOR TESTS STUDIED AT THE CLINICAL BIOCHEMISTRY LABORATORY OF BEZMIALEM FOUNDATION UNIVERSITY FACULTY OF MEDICINE HOSPITAL

C. Yildiz [2], Ö.F. Özer [1]
[1]*Department of Medical Biochemistry, Bezmialem Foundation University Faculty of Medicine, Istanbul, TÜRKİYE*
[2]*Department of Medical Biochemistry, IGDIR DR. NEVRUZ EREZ STATE HOSPITAL, IGDIR, TÜRKİYE*

BACKGROUND-AIM

The increasing number of laboratory tests requested by physicians has significantly raised the workload of laboratory specialists. Autoverification systems (AVS) offer a potential solution by automating the approval of test results based on predefined rules. This study aimed to develop and validate an AVS for laboratory tests performed at the Clinical Biochemistry Laboratory of Bezmialem Foundation University.

METHODS

An AVS was developed as a middleware system integrated with the Laboratory Information System (LIS). Rules and algorithms were created for biochemistry, immunochemistry, coagulation, complete blood count, and urine analysis tests. The AVS was initially validated using virtual patient data and then using real patient data from November 2023, encompassing 724,692 test results. Finally, the concordance between AVS and manual verification was evaluated.

RESULTS

Autoverification rates differed by test category, ranging from 55.6% to 94.6% for biochemistry tests, 39.6% to 95.5% for immunochemistry tests, 38.4% to 90.9% for coagulation tests, 80.2% for complete blood count, and 49.6% for urine analysis. The highest degree of concordance was observed between AVS and verification done by a resident doctor.

CONCLUSIONS

This study demonstrates the successful implementation of an AVS for laboratory tests. This system has the potential to reduce workload, expedite early error detection, and standardize the verification process. While the initial autoverification rates are promising, further multidisciplinary research is needed to optimize the performance of the AVS.

Artificial intelligence, clinical decision support and machine learning

P0233

ENHANCING HEMOGLOBINOPATHY DIAGNOSIS IN MULTI-ETHNIC POPULATIONS USING MACHINE LEARNING: A RANDOM FOREST-BASED APPROACH

L. Martinez Carreras [1], L.D. Andres Garrido [1], S. Martinez Martin [1], C. Avivar Oyonarte [1]
[1]*Clinical Laboratory Unit. Poniente University Hospital (Almeria, Spain)*

BACKGROUND-AIM

Hemoglobinopathies, including thalassemias and hemoglobin variants (Hb AS, Hb AC), affect approximately 7% of the global population. Their prevalence in regions with high immigration, such as southeastern Spain, highlights diagnostic challenges due to genetic and clinical heterogeneity. Traditional diagnostic methods, reliant on hematological and biochemical analyses, face limitations in diverse populations. This study explores a Random Forest (RF) machine learning modelto enhance diagnostic precision using advanced hematological parameters and sociodemographic data.

METHODS

This cross-sectional study analyzed data from 395 patients diagnosed atthe Hospital Universitario de Poniente (2023–2024) using high-performance liquid chromatography (HPLC) and capillary electrophoresis. Variables included red blood cell indices (MCV, MCH, RDW), reticulocyte parameters (VRM, ADR, UGC), and biochemical markers. The dataset was split into training (70%) and testing (30%) sets, and the RF model was optimized through cross-validation and hyperparameter tuning using KNIME software.

RESULTS

The RF model demonstrated >90% sensitivity and specificity in detecting beta-thalassemia and Hb AS. Parameters such as UGC effectively identified target cells, corroborating prior studies, while VMCE differentiated beta-thalassemia from other microcytic anemias. Sociodemographic factors, including ethnicity and age, significantly influenced the distribution of hemoglobinopathies, with 92% of cases involving immigrants, primarily from Sub-Saharan Africa and the Maghreb. Molecular techniques, including PCR for alpha-thalassemias, complemented the hematological findings, enhancing diagnostic accuracy.

CONCLUSIONS

The integration of a Random Forest-based predictive model improves diagnostic precision in hemoglobinopathies, surpassing traditional methods like HPLC alone. This approach addresses diagnostic gaps in multi-ethnic populations and aligns with personalized medicine objectives. By incorporating advanced laboratory parameters and sociodemographic data, this study lays the groundwork for integrating artificial intelligence into routine clinical workflows. Future research should explore scalability and application in broader healthcare settings.

Artificial intelligence, clinical decision support and machine learning

P0234

BARRIERS OF ARTIFICIAL INTELLIGENCE IMPLEMENTATION IN MEDICAL LABORATORIES ACCORDING LABORATORY PROFESSIONALS' PERCEPTIONS

H. Lame [2], E. Refatllari [2], N. Heta [2], I. Korita [2], V. Tole [2], A. Coraj [2], M. Lika [2], A. Bulo [1]

[1]*Laboratory Department, Faculty of Medicine, University of Medicine, Tirane*

[2]*Laboratory Department, Faculty of Medicine, University of Medicine, Tirane; Laboratory Networks, Mother Theresa University Hospital Centre*

BACKGROUND-AIM

Artificial Intelligence (AI) poses various challenges and barriers to implementation in medical laboratories that need to be addressed. As laboratories undergo digitalization and automation, lab professionals will likely be confronted with the challenges associated with AI. It is important to understand lab professionals' perceptions toward AI and address their concerns regarding these barriers.

METHODS

Lab professionals' perceptions on AI implementation barriers were investigated through a cross-sectional survey. The survey was designed on Google Forms and distributed via WhatsApp and Email to every member of Albanian Society of Laboratory Medicine (ASoLaM) and their staffs. Answers were accepted from 16th July to 13th August 2024. Data were processed and analysed on Excel and SPSS IBM.

RESULTS

The survey was successfully submitted by 198 female and 22 male professionals, with an average age 38 years old. 31% of the respondents were Lab Doctors, 60% were lab technicians and the rest microbiologists, laboratory residents and molecular biologists. 55% of the participants worked in ISO 15189 accredited laboratories, while 33.6% in labs undergoing accreditation process. 71% of the respondents were employed in public labs, while 29% in private laboratories. Survey participants identified multiple barriers that impede AI implementation in medical labs. In order of frequency, laboratory professionals believed that AI implementation barriers were: High costs of initial implementation (55.9% of the participants), Need of appropriate IT support (40.5% of the participants), Lack of appropriate infrastructure (34.5% of the participants), Lack of specialized staff (32.7% of the participants), Ethical considerations on data safety, transparency, fairness and human dignity (24.1% of the participants) and Usage difficulties (9.5% of the participants). Barriers' importance evaluation according lab professionals was Minimal (22.3%), Moderate (38.6%), Important (31.4%) and Maximal (7.7%).

CONCLUSIONS

Laboratory professionals identified as key barriers to AI application in medical laboratories high costs of initial implementation and the need for appropriate IT support. According to the majority of the survey participants these barriers were evaluated as moderate and important.

Artificial intelligence, clinical decision support and machine learning

P0235

MACHINE LEARNING COMBINED WITH A MULTICENTER COHORT STUDY IDENTIFIED A 2-PIRNAS DIAGNOSIS PANEL FOR EARLY STAGE LUNG ADENOCARCINOMA

s. Liang [1], z. Zhou [2], w. Cui [1]

[1]*Department of Clinical Laboratory, State Key Laboratory of Molecular Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China*

[2]*Department of Thoracic Surgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China*

BACKGROUND-AIM

Lung adenocarcinoma (LUAD) is the most common histological subtype of non-small cell lung cancer (NSCLC). Early diagnosis and timely treatment of LUAD are of crucial clinical significance. However, due to the lack of non-invasive early diagnostic markers, most patients are diagnosed at advanced stages. Piwi-interacting RNAs (piRNAs) are abundant in peripheral blood, stably expressed, and not easily degraded. We aimed to discover novel and tumor-specific piRNAs in serum. Besides, machine learning analysis and multi-center validation were applied to develop a piRNAs-based signature for early diagnosis.

METHODS

A total of 1545 participants from 3 medical centers were enrolled, including 1049 patients with LUAD, 109 patients with benign pulmonary nodules (BPNs), and 387 healthy control donors (HCs). We identified potential up-regulated piRNAs by applying LUAD tissue (n=96) and blood (n=10) non-coding RNA sequencing by high throughput profiling. Candidate piRNAs were confirmed in paired tissue specimens (n=96) and serum samples (n=48) by reverse transcription-quantitative polymerase chain reaction (RT-qPCR). A diagnostic model was identified by applying a random forest algorithm, and we tested 2-piRNAs based model in a training cohort (n = 866) and internal test cohort (n=371), followed by a multicenter external validation cohort (n = 308).

RESULTS

A diagnostic panel comprising 2 tumor specific piRNAs (piR-hsa-8393202, piR-hsa-8429916) was developed for the early detection.
2-piRNAs based signature exhibited high precision and accuracy with area under the curve (AUC) values of 0.923, 0.910 and 0.897 in the training, internal validation and external validation cohorts.
Besides, experiments showed that piR-hsa-8393202 and piR-hsa-8429916 expression levels in the serum samples were positively correlated (R2=0.287, p<0.0001; R2=0.190, p=0.002) with the size of the tumor tissue, and the marker levels decreased significantly after surgical resection of tumor. Expression level of 2 piRNAs could remain stable and undegraded within 4 repeated freeze-thaw cycles, demonstrating a certain degree of stability.

CONCLUSIONS

Our current study identified a 2-piRNAs panel as a robust, accuracy and non-invasive serum biomarker for early diagnosis of LUAD patients.

Artificial intelligence, clinical decision support and machine learning

P0236

A PARAMETRIC EMPIRICAL BAYES APPROACH TO PERSONALIZED REFERENCE INTERVALS AND REFERENCE CHANGE VALUES

E.Å. Røys [6], K. Viste [6], R. Kellmann [6], N.A. Guldhaug [6], B. Alaour [7], M.S. Sylte [4], J. Torsvik [5], H. Strand [8], M. Marber [7], T. Omland [2], E. Theodorsson [1], G.R.D. Jones [3], K.M. Aakre [6]

[1]*Department of Biomedical and Clinical Sciences, Division of Clinical Chemistry and Pharmacology, Linköping University, Linköping, Sweden*

[2]*Department of Cardiology, Division of Medicine, Akershus University Hospital, Lørenskog, Norway*

[3]*Department of Chemical Pathology, SydPath, St. Vincent's Hospital, Sydney, Darlinghurst, NSW, Australia.*

[4]*Department of Medical Biochemistry and Pharmacology, Haukeland University Hospital, Bergen, Norway*

[5]*Gade Laboratory for Pathology, University of Bergen, Norway*

[6]*Hormone Laboratory, Department of Medical Biochemistry and Pharmacology, Haukeland University Hospital, Bergen, Norway*

[7]*King's BHF Centre of Research Excellence, School of Cardiovascular Medicine and Sciences, King's College London, United Kingdom*

[8]*Multidisciplinary Laboratory Medicine and Medical Biochemistry, Akershus University Hospital, Lørenskog, Norway*

BACKGROUND-AIM

Laboratory medicine has long relied on population-wide reference intervals (popRI). This one-size-fits-all approach produces thresholds that are too wide compared to the tightly regulated biomarker ranges around the individual homeostatic set points. By establishing personalized reference intervals (prRI) based on these set points, we aimed to enhance diagnostic sensitivity and specificity. A parametric empirical Bayes (PEB) approach offers potential progress in this research field. By balancing individual measurement with population priors, the PEB derives stable prRIs from a limited number of individual results.

METHODS

We employed the PEB method, using routine data extracted from our Laboratory Information System (LIS), to estimate prRIs for albumin, calcium, creatinine, phosphate, cortisone, cortisol, testosterone, androstenedione, 17-hydroxyprogesterone, and 11-deoxycortisol. A robust regression framework established the PEB priors. We assessed the false positive rate of the resulting prRIs at a 5% significance level using serial results from healthy individuals and compared these rates with those obtained via conventional popRIs.

RESULTS

The established prRIs were consistently narrower than popRIs while exhibiting a false positive rate below the expected 5%. For instance, the false positive rate for albumin was 1.2% using prRIs versus 10.4% for popRIs. Calcium and creatinine each demonstrated a prRI false positive rate of 1.6% (compared to 2.2% and 1.9% under popRIs, respectively). Although 17-hydroxyprogesterone showed a slightly higher rate for prRIs (4.4%) than popRIs (1.4%), it remained under the 5% threshold.

CONCLUSIONS

The parametric empirical Bayes prRIs can preserve diagnostic specificity despite increased sensitivity. The approach may provide clinically relevant personalized thresholds based on LIS data.

Artificial intelligence, clinical decision support and machine learning

P0237

LABRI METHOD SHINY APPLICATION: INDIRECT ESTIMATION AND VERIFICATION OF POPULATION-BASED REFERENCE INTERVALS USING UNSUPERVISED MACHINE LEARNING AND GAUSSIAN MIXTURE DECONVOLUTION

A.C. Dias [1], L.M.F. Vieira [2]
[1]*Grupo LabR - Sabin*
[2]*LabR - Dasa*

BACKGROUND-AIM

The estimation or verification of clinical laboratory Reference Intervals (RI) is a requirement of ISO 15189. The objective was to develop a Shiny application with an interactive and user-friendly interface, designed to be accessible to users without expertise in R programming, thereby facilitating the estimation and indirect verification of RIs by clinical laboratory professionals.

METHODS

The RMarkdown version was completed in 2022, followed by the Shiny version in 2024, available on https://github.com/labrgrupo/LabRI_shiny.git. The LabRI method consists of two modules: the estimation module, which uses an adaptive, multi-criteria approach that integrates data cleaning, transformation, clustering, and the refineR, reflimR, and Expectation-Maximization (EM) algorithms, combining parametric and non-parametric approaches based on identified clusters in the truncated distribution; and the verification module, which performs a three-level analysis for RI equivalence, evaluating statistical uncertainty at the first level, testing practical significance at the second level by comparing estimated and reference RIs, and assessing concordance at the third level using Fleiss' Kappa, Lin's Concordance Correlation Coefficient, and flagging rates. Performance was evaluated using datasets from Omuse et al. (2020), Abbam et al. (2021), and simulated data (Testcase1, Testcase2, and Testcase4) from the refineR package.

RESULTS

The performance was consistent with direct sampling results from Omuse et al. (2020) and Abbam et al. (2021) and with simulated indirect sampling.

CONCLUSIONS

This hybrid approach with a user-friendly interface enables use by non-experts and addresses the limitations of the refineR, reflimR, and EM algorithms currently available.

Artificial intelligence, clinical decision support and machine learning

P0238

CIRCULATING CYTOKINE PROFILING AND CLUSTERING IDENTIFY BIOMARKER PREDICTING EFFICACY OF ICI IN COMBINATION WITH CHEMOTHERAPY

L. Yiling [1], X. Xiaotian [1], C. Wei [1]

[1]*Department of Clinical Laboratory, State Key Laboratory of Molecular Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China*

BACKGROUND-AIM

Cytokines are essential in tumor immunity, yet individual cytokine effect is limited due to functional complexity. Few research exists on the predictive capacity of multi-cytokine profiles for immunotherapy efficacy using artificial intelligence. Further investigation is needed to understand systemic inflammation patterns based on multi-cytokine profiles and explore clinical significance in cancer.

METHODS

We analyzed 1331 plasma samples from 1025 pan-cancer patients and 306 healthy controls, including 238 who received immune checkpoint inhibitors (ICI) combined with chemotherapy. Cytokine clusters were identified using non-negative matrix factorization algorithm. The clusters' effects on early-term immunotherapeutic response and progression-free survival (PFS) were evaluated, and a cytokine-based ICI Survival Index (CISI) was constructed for outcome prediction.

RESULTS

Three inflammatory classifications were identified. Cluster 1 showed elevated IFN-γ, IL-8, and IL-1β with proinflammatory traits. Cluster 2 had high IL-6 level. And Cluster 3 displayed increased IL-5 and IL-17, suggesting Th2 cell activation. Cluster 3 was associated with better PFS [hazard ratio: 2.44 and 3.84, P=0.00011] and objective remission rates (54.33%, 61.90% and 85.42%, P=0.00075) than Cluster 1 and Cluster 2. Higher IL-8 and IFN-γcorrelated poor ICI prognosis and an immunosuppressive microenvironment. The CISI model, incorporating cytokine clusters and clinical variables (treatment, IL-10, monocyte-to-lymphocyte ratio, and M stage), outperformed conventional biomarkers PD-L1 and IL-8 in predictive efficiency (C-index = 0.75 vs. 0.55 and 0.56).

CONCLUSIONS

Our cytokine clustering based on multi-cytokine profiles and CISI model predicted prognosis and immunotherapeutic response in tumor patients, providing new insights into personalized cancer therapy strategies.

Artificial intelligence, clinical decision support and machine learning

P0239

APPLICATION OF FEDERATED LEARNING TO PRESERVE THE PERFORMANCE OF AUTOMATIC CLASSIFIERS OF BLOOD IMAGES OBTAINED IN DIFFERENT CLINICAL LABORATORIES

K. Barrera [2], L. Natali [2], A. Merino [1], J. Rodellar [2]

[1]*CORE Laboratory. Biochemistry and Molecular Genetics Department. Biomedical Diagnostic Center. Hospital Clinic. Barcelona. Spain*

[2]*Department of Mathematics. Technical University of Catalonia. Barcelona. Spain*

BACKGROUND-AIM

Variations in staining across different laboratories may affect the performance of AI models for classification of blood cell images that were trained with images of a single reference center (RC). To address this issue, this study proposes the use of federated learning (FL), a strategy to train a generalizable model that keeps its performance when images from other centers are analyzed.

METHODS

A set was available with 10298 blood cell images of the five types of leucocytes from the Core Lab of the Hospital Clínic of Barcelona (RC). Four public sets were used: C1 (14514 images), C2 (2513), C3 (5000) and C4 (11353), In a first step, a VGG16 convolutional neural network was trained with RC images, achieving an accuracy of 99.4%. However, the performance dropped significantly when evaluated with the public sets: 58.6% for C1, 93.2% for C2, 60.30% for C3 and 69.82% for C4. FL was applied to address this issue. The first three convolutional blocks of the RC-trained model were frozen and fine-tuning was performed on the remaining blocks using 10% of the data from each set. The resulting weights from each set were combined to build a new refined Final Global Model.

RESULTS

After the FL training, the model was evaluated with all the images and performance metrics were compared to the pre-FL results. For the set C1, the accuracy was 95.22% (62.49% improvement), with precision of 0.964, recall of 0.952, specificity of 0.992, and F1 score of 0.958. For C2 accuracy was 99.20% (6.44% improvement), with precision, recall, specificity and F1 score of 0.992, 0.992, 0.998, and 0.992, respectively. The accuracy for C3 was 97.30% (61.36% improvement), with a precision of 0.974, recall of 0.973, specificity of 0.993, and F1 of 0.973. With C4, accuracy was 88.81% (27.20% improvement), with precision, recall, specificity and F1 of 0.884, 0.888, 0.992, and 0.885, respectively.

CONCLUSIONS

Federated learning FL is a solution to fine-tune automatic classifiers in multicenter scenarios, maintaining high performance despite variability in image sources. This study is one of the first to successfully apply FL to the classification of peripheral blood cells, demonstrating its potential to address multicenter challenges in this specific field.

Artificial intelligence, clinical decision support and machine learning

P0240

BIAS ASSESSMENT OF MEASURED LDL-C AND ITS ESTIMATION USING THE FRIEDEWALD, MARTIN/HOPKINS, AND SAMPSON EQUATIONS, AND OPTIMIZATION THROUGH THE IMPLEMENTATION OF SUPERVISED MACHINE LEARNING MODELS

A. Reche Martinez [2], N. Rico Rios [1], S. GarcÍa PinteÑo [2], S. Ochoa Guzman [2], J. DomÍnguez GarcÍa [2], E. Melguizo Madrid [2], M. Mayor Reyes [2], J. Corral Perez [2], C. CaÑavate Solano [2], J.D. Santotoribio Camacho [2]
[1]*Laboratory Medicine (Clinical Analysis), Puerta del Mar University Hospital, Cadiz, Spain*
[2]*Laboratory Medicine (Clinical Biochemistry), Puerto Real University Hospital, Cadiz, Spain*

BACKGROUND-AIM

LDL-C is crucial for assessing cardiovascular risk. While direct measurement is available, many laboratories rely on equations. The aim is to determine when to report measured LDL-C vs estimating it using Friedewald, Sampson and Martin-Hopkins equations, and to apply Artificial Intelligence (AI) techniques to identify superior predictive models.

METHODS

A multicenter retrospective study was conducted with 65480 samples collected between Jan 2022-March 2023 from 6 public hospitals in Spain. Data included measured LDL-C, total cholesterol, HDL-C and triglycerides (TG); and they were analyzed using Alinity-c analyzer (Abbott). Statistical analysis was performed using RStudio on the full sample and a subgroup of 56587 samples with TG ≤4.5 mmol/L. Biases between measured and estimated LDL-C values were calculated across 14 TG ranges using Friedewald, Sampson and Martin-Hopkins equations. Desirable bias was set at 5.4% based on biological variation. Sensitivity and specificity of each formula were assessed at clinical decision limits for LDL-C (1.40, 1.80 and 2.60 mmol/L). Several machine learning models including Gradient Boosting (GB) were applied to predict LDL-C, with performance evaluated by coefficient of determination (R^2) and bias, compared to classic equations.

RESULTS

Friedewald showed the poorest performance, failing to meet desirable bias at TG concentrations >1.1 mmol/L, with biases as high as -50.9%. Sampson and Martin met desirable bias up to TG of 2.3 and 4.5 mmol/L respectively, with biases exceeding these limits up to -16.6% and +10.4%. Among the 3 equations only Martin achieved >90% sensitivity and specificity. GB outperformed all traditional equations, meeting desirable biases across the entire range of TG concentrations, up to 28.5 mmol/L. GB showed R^2 of 0.998–0.999 across all TG ranges, while above 4.5 mmol/L , the best R^2 from classic equations ranged from 0.510 to 0.831.

CONCLUSIONS

Our results suggest greater compatibility between Martin-Hopkins and measured LDL-C, supporting a combined approach. The GB model indicates that AI can overcome the limitations of traditional equations, providing a promising alternative to replace both direct LDL-C measurements and classic estimates in all TG scenarios. Further validation in new populations is needed before clinical implementation.

Artificial intelligence, clinical decision support and machine learning

P0241

IMPLEMENTATION OF A CLINICAL DECISION SUPPORT SYSTEM FOR EARLY DETECTION AND MONITORING OF CHRONIC KIDNEY DISEASE IN AT-RISK PATIENTS

R.M. Murria Estal [1], A. Pozo Giráldez [1], E. Giménez Civera [2], F. Moncho Francés [2], A. Carratalá Calvo [1], J.L. Górriz Teruel [2], E. Rodríquez Borja [1]

[1]*Clinical Biochemistry and Molecular Pathology Laboratory. University Clinical Hospital of Valencia. Spain*

[2]*Nephrology Department. University Clinical Hospital of Valencia. Spain*

BACKGROUND-AIM

Chronic kidney disease (CKD) poses a growing global challenge due to an aging population and the increasing prevalence of type 2 diabetes, hypertension (HBP), and obesity. Early detection is essential for effective treatment and altering disease progression. However, CKD detection in high-risk groups remains suboptimal in clinical practice despite established guidelines.

METHODS

We implemented an early CKD detection program in our Health Department using a clinical decision support system (CDS) based on the "CDS-Ripple Down- Abbott Diagnostics" middleware. This system is integrated into the laboratory electronic ordering platform and patient medical records. When a primary care (PC) physician orders laboratory tests, the CDS identifies high-risk patients (aged 65–90 years, diabetes mellitus, HBP or obesity) and automatically adds serum creatinine, estimated glomerular filtration rate (eGFR), urine albumin/creatinine ratio (ACR), and urinary strip tests (with flow citometry if necessary). The CDS can also autonomously apply screening criteria and initiate testing without physician input (sentinel program).
Patients are staged according to KDIGO guidelines using eGFR and ACR. The system monitors disease progression, referring only those with ACR >300 mg/g, eGFR <30 ml/min/1.73 m^2, or significant progression to nephrology.

RESULTS

In the first year of the route implementation, 32,618 high-risk patients were screened, with 99.1% identified via the sentinel program. Among them, 25,806 (79.1%) had no CKD; 6,050 (18.5%) had CKD without nephrology referral criteria (moderate risk); and 762 (2.3%) had CKD meeting nephrology referral criteria (high risk). Only 1% of patients required in-person nephrology consultations, while 1.3% were managed with PC follow-up recommendations. Additionally, 2% of patients who were previously under follow-up showed progression of their disease.

CONCLUSIONS

This CDS-based program effectively automates CKD identification, monitors disease progression, and provides tailored recommendations to PC physicians, enhancing referral efficiency. As the first population-wide CKD screening initiative leveraging a CDS, it demonstrates promising results as a scalable, accessible, and potentially cost-effective method. The project has been granted intellectual property certification.

Artificial intelligence, clinical decision support and machine learning

P2756

COMPARISON BETWEEN TOTAL MAGNESIUM AND IONIZED MAGNESIUM MEASUREMENT: CLINICAL IMPACT AND POTENTIAL OF ARTIFICIAL INTELLIGENCE

N.H. Cahuana Santamaria [1], B.A. Lavin Gomez [1], A.R. Guerra Ruiz [1], N.L. Santamaria Urey [1], L. Carrera Ceballos [1], M. Fernandez Cuetara [1], S. Hernandez Vicente [1], G. Manso Gil [1]
[1]*HOSPITAL UNIVERSITARIO MARQUES DE VALDECILLA*

BACKGROUND-AIM

Magnesium is essential for several physiological processes, such as muscle function, nerve transmission, and protein synthesis. Its measurement is crucial, especially in patients undergoing hemodialysis or with conditions requiring strict fluid balance control. There is a distinction between total magnesium, which measures all magnesium in the sample, and ionized magnesium, which reflects the biologically active fraction. This difference impacts clinical interpretation and treatment choices. Additionally, while measuring ionized magnesium is more expensive, the use of artificial intelligence and predictive algorithms could reduce these costs without compromising clinical accuracy.

METHODS

A retrospective study analyzed 2,200 total magnesium samples from MODULAB, applying three formulas to calculate ionized magnesium Three formulas were applied to calculate ionized magnesium from the total magnesium values. Statistical tests compared total and ionized magnesium values. The resulting data were used to train an AI model with supervised learning techniques, employing algorithms like linear regression and neural networks. Model accuracy was assessed using cross-validation and metrics such as mean squared error (MSE) and R^2.

RESULTS

The Pearson correlation coefficient between total and ionized magnesium is 0.99, indicating a strong positive correlation. Significant differences were found between total magnesium levels and ionized magnesium estimates using three formulas (A, B, and C), with statistical significance ($p < 0.05$). The AI model showed excellent predictive ability, with a 99.9% accuracy ($R^2 = 1.64$), indicating strong correlation between predicted and actual ionized magnesium levels.

CONCLUSIONS

The results emphasize the importance of accurate magnesium measurement in clinical settings, especially for patients requiring careful fluid and electrolyte balance. Understanding the differences between total and ionized magnesium is crucial for selecting the appropriate measurement method. AI's potential to estimate ionized magnesium from total magnesium data could offer a cost-effective alternative, leading to significant long-term savings. These findings have both clinical and economic implications.