DEVELOPMENT AND VALIDATION OF NEW SCORING SYSTEM FOR PREDICTING ENDOSCOPIC FAILURE IN ADULT PATIENTS WITH UPPER GASTROINTESTINAL BLEEDING ULCER IN HOSPITAL SULTANAH AMINAH JOHOR BAHRU

By

KRISHANTHINI A/P R JEYABALAN

Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science

DECEMBER 2024

ACKNOWLEDGEMENT

First and foremost, I would like to extend my gratitude to my main supervisor, Dr Anis Kausar Ghazali for her constant guidance, help and motivation in every step of conducting this study. Thank you to my co-supervisor, Assoc. Prof Dr Najib Majdi Yaacob for always suggesting improvements to the study and also encouraging us to learn new methodologies.

To my field supervisor, Mr Hans Alexander Mahendran, thank you so much for always being there regardless of time from the moment we started discussing research ideas until getting my very first oral presentation prize. I am truly indebted to you for all the discussions, encouragement and support you have given in completing my thesis.

My heartfelt gratitude to Prof Dr Norsa'adah Bachok, Assoc. Prof Dr Sarimah Abdullah, Assoc. Prof Dr Najib, Assoc Prof Dr Erica Kueh Yee Cheng, Dr Wan Nor Arifin and Dr Siti Azrin for all the guidance given throughout this program. Thank you to my classmates, especially Zakiah, Mastura, Farhana, and Pei Ling for all the last-minute discussions and sharing of resources.

To Caryn, thank you for always being there for me both as a colleague and a friend. Completing this program successfully would not have been possible without immense support from my family members especially my husband. To my kids, thank you for being so patient and understanding.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

AUC	Area under ROC curve
CART	Classification and Regression Tree
CKD	Chronic Kidney Disease
HSAJB	Hospital Sultanah Aminah Johor Bahru
IHD	Ischemic Heart Disease
IQR	Interquartile range
NPV	Negative predictive value
NSAIDs	Non-steroidal Anti-Inflammatory Drugs
PPV	Positive predictive value
ROC	Receiver operating curve
SD	Standard deviation
SVM	Support Vector Machine
UGIB	Upper gastrointestinal bleeding
VIF	Variation inflation factor

LIST OF APPENDICES

Page

PEMBANGUNAN DAN PENGESAHAN SKOR PENILAIAN BARU UNTUK MERAMALKAN KEGAGALAN ENDOSKOPI DI KALANGAN PESAKIT DEWASA YANG MENGHIDAPI PENDARAHAN GASTROINTESTINAL BAHAGIAN ATAS DI HOSPITAL SULTANAH AMINAH JOHOR BAHRU

ABSTRAK

Pengenalan: Insiden pendarahan gastrointestinal bahagian atas adalah 72 bagi setiap 100 000 penduduk di Malaysia. Walaupun kadar kejayaan rawatan endoskopik adalah tinggi, sebanyak 6.4% hingga 19.2% pesakit mengalami pendarahan semula, yang meningkatkan risiko kematian sebanyak 4-5 kali ganda.

Objektif: Untuk mengenal pasti faktor-faktor yang berkaitan dengan kegagalan endoskopi dan membangunkan sistem penilaian yang divalidasi untuk meramalkan kegagalan endoskopi di kalangan pesakit dewasa yang menghidapi pendarahan gastrointestinal bahagian di HSAJB

Metodologi: Kajian ini menggunakan analisis sekunder dari rekabentuk kes-kawalan di mana 98 kes mengalami pendarahan semula dalam tempoh 30 hari selepas endoskopi awal berjaya dan 345 kes sebagai kawalan. Faktor-faktor yang berkaitan dengan pesakit, ciri-ciri klinikal dan endoskopi telah dibandingkan. Regresi logistik digunakan untuk mengenal pasti faktor yang berkaitan dengan kegagalan endoskopi, dan pemeramal dari model regresi logistik akhir digunakan untuk membangunkan sistem penilaian. Pengesahan dalaman dilakukan melalui bootstrap dan model tersebut dibentangkan sebagai carta skor.

Keputusan: Regresi logistik mengenal pasti penyakit buah pinggang kronik (CKD) (*adjusted* OR=2.08, 95%CI:1.09,3.88, p=0.025), ulser duodenum (*adj. OR*=2.91, 95%CI:1.67,5.30, p<0.001), dan saiz ulser lebih dari 2 cm (*adj. OR*=2.56, 95%CI:1.55,4.22, p<0.001) sebagai peramal signifikan kegagalan endoskopi. Pemeramal-pemeramal ini dimasukkan ke dalam

sistem penilaian, yang menjalani pengesahan dalaman melalui kaedah *bootstrapping*. Sistem penilaian ini telah menunjukkan ketepatan sebanyak 70.1%, dengan spesifisiti sebanyak 84% dan sensitivity sebanyak 47%, menyokong potensi utilitinya dalam amalan klinikal.

Kesimpulan: Ulser duodenum, ulser berdiameter lebih daripada 2cm, dan CKD adalah peramal penting dalam kegagalan endoskopi. Kemasukan faktor klinikal dan endoskopik dalam sistem penilaian yang dibangunkan dengan menggunakan data tempatan ini menjadikan sistem ini sebagai satu alat penilaian risiko yang lebih komprehensif**.**

DEVELOPMENT AND VALIDATION OF NEW SCORING SYSTEM FOR PREDICTING ENDOSCOPIC FAILURE IN ADULT PATIENTS WITH UPPER GASTROINTESTINAL BLEEDING ULCER IN HOSPITAL SULTANAH AMINAH JOHOR BAHRU

ABSTRACT

Introduction: Upper gastrointestinal bleeding (UGIB) has an incidence of 72 per 100 000 population in Malaysia. Despite high success rates of endoscopic treatment, 6.4% to 19.2% of patients develop rebleeding which increases the risk of mortality by 4-5 times.

Objective: To identify factors associated with endoscopic failure and develop a validated scoring system to predict endoscopic failure among adult patients with upper gastrointestinal bleeding ulcers in HSAJB

Methodology: This is a case-control study using secondary data of 443 cases where 98 cases had rebleeding within 30 days of successful initial endoscopy and 345 were controls. Patient-related, clinical and endoscopic findings were compared. Logistic regression was used to identify factors associated with endoscopic failure, and the predictors from the final logistic regression model were used to develop the scoring system. Internal validation was done via bootstrapping and the model was presented as a score chart.

Results: Logistic regression identified chronic kidney disease (CKD) (adjusted OR=2.08, 95%CI:1.09,3.88, p=0.025), duodenal ulcer (adj. OR=2.91, 95%CI:1.67,5.30, p<0.001), and ulcer size more than 2 cm (adj. OR=2.56, 95%CI:1.55,4.22, p<0.001) as significant predictors of endoscopic failure. These predictors were incorporated into a scoring system, which underwent internal validation using bootstrapping. The scoring system demonstrated an

accuracy of 70.1%, with a specificity of 84% and a sensitivity of 47%, supporting its potential utility in clinical practice.

Conclusion: Duodenal ulcers, ulcers more than 2cm in diameter, and CKD are significant predictors of endoscopic failure. The inclusion of both clinical and endoscopic factors in this locally developed scoring system provides for a more comprehensive risk assessment tool.

CHAPTER 1

INTRODUCTION

1.1 Introduction

Upper gastrointestinal bleeding (UGIB) is a common medical condition that can present with frank hematemesis, coffee ground vomitus, melena and, in a minority of cases, as bleeding per rectum (Rockey, 2022). Annual incidence rates of UGIB in Western populations, such as in the USA are approximately 67 cases per 100 000 individuals (Wuerth & Rockey, 2018). In Asia, the incidence of UGIB tends to be higher. For example, the hospitalized incidence rate in Thailand is around 166.3 per 100 000 population(Sangchan *et al.*, 2012), significantly exceeding the rates reported in Western countries. The incidence of UGIB in Hong Kong is reported as 3.26 cases per 100,000 person-months(Guo *et al.*, 2021). According to available Malaysian data, the UGIB incidence in the year 2001 was around 72 cases per 100 000 population (Ministry of Health, 2003).

Understanding the causes of UGIB is essential to addressing its management challenges. Causes of UGIB can be divided based on their pathophysiology into ulcerative and erosive lesions, vascular lesions, mass lesions and traumatic lesion all which have different risk factors (Rockey, 2022). Both in western and Asian populations, the predominant cause of UGIB are ulcerative and erosive lesions with peptic ulcer disease being the main cause at around 43% to 47% of UGIB (Wuerth & Rockey, 2018; Kiat Koh *et al.*, 2021). Mortality due to upper gastrointestinal bleeding ulcer ranges from 5.8% to 11.4% globally (Lau *et al.*, 2011), while rates in Thailand and Malaysia are comparatively lower, at 2% (Thanapirom *et al.*, 2016) and 4.7% (Iii *et al.*, 2000), respectively.

As part of standard management, endoscopy is done immediately or within 24 hours in patients presenting with acute UGIB, depending on severity, for the purpose of diagnosis, risk

assessment and treatment (Ministry of Health, 2003; Gralnek *et al.*, 2021; National Institute for Health and Care Excellence, 2021). While endoscopy remains the cornerstone of UGIB management, its limitations in achieving sustained haemostasis highlights the need for further research. Despite advancements in endoscopic techniques, rebleeding, which indicates endoscopic failure, occurs in 6.4% to 19.2% (Jairath *et al.*, 2012; Quan *et al.*, 2014; Dango *et al.*, 2017). Identifying patients at risk of endoscopic failure remains a significant clinical challenge, underscoring the importance of developing reliable predictive tools.

1.2 Problem statement

Risk scoring systems play an important role in the management of UGIB as these provide a consistent manner of risk stratification and standardises the risk communication among healthcare professionals (Barkun *et al.*, 2019). Although several risk scoring systems such as the Rockall, Glasgow-Blatchford and CANUKA scores exist (Rockall *et al.*, 1996; Blatchford, Murray & Blatchford, 2000; Oakland *et al.*, 2019), these are primarily designed for general UGIB risk stratification and do not comprehensively account for endoscopic and clinical factors. Additionally, most available models are not validated in the Malaysian population, limiting their applicability in local clinical practice. For instance, the scoring model by Lai et al. (2022) focuses exclusively on upper gastrointestinal bleeding ulcer patients with high risk of bleeding but excludes critical clinical variables such as anticoagulant or Non-Steroidal Anti-Inflammatory Drugs (NSAID) use, and endoscopic factors such as ulcer site and size. This omission limits its utility in diverse clinical settings where such variables are common predictors of endoscopic failure. While the Forest classification remains a valuable tool for identifying rebleeding risk, it is heavily reliant on operator-dependent endoscopic findings, overlooking clinical parameters that can significantly influence outcomes (Ministry of Health, 2003). This underscores the need for a more holistic approach that integrates both endoscopic and clinical predictors.

1.3 Study rationale

Rebleeding after successful endoscopy is associated with a four-to-five-fold increase in mortality risk (Jairath *et al.*, 2012) and prolonged hospitalisation (Baradarian *et al.*, 2004). Given the strain on Malaysia's healthcare system, a predictive model for endoscopic failure could significantly improve resource allocation, reduce complications and enhance patient outcomes. This study aims to address the limitations of existing scoring systems by incorporating both clinical and endoscopic parameters into a predictive model for endoscopic failure. By focusing on upper gastrointestinal bleeding ulcer patients, it seeks to provide a practical tool that can be applied during the index endoscopy, reducing the need for repeat procedures and associated complications.

1.4 Research question

1. What are the patient-related, clinical, and endoscopic factors associated with endoscopic failure?
2. What is the internal validity of the newly developed scoring system in predicting endoscopic failure?

1.5 Research objective

1.5.1 General objective

To develop a validated scoring system to predict endoscopic failure among adult patients with upper gastrointestinal bleeding ulcers in HSAJB.

1.5.2 Specific objectives

1. To determine the patient-related, clinical, and endoscopic factors associated with endoscopic failure among adult patients undergoing endoscopy for upper gastrointestinal bleeding ulcers in HSAJB from 2018 to 2023.

2. To develop and determine the internal validity of a scoring system for predicting endoscopic failure in the same population and time frame.

1.6 Research hypothesis

1. Patient-related, clinical and endoscopic factors are significantly associated with endoscopic failure among adult patients with upper gastrointestinal bleeding ulcer undergoing endoscopy in HSAJB.
2. The scoring system developed in this study is valid for predicting endoscopic failure among these patients from 2018 to 2023.

CHAPTER 2

LITERATURE REVIEW

2.1 Literature search strategy

Literature search for this study was done in 3 parts. The first part was looking for articles with regards to factors associated with endoscopic failure. The second part was looking for articles regarding prediction scores that were currently available and the last part was looking for articles regarding clinical prediction model development and validation. Keywords utilized for the literature search are as listed in Table 2.1. We limited our search results to articles that were published in the English language from January 2000 to February 2025. The type of studies included in our search were systematic reviews, meta-analysis, randomised controlled trials, observational studies and guidelines. The various search engines used in this study were PubMed, Scopus, Google Scholar, and Science Direct. In addition to this, reference list of articles was also combed through to find articles relevant to this study. Mendeley Reference Manager was used to store the retrieved articles and aid with citation.

Table 2.1 Keywords used for literature search and number of articles retrieved

Keywords with Boolean Operators	Search engine			
	PubMed	Google Scholar	Scopus	Science Direct
	126	2140	211	136
("Upper Gastrointestinal Bleeding" OR "Non-Variceal Bleeding" OR "Non-Variceal Gastrointestinal Bleeding" OR "Non-Variceal Haemorrhage" OR "Non-Variceal Gastrointestinal Haemorrhage" OR "Bleeding peptic ulcer" OR "Peptic ulcer bleed") **AND** ("Risk Factors" OR "Factor associated" OR "Predicting factors" OR "Predisposing factors") **AND** ("Endoscopic failure" OR "Rebleeding")				

("Upper Gastrointestinal Bleeding" OR "Non-Variceal Bleeding" OR "Non-Variceal Gastrointestinal Bleeding" OR "Non-Variceal Haemorrhage" OR "Non-Variceal Gastrointestinal Haemorrhage" OR "Bleeding peptic ulcer" OR "Peptic ulcer bleed") AND ("Scoring" OR "prediction scores" OR "risk scoring system") AND ("Endoscopic failure" OR "Rebleeding")	83	1250	77	94

2.2 Upper gastrointestinal bleeding

Gastrointestinal bleeding indicates bleeding that occurs in the gastrointestinal tract and they can be classified into upper and lower gastrointestinal bleed based on the location of bleed with regards to Ligament of Treitz. Lower gastrointestinal bleeding are bleeds that occur distal to the ligament, involving the colon and rectum (Strate & Gralnek, 2016). UGIB involves intraluminal bleeding in the gastrointestinal tract proximal to the ligament, occurring from the mouth to the duodenum (Tielleman, Bujanda & Cryer, 2015; Nelms & Pelaez, 2018; Stanley & Laine, 2019). UGIB is a significant medical emergency with high morbidity and mortality rates. It requires timely diagnosis and effective management to prevent complications, including rebleeding and death.

There are many causes of UGIB with the commonest cause being gastric and duodenal ulcers. Other causes are oesophageal ulcers, gastroesophageal varices, gastrointestinal tract erosions, Mallory-Weiss tears and Dieulafoy lesions. Malignancies and aorto-duodenal fistulas are some of the less common causes of UGIB (Feinman & Haut, 2014; Tielleman, Bujanda & Cryer, 2015). Risk factors of peptic ulcer bleeds include H.pylori infection, use of medications such as NSAIDs / antiplatelets and anticoagulants, and older age (Tielleman, Bujanda & Cryer, 2015; Costable & Greenwald, 2021). Understanding the aetiology and risk factors of UGIB is

critical for developing predictive tools that can stratify patients by their risk of adverse outcomes, such as endoscopic failure.

2.3 Endoscopic failure

Endoscopic failure, marked by rebleeding after achieving initial haemostasis in the index scope, is a significant predictor of poor outcomes, including prolonged hospitalisation, increased morbidity, and higher mortality rates. Presence of new onset hematemesis, melena, hemodynamic instability or a fall in haemoglobin level (>2g/dL in 24 hours post endoscopy) are considered as rebleeding (Harjit, Kandasami & Hanafiah, 2002; Parente *et al.*, 2005; Gralnek *et al.*, 2021).

2.4 Factors associated with endoscopic failure

2.4.1 Demographic factor

Older age is a significant predicting factor for rebleeding where there is almost an eight-fold increase in rebleeding risk in those aged more than 60 years (Bini, Cohen & York, 2003; Kim *et al.*, 2018; Hajiagha Mohammadi & Reza Azizi, 2019). Presence of age-related physiological changes such as impaired healing, comorbidities including cerebrovascular disease, renal disease and cardiovascular disease and the concurrent use blood-thinning medications contribute to the rebleeding risk among the elderly.

2.4.2 Clinical factors

Growing number of elderly populations compounded by increasing prevalence of non-communicable diseases have contributed to the high proportions of patients with chronic kidney disease (CKD) and end stage renal failure. This in turn is associated with increased risk of rebleeding with odds between 3.9 to 10.3 compared to those without renal disease (Cheung *et al.*, 2010; Lee *et al.*, 2013; Ogiyama *et al.*, 2021).

Patients with Ischemic Heart Disease (IHD) have higher risk of peptic ulcer bleeds as they are usually prescribed with either single or double antiplatelet agents and they also may have impaired cardiovascular function which hinders healing of peptic ulcers predisposing them to rebleeding episodes. The K-PUB study which analysed factors that were associated with rebleeding showed that comorbidities such as IHD had 2.95 times higher risk of rebleeding (Kim *et al.*, 2018).

A higher usage of NSAIDs is proven to increase the risk of rebleeding (Camus *et al.*, 2016; Dango *et al.*, 2017). Patients who took NSAIDs had 4.25 times higher risk of rebleeding (Kim *et al.*, 2016). The high prevalence of cardiovascular and cerebrovascular diseases has increased the usage of anticoagulants and antiplatelet agents. These 2 groups of drugs are also significant predictors for rebleeding (Jairath *et al.*, 2012; Dango *et al.*, 2017). Patients who resumed the usage of these drug groups after achieving successful endoscopic haemostasis had 3 times more risk for rebleeding compared to those who did not (Sostres *et al.*, 2019).

The interplay between chronic diseases, medication use and impaired healing processes explains the increased rebleeding risk in patients with comorbidities such as IHD and CKD. Including these factors into predictive models could improve risk stratification by providing a more precise assessment of risk rather than looking at individual factors only.

2.4.3 Endoscopic factor

Stigmata of recent haemorrhage has been proven to be an important predictor for rebleeding in patients with bleeding peptic ulcer. Those with active bleeding seen in endoscopy which includes spurting and oozing have higher risk of rebleeding. A meta-analysis review showed those with active bleeding have 1.7 times more risk compared to those who do not have active bleeding. (Chung *et al.*, 2001; García-Iglesias *et al.*, 2011; Nam *et al.*, 2017).

Another crucial endoscopic predictor for rebleeding is the size of the ulcer. Ulcers with size > 1cm have 2.81 times greater risk of rebleeding compared to smaller ulcers (Chung *et al.*, 2001; García-Iglesias *et al.*, 2011). This prospective study done in USA also showed that for every 10% increase in size of ulcer, the odds for rebleeding increases by 6% (Camus *et al.*, 2016).

The third endoscopic factor that has risk for rebleeding is the site of the ulcer. Ulcers situated in the posterior duodenal bulb and high gastric lesser curvature have 2-3 times higher odds of rebleeding compared to other sites (Chung *et al.*, 2001; García-Iglesias *et al.*, 2011). Larger ulcers may indicate more extensive mucosal damage, while specific sites, such as the posterior duodenal bulb, maybe associated with higher vascularity, increasing the rebleeding risk. Most of the studies done looking at rebleeding risk focus mainly on the stigmata of recent haemorrhage. In this study, we would like to highlight the role of other endoscopic factors such as site and size of ulcers as clinically important predictors of endoscopic failure.

2.5 Scoring system for predicting endoscopic failure

Risk scoring systems provide an objective method to assess risk of an outcome in a patient and aids the clinician in decision making. In UGIB, these scores are able to help clinicians identify patients who are at higher risk of complications or negative outcomes such as rebleeding and death (Monteiro, 2016; Strate & Gralnek, 2016). There are various scoring methods available for the prediction of rebleeding, some of which are as described in this section.

2.5.1 Forrest classification

2.5.1.1 Outcome predicted

This classification was used to predict rebleeding rate based on the 5 stages of ulcers where Stage Ia has the highest rebleeding rate and stage III has the lowest rebleeding rate

(Forrest, Finlayson & Shearman, 1974; Ministry of Health, 2003; Alzoubaidi, Lovat & Haidry, 2019).

2.5.1.2 Variables included

Only stigmata of recent haemorrhage based on endoscopy was used in this classification.

Table 2.2 Forrest classification

Stage	Characteristics	Rebleeding
I a	Spurting bleeding	60-100%
I b	Non spurting active bleeding	50%
II a	Visible Vessel (No active bleeding)	40-50%
II b	Non bleeding with overlying clot (No visible vessel)	20-30%
II c	Ulcer with hematin covered base	7-10%
III	Clean ulcer ground (No clot, no vessel)	3-5%

2.5.1.3 Strength and limitations

This classification was useful in the early days of endoscopy to create a common classification nomenclature that can be used by all surgeons, however in modern medicine, it is no longer sufficient to predict rebleeding risk based on a single factor.

2.5.2 Baylor Bleeding Score

2.5.2.1 Outcome predicted

In Baylor Bleeding Score, the scoring was used to stratify the patients who had major gastrointestinal haemorrhage secondary to ulcer, and had successful endoscopic treatment into low risk and high risk for rebleeding.

2.5.2.2 Variables included

This scoring had 3 parts to it: Pre endoscopy score, endoscopy score and post endoscopy score. Pre endoscopy score looked at clinical parameters such as age, number and severity of concurrent illness, endoscopy score was based on site and stigma of bleeding and post endoscopy score was the sum of the pre-endoscopy and endoscopy. Each parameter was

given a score ranging from 0 to 5. Pre endoscopy score of ≤ 5 and post endoscopy score of ≤ 10 were considered as low risk where else pre-endoscopy score ≥ 5 or pre-endoscopy score ≤ 5 with post endoscopy score ≥ 10 were considered as high risk for rebleeding (Saeed *et al.*, 1995).

2.5.2.3 Strength and limitations

This scoring included both endoscopic and clinical factors. However, this scoring did not look at comorbidities individually but as total number of comorbidities present instead.

2.5.3 Glasgow-Blatchford Score (GBS)

2.5.3.1 Outcome predicted

Glasgow-Blatchford Score (GBS) was initially developed based on a UK cohort to predict the need for treatment which was a composite outcome of requirement for blood transfusion, surgical / endoscopic intervention, rebleeding, drop in haemoglobin level and death.

2.5.3.2 Variables included

Score ranging from 1 to 6 is given to variables such as blood urea level, haemoglobin level, systolic blood pressure, pulse rate, presentation with melena, presentation with syncope, presence of hepatic disease and cardiac failure (Blatchford, Murray & Blatchford, 2000). Scores ≤ 1 are lower risk for rebleeding and these patients do not required admission for further treatment (Sung *et al.*, 2018; Gralnek *et al.*, 2021; National Institute for Health and Care Excellence, 2021).

2.5.3.3 Strength and limitations

GBS was developed based on patients that presented with acute gastrointestinal haemorrhage regardless of pathology and is able to predict risk not only for rebleeding but for mortality as well. However, this scoring did not include endoscopic findings.

2.5.4 Rockall scoring

2.5.4.1 Outcome predicted

The Rockall scoring on the other hand uses both clinical parameters and endoscopic findings for the prediction of mortality and rebleeding.

2.5.4.2 Variables included

The pre-endoscopy scoring is based on age, degree of shock and comorbidity where a maximum score of seven is given. Diagnosis and stigmata of recent haemorrhage is added in the post endoscopy scoring increasing the maximum score to eleven (Rockall *et al.*, 1996). A score of ≤ 2 had 0.1% mortality and 4.3% rebleeding but score > 8 had higher mortality at 41% and higher rebleeding at 42.1% (Ministry of Health, 2003).

2.5.4.3 Strengths and limitations

This scoring method was developed in UK based on patients who presented with acute gastrointestinal haemorrhage regardless of cause and can be used pre-endoscopy as well as post endoscopy. Limitation of this scoring method is that endoscopic factor was limited to stigmata of recent haemorrhage and comorbidities were grouped instead of being given an individual score.

2.5.5 CANUKA scoring

2.5.5.1 Outcome predicted

The aim of this scoring was to predict adverse outcomes including mortality and rebleeding and identify low risk patients who can be managed as outpatients

2.5.5.2 Variables included

CANUKA only looks at clinical parameters such as age, melena, hematemesis, syncope, liver disease, malignancy, heart rate, systolic blood pressure, haemoglobin and urea levels. A score of ≤ 1 indicates lower risk for mortality and rebleeding (Oakland *et al.*, 2019).

2.5.5.3 Strengths and limitations

Strength of CANUKA is that it was developed using a large, pooled variceal and non-variceal bleeding datasets from 3 Western populations (Canada, United Kingdom and Australia). However, it does not include endoscopic factors in its scoring.

2.5.6 Normogram

2.5.6.1 Outcome predicted

This Chinese study developed a model to predict rebleeding within 3 days of achieving endoscopic haemostasis specifically in high-risk peptic ulcer bleeding patients (Forrest Ia – Forrest IIb) (Lai *et al.*, 2022)

2.5.6.2 Variables included

Predictors such as albumin level, prothrombin time, presence of shock, presence of hematemesis or melena and Forrest classification were used to generate a nomogram. Each of these factors were allocated points ranging from 0 to 100. The probability of rebleeding increases as the total points increase.

2.5.6.3 Strength and limitations

Strength of this study is that both clinical and endoscopic factors were included and the ability to predict rebleeding within 3 days instead of the 30 days done by other scoring methods. But this scoring was developed based on data from a single centre, only internal validation has been done and interpreting a nomogram in a busy clinical setting may not be easy.

Table 2.3 Comparison of various scoring models

Risk Scoring Models	**Key Features**	**Strengths**	**Limitations**
Forrest	Utilises single endoscopic finding	Used as common classification system	Description of findings from a single centre

Baylor Bleeding Score	Utilises both endoscopic and clinical findings	Can be used to predict rebleeding based on pre-endoscopy or post endoscopy score	Comorbidities were not analysed individually
Glasgow Blatchford Score	Utilises only clinical factors	Can be used to predict both rebleeding risk and mortality	Scoring was developed not specific for peptic ulcer bleeds
Rockall	Utilises both endoscopic and clinical findings	Can be used to predict both rebleeding risk and mortality	Comorbidities were not analysed individually and only stigmata of haemorrhage was included in the scoring
CANUKA	Utilises only clinical factors	Developed using large dataset from 3 different Western populations	Does not include endoscopic findings
Normogram	Utilises both endoscopic and clinical findings	LASSO method was used for predictor selection to minimise overfitting	Only validated in a single centre

Using only endoscopic findings or clinical parameters to predict rebleeding is not appropriate as both factors are equally important. Despite the proven utility of all the models above, they often exclude critical factors, such as NSAIDs or antiplatelets or anticoagulant use, site and size of ulcer and also there is lack of external validation, limiting their applicability in diverse clinical settings. It is prudent to develop a more comprehensive predictive tool that incorporates these limitations and is also validated for use in our local population.

2.6 Clinical prediction model development

Clinical prediction models are very useful in the clinical field to aid with diagnosis and prognosis. However, developing these models are not straightforward and require multiple steps such as having a clear problem definition, choosing appropriate variable for the model selection, coding of the predictors, evaluating the model performance, validating the model, and finally presenting the model in a manner that can be used in the clinical field (Steyerberg & Vergouwe, 2014).

Selection of appropriate variables to be included in the model development can be done either using a parametric approach such as Linear regression, Logistic regression, Poisson regression and Discriminant Analysis or a non-parametric approach using machine learning methods such as Support Vector Machine (SVM), Classification and Regression Tree (CART), and Neural Network. There will be no regression coefficients available when non parametric methods are used and this will make interpretation of the final model harder. Model developed using CART maybe easier to interpret, but these models usually require a large dataset and there is always the risk of loss of information due to the need to recategorize continuous data (Steyerberg, 2019; Zhou *et al.*, 2019). It is also important to select the least number of predicting variables for optimum resource usage and higher likelihood for use in a busy clinical setting (Richter & Khoshgoftaar, 2018).

Logistic regression is commonly for prediction of binary outcome as it is a flexible model that can accommodate categorical variables, continuous variables, and interactions terms. This analysis is also able to generate coefficients adjusted to the other predicting factors in the model and these coefficients can be transformed into a scoring system for ease of interpretation (Shipe *et al.*, 2019; Steyerberg, 2019; Zhou *et al.*, 2019).

The next important aspect of creating a prediction model is the validation of the model. Ideally both internal and external validation are required. Internal validation looks at the

reproducibility of the model and external validation looks at the generalizability of the model (Zhou *et al.*, 2019). There are a few methods available for internal validation such as apparent validation, split-sample validation, cross validation, and bootstrap validation. Apparent validation is where we use the same development dataset to validate the model. This method produces a more optimistic estimate but it is stable (Steyerberg, 2019). Split-sample validation is a method where the dataset is split 50:50 or 2:1 into the development set and validation set. Since only a portion of data is used for the development, there is lower stability and not suitable for small sample size studies. Cross validation is a method where the dataset is divided into several subsets, for example 10 subsets, where 9 subsets will be used for the model development and 1 subset will be used for the model validation. This process is repeated with a different subset as the validation subset and the remaining subsets as the development subset until all 10 subsets have been used at least once as the validation subset. Although this method is more stable compared to the split-sample method, it underestimates the variability of the model (Zhou *et al.*, 2019).

Bootstrap validation is becoming a more popular option especially with more powerful computers available. In this method, bootstrap samples are sampled from the original dataset to create the bootstrap dataset. The same model development steps done in the original dataset will be repeated in the bootstrap dataset to create a new model. The difference in performance measures between the original dataset and bootstrap dataset will be calculated. This process is repeated 200 to 500 times using a different bootstrap dataset sampled from the original dataset with replacement. The average difference in performance measures indicates the optimism of the original model. Bootstrap validation is more stable, suitable for studies with small sample size or large number of predictor variables (Shipe *et al.*, 2019; Steyerberg, 2019).

2.7 Conceptual framework of the study

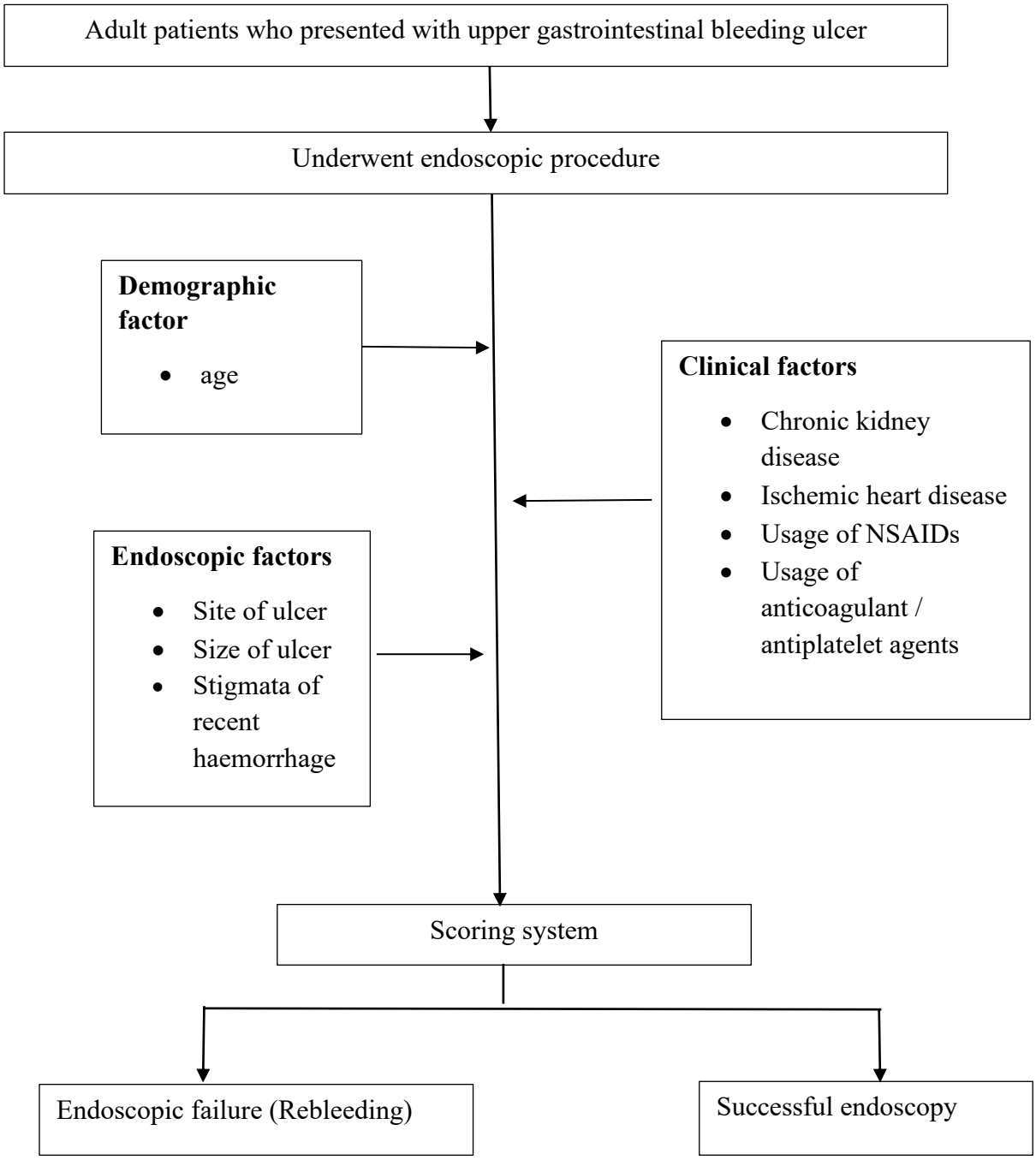

Figure 2.1 Conceptual framework of the study

CHAPTER 3

METHODOLOGY

3.1 Study design

A case-control design with 1:4 ratio was used for this study as the local prevalence of endoscopic failure was low, around 3-5%. Cases are those with endoscopic failure and controls are those without endoscopic failure. Endoscopic failure is defined as rebleeding within 30 days of successful initial therapeutic endoscopy.

3.2 Study period

The research was carried out from July 2023 till December 2024 where the data collection period was from December 2023 till May 2024.

3.3 Study location

Study was conducted in Department of General Surgery in HSAJB which is a tertiary public hospital. This department includes the Upper Gastrointestinal Surgery Unit which is a referral centre for all upper gastrointestinal related disease in Johor.

3.4 Reference population

Reference population for this study were adult patients who presented with upper gastrointestinal bleeding in Johor.

3.5 Source population

Source population were adult patients presenting upper gastrointestinal bleeding to HSAJB.

3.6 Study population

Adult patients with upper gastrointestinal bleeding who underwent endoscopy in Department of General Surgery, HSAJB from 2018 to 2023.

3.7 Sampling frame

Sampling frame were made up of adult patients who underwent endoscopy for upper gastrointestinal bleeding ulcer in Department of General Surgery from 1st January 2018 until 31st December 2023 and fulfilled the inclusion criteria and do not meet any exclusion criteria.

3.8 Study participants

3.8.1 Inclusion criteria for both cases and controls:

Table 3.1 Inclusion criteria of study participants

Cases	Control
Adults >18 years of age	Adults >18 years of age
Presenting with non-variceal bleeding in Department of General Surgery, HSAJB from 2018 – 2023	Presenting with non-variceal bleeding in Department of General Surgery, HSAJB from 2018 – 2023
Had rebleeding after successful endoscopic treatment	Did not have rebleeding after successful endoscopic treatment

3.8.2 Exclusion criteria for both cases and controls:

- History of previous gastrointestinal surgery or radiological intervention procedure
- Tumour or clearly malignant ulcers at endoscopy (i.e patients with large flat, plaque like, ulcerated tumours)
- Non ulcer lesions such as Dieulafoy's lesions and Mallory-Weiss tears
- ASA grade 5 (due to severe clinical condition)
- Pregnant

- Data missing for > 20% of the variables

3.9 Sample size determination

Sample size calculations were tailored to each objective to ensure adequate power for the predictors. The largest sample was used as the study's sample size (Dupont & Plummer, 1990).

Calculation for the predictors (Table 3.2) in primary objective 1 was done using Epi Info software version 7.2.5.0 (Dean *et al.*, 2011) using the unmatched case control study design. The Type I error was set at 5%, power was 80% and the ratio of controls to cases was 4:1. Medium effect size, Cohen's d=0.5 (Chen, Cohen & Chen, 2010) which corresponded to odds ratio of 2.5 after conversion using effectsize package in R studio was used while comparing proportions between both the arms. Medium effect size was chosen as it is able to estimate reasonable sample size required to produce a clinically significant difference between the case and control group (Aarts, Van Den Akker & Winkens, 2014).

Table 3.2 Sample size calculation for independent variables for objective 1

Factor	Proportion (%)	Alpha error	Power	Odds ratio	Sample for case	Sample for control	Total
Ischemic heart disease[a]	20.9	0.05	80%	2.5	61	243	304
Chronic kidney disease[a]	11.6	0.05	80%	2.5	87	347	434
NSAIDs /anticoagulant / antiplatelet use[b]	72	0.05	80%	2.5	88	351	439
High risk stigmata[b]	48	0.05	80%	2.5	56	222	278
Site of ulcer(duodenal)[b]	50	0.05	80%	2.5	57	226	283
Size of ulcer[c]	30.5	0.05	80%	2.5	54	214	268

[a](Jairath *et al.*, 2012)
[b](Laursen *et al.*, 2022)
[c](Bratanic *et al.*, 2013)

Sample size calculation for primary objective 2 was done using pmsampsize package in R Studio based on calculation of minimum sample size for developing prediction model by (Riley *et al.*, 2019) . C statistic of 0.85 was based on the comparison of different risk scores for predicting rebleeding in UGIB (Stanley *et al.*, 2017). Prevalence of rebleeding used is 0.2 as case control study design with 1:4 ratio is being used.

```
library(pmsampsize)
pmsampsize(type="b",cstatistic=0.85,parameters=14,prevalence=0.2)
```

```
## Given input C-statistic = 0.85  & prevalence = 0.2
## Cox-Snell R-sq = 0.2495
##
## NB: Assuming 0.05 acceptable difference in apparent & adjusted R-squared
## NB: Assuming 0.05 margin of error in estimation of intercept
## NB: Events per Predictor Parameter (EPP) assumes prevalence = 0.2
##
##            Samp_size Shrinkage Parameter CS_Rsq Max_Rsq Nag_Rsq  EPP
## Criteria 1       432     0.900        14 0.2495   0.632   0.395 6.17
## Criteria 2       378     0.888        14 0.2495   0.632   0.395 5.40
## Criteria 3       246     0.900        14 0.2495   0.632   0.395 3.51
## Final            432     0.900        14 0.2495   0.632   0.395 6.17
##
##  Minimum sample size required for new model development based on user inputs = 432,
##  with 87 events (assuming an outcome prevalence = 0.2) and an EPP = 6.17
##
##
```

Figure 3.1 Sample size calculation for objective 2

The minimum number of participants required to answer the objectives was 439.

3.10 Sampling method

The estimated number of adult patients who have undergone endoscopy for upper gastrointestinal bleeding ulcer in Department of General Surgery, HSAJB during the study period is 500. The minimum sample size required is 439 participants. In view of these factors and considering the inclusion and exclusion criteria, all eligible participants were included. No sampling method was applied.

3.11 Research tool

Retrospective data was obtained from the medical records of selected participants. Permission was obtained from Head of Department, Department of General Surgery, HSAJB to access the medical records. Data were extracted using a standardised collection form (Appendix 1), ensuring consistency across the variables of interest. Ideally the model development and validation should be done using different datasets. However, as this study utilised limited data from a single centre, the same dataset was used for both development and validation of the scoring system. In HSAJB, endoscopes are being done by gastroenterologists, general surgeons, and upper gastrointestinal surgeons. Only data from endoscopes done by the upper gastrointestinal surgeons and general surgeons during the study period was used for this study.

3.12 Data collection

The list of patients who underwent endoscopy during the study period was obtained from the Department of General Surgery, HSAJB. Initial screening was done based on indication for scope to exclude those who had endoscopy done for indications other than bleeding. The clinic cards of the remaining patients were traced and the cards that fulfil the inclusion and exclusion criteria were retained for the study data to be extracted. Any discrepancies were discussed within the study team which included a consultant upper gastrointestinal surgeon and a final conclusion were made regarding the selection of the participant.

3.13 Data management

Each selected patient was given a unique identification number and all personal identifiers such as full name and identity card number were removed. Data of all the variables of interest were extracted from clinic cards, coded and documented into the excel sheet. Duplicates were checked by using identification number and removed. Inconsistencies in the data were checked and corrected by referring back to the relevant clinic cards. There were no outliers. Patients who had data missing for more than 20% of the variables were excluded from the analysis as including these patients may affect the validity of the analysis(Schulz & Grimes, 2002)

Once data cleaning was done, recoding was done for certain variables. Age in years which was collected in a numerical form was categorized into those above 65 years old and those aged 65 years and below as international papers generally present using cut off value of 65 years. (Cheng *et al.*, 2010). Presence of either chronic kidney disease or End Stage Renal Failure were combined into a single variable, presence of chronic kidney disease with values of either yes or no. Next, use of NSAIDs, anticoagulant or antiplatelet drug were combined together to form the variable offending drug with the value of either yes or no. Site of ulcer were recategorized into either gastric ulcer or duodenal ulcer. Size of ulcers were divided into those with ulcer size more than 2cm and those with ulcer size of 2cm and below. Lastly, for stigmata of recent haemorrhage, they were divided into low risk (Forrest IIc and III) and high risk (Forrest Ia, Ib, IIa, IIb) (Monteiro, 2016).

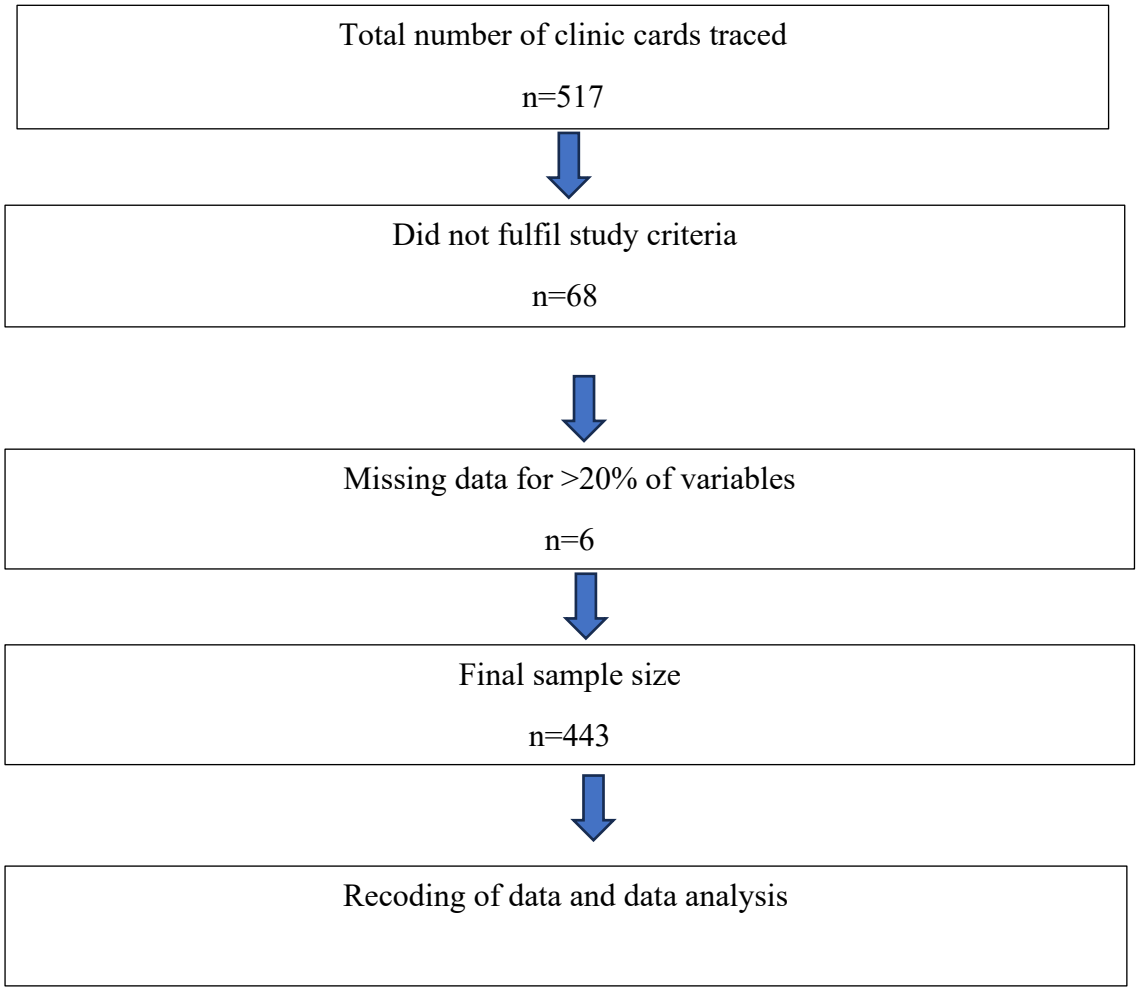

Figure 3.2 Flowchart for data management

3.14 **Operational definitions**

These were the operational definitions used for the variables:

Table 3.3 Operational definition

Variable	**Definition**
Ischemic heart disease	History of acute coronary syndrome or angina
Chronic kidney disease	eGFR < 60ml/min/1.73m that is present > 3months(Ministry of Health, 2018)
End stage renal failure	GFR < 15ml/min/1.73m^2 (Levin & Stevens, 2013)
Endoscopic failure	Rebleeding within 30 days of successful initial therapeutic endoscopy