

Leveraging Machine Learning to Aid in the Utilization of Diagnostic Testing in Thrombotic Thrombocytopenic Purpura

Avery L. Teman, Seanna C. Clark, Tatiana T. Vu, Zaraith S. Oviol, Maximo J. Marin

College of Medicine, University of Florida

Maximo Marin, Department of Pathology, Immunology, and Laboratory Medicine

Abstract

Artificial intelligence (AI) has the potential to revolutionize the medical field with machine learning utilization, improving patient outcomes. Thrombotic thrombocytopenic purpura (TTP) is a lifethreatening, blood clotting disorder which is confirmed by the ADAMTS13 activity assay. The improper usage of ADAMTS13 and constrained resources in laboratories leads to inefficient patient care. This research project will result in a decision tree (DT) algorithm, aiding in efficiently diagnosing TTP. This machine learning (ML) support tool would reduce the over-utilization of ADAMTS13 testing and save lives. In Phase 1, the principal investigator coded the ML algorithm, which was developed by training and testing with preliminary data, producing an overall accuracy of 81%. Phase 2 curates a collection of patient data using the UF Health electronic health record for validation of the algorithm. Phase 3 includes additional testing with new data, while Phase 4 requires review of guidelines for implementation into the laboratory. This knowledge will help close the mortality gap for TTP and provide the framework to advance the development of AI support tools for various diseases. The overarching mission is to create the lab of the future where AI-generated decision support tools guide better diagnostic testing to aid clinicians in improving patient care.

Keywords: artificial intelligence, thrombotic thrombocytopenic purpura, machine learning, clinical laboratory

Introduction

Ever since the birth of AI in 1950, the technology has spurred countless innovations, created jobs that never existed prior, and introduced the possibilities of automation. With the rapid onset of artificial intelligence technology, a sort of "arms race" has begun within all industries in a multitude of ways. Not only is there a competition for efficiency and modernization, with AI being used to automate daily processes and alleviate costs, but also there is a race to produce the newest and most enticing technology for customer usage (Chen & Chen,

2022). Additionally, there is competition between the creators of these new AI platforms and governments, with growing fears concerning the privacy, safety, and security of the users. Despite potential misinformation and hesitation on implementation of these tools, there are insurmountable benefits with the application of new ML technology in many different fields. One area of focus is how AI has the potential to revolutionize the medical field (Haymond $\&$ McCudden, 2021).

Throughout history, greater access to information has led to pivotal discoveries in medicine, from cutting edge surgical developments to unimaginable lab testing methods. AI has the potential to create the lab of the future by optimizing clinical laboratory testing, drastically improving efficiency, and reducing resource misuse. ML algorithms are on the cusp of being implemented as tools to aid in the diagnosis of various diseases (Herman et al., 2021). There are different types of ML algorithms, such as decision trees (DTs). DTs are a form of supervised learning which makes predictions and categorizations based on a series of questions answered by the user. Just like how a tree branches off, the answer to each question leads to another question until a single outcome is reached. The benefit of using a DT, especially in medical-related topics, is that this ML algorithm is transparent and may provide flexibility in its predictive capabilities. The creators of these algorithms can easily see how a decision was made and make improvements at each iteration (Punchoo et al., 2021). Further, physicians with minimal experience in coding or with ML algorithms can view the inner workings of the DT; thus, allowing them to build confidence in the ML decision outcome. In contrast, other ML algorithms, called "black box" ML algorithms, are often difficult to understand for various reasons, including manufacturer propriety and complex mathematics. The accuracy of a DT is measured by the algorithm's ability to properly predict the desired result when new data is provided (Poon & Sung, 2021).

Following the COVID-19 pandemic, clinical laboratories continue to feel the impact of supply chain issues, reagent shortages, and staffing level concerns. Even though, to a certain degree, these problems existed prior to 2020, the pandemic only accelerated and magnified these issues. All across the United States, labs are short-staffed, and the current employees are facing more burnout and heavy workloads than ever before (Leber et al., 2022). Additionally, labs continue to face another dilemma: over-testing. Tests are ordered unnecessarily, which leads to further stretching of limited resources (Cadamuro et al., 2018). To mitigate these ongoing issues, the clinical labs have begun to implement automated instruments, leverage information technology, and incorporate AI to increase productivity (Haymond & McCudden, 2021). This project analyzes the potential of these advancements for the ADAMTS13 activity assay, which is overused in seeking to confirm the diagnosis of thrombotic thrombocytopenic purpura (TTP).

Thrombotic thrombocytopenic purpura is a rare thrombotic microangiopathy (TMA) and a medical emergency. TTP is a life-threatening blood clotting disorder in which the clots that form in blood vessels limit the flow of blood to vital organs, leading to severe organ damage, long-term health problems, and eventual death if left untreated. There are two forms of TTP, congenital and acquired. The focus of this project will be on acquired TTP because this is the most common form, comprising approximately 95% of all cases (Herrera Rivera et al., 2023). Generally, TMAs present with three characteristics: a drop in the platelet count, presence of red blood cell schistocytes (i.e., fragmented red blood cells) in the peripheral blood, and blood vessels with fibrin clots. This common presentation leads to a difficulty in diagnosing the appropriate TMA, which can potentially delay proper treatment. The etiology and pathophysiology of each TMA are different (Zheng et al., 2020).

Acquired TTP involves an autoantibody that binds to the ADAMTS13 enzyme which inhibits its activity to breakdown von Willebrand (vWF) molecules. vWF is involved with primary hemostasis. In other words, it is involved with the initiation of clot formation using platelets. Normally, when blood vessels are injured, they expose collagen, which provides a binding site for platelets directly and more importantly, indirectly via vWF. vWF is a long biomolecule that binds to collagen and has binding sites for platelets. When platelets bind to vWF, they become activated. Platelet activation leads to platelet aggregation and initiates the coagulation cascade (secondary hemostasis). However, in order not to have excessive clot formation, there are multiple regulatory molecules. ADAMTS13 regulates vWF by cleaving it. However, in TTP, the autoantibody prevents ADAMTS13 from cleaving vWF leading to unregulated clot formation and the observed common clinical presentation seen in TMAs. Thus, the importance of ADAMTS13 testing is to distinguish between the TMAs (Sukumar et al., 2021).

The ADAMTS13 assay is not commonly done in hospital labs since it is a semi-manual test that requires maintenance, expert technical skills, and interpretation (Paydary et al., 2020). Hospital laboratories that do run these tests rely heavily on appropriate requests from the

providers to ensure unnecessary testing is not being performed due to the significant amount of required resources. There are several assays available. A common test method measures a patient's ADAMTS13 activity by providing a synthetic vWF molecule that serves as a substrate and contains the specific cleavage site for the ADAMTS13. Once this molecule is cleaved by the ADAMTS13 in the sample, fluorescence is released, which is then measured and compared to determine the ADAMTS13 activity level from normal control samples, respectively (Sukumar et al., 2021). If the assay yields an activity level >20%, the patient is not considered for a diagnosis of TTP. If the activity level is <20%, additional reflex testing is required. There are generally two common explanations as to why a patient would have reduced activity levels. Either there is no ADAMTS13 enzyme, or there is activity inhibition by specific antibodies that are anti-ADAMTS13. Other explanations exist for a low ADATMS13 level, but these are complex and are beyond the scope of this article. The additional reflex testing details the relative amount of potential autoantibodies, which would allow for the determination of the source of the reduced activity. In cases of TTP, the assay generally reveals activity levels less than 10-13%. With this disease, patients who go untreated may potentially die within 24 hours if found to have TTP, and testing takes a significant amount of time to be performed when it is sent out of the local treatment facility to be performed by a reference laboratory (Herrera Rivera et al., 2023). This is the main reason why the PLASMIC score was developed.

The PLASMIC score serves as a rapid assessment tool to evaluate the potential diagnosis of TTP, allowing patients to begin receiving treatment while awaiting the results of the ADAMTS13 test (Li et al., 2018). The PLASMIC score has seven features: platelet count, hemolysis, absence of active cancer, absence of transplant history, mean corpuscular volume (MCV), prothrombin time international normalized ratio (PT-INR), and creatinine (Cr). In calculating the score, each feature is assigned one point, and there is a certain risk level for TTP associated with each total score out of seven points (Upadhyay et al., 2019). A higher PLASMIC score is proportional to the risk of having TTP. Although the PLASMIC score may be effective for patient management, lessening costs, and reducing over-utilization of tests, it is not always used properly for a multitude of reasons. Physicians may be unaware of how to accurately apply this probability assessment tool and the PLASMIC relies heavily on the user to input the appropriate variables. For example, in the context of the timing, scores can be falsely calculated if the inputted laboratory results are not within the required time frame, which is based on the

initial suspicion for TTP. Therefore, scores can be overestimated or potentially underestimated, leading to a misdiagnosis of TTP (Paydary et al., 2020). The goal of the new ML algorithm from this research project is to provide a tool that is less dependent on the user to guide appropriate diagnostic testing.

In Phase 2 of the project, there are 21 data points being collected: patient initials, medical record number (MRN), age, ethnicity, sex, if the patient is deceased, if the patient came through the emergency room, if the patient had labs from less than 72 hours before the ADAMTS13 test, if the patient was given blood products before the ADAMTS13 test, platelet count, if the schistocytes were present and in what quantity, lactate dehydrogenase (LDH), d-dimer, reticulocyte count, haptoglobin, indirect bilirubin, if the patient has active cancer, if the patient has a transplant history, mean corpuscular volume (MCV), prothrombin time international normalized ratio (PT-INR), Creatinine (Cr), and the value of the ADAMTS13 test in percent form. It is important to note whether or not the patient was given blood products before the ADAMTS13 activity assay because this would provide a platelet count that is not truly representative. Platelets are an extremely important indicator of TTP. If the platelet count is very low, less than $30,000/\mu L$, this counts as one point for the PLASMIC score. Hemolysis also counts as one point. The three data points that would signal hemolysis in a patient are schistocytes, undetectable haptoglobin, and indirect bilirubin greater than 2.0 mg/dL. Schistocyte presence is quantified as: few, many, moderate, or none. If haptoglobin is less than 30, this is considered "undetectable." The LDH and d-dimer values are collected because certain levels are characteristic findings in patients with TTP. Reticulocyte count (auto-corrected) greater than 2.5% is also indicative of hemolysis. If the patient has active cancer or a history of any transplants, these would both count as zero points since these events may have an effect on the lab values. MCV less than 90 fL, PT-INR less than 1.5, and Cr less than 2 are all indicative of TTP in a patient as well (Li et al., 2018). Lastly, the ADAMTS13 percentage is collected to check against the PLASMIC score as well as the ML algorithm for accuracy. The 72-hour window for collecting data is important because some of the variables and lab values being collected are not done on a daily basis. By having a set limit of 72 hours before the collection instant, this incorporates as much data as possible that is close to when the ADAMTS13 testing was performed, but not too distant that the values are not reflective of the patient's current condition (Herrera Rivera et al., 2023).

This project is currently divided into four phases which include Phase 1: Initial algorithm creation, Phase 2: Data collection, Phase 3: Algorithm validation testing, Phase 4: Implementation into the clinical laboratory. The goal of this research is to relieve clinicians and physicians from the hindrances posed by over-testing, inefficiency, and resource-waste. With the crucial findings of this project, the research team aims to create the lab of the future by incorporating artificial intelligence into everyday testing processes. Overall, the utilization of AI as a support tool will increase efficiency in clinical laboratories, improve positive patient outcomes, and potentially change how medicine is practiced.

Methods

Phase 1: Initial Algorithm Creation

In Phase 1 of this research project, the initial algorithm was developed. This DT ML model was trained and tested using in-house data as well as data from literature. The total dataset included 104 patients, with 30 in-house patients and 74 literature-derived patients. Within this original dataset, 52 patients had TTP and 52 did not have TTP. This DT algorithm was derived using the PLASMIC score variables serving as features. This initial algorithm had an overall accuracy percentage of 81% while maintaining a high negative predictive value (low number of false negatives).

Phase 2: Data Collection

The patient data was collected through EPIC, a software platform built for hospitals that includes all patients' medical records including tests and charting notes. One of the researchers is a student, so obtaining access to EPIC provided a limitation. Within the department in which the principal investigator on this project works, there is not a well-known process that allows students to view these confidential records. After several weeks, multiple training modules, and

compliance agreements, the student researcher had limited, but functional access to the necessary

patient data and could begin capturing the data.

The initial data points being collected in this experiment were: patient initials, MRN, age, ethnicity, sex, if the patient is deceased, if the patient came through the emergency room, if the patient had labs from less than 72 hours before the ADAMTS13 test, if the patient was given blood products before the ADAMTS13 test, platelet count, if the schistocytes were present and in what quantity, LDH, d-dimer, reticulocyte count, haptoglobin, indirect bilirubin, if the patient has active cancer, if the patient has a transplant history, MCV, PT-INR, Creatinine, and the value of the ADAMTS13 test in percent form. These data points all relate to the variables from the PLASMIC score system. After collecting each of these values for a patient, there is a designated "Y" for yes or "N" for no listed next to the values. This answer is based on whether or not the patient data meets the criteria for the PLASMIC score assignment. Afterwards, the student

Figure 1. Initial Data Points and Collection

researcher calculated the PLASMIC score for each patient in order to compare this manually

Plasmic Score Platelet count Less than 30,000/microL - 1 point Greater than 30,000/microL - 0 points Hemolysis If yes - 1 point If no - 0 points Active cancer If yes - 0 points If no - 1 point Transplant If yes - 0 points If no - 1 point **MCV** If less than 90 fL - 1 point If not less than 90 fL - 0 points Creatinine If less than 2 - 1 point If not less than 2 - 0 points PT INR If less than 1.5 - 1 point If not less than 1.5 - 0 points Total plasmic score

calculated score to the DT algorithmic score in the future phase of the research project.

Time is a vital component in this project. The data values that are collected must be within 72 hours before the collection instant of the ADAMTS13 test. For the values collected from a non-emergency room hospital visit (i.e., transferred from an outside hospital), the goal was to collect the data close to the time of specimen collection for the ADAMTS13. However, for the values collected from the emergency room visits, the goal was to collect the earliest possible value from the patient's initial admission into the ED. Even if one of the data values was collected one second after the ADAMTS13 collection, it was not included.

With the expansion of this project, the data points being collected have grown to include AST, ALP, if the patient has liver disease, and if the patient has a form of TMA. As this phase of the project has progressed, the researchers have noticed a possible relation among these new data points and TTP. There are plans to delve further into this hypothesis in the next phase of the

Figure 2. PLASMIC Score Calculation

project. This data collection process is ongoing. There are still patients who are being tested for

ADAMTS13 and being classified as "TTP" and "non-TTP" patients. New patients will continue to be added to the dataset until the deadline of May 2024.

AST	ALT	Alkaline Phosphatase (ALP)	Liver disease?
48	50	37	N
66	35	161	N
23	8	59	N
21	9	51	N
19	13	75	N
106	4	N/A	N
N/A	N/A	N/A	N
1781	542	27	N
11870	4384	235	N
N/A	N/A	N/A	N
3935	1777	267	DCK LIVER 9/3/2
19	17	67	N
19	45	96	N
64	14	52	N
N/A	N/A	N/A	Ν
10460	836	186	SHOCK LIVER

Figure 3. New Data Points and Collection

Results

With the current phase of the project being data collection, results are pending. From Phase 1, the DT ML model was found to have overall accuracy of 81%, sensitivity of 100%, specificity of 69%, positive predictive value of 67%, and negative predictive value of 100%. In the creation of this algorithm, the main goal was to not restrict testing or miss any potential cases of TTP (false negatives). Thus, higher amounts of false positive cases were allowed in exchange for a higher negative predictive value (i.e., reduction of false negatives). With this preliminary algorithm testing, there is the possibility to rule out suspected TTP patients. After applying and testing the ML model on the new data being collected in Phase 2, the goal of Phase 3 is to improve overall accuracy and continue to maintain the high negative predictive value. In Phase 4, the model will be shadow tested in a clinical environment to ensure its robustness in assigning TTP or assigning non-TTP. Finally, this phase also contains one of the greatest barriers: the official process of implementation into the clinical laboratory setting. This will require resources

to evaluate the DT ML model's regulatory status which includes medical device and Clinical Decision Support Software guidelines.

Discussion

The ML model should only be used as a decision support tool for physicians and does not aim to replace their clinical expertise. Although AI has revolutionary applications in healthcare, uncertainties surrounding the ethics of ML systems have arisen. Some providers fear that there may be risks associated with allowing AI to aid in diagnosing patients, potentially leading to errors in management and treatment. With any field of research, there can be no exact guarantees of efficacy; however, the studies referenced in the introduction have demonstrated the success of AI systems in other disciplines. With the implementation of ML into the clinical laboratory, one of the greatest challenges to overcome will be resistance by those who are unfamiliar with AI tools. It is vital that procedures are also established to educate clinical laboratory staff on these new technologies and ensure that clinicians feel comfortable using these tools on a daily basis.

TTP is a rare, yet fatal disease, therefore the development of a diagnostic tool to increase efficiency and positive patient outcomes is crucial. Additionally, the ADAMTS13 test is scarce among hospitals and other clinical laboratory facilities. The implications from the findings of this project will impact many lives and will be applied to a wide variety of other diseases. Currently, there are very few projects relating to applying AI to TTP. This research is unique because of the type of machine learning algorithm, the features, and the binary system being utilized to leverage transparency. Many clinicians are wary of ML and its lack of transparency. The algorithm utilized in this project, DT, is crafted in a manner that allows users to see every single decision and outcome made by the system.

The ultimate goals for this project are to provide the following: develop a pipeline system to integrate ML learning to assist clinical teams with decisions that can improve patient outcomes, cost-effectiveness, reduce over-utilization of valuable laboratory resources and inefficiency. Implementing and improving decision support tools for clinical teams can alleviate decision fatigue, improve workflow and assist in difficult clinical cases with regard to guiding appropriate diagnostic testing selection. We hope to push the boundaries and create the lab of the future by

incorporating AI into testing procedures. Specifically, if this research project can work towards reducing the mortality gap with AI for TTP, then the application to other diseases is limitless.

References

- Cadamuro, J., Gaksch, M., Wiedemann, H., Lippi, G., von Meyer, A., Pertersmann, A., Auer, S., Mrazek, C., Kipman, U., Felder, T. K., Oberkofler, H., & Haschke-Becher, E. (2018). Are laboratory tests always needed? Frequency and causes of laboratory overuse in a hospital setting. *Clinical biochemistry*, *54*, 85–91.<https://doi.org/10.1016/j.clinbiochem.2018.01.024>
- Chen, R. H., & Chen, C. C. (2022). *Artificial intelligence: An introduction for the inquisitive reader* (First edition.). CRC Press. <https://doi.org/10.1201/9781003214892>
- Haymond, S., & McCudden, C. (2021). Rise of the machines: Artificial intelligence and the clinical laboratory. *The journal of applied laboratory medicine*, *6*(6), 1640–1654. <https://doi.org/10.1093/jalm/jfab075>
- Herman, D. S., Rhoads, D. D., Schulz, W. L., & Durant, T. J. S. (2021). Artificial intelligence and mapping a new direction in laboratory medicine: A review. *Clinical chemistry*, *67*(11), 1466– 1482. https://doi.org/10.1093/clinchem/hvab165
- Herrera Rivera, N., McClintock, D., Alterman, M., Alterman, T., Pruitt, H., Olsen, G., Harris, N., & Marin, M. (2023). A clinical laboratorian's journey in developing a machine learning algorithm to assist in testing utilization and stewardship. *Journal of Laboratory and Precision Medicine, 8.* doi: 10.21037/jlpm-23-11
- Leber, A. L., Peterson, E., Dien Bard, J., & Personnel Standards and Workforce Subcommittee of the American Society for Microbiology (2022). The hidden crisis in the times of COVID-19: Critical shortages of medical laboratory professionals in clinical microbiology. *Journal of clinical microbiology*, *60*(8), e0024122.<https://doi.org/10.1128/jcm.00241-22>
- Li, A., Khalighi, P. R., Wu, Q., & Garcia, D. A. (2018). External validation of the PLASMIC score: a clinical prediction tool for thrombotic thrombocytopenic purpura diagnosis and treatment. *Journal of thrombosis and haemostasis : JTH*, *16*(1), 164–169. <https://doi.org/10.1111/jth.13882>
- Paydary, K., Banwell, E., Tong, J., Chen, Y., & Cuker, A. (2020). Diagnostic accuracy of the PLASMIC score in patients with suspected thrombotic thrombocytopenic purpura: A systematic review and meta-analysis. *Transfusion*, *60*(9), 2047–2057.<https://doi.org/10.1111/trf.15954>
- Poon, A. I. F., & Sung, J. J. Y. (2021). Opening the black box of AI-medicine. *Journal of Gastroenterology and Hepatology, 36*(3), 581–584. https://doi.org/10.1111/jgh.15384
- Punchoo, R., Bhoora, S., & Pillay, N. (2021). Applications of machine learning in the chemical pathology laboratory. *Journal of clinical pathology*, *74*(7), 435–442. [https://doi.org/10.1136/jclinpath-2021-](https://doi.org/10.1136/jclinpath-2021-207393) [207393](https://doi.org/10.1136/jclinpath-2021-207393)
- Sukumar, S., Lämmle, B., & Cataland, S. R. (2021). Thrombotic thrombocytopenic purpura: pathophysiology, diagnosis, and management. *Journal of clinical medicine*, *10*(3), 536. <https://doi.org/10.3390/jcm10030536>
- Upadhyay, V. A., Geisler, B. P., Sun, L., Uhl, L., Kaufman, R. M., Stowell, C., Makar, R. S., & Bendapudi, P. K. (2019). Utilizing a PLASMIC score-based approach in the management of suspected immune thrombotic thrombocytopenic purpura: a cost minimization analysis within the Harvard TMA Research Collaborative. *British journal of haematology*, *186*(3), 490–498. https://doi.org/10.1111/bjh.15932
- Zheng, X. L., Vesely, S. K., Cataland, S. R., Coppo, P., Geldziler, B., Iorio, A., Matsumoto, M., Mustafa, R. A., Pai, M., Rock, G., Russell, L., Tarawneh, R., Valdes, J., & Peyvandi, F. (2020). ISTH guidelines for the diagnosis of thrombotic thrombocytopenic purpura. *Journal of thrombosis and haemostasis : JTH*, *18*(10), 2486–2495.<https://doi.org/10.1111/jth.15006>