VALIDATING A PREDICTION MODEL FOR IMMUNE THROMBOCYTOPENIA

DEVELOPING A PROTOCOL FOR THE EXTERNAL VALIDATION OF A CLINICAL PREDICTION MODEL FOR THE DIAGNOSIS OF IMMUNE THROMBOCYTOPENIA

By SYED MAHAMAD, B.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Master of Science

McMaster University

MASTER OF SCIENCE (2023)

Hamilton, Ontario, Canada (Medical Sciences)

TITLE: Developing a Protocol for the External Validation of a Clinical Prediction Model for the Diagnosis of Immune Thrombocytopenia

AUTHOR: Syed Mahamad, B.Sc.

SUPERVISOR: Donald Arnold, MD M.Sc. FRCP(C)

NUMBER OF PAGES: xi, 48

LAY ABSTRACT

There lack of a standardized test to diagnose immune thrombocytopenia (ITP) leads to delays in care, use of incorrect treatments, and increased patient anxiety. The Predict-ITP Tool was developed to classify patients as ITP or non-ITP using the following data: 1) platelet counts in the recent past; 2) the highest mean platelet volume; and 3) major bleeding at any time in the past. The preliminary internal validation study showed promise.

I developed a study protocol to externally validate the Predict-ITP Tool that will collect data from 960 patients from 11 clinics across Canada to see how accurately the tool would have performed to classify patients as ITP or non-ITP at the first hematology visit compared with the gold standard clinical diagnosis by the hematologist or an independent expert committee. A successful external validation that demonstrates the tool's predictive accuracy in an external population must be completed before widespread use.

ABSTRACT

Defined as a platelet count <$100x10^9$/L with no known cause, immune thrombocytopenia (ITP) is a diagnosis of exclusion, meaning other thrombocytopenic conditions must be ruled out before establishing the ITP diagnosis. This can lead to errors, unnecessary exposures to expensive and harmful treatments, and increased patient anxiety and distress. In the absence of a standardized diagnostic test, a clinical prediction model, called the Predict-ITP tool, was developed to aid hematologists in establishing the ITP diagnosis among patients who present with thrombocytopenia. Based on a cohort of 839 patients referred to an academic hematology clinic and using penalized logistic regression, the following predictor variables for the ITP diagnosis were identified: 1) high platelet variability index; 2) lowest platelet count; 3) highest mean platelet volume; and 4) history of a major bleed. Internal validation was completed using bootstrap resampling, and showed good discrimination and excellent calibration.

Following internal validation and prior to implementation, the Predict-ITP Tool must undergo external validation by evaluating the tool's performance in a different cohort. A study protocol was developed with the objective of externally validating the Predict-ITP Tool by collecting data from 960 patients from 11 clinics across Canada. The tool will compute the probability of ITP using information available at the time of the initial consultation, and results will be compared with either the local hematologist's diagnosis at the end of follow-up or the adjudicated diagnosis. Discrimination (the ability to differentiate between patients with and without ITP) and calibration (the agreement between predicted and actual classifications) of the tool will be assessed.

The Predict-ITP Tool must demonstrate good discrimination (c-statistic ≥ 0.8) and excellent calibration (calibration-in-the-large close to 0; calibration slope close to 1) to achieve

external validation. If implemented, this tool will improve diagnostic accuracy and reduce delays in diagnosis and unnecessary treatments and investigations.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my gratitude to my supervisor, Dr. Donald Arnold. Thank you for making the time to meet with me, share your knowledge, and provide your mentorship. It truly has been an honour and privilege to work with and learn from you. I would like to thank Dr. Alfonso Iorio, Dr. Sameer Parpia, and Dr. Na Li for their guidance throughout the process of developing and refining this study protocol. Their insight has been invaluable and instrumental to developing a rigorous study design. I would also like to thank Dr. Peter Gross for serving as my external examiner and providing feedback.

I would like to thank McMaster University and the Medical Sciences Graduate Program for the opportunity to pursue this master's degree in Medical Sciences. Completing my thesis would not have been possible without the support I received from various members of the Michael G. DeGroote Centre for Transfusion Research (MCTR). In particular, I would like to thank Ms. Shannon Lane, Ms. Erin Jamula, Ms. Melanie St. John, Ms. Jess Clarke, and Ms. Julie Carruthers for providing the support and feedback I needed to succeed. I also had the pleasure of being a part of an excellent group of graduate and undergraduate students, including Dimpy Modi, Saifur Rahman Chowdhury, Nelson Gutierrez-Cardona, Nadia Gabarin, Emily Sirotich, Shangi Vijenthira, Laura Venier, and Mike Radford. Thank you all for helping create a productive learning environment.

Finally, I would like to thank my family and friends for their love and encouragement through my master's degree.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

LIST OF ABBREVIATIONS

CI	Confidence Interval
CPM	Clinical Prediction Model
DVT	Deep Vein Thrombosis
eCRF	Electronic Case Report Form
EMR	Electronic Medical Record
HIV	Human Immunodeficiency Virus
IQR	Interquartile Range
ITP	Immune Thrombocytopenia
IVIG	Intravenous immunoglobulin
MPV	Mean Platelet Volume
PROBAST	Prediction model Risk Of Bias ASsessment Tool
PTS	Post-thrombotic Syndrome
REDCap	Research Electronic Data Capture
ROC	Receiver Operating Characteristic
TRIPOD	Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis

DECLARATION OF ACADEMIC ACHIEVEMENT

I declare that the Master's thesis contained herein is my own work. This document presents that work completed in developing a study protocol for the external validation the Predict-ITP Tool, a novel clinical prediction model designed by my supervisor, Dr. Donald Arnold, to help hematologists establish the diagnosis of immune thrombocytopenia. I reviewed relevant literature to obtain a strong understanding of the methodology of external validation studies. With the guidance of Dr. Arnold as well as my two committee members, Dr. Alfonso Iorio and Dr. Sameer Parpia, I developed the study protocol for the external validation of the tool. With the help of my research group, the Michael G. DeGroote Centre for Transfusion Research (MCTR), I recruited 11 participating centres. I led two investigator meetings to introduce the study to local centres and to describe the data collection process. I created the electronic case report forms using REDCap along with an operations manual to assist data abstractors. With support, I developed the outcome adjudication process, including the criteria for adjudication, number of required adjudicators, and meeting frequency.

1.0 INTRODUCTION

1.1 Immune Thrombocytopenia

Immune thrombocytopenia (ITP) is the most common acquired bleeding disorder worldwide, affecting 1 in 8,000 Canadians (1,2). It is an autoimmune disease that has a complex and heterogenous pathophysiology. Platelet autoantibodies result in platelet destruction via macrophages in the spleen and cause either the destruction or impairment of bone marrow megakaryocytes, leading to reduced platelet production (3). Other pathophysiologic pathways have been described, including direct cytotoxic T cell-mediated destruction (4), premature clearance of desialylated platelets (5), and complement-mediated platelet destruction (6). This wide scope suggests that ITP is more like a syndrome with several possible immunological mechanisms that all lead to a final common pathway of thrombocytopenia (7).

The accepted definition of ITP is a platelet count below 100 x 10^9/L (normal range 150 – 400 x 10^9/L) (8); however, this definition is non-specific and easily confused with other thrombocytopenic conditions (e.g., occult liver disease, myelodysplastic syndrome, hypersplenism, hereditary thrombocytopenia and others). Due to the potential of a great reduction in circulating platelets, patients with ITP are at risk of spontaneous bleeding (9). Major bleeding occurs in approximately 9.6% of adults with ITP [95% confidence interval (CI), 4.1–17.1%] and intracerebral hemorrhage occurs in 1.4% (95% CI, 0.9–2.1%) (10). In addition, ITP impacts patients' quality of life due to fatigue, anxiety and limitations on physical activities (11–14) similar to the impact of diabetes or cancer (15,16).

1.1.1 Current Approach to Diagnosing Immune Thrombocytopenia

Thrombocytopenia is one of the most frequent reasons for a hematology consultation in the clinic and it can be caused by many conditions, including ITP. As discussed, ITP has a complex and heterogenous pathophysiology. As a result, there is no standardized diagnostic approach for ITP, meaning other possible causes of thrombocytopenia must be excluded before arriving at the diagnosis of ITP (17). To rule out the other causes, patients with suspected ITP undergo a comprehensive history, physical examination, and the following laboratory studies: complete blood count, peripheral blood film, human immunodeficiency virus serology, hepatitis C serology, and comprehensive metabolic panel (including transaminases, bilirubin, and alkaline phosphatase) (18).

One characteristic feature that can distinguish ITP from other causes of thrombocytopenia is a platelet count increase after treatment with high-dose intravenous immunoglobulin (IVIG) or high-dose corticosteroids (19,20). Though the detailed mechanisms are not completely understood, corticosteroids suppress and IVIG modules the activity of immune cells (21). Thus, if the platelet count responds to corticosteroids or IVIG, it helps confirm that the cause is likely immune-related rather than due to other factors. Nevertheless, response to treatment is a suboptimal method for diagnosis, as it cannot be applied to new patients who have not yet received treatment.

1.1.2 Problems with Diagnosing by Exclusion

Diagnosing by exclusion involves several investigations and can lead to incorrect conclusions since many other thrombocytopenic disorders resemble ITP. In a study of 614 patients enrolled in the McMaster ITP Registry (from 2010 – 2016), 1 in 7 patients (15%) who

presented for a hematologist consultation for evaluation of thrombocytopenia were either mislabeled as ITP when they really had some other platelet disorder or mislabeled as some other platelet disorder when they really had ITP (22). In this study, the diagnosis of ITP was established by two experienced hematologists at a tertiary academic clinic. It is anticipated that the frequency of misdiagnosis would be substantially higher in non-specialized hematology clinics. Errors in diagnosis result in delays in access to proper care, unnecessary exposures to therapies, excess cost, and patient distress and anxiety (23–26).

For most adults, ITP is a chronic disease that requires ongoing monitoring and treatment by a hematologist. Patients with ITP often require several treatments to improve platelet count levels and reduce the risk of bleeding. First-line therapies are broadly immunosuppressive and include corticosteroids and IVIG (18,27–30). Use of these therapies results in relatively rapid, but transient responses (18,27–30). Second-line therapies include rituximab, thrombopoietin receptor agonists, and the spleen tyrosine kinase inhibitor fostamatinib (31,32). Due to error-prone approach of diagnosing by exclusion, many patients who receive these treatments do not end up having ITP at all, adding unnecessary risk and cost. The overuse of corticosteroids can lead to significant toxicities, including cognitive changes, obesity, hypertension, and diabetes (33). IVIG therapy must be used judiciously due to its high cost and looming shortage in Canada (34). Splenectomy and immune suppressant therapies are associated with an increased risk of infection and impaired vaccine responsiveness (35). Thrombopoietin receptor agonists are also very costly, ranging from $6,000 to $8,000 (36). In addition, reports have documented frequent unnecessary exposures to medications and splenectomies in patients suspected of ITP (7,37,38). It is also worth noting that these errors in diagnosis have likely resulted in case mix in clinical

trials, calling into question the validity of their findings. Thus, it is important that the ITP diagnosis can be established accurately early in the disease course.

1.1.3 Other Approaches to Diagnosing Immune Thrombocytopenia

Besides a treatment response, disease biomarkers have also been investigated in ITP, but none have been shown to be useful diagnostic tests. Platelet antibody testing has been studied over many years. In a systematic review (n=1,395), the overall sensitivity of platelet antibody testing was poor (53%; 95% CI, 44-61%) (39). Another marker is a high immature platelet fraction (>8.5%), which correlates with the more readily available mean platelet volume (9). The immature platelet fraction, while promising in small observational studies (n= 27 – 122 patients) (40–42), is neither useful on its own nor widely accessible in practice. Thus, there is currently no single laboratory test or specific investigation that can distinguish ITP from non-ITP in practice.

1.2 Clinical Prediction Models

A clinical prediction model (CPM) combines several characteristics to predict outcomes of individual patients to inform diagnosis or prognosis in clinical settings (43). CPMs can be developed for different purposes. They may combine predictor variables to assess the effects of predictor variables included in a model (43). CPMs can also be used to provide absolute risk estimates for individual patients (44). Logistic regression is often used to develop CPMs for diagnosis, as it allows for modelling of a binary outcome (disease vs. no disease) (45).

CPMs have become a part of routine care in various areas of medicine, including nephrology. For example, the Kidney Failure Risk Equation was developed to predict the

progression of chronic kidney disease to kidney failure (46). The equation was initially developed using data from patients who were referred to nephrology, however the equation is now used in primary care to help decide if the patient should be referred to nephrology (47). In hematology, the SOX-PTS score was developed to predict the occurrence post-thrombotic syndrome (PTS) after proximal deep vein thrombosis (DVT) based on the presence of selected risk factors (48).

1.2.1 Stages of Development of Clinical Prediction Models

Generally, the stages of development of CPMs can be separated into four stages: 1) derivation and internal validation; 2) external validation; 3) modification; and 4) impact evaluation.

Derivation and internal validation

Candidate predictor variables for the outcome of interest are identified based on expert opinion or a literature review. It is important consider the practicality of potential predictor variables. Predictor variables are most useful if they are clearly defined, objectively measured, and easily available, as it would allow the overall model to be generalizable (49). Another consideration is the correlation between predictor variables. If two predictor variables are highly correlated, only one should be included in the multivariable model (49). Both variables may be individually correlated to the outcome of interest, but when combined in the model, each becomes less predictive. Selecting one of the two variables requires consideration of their individual correlation with the outcome of interest, but also the practicality of obtaining the data for the variable (e.g., availability/accessibility, frequency of missing values, etc). Successful

CPMs are parsimonious, meaning the model is simplified (i.e., less predictor variables) but possesses great explanatory predictive power (50). Limiting the number of predictor variables improves the user experience and thus the likelihood of routine use in clinical practice. There are several methods of eliminating candidate predictor variables and selecting the variables to be included in the final model, including backward elimination, forward selection, stepwise selection, and machine learning techniques (45,50).

When deriving a CPM, the final model may be overfitted, meaning the model is too specific to the derivation dataset. This occurs because the random variation within the derivation dataset is included in the associations between the predictor variables and the outcome (49). As a result, the performance of a CPM in the dataset from which it was derived is likely to be overly optimistic. Internal validation addresses this issue by estimating the potential for overfitting the model and the optimism in the model's performance using the derivation dataset (50). There are various methods of internal validation, but the preferred approaches are bootstrap resampling or *k*-fold cross-validation (49).

External validation

The internal validation of a CPM demonstrates the model's reproducibility, which is its ability to perform accurately across new samples from the same population (51). Following the internal validation, an external validation should be conducted to demonstrate the model's transportability, which is its ability perform well across samples from different but comparable populations (51). When a CPM is applied to new patients (i.e., an external dataset), it performs worse relative to its performance on the derivation dataset, even after internal validation. Thus, if

a CPM is intended for wide use, it must be externally validated. External validation studies may be conducted retrospectively or prospectively.

Modification

The performance of the model in an external validation study dictates the next steps for the CPM. If the CPM is found to have a low predictive accuracy, it may be rejected or updated to improves its predictive accuracy (50). There are several ways to modify the model, including adjusting the baseline risk, adjusting all predictor variable weights, adjusting a single predictor variable weight, or adding a new predictor variable (50). After these modifications are made, additional testing and calibration may be required.

Impact evaluation

Following its successful external validation, the impact that the CPM has on clinical practice (healthcare providers and/or individual patients) should be evaluated in a clinical impact study (50). The preferred method of assessing impact is a randomized control trial, where individuals, healthcare providers, or centres are randomly assigned to either the intervention (i.e., using the CPM) or standard care (51).

1.2.3 A Clinical Prediction Model for Immune Thrombocytopenia

A reliable, accurate method of diagnosing ITP is lacking, even though patients with thrombocytopenia represent one of the most common reasons for referral to a hematologist. An improved diagnostic strategy is needed to avoid potentially harmful exposures to unnecessary treatments, avoid anxiety over a lack of diagnosis, and improve education about ITP.

To address this problem, a CPM for the diagnosis of ITP, called the Predict-ITP Tool, was designed for use by hematologists in their assessment of patients with thrombocytopenia in the outpatient setting (52). For the model, a platelet count response after high-dose corticosteroids or high-dose IVIG during follow-up was used as the definition of ITP. A platelet count response was defined as an increase in platelet count to 50 x 10^9/L or higher and doubling from baseline within 4 weeks (8). From a derivation cohort of consecutive patients with thrombocytopenia with (n=91) and without ITP (n=432), a penalized logistic regression model was used to identify independent baseline variables associated with the ITP diagnosis. The initial list of 19 variables was reduced to four independent variables after elastic net regulation and five-fold cross validation. The risk of bias was assessed using established criteria for CPMs (53), including consistency in the assessment of predictor variables and blinding of outcomes.

Variables in the final model were: 1) the platelet variability index (a measure of platelet fluctuations) (43); 2) the lowest platelet count value; 3) the maximum mean platelet volume (MPV); and 4) a history of major bleeding at any time. The platelet variability index is a novel score that was developed to capture the degree of platelet count fluctuation over time (the more fluctuations, the higher the score) (54). High fluctuations occur in patients with ITP (55,56) but not in non-ITP (57,58) or healthy individuals (54). Three or more sequential platelet count values in the preceding 5 years are needed to calculate the platelet variability index, which are readily available at the time of the initial hematology consultation. In the derivation cohort study, the median lowest platelet count level was 16 x 10^9/L (interquartile range [IQR] 3, 46) for ITP vs. 78 x 10^9/L (IQR 44, 113) for non-ITP (54). MPV is a continuous variable that is available on routine blood test results. A high MPV is an indicator of ‘young’ platelets, which are typically present in patients with ITP. It is correlated with high immature platelet fraction (9), which showed

potential to distinguish ITP from other thrombocytopenic conditions (40–42), but MPV is more accessible. Major bleeding is defined as the occurrence of any of the following: 1) >5 bruises with size >2 cm and/or diffuse petechiae; 2) multiple oral blood blisters and/or gum bleeding >5 min; 3) epistaxis >5 min (per episode); 4) gross blood loss from the gastrointestinal tract; 5) macroscopic hematuria; 6) menstrual bleeding (more than spotting) outside of the menstrual period or a very heavy period; 7) pulmonary hemorrhage; 8) retinal hemorrhage; or 9) intracranial hemorrhage, according to the ITP Bleeding Score (59). The independent association of each of the four predictor variables and the probability calculation for the tool are shown in Table 1.

In the internal validation study, the tool demonstrated good discrimination and excellent calibration (*see Section 3.5*) with an optimism-corrected c-statistic of 0.83, calibration slope of 0.88 and calibration-in-the-large for all performance measures <0.001 with standard error <0.001 [expected optimism was 0.016 (95% CI: 0.014, 0.018) for the c-statistic and 0.12 (95% CI: 0.11, 0.13) for the calibration slope] (52). In a preliminary analysis, various threshold values of the Predict-ITP Tool were considered to classify the patient as having ITP. Threshold values were determined using various approaches, including maximizing the Youden-index (calculated by sensitivity and specificity) (60), F1-score (calculated by sensitivity and positive predictive value) (61), and Cohen's Kappa (to measure the observed and expected agreements using confusion matrix) (62). Of the three approaches considered, maximizing the Cohen's Kappa was preferred as it resulted in the highest specificity. Prioritizing specificity positions the Predict-ITP Tool as a method of ruling-in ITP among patients with thrombocytopenia. Maximizing Cohen's Kappa resulted in a cut-off of 0.35, which yielded a sensitivity of 56% and specificity of 90% for the ITP diagnosis.

Having internally validated the Predict-ITP Tool, the next stage in its development is to assess the performance of the tool in an external dataset. This thesis focuses on the development of a study protocol for this external validation study.

2.0 OBJECTIVES

2.1 Overall Objective

The objective of the proposed study is to perform an external validation of a CPM for the diagnosis of ITP (the Predict-ITP Tool) in both academic and community outpatient settings for patients referred to a hematologist for evaluation of thrombocytopenia.

2.2 Specific Aims

1. To develop a study protocol for the external validation of the Predict-ITP Tool.
2. To determine the feasibility of the Predict-ITP Tool.
3. To develop the process for determining the clinical diagnosis of ITP by independent adjudication.
4. To develop, test, and refine the data collection process.

3.0 METHODS

3.1 Description of Study Population

The target population of the proposed study is consecutive patients who presented to a hematologist for the evaluation of thrombocytopenia in whom the diagnosis of ITP may be suspected. Patients from seven academic and four community-based hematology clinics (11 centres in total) will be included. Retrospective data will be collected for patients presenting for

evaluation between January 1, 2010 and December 31, 2021. This timeframe was selected to ensure that a sufficient number of patients are available from each clinic and that at least 8 months of follow-up data are available by the time data collection commences. Data on historical lab values (e.g., platelet count, mean platelet volume, etc.) may be collected up to 5 years prior to the initial hematology consultation, meaning the data collection window is January 1, 2005 to December 31, 2021.

Inclusion criteria

- Consecutive outpatients who were referred to a hematologist for evaluation of thrombocytopenia
- At least 1 platelet count measurement <100 x 10^9 /L at any time up to and including the initial hematology consultation
- At least 3 platelet count measurements in the 5 years leading up to and including the initial hematology consultation
- Initial hematology consultation was between January 1, 2010 and December 31, 2021

Exclusion criteria

- None

Identifying eligible patients

The following procedure may be used and/or adapted by the participating centres to screen and identify eligible patients:

1. Identify all patient visits to the outpatient hematology clinic with a billing code for thrombocytopenia that occurred between January 1, 2010 and December 31, 2021.
2. Identify the date of each patient's initial hematology consultation for thrombocytopenia in the outpatient clinic, and exclude all follow-up visits from the list.
3. Organize the medical record numbers in chronological order by the date of the initial hematology consultation for thrombocytopenia.
4. Manually review charts to identify patients who meet both of the following criteria:
    a. had at least 1 platelet count $<100 \times 10^9$/L at any time up to and including their initial hematology consultation visit; and
    b. had at least 3 platelet count measurements available in the 5 years leading up to and including their initial hematology consultation.

3.2 Sample Size

As part of the protocol development, the sample size for the proposed study was determined according to the recommendations by Riley et al. (63) to ensure precise discrimination and calibration estimates of the model. The sample size calculation is based on the targeted c-statistic, the standard error of the c-statistic, and the anticipated prevalence of the outcome event (i.e., ITP diagnosis) in the external validation cohort (63). In this study, the target c-statistic is 0.8, indicating good discrimination (43,64). The standard error of the c-statistic is 0.02, which is based on the findings from deriving and internally validating the Predict-ITP Tool (52). In the derivation cohort, the prevalence of ITP was 0.16 (52). It is possible that the prevalence of ITP will be lower in non-specialized clinics included in the external validation study. Thus, the anticipated prevalence of ITP in the study population is 0.15. Riley et al. (63)

used an iterative process to determine the sample size needed to satisfy the target c-statistic, its standard error, and the anticipated prevalence of ITP in the validation cohort. Using the provided R code and given a target c-statistic of 0.80 (standard error 0.02) and a prevalence of ITP of 0.15 in the study population, a sample size of 960 patients (estimated 144 patients with ITP) is required for this external validation study. With 11 centres and an 11-year time horizon, each participating centre will provide data for 80-100 eligible patients.

3.3 Study Design

The proposed study will be a multi-centre retrospective cohort study.

3.3.1 Data Collection

Data from medical records will be collected retrospectively. Data will be collected using electronic case report forms (eCRFs) in REDCap, an electronic data collection software. Table 2 outlines the data that will be collected from each patient chart. Data will be abstracted by a single data abstractor at each participating centre. If a centre opts to have more than one data abstractor, additional steps are required (*see Section 3.3.7*).

Data to collect from the initial hematology consultation

The following data will be collected from each patient's initial hematology consultation: baseline demographics (e.g., year/month of birth, sex, race, etc), duration of thrombocytopenia, diagnosis, pregnancy status, current and past medication use, past splenectomy, complete blood count values available in the 5 years leading up to and including the initial consultation, and the occurrence of a major bleed up to and including the initial consultation. A major bleed is defined

according to the ITP Bleeding Score (59) as any of the following: 1) more than >5 bruises with size greater than >2 cm and/or diffuse petechiae; 2) two or more multiple blood blisters and/or gum bleeding lasting more than >5 minutes; 3) epistaxis lasting more than >5 minutes (per episode); 4) gross blood loss from the gastrointestinal tract (grossly bloody or black stool); 5) macroscopic hematuria; 6) menstrual bleeding (more than spotting) not at time of the menstrual period or a very heavy period; 7) pulmonary hemorrhage; 8) retinal hemorrhage; or 9) intracranial hemorrhage.

The coordinating centre (McMaster University) will use the Predict-ITP Tool to calculate the predicted risk of ITP at the time of initial consultation for each patient using the following variables collected by the centres: 1) platelet variability index (calculated as previously described (54) based on 3 or more consecutive platelet count values); 2) lowest platelet count value; 3) highest mean platelet volume; and 4) history of a major bleed. Data from all patients will be used, even if one or more of the four variables is missing.

Data to collect from the follow-up visits

Data will also be collected from patient charts for follow-up hematology visits until December 31, 2021. The patient's diagnosis, pregnancy status, and other ITP related variables will be collected from each follow-up visit.

Pilot testing the data collection process

In preparation for conducting the proposed study, the data collection process has been tested internally using data from the McMaster University Medical Centre. Using the McMaster ITP Registry, 10 patients who were challenging to diagnose (i.e., their diagnosis changed at least

once during follow-up) were identified for testing purposes. Five patients had a final diagnosis of ITP, and five patients had a final diagnosis of non-ITP. The eCRFs were pilot tested by two research assistants who were previously unfamiliar with the study. They attended a training session, were provided with the operations manual, and instructed to complete the eCRFs using the electronic medical record (EMR). This process is identical to one that will be used with data abstractors at the participating centres. Reflecting on both the operations manual and the eCRFs, the research assistants were asked to consider the following while completing the testing process:

- Overall impressions
- The time required to complete the CRFs for each patient
- Issues/errors
- Unclear/confusing instructions
- Parts of the eCRF that required a lot of work to complete and/or data that was difficult to find in the EMR

3.3.2 Outcome Adjudication

A subgroup of patients included in the proposed study will have their diagnosis determined by a blinded independent outcome adjudication committee. This subgroup will consist of patients whose diagnosis (as determined by the local treating hematologist at the end of follow-up) is considered ambiguous as per one of the following criteria: 1) no clear cause of thrombocytopenia was identified at the end of follow-up (December 31, 2021); 2) the diagnosis changed from one visit to another; or 3) the thrombocytopenia occurred in the context of pregnancy. In a recent study, 16.5% of consecutive patients with thrombocytopenia (n=789) met these adjudication criteria (65). It is anticipated that 158 patients in the proposed study will meet

the adjudication criteria (16.5% of 960 target sample size). In addition, a random sample of 10% of patients from each centre who do not meet these criteria will also be adjudicated to verify the clinical diagnosis in routine practice (10% of 802 = 80 patients). Overall, the number of patients requiring adjudication is estimated to be 238.

All data relevant to adjudicating the clinical diagnosis will be provided to the adjudication committee for review, including investigations, laboratory tests, imaging studies, and diagnostic procedures. The adjudication committee will consist of a pool of five independent hematologists with experience in the management of patients with ITP. Individual patients will be adjudicated in duplicate/ triplicate based on information from the patient's medical chart and the diagnosis will be established independently and by consensus.

Adjudication rules

The following rules will be used to guide adjudicators for making the diagnosis of ITP in this proposed study:

1. An increase in baseline platelet count to 50 x 10^9/L or higher and doubling of the baseline platelet count level within 4 weeks of starting high-dose IVIG or high-dose corticosteroids.
2. ITP could be primary or secondary to an underlying cause (autoimmune disease, infection, lymphoproliferative disease).
3. Other causes of thrombocytopenia excluded.

The above rules and additional guidance on establishing the diagnosis of thrombocytopenic disorders (Table 3) were established based on a previous study (65). In this

previous study, the diagnosis could be resolved for 95.6% of patients (n=92) using this approach (65).

Patients will be classified as definite ITP, suspected ITP or non-ITP as defined below:

- Definite ITP:
    - Achieved a platelet count response within 4 weeks after the administration of high-dose IVIG (1 – 2 g/kg) or high-dose corticosteroids (prednisone 1mg/kg daily; or dexamethasone 40mg daily for 4 days) (8) and other possible causes of the thrombocytopenia were excluded
- Suspected ITP:
    - ITP is the most likely diagnosis (i.e., other possible causes of the thrombocytopenia were excluded), but patients never received a treatment trial of high-dose IVIG or high-dose corticosteroids
- Non-ITP:
    - Did not achieve a platelet count response within 4 weeks after the administration of high-dose IVIG (1 – 2 g/kg) or high-dose corticosteroids (prednisone 1mg/kg daily; or dexamethasone 40mg daily for 4 days) (8) or non-ITP is the most likely diagnosis (i.e., there is evidence indicating the thrombocytopenia is due to a non-immune cause), but the patient never received a treatment trial with high-dose IVIG or high-dose corticosteroids

Adjudication process

The adjudication will occur as follows:

1. Two adjudicators will review each patient's data independently and submit the diagnosis on REDCap.
2. Disagreements over the diagnosis will be resolved by consensus at a monthly virtual adjudication meeting, led by the Chair.
    a. The meetings will be approximately 2 hours long and involve the review of 15-20 patients. All adjudications are anticipated to be completed in 12-18 months.
3. The Chair will enter the final adjudicated diagnosis in REDCap.

Training and calibration of adjudication committee

The adjudication rules and process outlined above will be discussed in an initial training session with the adjudication committee. Prior to commencing independent adjudication, a calibration exercise will be conducted to reduce the variability in assessments among adjudicators. Independently and blinded to the diagnosis determined by the local treating hematologist and to each others' assessment, all five adjudicators will examine the relevant data of 20 patients who meet the criteria for adjudication. The 20 patients will be selected by the study's principal investigator as follows: 10 patients with ITP and 10 patients with non-ITP.

After the first 10 patients have been assessed, adjudicators will discuss their assessments in a meeting, identify any reasons for disagreement, and clarify the criteria for establishing the diagnosis (i.e., refining the adjudication rules and statements in Table 3). The remaining 10 patients will then be assessed independently, and a second meeting will be held to discuss disagreements and establish greater clarity on the diagnosing criteria. Since the adjudicators will complete their assessment independently, the outcomes assessed will be balanced (10 ITP vs. 10

non-ITP), and there are more than two adjudicators, Fleiss kappa is an appropriate method of assessing inter-rater reliability (66). Based on past studies (66,67), a Fleiss kappa of 0.8 will be used as the threshold level for excellent agreement among the adjudicators. If this value of 0.8 is achieved in the calibration exercise, the independent, duplicate adjudication can proceed.

3.4 Reference Standard

The predicted diagnosis determined by the Predict-ITP Tool will be compared with the Reference Standard, which is either the clinical diagnosis determined by the local treating hematologist at the end of follow-up, or the clinical diagnosis determined by the adjudication committee for patients who meet criteria for adjudication (Figure 1). In accordance with the PROBAST guidelines (53), the diagnosis predicted by the Predict-ITP Tool will be unknown to the research personnel and centre investigator when recording the diagnosis at the end of follow-up and will also be unknown to the adjudication committee. This will be accomplished by running the model for all patients centrally at the coordinating centre (McMaster University).

3.5 Assessment of Predictive Performance

The main study objective of the proposed study is the performance of the Predict-ITP Tool, which will be measured by calibration and discrimination compared with the clinical diagnosis. The tool performs well if its predicted diagnosis corresponds with the actual diagnosis. Calibration refers to the agreement between the predicted and actual diagnoses (43,64). It is assessed by estimating the calibration slope (i.e., the slope of the calibration plot) and calibration-in-the-large as well as visual inspection of the calibration plot. The calibration slope assesses how well the predicted probabilities align with the observed outcomes across

various risk levels (43,64). The calibration-in-the-large is a measure of the model's overall calibration when considering the average predicted probability across all observations (43,64). The closer the calibration slope is to 1 and the calibration-in-the-large is to 0, the better the model's calibration (43,68). The calibration plot graphs the actual probability on the y-axis and the predicted probability on the x-axis (68). A straight diagonal line down the middle represents a model with perfect agreement between predicted and actual outcomes (68). To assess the performance of the model in question, one examines how close the model tracks along the diagonal line. Researchers make conclusions about the agreement between predicted and actual outcomes by inspecting the calibration plot and considering the calculated calibration slope (i.e., closeness to 1) and calibration-in-the-large (i.e., closeness to 0).

Discrimination, the ability of the model to distinguish a patient likely to have ITP from a patient unlikely to have ITP (43,64), is another indicator of performance. The discriminative ability of a model is quantified by the concordance (c) statistic. Since the outcome is binary (i.e., ITP or non-ITP), the c-statistic is identical to the area under the receiver operating characteristic (ROC) curve (43,64). The ROC curve plots sensitivity on the y-axis and 1 – specificity on the x-axis (43,64). A c-statistic of 0.5 means the model does not discriminate any better than chance, and a c-statistic of 1 indicates perfect discrimination. The closer the c-statistic is to 1, the greater the model's discriminative ability (43,64). In the literature, the threshold c-statistic indicative of good discrimination ranges between 0.7 and 0.8 (64,68). For this reason, the target c-statistic for this external validation study is 0.8, which would indicate good discrimination according to the highest threshold.

3.6 Data Management

All data will be manually entered into an eCRF in REDCap by centre personnel. In order to detect discrepancies and to ensure the data are accurate and complete, computerized edit-checks and manual reviews will be conducted. Local study staff will be responsible for attending to any queries that may arise from participating centres. If any changes are made, the reason for these changes, the name of the person who conducted the changes, as well as the time and date these changes occurred will be documented in REDCap. All data changes in the database that occur after the first data entry will be marked with an electronic audit trail system.

3.7 Calibration of Multiple Data Abstractors

Data collection will be completed by a single data abstractor at each participating centre of the proposed study. If a participating centre chooses to use more than one data abstractor, data abstractors will complete a calibration exercise to ensure consistency. Data will be collected in duplicate for the first 10 patients. Discrepancies between abstractors for the first 10 patients must be reviewed and resolved by local study staff. This will require a detailed look at the patient chart. A third party may be consulted if there are unresolvable disagreements. Participating centres are encouraged to make changes to their centre-specific operating procedures to clarify and ensure consistent data collection going forward. Once the disagreements between abstractors are resolved, the centre is responsible for collecting data in duplicate for another 10 patients to ensure agreement between abstractors and thus consistency of data collection. This process is repeated until local data abstractors are fully calibrated, meaning there is 100% agreement between abstractors in all data points collected. At that point, duplicate data collection for the remaining patients is not required.

3.8 Data Analysis Plan

Descriptive statistics will be presented to compare the original sample for model derivation and the external sample for validation. Continuous variables will be summarized with means and standard deviations or medians and interquartile ranges. Categorical variables will be reported as frequencies and proportions. For each patient, the probability of the ITP diagnosis will be determined using the regression coefficients of the originally derived model. If the prevalence of ITP differs between the derivation cohort and external validation cohort, the intercept from the external validation cohort will be estimated and used.

The complete external dataset will be used to assess the discrimination and calibration of the model against the reference standard, (i.e., the clinical diagnosis determined by the local treating hematologist at the end of follow-up or the adjudicated clinical diagnosis). The primary analysis will include all patients to demonstrate discrimination of patients with ITP (combining definite ITP and suspected ITP) from patients without ITP. The secondary analysis will assess discrimination of patients with definite ITP from patients with non-ITP. Discrimination will be measured using the c-statistic, and calibration will be measured by estimating calibration slope, calibration-in-the-large and by visual inspection of the calibration plot. The ratio of expected to observed events will be calculated, which should be close to 1 if the model calibrates well in the validation dataset. Based on our experience with model derivation, <10% missing data is expected, given that the data required for the predictor variables and outcome assessments are routinely available for the study population.

Missing data will be imputed where possible by chained equations using the mice package in R (69). Variables in the multiple imputation models include platelet count, MPV, history of bleeding (required for the tool); and use of blood thinners and age (which influence

bleeding). Ten imputed datasets will be generated and measures of discrimination and calibration and will be pooled using Rubin's rules (70).

To mitigate spectrum bias, consecutive patients presenting for hematology evaluation will be enrolled, which will ensure a mix of patients with a more obvious and less obvious diagnosis. In addition, a sensitivity analysis that excludes patients at the extremes of platelet count levels will be conducted. Since these groups can be considered most likely (lowest platelet levels) or least likely (highest platelet levels) to have ITP, this will focus the analysis on the most ambiguous population, where a prediction model is particularly needed. Model performance in the subgroup of patients who never received a trial of IVIG or corticosteroids will also be evaluated.

There are known sex-differences in the prevalence of ITP, and minor sex-differences in platelet count levels and platelet parameters, including MPV (71,72). Platelet counts tend to be lower in pregnant females (73), and women experience more bleeding symptoms than men due to gynecological bleeding and pregnancy-related bleeding events (74). The impact of sex on model performance will be evaluated and disparate impact analysis will be conducted (75) to evaluate algorithm bias on sex groups. Specifically, prediction scores will be generated using the Predict-ITP Tool to facilitate a comparison of the performance measures of the model between sex groups. The impact of race on model performance will be evaluated since race may influence ITP prevalence and severity (76) and baseline platelet count levels (77). Neither gender nor ethnicity can be captured reliably in this retrospective study since it was not routinely recorded in patient charts during the time period of data collection.

3.9 Challenges and Mitigation Strategies

The participating centres that have been recruited to the study have indicated their ability to provide the requisite sample of 80-100 patients per centre. Nevertheless, the study team will be in regular communication with each participating centre as they identify eligible patients. If there is a risk of a lower than expected sample size, the study team is prepared to recruit additional clinics through the Canadian Hematology Society and Canadian Transfusion Research Network.

As described previously, a random sample of 10% of patients from each centre who do not meet the adjudication criteria will also be adjudicated to verify the accuracy of routine clinical diagnoses. A potential issue may occur if there is substantial disagreement between the routine clinical diagnoses and the adjudicated diagnoses. If this occurs, additional patients who did not meet the adjudication criteria will be adjudicated until satisfactory agreement has been reached.

If the model does not show good calibration in the external validation, the model can be adjusted accordingly to improve its calibration (50). Additional updates could be considered, including adjusting the predictor variable weights or adding a new predictor variable, however this would require a greater understanding of the how the predictor variables differ between the derivation and validation cohorts (50).

4.0 RESULTS AND PROGRESS

4.1 Development of the Study Protocol

A study protocol for the external validation of the Predict-ITP Tool has been developed based on the information presented in the previous sections. It was designed in accordance with

the TRIPOD Statement, which aims to improve the reporting of studies developing, validating, or updating a CPM (78), and PROBAST, which aims to assess the risk of bias and applicability of CPM studies (53).

4.2 Feasibility of the Predict-ITP Tool and an External Validation

Along with the development of the study protocol, a survey was conducted to assess the feasibility of the Predict-ITP Tool. The online survey was targeted at Canadian hematologists with expertise in managing patients with thrombocytopenia who will be participating in the external validation study. Respondents were asked about the feasibility of obtaining the predictor variables required for the tool at the time of the initial hematology consultation, including 3 or more platelet count measurements, MPV, and major bleeding events. The hematologists were also asked to identify current barriers for establishing the ITP diagnosis in practice.

The survey was completed by eight hematologists. Six respondents practiced at an academic hospital, and two practiced at a community clinic. All respondents treated adults, and one respondent treated both adults and children. Seven of the hematologists stated that obtaining at least 3 platelet count measurements were feasible and obtaining 8 platelet count measurements were feasible or somewhat feasible. All respondents stated that MPV was feasible or somewhat feasible, and bleeding assessments were feasible for gastrointestinal, ocular, and intracranial bleeding. However, four respondents reported that bruises and oral mucosal bleeding were not feasible. Respondents identified the following barriers to establishing the ITP diagnosis in practice: 1) the lack of a confirmatory diagnostic test; 2) the high number and wide variability of investigations needed to exclude other thrombocytopenic conditions; and 3) the undesirable need to use a treatment trial of IVIG or corticosteroids to confirm the diagnosis.

Overall, this study found that the variables required for the Predict-ITP Tool were readily available at the time of the initial hematology consultation, but that details of skin and oral mucosal bleeding were not always available when accessing medical charts retrospectively. Most importantly, all the hematologists felt that a CPM for the diagnosis of ITP would be useful in practice to improve accuracy, reduce unnecessary treatments, and cut down on costly investigations.

Hematologists with the capacity to participate in the external validation were recruited via email. The recruitment letter included an overview of the intended study as well as the expected number of participants required from each centre (i.e., 80-100 patients). Eleven participating centres have been recruited to date. All sites have fully executed data transfer agreements, and 10 of 11 have obtained ethics approval. A decision-support algorithm to identify eligible patients was developed and provided to research personnel at each site, allowing them to tailor the algorithm to their specific centre. Two investigator meetings were held to introduce the study to centre personnel and describe the data collection process. The progress made and level of engagement with each site are strong indicators that conducting an external validation study is feasible.

4.3 Operationalizing Data Collection

Data will be collected using eCRFs in REDCap. In preparation for conducting the proposed study, the eCRFs have been developed to capture demographic data and the necessary data from the patient's initial hematology consultation and all subsequent follow-up visits during the study period. An operations manual was created to support research personnel at participating

centres in completing various study procedures, including identifying eligible patients, assigning participant identification codes, and completing the eCRFs.

4.3.1 Developing the Outcome Adjudication Process

The proposed study will involve independent outcome adjudication (*see Section 3.3.2*). Developing the eCRF for adjudication involved identifying the various diagnoses that are made among patients who are referred to hematology for evaluation of thrombocytopenia (Table 3). Subsequently, the investigations (e.g., laboratory tests, imaging studies, etc) used to support each diagnosis were identified. The results of any such investigations conducted for each patient within the study period would need to be provided to the adjudication committee to allow them to make an informed decision regarding the patient's diagnosis. The eCRF that captures these data is currently under development.

4.3.2 Pilot Testing the Data Collection Process

In preparation for conducting the proposed study, the data collection process was tested by two research assistants who were previously unfamiliar with the study. They attended a training session and were given the operations manual to guide them in entering data into the eCRFs, which aligns with the intended process at each participating site. The research assistants were asked to reflect on the overall data collection process, including the operations manual and eCRFs, and provide feedback on their overall impressions, the time required to complete the eCRF, and any issues, errors, challenges, or sources of confusion.

The research assistants found the eCRFs to be very straightforward and simple to input data. They noted that the eCRFs had several built-in functions to ensure data was entered in the

appropriate format and no required data was missing. The real time feedback regarding any errors were helpful. They also found that the operations manual complemented the eCRFs well and clarified most issues. Both research assistants mentioned that it required approximately one hour to complete all of the eCRFs for each patient, with the exception of one patient who required close to three hours due to having 33 follow-up visits. The median number of follow-up visits in the sample of patients was 7 (IQR 2, 12).

There were discrepancies in the date when thrombocytopenia was first discovered in two patients, which was due to how and from where this information was determined. One research assistant relied on the clinic note associated with the initial hematology consultation, while the other reviewed the platelet count values reported in the EMR. This prompted a revision to both the eCRFs and operations manual, where it was clarified that the earliest of the two dates should be recorded.

The research assistants reported that the laboratory data were easy to find and record. They both mentioned that MPV values were not available for one patient. One source of confusion was the discrepancy in the units for hemoglobin between the eCRF and the EMR. A note was added to both the eCRFs and the operations manual to prompt data abstractors to verify that the units for all laboratory values matched before entering them into the eCRF.

As anticipated based on the feasibility study, the research assistants reported that bleeding events were rarely mentioned in the clinic notes and sometimes vague terminology was used, such as “heavy bleeding” or “significant bruising”. The eCRF asks the data abstractors, “Up to and including the initial hematology consultation, did the patient ever have any of the following characteristics?”, and then lists each element of ITP Bleeding Score (*see Section 3.3.1*). To address this issue, the following category was added to the eCRF, “Bleeding of significant nature

that was not captured above (including bleeding described as serious, severe, catastrophic, major or life-threatening, etc)".

One of the more challenging aspects of the eCRFs was recording the platelet count values that occurred four weeks before and after high-dose corticosteroids or high-dose IVIG. In some cases, a clinic note mentioned the platelet count increased but did not mention by how much. As a result, it is unknown if a platelet count response (defined as an increase in platelet count to 50 x 10^9/L or higher and doubling from baseline within 4 weeks) was achieved. This finding prompted a revision to the eCRF to capture verbatim what was said in the clinic note regarding the platelet count increase. This information will be included in the data shared with the adjudication committee to aid them in determining the diagnosis.

The research assistants observed that the hematologist did not come to a conclusive diagnosis until after several follow-up visits. During this time, the hematologist had multiple suspected diagnoses, which resulted in confusion as to what to record for the diagnosis at various follow-up visits. In response to this issue, the operations manual was revised to clarify that if one diagnosis is favoured among multiple suspected diagnoses, record the favoured diagnosis as the diagnosis in the eCRF. If there are two or more equally likely diagnoses, then record the diagnosis as unknown in the eCRF. Similarly, one research assistant expressed confusion over what to do if the hematologist is awaiting the results of an investigation to confirm a diagnosis. In this case, the operations manual was revised to clarify that the diagnosis should be recorded as unknown in the eCRF if the hematologist is speculating and is awaiting an investigation that confirms the diagnosis. However, if the hematologist favours a particular diagnosis and is not awaiting an investigation, the favoured diagnosis should be recorded as the diagnosis in the eCRF.

Overall, the pilot testing the eCRFs and operations manual demonstrated that they were well-designed and straightforward to follow. The average time required to complete the eCRFs was one hour, which aligned with the study team's expectations. Through the testing process, the research assistants identified potential sources of confusion that were addressed through additional instructions to both the eCRF and operations manual.

5.0 DISCUSSION

5.1 Summary of Progress

This thesis project focused on the development of the protocol for the external validation of the Predict-ITP Tool, a novel CPM for the diagnosis of ITP. The protocol was carefully developed to align with best practice guidelines, including the TRIPOD Statement (78) and PROBAST (53). A feasibility study was conducted, which demonstrated that the variables required for the Predict-ITP Tool were readily available at the time of the initial hematology consultation, indicating that its use in real-world settings is encouraging. This study also reported that there is a recognized need for a CPM to help hematologists diagnose ITP among patients presenting for evaluation of thrombocytopenia.

Participating centres were recruited and all but one site has completed study start-up activities. The eCRFs and operations manual have been developed, tested, and revised. Additional testing and revisions will be necessary for the eCRF capturing the data required for adjudication. The pilot testing of the eCRFs and operations manual were helpful in identifying potential sources of confusion and facilitated clearer instructions that will ultimately be shared to the data abstractors at the participating centres.

5.2 Strengths and Limitations of Study Protocol

Since the study protocol was intentionally developed to align with the TRIPOD Statement (78) and PROBAST (53), it possesses several strengths. The participants are selected in this study in a manner that does not introduce bias. Though the data will be collected retrospectively, the data source is appropriate, as the methods used to collect the data are defined and consistently applied for inclusion and exclusion criteria, assessment of predictor variables, and outcome determination across a predefined follow-up period (53). The inclusion and exclusion criteria are purposely broad to represent the intended target population, which are patients referred to hematology for evaluation of thrombocytopenia. Moreover, consecutive patients will be enrolled, which will ensure a mix of patients with a more obvious and less obvious diagnosis, thereby mitigating spectrum bias (79). The predictor variables include platelet count values, MPV, and major bleeding history. Both platelet count values and MPV are quantitative and reported in standard units, which means they will be assessed in a similar manner for all patients included in the study. Major bleeding history is assessed using the ITP Bleeding Score (59), which reduces the variability in classification. In addition, according to the results of the feasibility study and pilot testing, all three predictor variables are generally available at the time the Predict-ITP Tool is intended to be used (i.e., the initial hematology consultation). Importantly, the study is designed in a way that ensured the predictor variables are assessed without knowledge of the diagnosis as determined by the Predict-ITP Tool. The outcome will be the diagnosis determined by the local treating hematology at the end of the follow-up period. However, if there is ambiguity surrounding this diagnosis and the adjudication criteria are met (*see Section 3.3.2*), the outcome will be the diagnosis determined by the independent adjudication committee. Due to the training and calibration of the adjudicators, this

approach ensures the outcome is determined appropriately and consistently. Another strength of the protocol was using the most current method of determining an appropriate sample size for external validation of a CPM with a binary outcome (63).

Despite the notable strengths of the protocol, there are several limitations. For example, the inclusion of major bleeding data in the model may be a limiting factor. As indicated in the feasibility survey and pilot testing, there was some difficultly in assessing major bleeding retrospectively. However, a prior study has demonstrated that major bleeding can be distinguished from minor bleeding in ITP based on retrospective data (80). Thus, it remains to be seen how much of an issue collecting major bleeding data will be. Furthermore, the Predict-ITP Tool can still operate without this data. Another potential limitation is the inclusion criterion of at least 3 platelet count measurements within five years before the initial hematology consultation. In a previous study, 56.9% of patients (n=919) who were referred to hematology for evaluation of thrombocytopenia met this criterion, meaning the Predict-ITP Tool could not be used for 44.1% of patients. However, this criterion, though restrictive, is unavoidable as 3 platelet count measurements are necessary to run the CPM.

5.3 Future Directions

There are four stages of development of CPMs: 1) derivation and internal validation; 2) external validation; 3) modification; and 4) impact evaluation. This thesis focuses on developing the study protocol for the external validation of the Predict-ITP Tool (stage 2); however, the study team is committed to completing the remaining stages to fully realize the widespread implementation of the tool in practice.

5.3.1 Conducting the External Validation

Prior to starting the external validation study, the eCRFs that capture the data required by the adjudication committee needs to be completed and tested. Along with finalizing this remaining eCRF, the adjudication committee will be established, trained, and calibrated to ensure reduced variability in their assessments. Data collection will commence at each participating site once they have completed their training (i.e., review of updated eCRFs, operations manual, and study protocol). Data collection will likely take 6-8 months, and the adjudication process will begin approximately 2 months into data collection as the patients requiring adjudication are identified. Once data collection and outcome adjudication have been completed, the data will be analyzed and interpreted. The primary focus of this external validation study is the performance of the Predict-ITP Tool in this external cohort of patients. However, other outcomes will be evaluated, including the utility of the outcome adjudication process.

5.3.2 Model Modification and Clinical Impact Study

In the process of completing the external validation study, there is a possibility that new variables are identified, or existing predictor variables require modification. If this is the case, the new model will need to undergo external validation before examining the clinical impact of the Predict-ITP Tool.

The Predict-ITP Tool provides a probability of the ITP diagnosis for clinicians, which can be very informative. However, the threshold level that dictates whether treatment should be started, or additional investigations should be conducted still need to be established. Several factors influence this threshold level, including cost, consequences of false positives and false negatives, and patients' values and preferences. As such, establishing an appropriate threshold

level requires additional work and collaboration with patients and other stakeholders. Once a threshold level has been established, a clinical impact study can be conducted. Ideally, a randomized control trial will be conducted to evaluate the impact of model-driven treatment and management algorithm compared with standard of care on patient-important outcomes (51).

6.0 REFERENCES

1. Bennett D, Hodgson ME, Shukla A, Logie JW. Prevalence of diagnosed adult immune thrombocytopenia in the United Kingdom. Adv Ther. 2011;28(12):1096–104.

2. Terrell DR, Beebe LA, Neas BR, Vesely SK, Segal JB, George JN. Prevalence of primary immune thrombocytopenia in Oklahoma. Am J Hematol. 2012;87(9):848–52.

3. Jiang D, Al-Samkari H, Panch SR. Changing Paradigms in ITP Management: Newer Tools for an Old Disease. Transfus Med Rev [Internet]. 2022 Oct 1 [cited 2023 Jul 17];36(4):188–94. Available from: https://pubmed.ncbi.nlm.nih.gov/36273934/

4. Malik A, Sayed AA, Han P, Tan MMH, Watt E, Constantinescu-Bercu A, et al. The role of CD8+ T-cell clones in immune thrombocytopenia. Blood [Internet]. 2023 May 18 [cited 2023 Jul 17];141(20):2417–29. Available from: https://dx.doi.org/10.1182/blood.2022018380

5. Marini I, Zlamal J, Faul C, Holzer U, Hammer S, Pelzl L, et al. Autoantibody-mediated desialylation impairs human thrombopoiesis and platelet lifespan. Haematologica [Internet]. 2021 Jan 1 [cited 2023 Jul 18];106(1):196–207. Available from: https://pubmed.ncbi.nlm.nih.gov/31857361/

6. Najaoui A, Bakchoul T, Stoy J, Bein G, Rummel MJ, Santoso S, et al. Autoantibody-mediated complement activation on platelets is a common finding in patients with immune thrombocytopenic purpura (ITP). Eur J Haematol [Internet]. 2012 Feb [cited 2023 Jul 18];88(2):167–74. Available from: https://pubmed.ncbi.nlm.nih.gov/21985182/

7. Cines DB, Bussel JB, Liebman HA, Luning Prak ET. The ITP syndrome: Pathogenic and clinical diversity. Blood. 2009;113(26):6511–21.

8. Rodeghiero F, Stasi R, Gernsheimer T, Michel M, Provan D, Arnold DM, et al. Standardization of terminology, definitions and outcome criteria in immune thrombocytopenic purpura of adults and children: Report from an international working group. Blood. 2009 Mar 12;113(11):2386–93.

9. Bodrova V V., Shustova ON, Khaspekova SG, Mazurov A V. Platelet reticulated forms, size indexes, and functional activity. Interactions in healthy volunteers. Platelets. 2021;33(3):398–403.

10. Neunert C, Noroozi N, Norman G, Buchanan GR, Goy J, Nazi I, et al. Severe bleeding events in adults and children with primary immune thrombocytopenia: A systematic review. Journal of Thrombosis and Haemostasis. 2015;13(3).

11. Heitink-Pollé KMJ, Haverman L, Annink K V., Schep SJ, De Haas M, Bruin MCA. Health-related quality of life in children with newly diagnosed immune thrombocytopenia. Haematologica. 2014;99(9):1525–31.

12. Trotter P, Hill QA. Immune thrombocytopenia: improving quality of life and patient outcomes. Patient Relat Outcome Meas. 2018;9:369–84.

13. Grace RF, Klaassen RJ, Shimano KA, Lambert MP, Grimes A, Bussel JB, et al. Fatigue in children and adolescents with immune thrombocytopenia. Br J Haematol. 2020;191(1):98–106.

14. Kruse C, Kruse A, DiRaimo J. Immune thrombocytopenia: The patient's perspective. Ann Blood. 2021;6:9–21.

15. Kuter DJ, Mathias SD, Rummel M, Mandanas R, Giagounidis AA, Wang X, et al. Health-related quality of life in nonsplenectomized immune thrombocytopenia patients receiving romiplostim or medical standard of care. Am J Hematol. 2012;87(5):558–61.

16. McMillan R, Bussel JB, George JN, Lalla D, Nichol JL. Self-reported health-related quality of life in adults with chronic immune thrombocytopenic purpura. Am J Hematol. 2008;83(2):150–4.

17. Kelton JG, Vrbensky JR, Arnold DM. How do we diagnose immune thrombocytopenia in 2018? Hematology. 2018;2018(1):561–7.

18. Provan D, Arnold DM, Bussel JB, Chong BH, Cooper N, Gernsheimer T, et al. Updated international consensus report on the investigation and management of primary immune thrombocytopenia. Blood Adv. 2019;3(22):3780–817.

19. Salib M, Clayden R, Clare R, Wang G, Warkentin TE, Crowther MA, et al. Difficulties in establishing the diagnosis of immune thrombocytopenia: An agreement study. Am J Hematol. 2016;91(8):E327-9.

20. Gabe C, Sirotich E, Li N, Ivetic N, Nazy I, Smith J, et al. Performance characteristics of platelet autoantibody testing for the diagnosis of immune thrombocytopenia using strict clinical criteria. Br J Haematol. 2021 Jul;194(2):439–43.

21. Almizraq RJ, Branch DR. Efficacy and mechanism of intravenous immunoglobulin treatment for immune thrombocytopenia in adults. Ann Blood [Internet]. 2021 Mar 31 [cited 2023 Nov 23];6(0). Available from: https://aob.amegroups.org/article/view/6137/html

22. Arnold DM, Nazy I, Clare R, Jaffer AM, Aubie B, Li N, et al. Misdiagnosis of primary immune thrombocytopenia and frequency of bleeding: Lessons from the McMaster ITP Registry. Blood Adv. 2017;1(25):2414–20.

23. Cooper N, Kruse A, Kruse C, Watson S, Morgan M, Provan D, et al. Immune thrombocytopenia (ITP) World Impact Survey (iWISh): Patient and physician perceptions of diagnosis, signs and symptoms, and treatment. Am J Hematol. 2021 Feb 1;96(2):188–98.

24. Cines DB, Liebman H, Stasi R. Pathobiology of secondary immune thrombocytopenia. Semin Hematol. 2009 Jan;46(1 Suppl 2):S2–14.

25. Pettigrew M, Garces K, Deuson R, Kassis J, Laroche V. Comparative net cost impact of the utilization of romiplostim and intravenous immunoglobulin for the treatment of patients with immune thrombocytopenia in Québec, Canada. J Med Econ. 2013;16(2):318–26.

26. Xie F, Blackhouse G, Assasi N, Campbell K, Levin M, Bowen J, et al. Results of a model analysis to estimate cost utility and value of information for intravenous immunoglobulin in Canadian adults with chronic immune thrombocytopenic purpura. Clin Ther. 2009;31(5):1066–8.

27. Mithoowani S, Gregory-Miller K, Goy J, Miller MC, Wang G, Noroozi N, et al. High-dose dexamethasone compared with prednisone for previously untreated primary immune thrombocytopenia: a systematic review and meta-analysis. Lancet Haematol. 2016;3(10):e489–96.

28. Kovaleva L, Apte S, Damodar S, Ramanan V, Loriya S, Navarro-Puerto J, et al. Safety and efficacy of a 10% intravenous immunoglobulin preparation in patients with immune thrombocytopenic purpura: Results of two international, multicenter studies. Immunotherapy. 2016;8(12):1371–81.

29. Robak T, Mainau C, Pyringer B, Chojnowski K, Warzocha K, Dmoszynska A, et al. Efficacy and safety of a new intravenous immunoglobulin 10% formulation (octagam ® 10%) in patients with immune thrombocytopenia. Hematology. 2010;15(5):351–9.

30. Robak T, Salama A, Kovaleva L, Vyhovska Y, Davies S V., Mazzucconi MG, et al. Efficacy and safety of Privigen®, a novel liquid intravenous immunoglobulin formulation, in adolescent and adult patients with chronic immune thrombocytopenic purpura. Hematology. 2009;14(4):227–36.

31. Neunert C, Terrell DR, Arnold DM, Buchanan G, Cines DB, Cooper N, et al. American Society of Hematology 2019 guidelines for immune thrombocytopenia. Blood Adv [Internet]. 2019 Dec 10 [cited 2022 Apr 4];3(23):3829–66. Available from: https://ashpublications.org/bloodadvances/article/3/23/3829/429213/American-Society-of-Hematology-2019-guidelines-for

32. Boccia R, Cooper N, Ghanima W, Boxer MA, Hill QA, Sholzberg M, et al. Fostamatinib is an effective second-line therapy in patients with immune thrombocytopenia. Br J Haematol [Internet]. 2020 Sep 1 [cited 2023 Nov 23];190(6):933–8. Available from: https://pubmed.ncbi.nlm.nih.gov/33439486/

33. Cuker A, Liebman HA. Corticosteroid overuse in adults with immune thrombocytopenia: Cause for concern. Res Pract Thromb Haemost. 2021;5(6):e12592.

34. Office of the Auditor General of Ontario. Value-for-Money Audit: Blood Management and Safety [Internet]. 2020 [cited 2021 Oct 25]. Available from: https://www.auditor.on.ca/en/content/annualreports/arreports/en20/20VFM_02bloodmgmt.pdf

35. Nazi I, Kelton JG, Larché M, Snider DP, Heddle NM, Crowther MA, et al. The effect of rituximab on vaccine responses in patients with immune thrombocytopenia. Blood. 2013;122(11):1946–53.

36. Fust K, Parthan A, Li X, Sharma A, Zhang X, Campioni M, et al. Cost per response analysis of strategies for chronic immune thrombocytopenia. Am J Manag Care. 2018 Jul;24(8 Spec No.):SP294–302.

37. Kurihara Y, Taoka K, Takagi E, Toyama K, Nakazaki K, Kurokawa M. Treatment of secondary immune thrombocytopenia with non-hodgkin lymphoma: A case report and literature review. Internal Medicine. 2021;60(10):1583–8.

38. Kojouri K, Perdue JJ, Medina PJ, George JN. Occult quinine-induced thrombocytopenia. J Okla State Med Assoc. 2000;93(11):519–21.

39. Vrbensky JR, Moore JE, Arnold DM, Smith JW, Kelton JG, Nazy I. The sensitivity and specificity of platelet autoantibody testing in immune thrombocytopenia: a systematic review and meta-analysis of a diagnostic test. J Thromb Haemost. 2019;17(5):787–94.

40. McDonnell A, Bride KL, Lim D, Paessler M, Witmer CM, Lambert MP. Utility of the immature platelet fraction in pediatric immune thrombocytopenia: Differentiating from bone marrow failure and predicting bleeding risk. Pediatr Blood Cancer. 2018;65(2):e26812.

41. Ali I, Graham C, Dempsey-Hibbert NC. Immature platelet fraction as a useful marker in the etiological determination of thrombocytopenia. Exp Hematol. 2019;78:56–61.

42. Jeon K, Kim M, Lee J, Lee JS, Kim HS, Kang HJ, et al. Immature platelet fraction: A useful marker for identifying the cause of thrombocytopenia and predicting platelet recovery. Medicine (Baltimore). 2020;99(7):e19096.

43. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J [Internet]. 2014 Aug 1 [cited 2022 Jan 31];35(29):1925. Available from: /pmc/articles/PMC4155437/

44. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? BMJ. 2009 Feb 23;338:b375.

45. Chowdhury MZI, Turin TC. Variable selection strategies and its importance in clinical prediction modelling. Fam Med Community Health. 2020;8(1):e000262.

46. Tangri N, Stevens LA, Griffith J, Tighiouart H, Djurdjev O, Naimark D, et al. A predictive model for progression of chronic kidney disease to kidney failure. JAMA [Internet]. 2011 Apr 20 [cited 2023 Feb 20];305(15):1553–9. Available from: https://pubmed.ncbi.nlm.nih.gov/21482743/

47. Tangri N, Major RW. Risk-based triage for nephrology referrals: The time is now. Kidney Int Rep. 2021 Aug;6(8):2028–30.

48. Rabinovich A, Ducruet T, Kahn SR, Shapiro S, Tagalakis V, Johri M, et al. Development of a clinical prediction model for the postthrombotic syndrome in a prospective cohort of patients with proximal deep vein thrombosis. J Thromb Haemost [Internet]. 2018 Feb 1 [cited 2023 Feb 20];16(2):262–70. Available from: https://pubmed.ncbi.nlm.nih.gov/29193770/

49. Grant SW, Collins GS, Nashef SAM. Statistical Primer: developing and validating a risk prediction model. Eur J Cardiothorac Surg. 2018 Aug 1;54(2):203–8.

50. Shipe ME, Deppen SA, Farjah F, Grogan EL. Developing prediction models for clinical use using logistic regression: an overview. J Thorac Dis. 2019 Mar;11(Suppl 4):S574–84.

51. Debray TPA, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KGM. A new framework to enhance the interpretation of external validation studies of clinical prediction models. J Clin Epidemiol [Internet]. 2015 Mar 1 [cited 2022 Jan 31];68(3):279–89. Available from: https://pubmed.ncbi.nlm.nih.gov/25179855/

52. Li N, Mahamad S, Parpia S, Iorio A, Foroutan F, Heddle NM, et al. Development and internal validation of a clinical prediction model for the diagnosis of immune thrombocytopenia. J Thromb Haemost [Internet]. 2022 Dec 1 [cited 2022 Dec 6];20(12):2988–97. Available from: https://pubmed.ncbi.nlm.nih.gov/36121734/

53. Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration. Ann Intern Med [Internet]. 2019 Jan 1 [cited 2022 Jan 31];170(1):W1–33. Available from: https://pubmed.ncbi.nlm.nih.gov/30596876/

54. Li N, Heddle NM, Nazy I, Kelton J, Arnold D. Platelet variability index: A measure of platelet count fluctuations in patients with immune thrombocytopenia. Blood Adv. 2021 Sep 13;5(20):4256–64.

55. Tarantino MD, Fogarty PF, Shah P, Brainsky A. Dental procedures in 24 patients with chronic immune thrombocytopenia in prospective clinical studies of eltrombopag. Platelets. 2015;26(1):93–6.

56. Tsai CH, Bussel JB, Imahiyerobo AA, Sandler SI, Ogunnaike BA. Platelet count control in immune thrombocytopenic purpura patient: Optimum romiplostim dose profile. J Process Control. 2016;45:76–83.

57. Gschwantler M, Vavrik J, Gebauer A, Kriwanek S, Schrutka-Kölbl C, Fleischer J, et al. Course of platelet counts in cirrhotic patients after implantation of a transjugular intrahepatic portosystemic shunt - A prospective, controlled study. J Hepatol. 1999;30(2):254–9.

58. Al-Huniti A, Kahr WH. Inherited platelet disorders: Diagnosis and management. Transfus Med Rev. 2020;34(4):277–85.

59. Page LK, Psaila B, Provan D, Michael Hamilton J, Jenkins JM, Elish AS, et al. The immune thrombocytopenic purpura (ITP) bleeding score: assessment of bleeding in patients with ITP. Br J Haematol. 2007 Jul;138(2):245–8.

60. Schisterman EF, Perkins NJ, Liu A, Bondell H. Optimal cut-point and its corresponding Youden Index to discriminate individuals using pooled blood samples. Epidemiology. 2005 Jan;16(1):73–81.

61. Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics. 2020 Jan 2;21(1):6.

62. Chang CH. Cohen's kappa for capturing discrimination. Int Health. 2014 Jun;6(2):125–9.

63. Riley RD, Debray TPA, Collins GS, Archer L, Ensor J, van Smeden M, et al. Minimum sample size for external validation of a clinical prediction model with a binary outcome. Stat Med. 2021;40(19):4230–51.

64. Alba AC, Agoritsas T, Walsh M, Hanna S, Iorio A, Devereaux PJ, et al. Discrimination and Calibration of Clinical Prediction Models: Users' Guides to the Medical Literature. JAMA [Internet]. 2017 Oct 10 [cited 2021 Dec 7];318(14):1377–84. Available from: https://jamanetwork.com/journals/jama/fullarticle/2656816

65. Gabe C, Mahamad S, St. John M, Duncan J, Kelton JG, Arnold DM. Adjudicating the diagnosis of immune thrombocytopenia in a clinical research study. TH Open. 2023;

66. Nichols TR, Wisner PM, Cripe G, Gulabchand L. Putting the Kappa Statistic to Use. The Quality Assurance Journal. 2010 Jul 12;13(3–4):57–61.

67. Salib M, Clayden R, Clare R, Wang G, Warkentin TE, Crowther MA, et al. Difficulties in establishing the diagnosis of immune thrombocytopenia: An agreement study. Am J Hematol. 2016;91(8):E327-9.

68. Stevens RJ, Poppe KK. Validation of clinical prediction models: what does the "calibration slope" really measure? J Clin Epidemiol. 2020 Feb 1;118:93–9.

69. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011 Feb 20;30(4):377–99.

70. Rubin D. Multiple Imputation for Nonresponse in Surveys. New York: Wiley; 2004.

71. Bain BJ. Platelet count and platelet size in males and females. Scand J Haematol [Internet]. 1985 Jul 1 [cited 2023 Feb 20];35(1):77–9. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1600-0609.1985.tb00804.x

72. Zheng M, Chen Y, Chen C, Gopal N, Jiao J. Clinical characteristics of platelets and its possible gender dimorphism in patients with nonalcoholic fatty liver disease (NAFLD). Postgrad Med [Internet]. 2021 [cited 2023 Feb 20];133(3):299–306. Available from: https://pubmed.ncbi.nlm.nih.gov/32921191/

73. Juan P, Stefano G, Antonella S, Albana C. Platelets in pregnancy. J Prenat Med. 2011 Oct;5(4):90–2.

74. Piel-Julian ML, Mahévas M, Germain J, Languille L, Comont T, Lapeyre-Mestre M, et al. Risk factors for bleeding, including platelet count threshold, in newly diagnosed immune thrombocytopenia adults. J Thromb Haemost. 2018 Sep;16(9):1830–42.

75. Friedler SA, Choudhary S, Scheidegger C, Hamilton EP, Venkatasubramanian S, Roth D. A comparative study of fairness-enhancing interventions in machine learning. In: Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19) [Internet]. New York: Association for Computing Machinery, Inc; 2019 [cited 2023 Mar 6]. p. 329–38. Available from: https://dl.acm.org/doi/10.1145/3287560.3287589

76. Kim TO, Grimes AB, Kirk SE, Gilbert MM, Reed HD, Staggers KA, et al. Racial variation in ITP prevalence and chronic disease phenotype suggests biological differences. Blood. 2020 Jul 30;136(5):640–3.

77. Takami A, Watanabe S, Yamamoto Y, Kondo H, Bamba Y, Ohata M, et al. Reference intervals of red blood cell parameters and platelet count for healthy adults in Japan. Int J Hematol. 2021 Sep;114(3):373–80.

78. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. Ann Intern Med. 2015;162(1).

79. Hall MK, Kea B, Wang R. Recognising Bias in Studies of Diagnostic Tests Part 1: Patient Selection. Emerg Med J. 2019 Jul;36(7):431–4.

80. Mithoowani S, Cervi A, Shah N, Ejaz R, Sirotich E, Barty R, et al. Management of major bleeds in patients with immune thrombocytopenia. J Thromb Haemost. 2020;18(7):1783–90.

7.0 APPENDIX

7.1 Tables

Table 1. The independent association of each of the predictor variables for the ITP diagnosis and the equation for the Predict-ITP Tool (from Li N et al. *J Thromb Haemost*. 2022 Dec 1;20(12):2988–97).

Predictor	**β coefficient (95% CI)**	**Odds ratio (95% CI)**	**P-value**
Intercept (constant β_0)	-5.18 (-7.46, -3.06)	-	<0.001
Lowest platelet count (by 10^9/L; β_1)	-0.03 (-0.05, -0.02)	0.97 (0.95, 0.98)	<0.001
Platelet variability (log scale; β_2)	0.15 (0.03, 0.28)	1.16 (1.03, 1.32)	0.016
Maximum mean platelet volume (by 1 fL; β_3)	0.22 (0.09, 0.36)	1.25 (1.09, 1.43)	0.001
Major bleed ever (Yes vs No; β_4)	0.87 (0.35, 1.41)	2.40 (1.42, 4.09)	0.001

Using the tool, one can determine the probability of the ITP diagnosis by logistic regression:

$$\text{Probability (ITP)} = \frac{\exp\left(\beta_0 + \beta_1 X_{\text{lowest pltct}} + \beta_2 X_{\text{pltct varibility}} + \beta_3 X_{\text{max MPV}} + \beta_4 X_{\text{major bleed}}\right)}{1 + \exp\left(\beta_0 + \beta_1 X_{\text{lowest pltct}} + \beta_2 X_{\text{pltct varibility}} + \beta_3 X_{\text{max MPV}} + \beta_4 X_{\text{major bleed}}\right)}$$

where pltct = platelet count; MPV = mean platelet volume*;* β_0 is the intercept and $\beta_{1, \ldots, 5}$ are the respective coefficients of the variables in the model.

Table 2. Data collected from each patient chart.

Time point	**Data collected**

Initial hematology consultation	• Year of birth
	• Month of birth
	• Sex
	• Race
	• Date of first recognition of thrombocytopenia
	• Date of initial hematology consultation at clinic
	• Diagnosis as determined in clinic
	• Pregnancy status
	• Current and past medications use (including treatments for ITP, anticoagulants/antiplatelets, and hormone therapy)
	• Platelet counts pre- and post-treatments for ITP
	• Past surgeries (including splenectomy)
	• Complete blood counts in the 5 years leading up to and including the initial consultation (along with date)
	• To assess major bleeding according to the ITP bleeding score (59)
	  ○ Based on physical examination
	  ▪ Skin (>5 bruises with size >2 cm and/or diffuse petechiae)
	  ▪ Oral (multiple blood blisters and/or gum bleeding)
	  ○ Based on patient history
	  ▪ Skin (>5 bruises with size >2 cm and/or diffuse petechiae)
	  ▪ Oral (Multiple blood blisters and/or gum bleeding >5 min)
	  ▪ Epistaxis (bleeding >5 minutes per episode)
	  ▪ Gastrointestinal (gross blood, bloody stool)
	  ▪ Urinary (macroscopic hematuria)
	  ▪ Gynecological (bleeding >spotting not at time of period or very heavy period)
	  ▪ Pulmonary hemorrhage
	  ▪ Intracranial hemorrhage

	▪ Eye bleeding with vision loss (e.g., retinal hemorrhage) • All relevant ITP-related investigations[a] and clinical events to adjudicate the outcome
Follow-up visits	• Diagnosis as determined in clinic • Pregnancy status • Complete blood counts • Treatments for ITP • Platelets count pre- and post-treatments for ITP • All relevant ITP-related investigations[a] and clinical events to adjudicate the outcome

[a] This can include blood tests, procedures, and diagnostic imaging studies to rule-out other thrombocytopenia disorders, consistent with current practice guidelines (31).

Table 3. Criteria for adjudicating the diagnosis of thrombocytopenic disorders.

Diagnosis	**Adjudication criteria**
Mild thrombocytopenia	Platelet count consistently between 100-150 x 10^9/L. For pregnant patients, consider gestational thrombocytopenia. Mild thrombocytopenia supersedes other diagnoses (e.g., hepatitis C or family history of thrombocytopenia)
Primary ITP	Patients with platelets <100 x 10^9/L and no other diagnosis; and a platelet count response* to corticosteroid or high dose IVIG. If platelet counts improve > 100, the diagnosis should remain primary ITP.
Secondary ITP • Antiphospholipid syndrome • Chronic lymphocytic leukemia • Common variable immune deficiency • Evan's syndrome • Helicobacter pylori (*H. pylori)* • Hepatitis C • HIV • Non-specific infection • Pregnancy-associated ITP • Lymphoma • Systemic lupus erythematosus • Sarcoidosis • Other autoimmune disease	Once a diagnosis of secondary ITP is made it should remain as such even if the underlying disorder is treated and the ITP persists. For pregnancy-associated ITP platelets should improve with ITP treatments. If the ITP persists post-partum or pre-dates pregnancy, the diagnosis should be primary ITP. For *H. pylori*-associated ITP, there should be evidence of *H. pylori* eradication and improvement in platelet count (definite); or evidence of active *H. pylori* infection and improvement of platelet count with treatment, even without evidence of eradication (probable).
Drug-induced ITP	Onset of thrombocytopenia is typically 5-10 day after initial drug exposure and platelet count recovery typically occurs after discontinuing the drug, with no other drugs implicated. Confirmation requires either a drug challenge or the demonstration of drug-induced platelet antibodies.

Non-immune thrombocytopenia • Alcohol related • Familial • Incidental thrombocytopenia in pregnancy (gestational thrombocytopenia) • Hypertensive disorders of pregnancy • Liver disease • Splenomegaly/ hypersplenism • Myelodysplastic syndrome • Pseudothrombocytopenia • Drug-induced bone marrow suppression • Thrombocytopenia associated with malignancy including aplastic anemia	For familial thrombocytopenia, platelet count should be below 100 x 10^9/L (otherwise, classify as mild thrombocytopenia), with a family history in first-degree relatives. For incidental thrombocytopenia of pregnancy (gestational thrombocytopenia), platelet count is typically above 70 x 10^9/L during pregnancy, normalization of platelet count post-delivery, no history of thrombocytopenia (except during a prior pregnancy), and no thrombocytopenia in the fetus or newborn. Fatty liver disease alone (without other stigmas of chronic liver disease) is not a cause of non-immune thrombocytopenia. For patients with spleen enlargement, classify as splenomegaly unless the patient has liver cirrhosis or portal hypertension, in which case classify as liver disease. The diagnosis of myelodysplastic syndrome can be presumed even without bone marrow evaluation when other features are present (e.g., variable sized platelets, hypogranular platelets, and hypolobulated neutrophils on the peripheral blood smear). For pseudothrombocytopenia, platelet clumping observed in the peripheral blood smear with a routine complete blood count, and the platelet count normalizes when citrate or heparin is used in the blood collection tube.

Other thrombocytopenia disorders • Cyclical thrombocytopenia • Heparin induced thrombocytopenia • Thrombotic microangiopathies	For cyclical thrombocytopenia, there should be evidence of large platelet count fluctuations independent of treatment. If fluctuations resolve but platelet count stays low, the diagnosis of ITP should be considered.
Unknown	Does not meet criteria for any category, or meets criteria for more than one category, or data is not available.

ITP, immune thrombocytopenia; HIV, human immunodeficiency virus. *Response is defined as an increase in baseline platelet count to 50 x 10^9/L or higher and doubling of baseline platelet count within 4 weeks.

7.2 Figures

Figure 1. Process of determining the actual diagnosis of patients.