Impact of Blood Pressure Measurement Errors on CVD Risk Prediction in New Zealand

A thesis submitted to Auckland University of Technology
in fulfilment of the requirements for the degree of
Doctor of Philosophy

Supervisors:
Dr. Victor Miranda
Prof. Andrew Lowe
Dr. Tet Chuan Lee

By
Tanvi Chandel

2025

School of Engineering, Computer and Mathematical Sciences
Auckland University of Technology

Abstract

Cardiovascular disease (CVD) includes conditions affecting the heart and blood vessels, such as coronary artery disease, stroke, and heart failure. It is a leading global cause of death, increasing healthcare costs and reducing quality of life. CVD develops due to genetic, lifestyle, and environmental factors, making early identification and prevention essential. Accurate CVD risk prediction helps identify high-risk individuals, enabling timely interventions. Risk models estimate cardiovascular event likelihood based on factors like age, sex, smoking, diabetes, blood pressure (BP), and cholesterol, guiding clinical decisions and treatment strategies.

BP is a key CVD risk factor and is widely used in risk prediction equations due to its strong association with cardiovascular events. Accurate BP measurement is essential, as errors can misclassify individuals into incorrect risk categories. These errors stem from factors such as improper cuff size, white-coat hypertension, observer bias, and device inaccuracies. Previous studies found that rounding BP to zero end-digits mostly led to overtreatment. New Zealand introduced the PREDICT-1 equation in 2018, but the impact of rounding on this model remains unknown.

Automated BP measurement devices, though widely used, have inherent inaccuracies. International standards such as ISO and ANSI/AAMI SP10 acknowledge these errors and set acceptable thresholds for reliability. However, some degree of error remains unavoidable, and its impact on CVD risk prediction is not fully understood. Additionally, many clinical studies on device accuracy do not adhere to international standards, either exceeding allowable error margins or using insufficient sample sizes ($N<85$). The implications of using smaller samples remain unexplored, leading to inconsistencies in adherence to these criteria.

Objective: The main objective of this research is to evaluate the impact of BP measurement errors, specifically rounding to the nearest zero end-digit and device inaccuracies, on CVD risk prediction in the NZ population. This study also aims at examining how BP measurement errors affect risk estimation for Māori and Pacific individuals compared to Europeans. Furthermore, this research aims at providing a new statistical framework to assess the impact of not adhering to the international standards and necessary modifications to be adapted to use smaller sample sizes.

Findings: The results show that even though there were slight variations in the overall Cox PH model predictability, there were notable changes in the risk classification. With just rounding approximately 4.24% of high-risk men and 3.21% of high-risk women are misclassified into lower-risk categories. Additionally, 1.19% of men and 0.62% of women are overclassified into the moderate-risk group, while 0.47% of men and 0.20% of women are overclassified into the high-risk group. Over 5 years, these misclassifications could cost the healthcare system approximately NZD $1.57 million, with potential expenses reaching up to NZD $8.2 million.. After adjusting the BP readings for the device inaccuracies within the acceptable range, the maximum observed misclassification rates showed that up to 7.50% of men were overclassified into higher-risk categories, which is around 5.65% for women. The findings also suggest disparities among ethnic groups, with Māori exhibiting the highest rates of misclassification, hence increasing the risk of both undertreatment and overtreatment. Additionally, the results also highlight the changes in the acceptance region for different sample size to ensure adherence to the existing standards. The use of proposed methodology resulted in identifying various clinical studies not adhering to the criteria.

Conclusion: While risk equations like PREDICT-1 play a crucial role in guiding clinical decisions and improving health outcomes, the continued use of measurement practices with known errors, without fully understanding their consequences, remains a concern. The results confirm that even small inaccuracies in BP measurements can lead to notable misclassification, resulting in both overtreatment and undertreatment, with direct implications for healthcare costs. These results emphasise the importance of improving BP measurement practices to enhance the accuracy of CVD risk prediction and ensure equitable healthcare outcomes.

Table of Contents

Attestation of authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the qualification of any other degree or diploma of a university or other institution of higher learning.

Tanvi Chandel

February, 2025

Co-authorship contribution

STUDENT AND SUPERVISOR APPROVALS

By signing you are confirming that the co-author contributions stated in the table(s) below are accurate.

Student Name	Tanvi Chandel	**Signature**		**Date**	17/02/2025
Supervisor Name	Dr. Victor Miranda	**Signature**		**Date**	20/Feb/2025

Chapter Number:	Chapter 3
Manuscript Title:	Zero End-Digit Preference in Blood Pressure and Implications for Cardiovascular Disease Risk Prediction—A Study in New Zealand
Publication Status:	Published
Reference if published:	https://doi.org/10.3390/jcm13226846

AUTHOR SURNAME: (order as per manuscript)	**CONTRIBUTION** (May copy from the guidelines above)
Chandel	Conceptualisation, Methodology, Formal Analysis, Investigation, Writing – original draft presentation, Visualisation
Miranda	Methodology, Validation, Writing—review and editing, Supervision
Lowe	Methodology, Validation, Writing—review and editing, Supervision, funding acquisition
Lee	Methodology, Validation, Supervision

Chapter Number:	Chapter 6
Manuscript Title:	Blood Pressure Measurement Device Accuracy Evaluation: Statistical Considerations with an Implementation in R
Publication Status:	Published
Reference if published:	https://doi.org/10.3390/technologies12040044

AUTHOR SURNAME: (order as per manuscript)	**CONTRIBUTION** (May copy from the guidelines above)
Chandel	Conceptualisation, methodology, formal analysis, writing—original draft preparation, visualization
Miranda	Conceptualisation, methodology, validation, supervision
Lowe	validation, writing—review and editing, supervision, funding acquisition
Lee	writing—review and editing, supervision

Acknowledgements

I would like to express my deepest gratitude to my primary supervisor, Dr. Victor Miranda for his unwavering support, invaluable guidance, and encouragement throughout this research. His expertise, insightful feedback, and commitment to excellence have been instrumental in shaping this thesis. I also extend my sincere appreciation to my secondary supervisors, Prof. Andrew Lowe and Dr. Tet Chuan Lee, for their valuable insights and constructive feedback. Their support and expertise have played a crucial role in strengthening this research, and I am thankful for their guidance throughout this journey.

A special thanks to Prof. Rod Jackson and the VAREANZ research team for providing access to the PREDICT cohort data and for their guidance on the relevant aspects of this work. Their contributions were essential in enabling this research, and I truly appreciate their support.

I am also incredibly grateful to my Manager Mrs. Joyce D'Mello for always being a constant source of support, both professionally and personally. Your ability to ensure that everything runs smoothly has made this journey significantly easier, and I sincerely appreciate all your efforts.

Lastly, I would like to thank my parents, my sister for their unconditional love, patience, and constant encouragement. Your belief in me has been my greatest source of strength. To my friends, thank you for your support, motivation, and for always being there when I needed it most. This journey would not have been the same without all of you.

List of Abbreviations

ANSI	American National Standards Institute
AAMI	Association for the Advancement of Medical Instrumentation
ADA	American Diabetes Association
ACCORD	Action to Control Cardiovascular Risk in Diabetes
BMI	Body Mass Index
BP	Blood Pressure
BHS	British Hypertension Society
CVD	Cardiovascular Disease
CHD	Coronary Heart Disease
Cox PH	Cox Proportional Hazard Model
C-INDEX	Concordance Index
DBP	Diastolic Blood Pressure
ESH	European Society of Hypertension
FRS	Framingham Risk Score
GP	General Practitioner
HR	Hazard Ratio
IHD	Ischemic Heart Disease
ICU	Intensive Care Unit
ISO	International Organization for Standardization
ICD	International Classification of Diseases
IQR	Interquartile Range
JBS	Joint British Societies
MI	Myocardial Infarction
MRFIT	Multiple Risk Factor Intervention Trial
MELAA	Middle Eastern/Latin American/African
MLE	Maximum Likelihood Estimate
MOH	Ministry of Health
NZ	New Zealand
NZD	New Zealand Dollar

NSTEMI	non-ST-elevation myocardial infarction
NZDEP	New Zealand Deprivation Index
NIBP	Non-Invasive Blood Pressure
PCE	Pooled Cohort Equation
PREVENT	Predicting Risk of CVD Event
PWA	Pulse Wave Analysis
PTT	Pulse Transit Time
PPG	Photoplethysmography
PHARMAC	Pharmaceutical Management Agency
PAT	Pulse Arrival Time
QRISK	QRESEARCH Cardiovascular Risk
ROC	Receiver Operating Characteristic
SCORE	Systemic Coronary Risk Evaluation
SBP	Systolic Blood Pressure
SD	Standard Deviation
SPRINT	Systolic Blood Pressure Intervention Trial
SNS	Sympathetic Nervous System
TC:HDL	Total Cholesterol to High Density Lipoprotein Cholesterol
TIA	Transient Ischemic Attack
US	United States
USD	United States Dollar
WHO	World Health Organisation

List of Figures

List of Tables

CHAPTER 1. INTRODUCTION

1.1 Background

1.1.1 Global Impact of Cardiovascular Disease

Cardiovascular Disease (CVD) refers to a group of conditions affecting the heart and blood vessels, including conditions such as coronary heart disease (CHD), cerebrovascular disease, peripheral artery disease, heart failure, congenital heart disease, and rheumatic heart disease [1]. CVD does not have a specific aetiology. The principal cause of CVD is atherosclerosis, characterised by the accumulation of fatty plaques on arterial walls, which consist of fat, cholesterol, calcium and various other components [2]. As time progresses, the plaques harden, constricting the opening of the arteries and impeding blood circulation as represented in Figure 1. The plaques tend to rupture over time, resulting in a thrombus (blood clot) that can restrict or completely obstruct blood flow throughout the body [3].

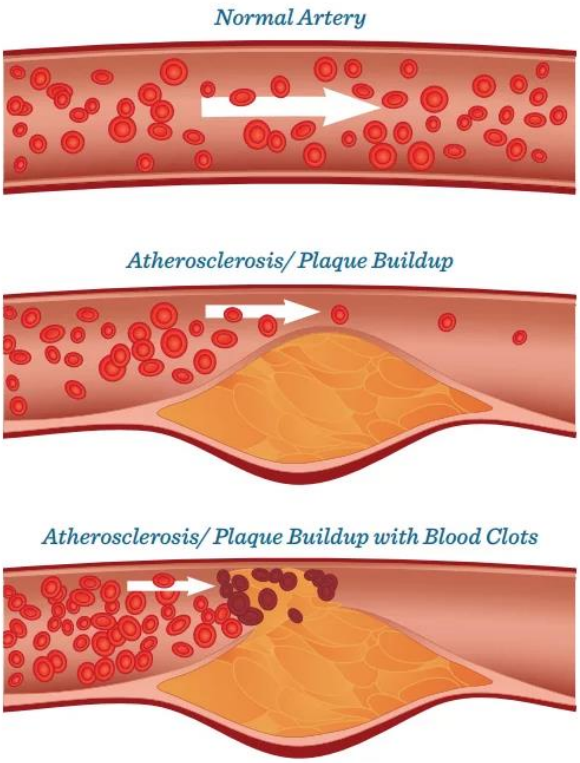


Figure 1: Progression of atherosclerosis, the primary cause of CVD. Source: The Heart Foundation, Atherosclerosis (2018).

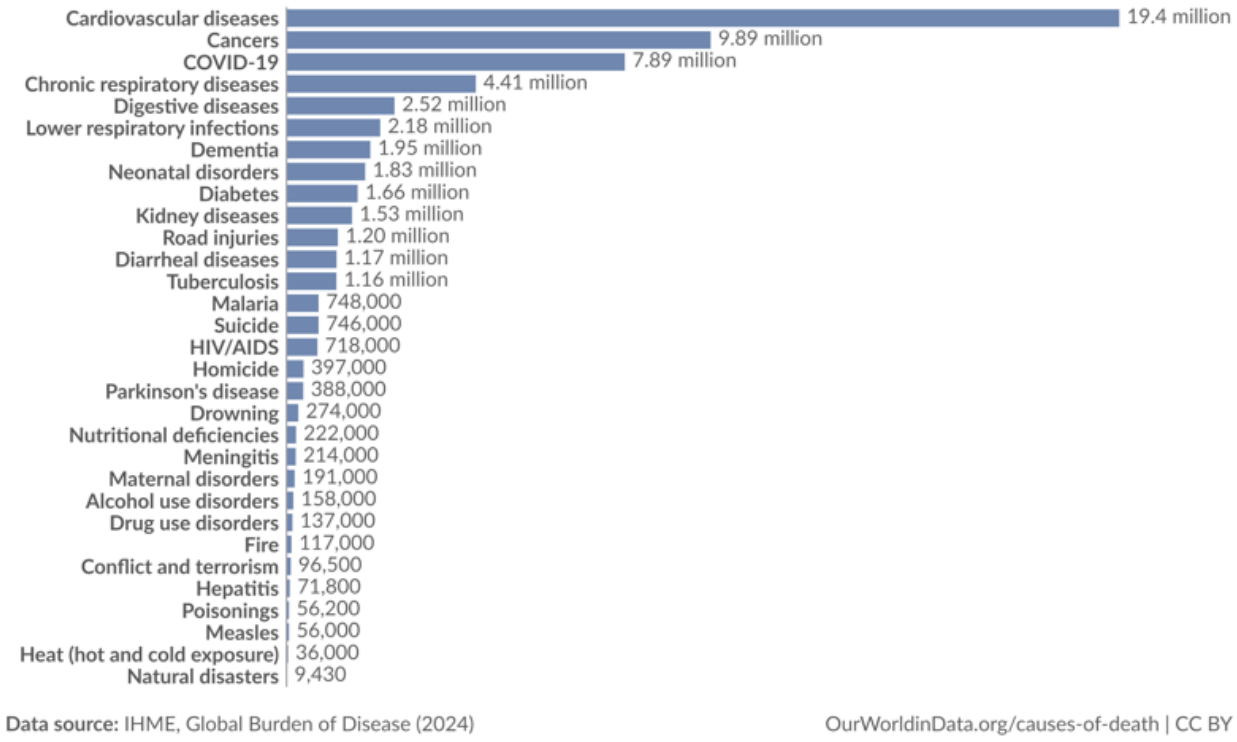


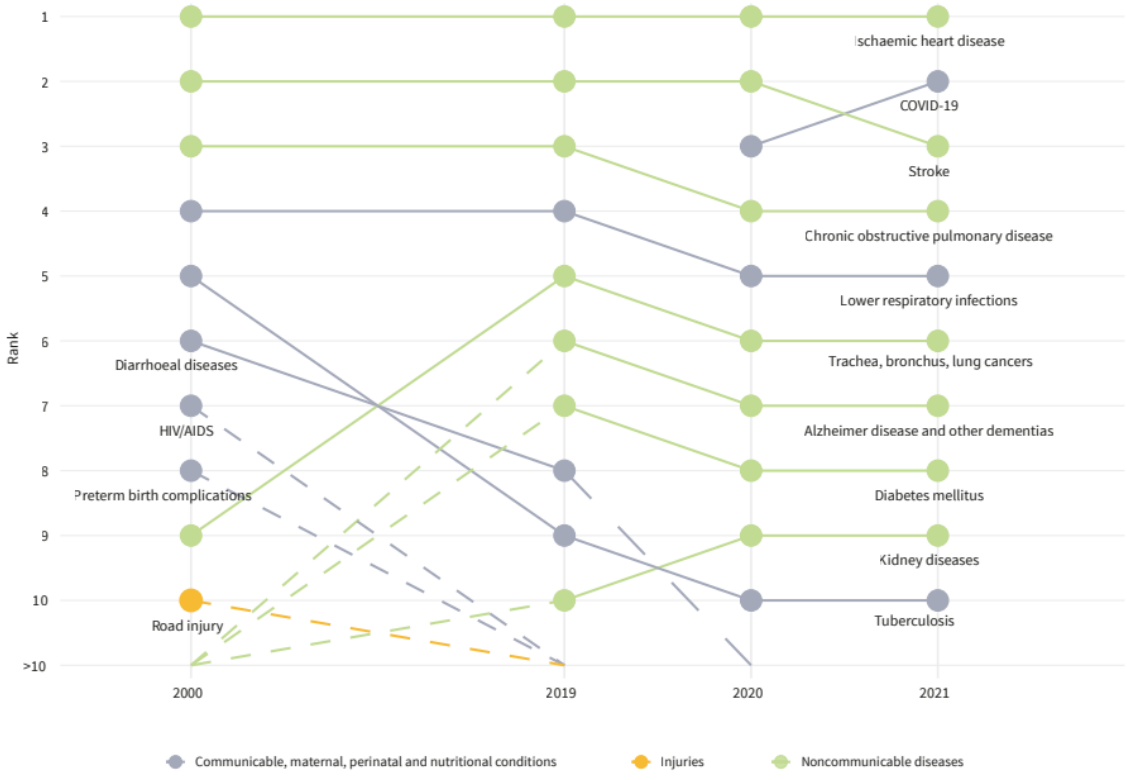
Figure 2: Leading global causes of death in 2024, highlighting cardiovascular diseases as the top cause with 19.4 million deaths, followed by cancers and COVID-19. Data source: IHME, Global Burden of Disease (2024).

CVDs are the predominant cause of premature death and morbidity among non-communicable diseases globally. Figure 2 illustrates the global burden for the year 2024, which highlights CVD being the leading cause of worldwide mortality, followed by cancers, COVID-19 and respiratory illnesses among others [4].

More than three quarters of global fatalities result from CVDs, predominantly in low- and middle-income nations, with an anticipated increase to over 23 million by 2030 from the current annual total of over 17 million deaths worldwide [5]. This difference in low- and middle-income nations is influenced by several variables, including restricted access to healthcare, less awareness of CVD risk factors, and an increased prevalence of risk factors such as hypertension, diabetes, and tobacco consumption [6]. Moreover, most CVD fatalities result from heart attacks and strokes, accounting for about 85% of deaths [7][5].

According to the 2021 World Health Organisation (WHO) factsheet, ischemic heart disease (IHD) and stroke have been the predominant causes of mortality for the past 21 years (Figure 3) [8], both of which are types of CVD. In the United States (US), it is estimated that around 1.54 million

individuals over the age of 20 years have IHD, constituting 6.4% of the population (7.9% of men and 5.1% of women). CVD is estimated to account for 4 million deaths annually in Europe, representing 47% of total fatalities [9]. Globally there are 12.2 million new strokes annually, which is equivalent to one stroke occurring every 3 seconds [10].



[1] Figure 3: Trends in the global top 10 causes of death from 2000 to 2019, illustrating the shift from communicable diseases. Source: WHO

CVD is also a major factor in chronic disability. Individuals who survive CVD events frequently have prolonged health repercussions, encompassing diminished quality of life and increased susceptibility to subsequent occurrences [6]. The chronic nature of CVD places a significant strain on healthcare systems, requiring continuous treatment, medication, and rehabilitation services, especially in ageing populations with a higher prevalence of CVD [11].

A study conducted in 2011 by the World Economic Forum and Harvard School of Public Health emphasises the significant economic burden of CVD, which incurs substantial healthcare expenditures worldwide like hospital care and medications. In 2010, the global expenditure on CVD

[1] Solid lines represent movement within the top 10 causes of death. Dashed lines represent movement in or out of the top 10 causes of death.

was estimated at $863 billion US Dollars (USD) with projections indicating a 22% increase to $1,044 billion USD by 2030, disproportionately impacting low- and middle-income countries, where healthcare resources are more limited, resulting in greater challenges in managing the disease and its long-term consequences. The economic impact of CVD transcends healthcare expenditures, influencing overall economic stability and growth, especially in countries with limited healthcare resources to address the rising incidence of the disease [12].

CVD also poses a social and psychological influence. Chronic diseases such as heart disease and stroke frequently induce significant emotional and psychological distress, affecting both patients and their caregivers, leading to anxiety, depression, and social isolation[13]. A study examining the relationship between psychological distress and health perception in patients with a previous myocardial infarction (MI) involved 2436 patients with a history of MI or stroke [14]. Their findings indicate a significant association between psychological distress and poor health perception in survivors of CVD. Patients with a history of previous MI or stroke, who experience anxiety or depression have lower odds of reporting a positive perceived health status compared to their counterparts without these mental health conditions. The likelihood of positive health perception was 48% lower for those with anxiety and 55% lower for those with depression in post-MI patients. Similarly, post-stroke patients with anxiety had 39% lower odds, and those with depression had 63% lower odds of positive health perception [13].

1.1.2 Burden of CVD: The New Zealand perspective

In New Zealand (NZ) CVD is one of the leading causes of mortality [15]. Almost one in 21 adults suffer from CVD [16] and every 90 minutes, a New Zealander succumbs to the condition[15]. As of 2023, approximately 180,000 people are presently living with heart disease [17]. Figure 4 illustrates the ten leading causes of mortality in NZ. IHD, a type of CVD, is the primary cause of death in the country, followed by stroke. Recent mortality numbers indicate that in 2017, IHD accounted for 41% of female deaths and 59% of male deaths [18].

CVD has been increasing in recent years, partly due to the growing prevalence of lifestyle-related conditions such as obesity, diabetes, hypertension and smoking, which are some of the key risk

factors for CVD[19]. Approximately 1.26 million adults in NZ are classified as obese defined as having a Body Mass Index (BMI) of 30 or more [16].

Obesity is a primary contributor to insulin resistance, often leading to the development of type 2 diabetes, which substantially increases the risk of CVD [20][21]. Currently, around 5% of the overall population is affected by type 2 diabetes, with projections indicating an increase to 7% by 2040, equivalent to an estimated 430,000 individuals with the condition [22]. The prevalence of diabetes is

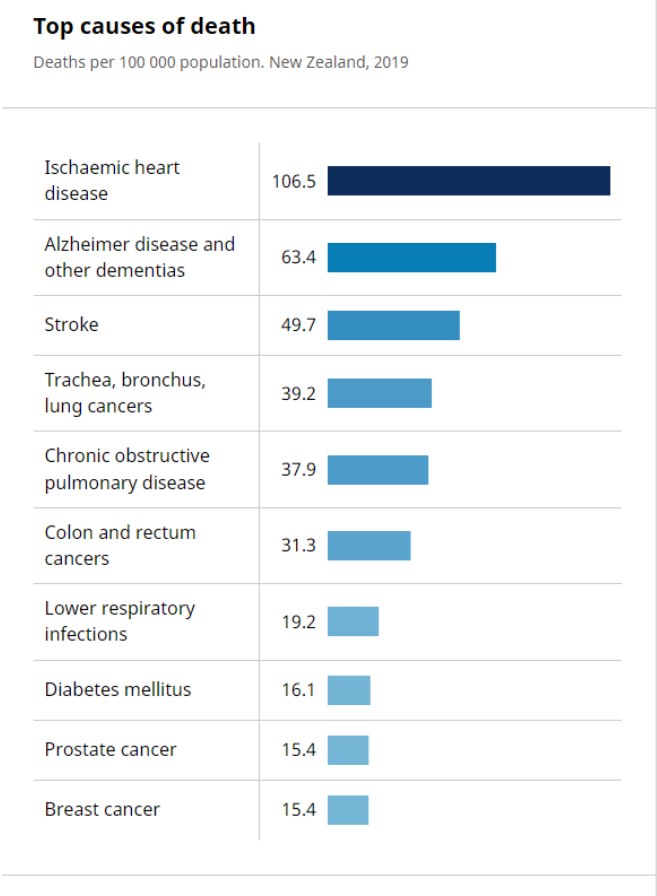


Figure 4: Top causes of death in 2019 in New Zealand. Source: WHO.

highest in older demographics, approximating 15 – 20% among individuals over 65 years; yet there is a rising incidence among younger populations in NZ [23]. Data from the Ministry of Health indicates rise in the incidence of type 2 diabetes among those aged 30 to 39 years, with rates nearly doubling from 2006 to 2018. Māori, Pacific and South-Asian ethnicities, as well as individuals from lower socioeconomic backgrounds, experience disproportionately higher rates of obesity and type 2 diabetes. The prevalence of diabetes among Māori and Pacific individuals aged 25 to 39 years is almost two to three times greater than Europeans [24].

Active smoking and exposure to second-hand smoke determine more than 30% of CHD mortality. Approximately 5,000 people die from smoking every year in NZ [25]. Female smokers have a higher mortality rate from CVD compared to male smokers, and they possess a 25% increased risk of developing CHD, relative to males with equivalent tobacco exposure [26]. In addition, CVD is a significant contributor to declining health among Māori population and contributes to health disparities in NZ concerning both ethnicity [27], [28] and socioeconomic status [29].

High blood pressure (BP) is a major issue affecting 1 in 5 adults in NZ causing long-term arterial damage. It is a major risk factor for heart conditions and CVDs like CHD, stroke, MI, diabetes and heart failure [30].

In 2022, CVD accounted for roughly 11,900 fatalities and 183,000 disability-adjusted life years in NZ [31]. The financial impact of CVD on the healthcare system is expected to be roughly 3.3 billion New Zealand Dollar (NZD) each year. The economic loss attributable to CVD in NZ is estimated at NZD 614.88 million annually, representing 15.6% of the overall income loss from all diseases [32] .

1.2 Prevention of CVD

1.2.1 Lifestyle change via CVD risk prediction equations

An individual's risk of future cardiovascular events can be reduced through lifestyle changes and preventive medical interventions. Lifestyle modifications may encompass reduction of smoking, adherence to a nutritious diet, and consistent physical activity, among others [33], while preventive medical interventions may encompass statins, low-dose aspirin, and antihypertensive therapy, among others. Consequently, predicting an individual patient's risk is crucial for determining the appropriate lifestyle modifications and preventative medical interventions [34].

Various risk models have been established to predict the CVD risk of individual patients. Examples include the Framingham Risk Score (FRS), Pooled Cohort Equations (PCE), Systematic Coronary Risk Evaluation (SCORE) and QRESEARCH cardiovascular risk algorithm (QRISK) [35], [36], [37]. In 2018, NZ formulated its own CVD risk prediction equation, known as PREDICT-1 equation which assesses an individual's likelihood of experiencing a CVD event within five years[38]. These

equations were developed using the PREDICT cohort, an ongoing open cohort study conducted in NZ. The cohort enlists people following the completion of standardised CVD risk assessments and the equations were formulated via a survival analysis approach, accounting for several risk factors such as:

1) Age
2) Ethnicity
3) New Zealand Deprivation Index (NZDep)
4) Ex-Smoker
5) Current Smoker
6) Family History of premature CVD
7) Atrial Fibrillation
8) Diabetes
9) Systolic Blood Pressure (SBP)
10) Total Cholesterol to High-Density Lipoprotein Cholesterol (TC: HDL)
11) BMI
12) On BP lowering medication
13) On lipid lowering medication
14) On either antiplatelet or anticoagulant medications.

1.2.2 Blood Pressure measurement

1.2.2.1 Importance of Blood Pressure Measurement

To date, all CVD risk equations published by different countries for various ethnically diverse populations have found BP as a significant variable in predicting CVD risk [39], [40], [41]. BP measurement returns two values, for instance, 130/90 mmHg. These values are systolic blood pressure (SBP) and diastolic blood pressure (DBP) values, respectively. SBP is deemed a better predictor as it is highly correlated with mortality rate due to heart diseases and stroke [42].

SBP is also considered a significant predictor in the PREDICT-1 CVD risk equations, developed in NZ. It is included as a continuous variable, with Hazard Ratios (HRs) calculated per 10 mmHg

increase. SBP also interacts significantly with age and BP lowering medication use in the PREDICT-1 equation [38].

1.2.2.2 BP measurement Techniques

The techniques of BP measurement can be divided into two categories: Direct (invasive) and indirect (non-invasive) [43]. The invasive intra-arterial cannulation technique yields the most accurate beat-to-beat BP measurements, considered the gold standard in BP monitoring. This approach requires an invasive catheter to directly engage with blood flow, while the pressure transducer transforms the mechanical signal generated by BP into a real-time electrical signal. Ultimately, the electrical signal is conveyed and presented as a pulse waveform to the physicians [44]. Although the invasive approach effectively and reliably reflects real-time intravascular pressure, its high cost, professional operational demands, and potential risks to patients restrict its applicability. It is particularly inappropriate for home-based surveillance. Consequently, it is typically utilised in Intensive Care Unit (ICU) environments[43].

A sphygmomanometer, also referred to as a BP monitor or BP gauge, is a traditional cuff-based device, which is non-invasive, for measuring BP. The traditional manual auscultation method utilises three elements: an inflating cuff, a stethoscope, and a measuring device (either aneroid gauge or mercury manometer). The operational principle is inflating the cuff encircling the artery (in the arms or legs) to obstruct blood flow. During the gradual deflation of the cuff, the SBP is identified at the initial occurrence of the Korotkoff sound, while the DBP is conventionally noted at the point where the fourth Korotkoff sound is faintly discernible to skilled professionals using a stethoscope [45]. This conventional manual sphygmomanometer remains utilised in certain clinical contexts owing to its adequate accuracy, lower cost, and non-invasive nature. The necessity for skilled staff to operate the sphygmomanometer restricts its application in domestic environments [46].

Although invasive BP measurement techniques are deemed more precise, non-invasive methods are favoured in most clinical environments [47]. However, this method introduces a degree of error in BP measurement, such as overestimating BP in patients with higher BMI due to improper cuff size [48].

As a result, established standards, such as American National Standards Institute/ Association for the Advancement of Medical Instrumentation (ANSI/AAMI) SP10 and the International Organization for Standardization (ISO) [49], outline the performance requirements and validation procedures for non-invasive blood pressure measurement devices. These criteria outline an acceptable error range and establish a threshold for the number of individuals required to assess the inaccuracy of automated cuff-based BP devices. These BP errors are in the form of mean error and standard deviation (SD) of the error. The standard assumes that errors follow a normal distribution, with the mean error representing the average deviation from the true value. The maximum allowable SD can range up to 6.95 mmHg when the mean error is 0 mmHg, indicating the acceptable level of variability in the device's measurements. This threshold ensures that the BP device provides accurate and consistent results within the specified limits.

1.3 Research gaps

1.3.1 Impact of different sources of BP errors

BP measurements are subject to errors such as inappropriate patient positioning, white coat effect, improper cuff-size, excessive deflation rate and rest period before measurement [50]. As a result, there are procedures in place which must be considered by an expert professional before measuring a patient's BP. Despite these procedures and acceptable accuracy ranges for BP measurement devices, outlined by the international standards the inaccuracies in BP measurement cannot be entirely avoided. Consequently, even if the measurement errors are within the acceptable range, they can still have significant implications for clinical decision-making and patient outcomes. The impact of these errors that arise due to the use of automated cuff-based BP devices, particularly on CVD risk prediction in NZ, remains largely unexplored, and further investigation is needed to understand their role in clinical practice.

Another common source of error observed while measuring BP is the zero end-digit preference, while using a manual sphygmomanometer [51]. It is a tendency to round the BP measurement to the nearest zero end-digit, which is a common practice in the clinical setting [52]. Previous work on the impact of zero end-digit preference on classification of patients for management in primary care in

NZ have utilised three CVD risk equations from other countries. However, the extent to which errors in BP measurements lead to shifts in predicted CVD risk or their impact on CVD outcomes, when using the PREDICT-1 equation specifically tailored to the NZ population, remains unknown. There is also a lack of comprehensive analysis regarding the resulting healthcare burden due to these misclassifications.

The PREDICT-1 equation is currently being used to stratify patients into three risk categories (<5%, 5-15%, and >15%) as established by the Ministry of Health [53]. These risk categories guide clinicians in recommending appropriate interventions, setting goals, and planning follow-up care based on the assessed cardiovascular risk. Thus, this research also examines the impact of BP measurement inaccuracies on the misclassification of patients into various CVD risk groups within New Zealand's diverse population.

1.3.2 Non-Adherence to the existing standards

While reviewing the range of BP device inaccuracies, it was identified that clinical studies frequently fail to comply with the standards that outline requirements for device testing [54]. However, currently, formal clinical validation is not mandated in all countries for BP measurement devices, with evidence indicating that less than 20% are validated according to established methods [55]. The growing use of out-of-office BP monitoring in clinical practice, predominantly conducted with electronic devices, alongside the gradual substitution of auscultatory devices in offices or clinics with electronic alternatives, has crucially rendered clinical validation more essential nowadays [56], [57].

According to ISO, validation of BP measurement device requires a minimum sample size of 85 participants to evaluate the inaccuracy of BP devices. Also, it mandates that at least 85% of BP errors must lie within −10 to 10 mmHg. This range is also termed as a *tolerable error limit* and includes procedures such as using appropriately sized cuffs and ensuring correct patient posture during measurements. Keeping these protocols under consideration requires that the mean error and SD of the error must be within an acceptable limit as specified by the standard.

Examining clinical studies on BP measurement device errors revealed variations in reported inaccuracies. Some studies applied acceptable limits set for 85 or more samples, despite using

smaller sample sizes. For example, a novel calibration procedure using pulse transit time based BP measurement provides device inaccuracy within the acceptable range but using 10 participants. Still, these studies are considered in clinical studies because of the lack of information on variations in the acceptance limits for varied sample sizes. In some cases, collecting 85 samples may not be feasible, such as when working with populations affected by a rare disease or a specific age group [58].

To facilitate comparisons between studies using the required sample size (N=85), the standard provides the probability of accepting a device based on different combinations of mean and SD of errors. The probability of accepting a device in this context refers to the probability that a device will pass the performance criteria specified by the standard, under the assumption that all the procedural requirements of BP measurement are met. For a given sample size N=85, this involves determining the probability that at least 85% of the BP measurement errors fall within the range of -10 mmHg to 10 mmHg.

This probability is determined using uncertainty analysis via quadrature. Quadrature is a method used to approximate the combination of uncertainties from independent variables by summing their variances. In the standard, quadrature is applied to evaluate the SD of the probability distribution for tolerable error by combining the variances of the mean error and variance components. The result is expressed as the square root of the sum of their variances, weighted by partial derivatives, providing a measure of the overall uncertainty in the estimated probability of a tolerable error. This SD is expressed as:

$$SD_{ISO} = \frac{1}{\sqrt{2\pi(N-1)}} \quad (1)$$

However, these evaluations were based on approximations, that solely relied on the sample size without accounting for the influence of the mean and the SD of the BP errors.

1.4 Research objectives

The main goal of this research is to investigate the influence of inaccuracies in BP measurement on CVD risk prediction in NZ focussing on two prevalent sources of error: rounding errors and BP measurement device inaccuracies. Rounding errors have been previously examined in NZ using CVD risk equations developed in other countries. Hence, this study expands upon earlier research by assessing the impacts of rounding errors on the CVD risk prediction in NZ using the PREDICT-1 equation, specifically designed for the NZ population. The BP device inaccuracies, on the other hand, were selected due to the existence of international standards that specify the acceptable error limits for device inaccuracy. However, the impact of these acceptable errors on the overall CVD risk prediction and consequently on treatment decisions in NZ, remains unknown. To assess the economic burden resulting from these sources of BP errors, the research aims to evaluate the cost of CVD risk assessments specific to NZ. The results highlight the significance of precise BP readings in improving patient treatment associated with CVD mismanagement.

This research also investigates the impact of varying sample sizes on the acceptance limits for BP devices, particularly in cases where the standard requirement of N=85 participants is not met. The study introduces a comprehensive methodology aimed at enhancing the evaluation of the probability of accepting a device. This approach incorporates both the mean error and the SD of BP measurement errors, allowing for a more nuanced comparison of studies that report differing device inaccuracies across varying sample sizes.

The following objectives support the main goal of the study:

1. To analyse the impact of rounding BP measurements to the nearest zero end-digit on the CVD risk prediction in NZ
2. To analyse the impact of BP measurement device inaccuracy on the CVD risk prediction in NZ
3. To describe the potential effect of the prevalence of errors in SBP on the adjusted risk of different ethnic groups (European, Māori and Pacific).
4. To establish a framework to enable comparisons of different methodologies/algorithms being used to develop BP devices.

5. To assess how the BP device acceptability varies for different BP measurement devices and offer recommendations to perform adjustments to adhere to the standards.

1.5 Thesis synopsis

The thesis is composed of seven chapters. Chapter 2 systematically reviews previous CVD risk equations along with the one developed in NZ, which is currently being used for CVD management, highlighting the guidelines and interventions proposed by the government health body. It also highlights the importance of BP in CVD management and reviews previous research on different sources of BP measurement errors and their potential impacts along with a detailed summary of the international guidelines for clinical validations involving BP measurement devices.

Chapter 3 analyses the impact of rounding BP error values on the prediction of CVD risk within the NZ population. By comparing datasets with original and rounded SBP values, it evaluates the misclassification rates, model estimates in the form of HRs and their implications for clinical decisions. This chapter also includes a detailed financial cost analysis detailing an estimated burden on the healthcare sector due to these misclassifications. The findings emphasise the importance of precise SBP measurements in reducing risk misclassification and ensuring efficient resource allocation for targeted interventions and optimal patient management.

This chapter has been published under the name "*Zero End-Digit Preference in Blood Pressure and Implications for Cardiovascular Disease Risk Prediction—A Study in New Zealand*" in the journal "*Journal of Clinical Medicine*" (Q1, Impact Factor: 3.0, Database: Clarivate, Metric: JCR and Journal Category: Medicine).

Chapter 4 explores the impact of inaccurate BP measurement devices, within the allowable standards, on the CVD risk prediction in NZ. Simulations are performed for different ranges of BP errors, which are assumed to be normally distributed. The analysis results in comparison of misclassification rates arising from errors in observed BP due to device inaccuracies. Financial consequences due to these misclassifications are also studied, emphasising the need for stricter device standards.

Chapter 5 performs analysis to evaluate the impact of BP measurement errors from inaccurate devices on CVD risk prediction across Māori, Pacific, and European populations. Māori and Pacific individuals, being at higher risk for CVD, are particularly affected by misclassification due to these errors. A simulation study was conducted using the highest allowable SD of 6.95 mmHg to quantify misclassification rates across three risk groups and assess changes in CVD risk for various risk factors. The results emphasise disparities in misclassification and its implications for different ethnic groups.

Chapter 6 introduces a new methodology based on the *sampling distribution of sampling proportions* to study the impact of BP measurement errors for different sample sizes. This approach presents the fundamentals to compute the probability of a device meeting international standards, allowing for comparisons between technologies used in studies with smaller sample sizes. It helps in understanding how different sample sizes affect the likelihood of meeting acceptable BP error criteria, offering insights into device performance in real-world settings.

This chapter has been published under the name "*Blood Pressure Measurement Device Accuracy Evaluation: Statistical Considerations with an Implementation in R" in the journal "Technologies"* (Q1, Impact Factor: 4.2, Database: Clarivate, Metric: JCR and Journal Category: Engineering).

Chapter 7 summarises the academic contribution of this thesis and discusses potential future work with an emphasis on exploring other sources of BP errors and ways to mitigate them for better CVD outcomes for the NZ population.

It has to be noted that there will be some inevitable replication of information such as international standards statistical considerations and CVD guidelines from the Ministry of Health, due to the fact that this thesis is written in the paper-based format with Chapter 3 and Chapter 6 already published and Chapter 4 and Chapter 5 ready to be submitted.

CHAPTER 2. LITERATURE REVIEW

2.1 Overview

This chapter provides the foundations of CVD risk prediction and its relationship with BP measurement errors. It commences with a review of CVD risk prediction equations developed in other countries and those currently being employed in NZ.

The modelling relationship between BP and CVD is examined in detail, followed by a comprehensive explanation of various sources of BP measurement errors. The primary emphasis is on the errors arising from automated cuff-based BP measurement device inaccuracies and due to rounding to the nearest zero end-digit. However, additional sources of BP measurement errors are briefly addressed to provide broader context. The section also discusses international standards for allowable error ranges, a brief review of previous studies that fail to meet these criteria, and the implications of these errors for healthcare.

2. 2 CVD Risk Prediction Equations

Prediction models support decision-making and clinical reasoning by providing risk estimates based on patient-specific factors, enabling personalized care, risk stratification [59]. Treatment decisions are built based on the probability of an individual experiencing a cardiovascular event within a specified timeframe [60]. Understanding the risk level enables an individual and their practitioners to make informed decisions regarding the prevention and treatment of CVD, encompassing lifestyle recommendations, diabetes management, the prescription of lipid-modifying and antihypertensive medications, and/or post-MI or ischaemic stroke pharmacotherapy [61].

For over twenty years now, risk prediction equations have served as the foundation for CVD prevention programmes and have been utilised to transform data on various risk variables into a single estimate of a person's chance of having a heart attack or stroke during a certain time period [62]. Expert systems [63], regression analysis [64], fuzzy logic and rough sets [65], data mining [66],

and other approaches have been used to generate CVD risk prediction models. with regression analysis being the most common approach. Linear regression, binary logistic regression, and polynomial regression models are only a few of the methodologies accessible but in the field of clinical studies, survival analysis is a widely used statistical tool for predicting CVD [64].

Table 1: Region-specific CVD risk assessment tools and corresponding risk stratification criteria.

Region	Risk Assessment Tool	Risk Categories	Reference
United States	• Framingham Risk Score • Pooled Cohort Equations • PREVENT Equations	<10% 10-20% >20%	[67], [68], [69]
Europe	SCORE	<1% 1-5% 5-10% >10%	[37]
United Kingdom	QRISK	<10% 10-20% >20%	[70]
New Zealand	PREDICT-1	<5% 5-15% >15%	[38]

The primary application of CVD risk prediction equations is to classify individuals into distinct risk groups, which helps guide preventive and treatment strategies. Across countries, such risk groups are defined distinctly based on the estimated probability of experiencing a cardiovascular event within a specified timeframe. The thresholds for these groups vary depending on each country's assessment strategies, population characteristics, and healthcare priorities. Table 1 summarises the CVD risk categories established by various countries, highlighting the diversity in approaches to

managing cardiovascular risk globally. A detailed explanation of these equations is provided in subsequent subsection.

2.2.1 Framingham Risk Score

The original Framingham Risk Scores (FRS) were developed in 1998 using a population-based sample of the original Framingham Cohort, conducted in the US during a span of 10 years. The FRS were used to predict the 10-year CHD of an individual using the risk factor categories and were included in the Adult Treatment Panel's Third Report of the National Cholesterol Education Program (Adult Treatment Panel III), establishing risk assessment for primary prevention. In 2001, FRS became an essential component for risk prediction [71], although this framework was not widely adopted in primary care settings due to the lack of a single equation that could assess the patient's risk for all CVD events. This led to the development of an updated Framingham Risk equation in 2008 using Cox Proportional Hazard (Cox PH) models. The updated equation provides a single multivariate risk assessment tool to assess 10-year CVD, cerebrovascular events, peripheral artery disease, and heart failure [72]. The model included SBP and hypertension therapy as risk factors, both of which were highly significant in the overall prediction. Initially, DBP was employed, but its insignificant effects in the overall model led to the exclusion of DBP from further analysis.

2.2.2 Pooled Cohort Equation

The Framingham Risk Equation was developed using a population sample of middle-aged white Americans. In 2013, another CVD risk prediction equation, referred to as the Pooled Cohort Equation (PCE), was developed using cohorts predominantly representative of the whites and African American populations in the US. PCE are nowadays employed to predict the 10-yearAtherosclerotic CVD , defined as the first incidence of non-fatal MI, CHD related death, or fatal or nonfatal stroke in the US. Although, PCE was built on the previously developed FRS, they also include several ethnically and geographically varied datasets to improve the overall model's predictive capability [36].

While numerous studies have evaluated the performance of PCE, the results have not been conclusive. Some found acceptable calibration, and others found overprediction [73]. These studies

suggested that the PCE equations may incorrectly estimate risk in the non-US population [74]. Additionally, even within the US, the effectiveness of the PCE equations can be influenced by ethnic diversity, potentially leading to less precise predictions for certain demographic groups.

2.2.3 PREVENT Equation

The PREVENT equation also known as the Predicting Risk of CVD EVENTs equation is a new approach to CVD risk assessment. Unlike PCE which primarily aim to predict 10-year Atherosclerotic CVD risk in adults aged 40 to 79, the PREVENT equations broaden the risk prediction to encompass younger adults, starting from age 30, providing both 10-year and 30-year CVD risk assessments. The development and validation of the PREVENT equation included almost 6 million adults from both traditional cohort studies as well as real-world clinical data derived from electronic medical records[75].

The PREVENT equation incorporates a revised and more comprehensive group of risk factors for the sex-specific base models. Moreover, the PREVENT equations allow for the incorporation of glycated haemoglobin (HbA1c) and urine albumin-creatinine ratio, along with a social deprivation index in optional models when such data is accessible. These additional variables have improved the accuracy of risk prediction, offering a more comprehensive assessment of CVD risk. Over a median follow-up of approximately 5 years, investigators found that the base model for PREVENT equations had good discrimination for total CVD events in women ([2]*C*-statistic=0.794) and men (*C*-statistic=0.757), and marginally better discrimination relative to PCE for women (ΔC-statistic=0.005) and men (ΔC-statistic=0.010)[69].

However, these three equations were developed with the US population in consideration, potentially leading to inaccurate risk estimations for population with differing ethnic backgrounds. As a result, it was recommended that other countries develop their own risk score algorithms tailored to their

[2] The C-statistic (also known as the Concordance Index or C-Index) is a measure used to evaluate the discriminatory ability of a survival model. It quantifies the model's ability to correctly rank individuals based on their predicted risk of an event, such as a cardiovascular event. The C-statistic reflects the probability that, for a randomly selected pair of individuals, the one with a higher predicted risk will experience the event first.

population and ethnic diversity [76]. This contributed to the development of region-specific modifications, such as SCORE in Europe, QRISK in the UK, and PREDICT-1 in New Zealand.

2.2.4 SCORE model

The SCORE (Systematic Coronary Risk Evaluation) model project was initiated by the European Society of Cardiology to develop a risk assessment framework for managing CVD in European clinical settings. The data from twelve European cohort studies was analysed involving 205,178 individuals and 2.7 million person-years of follow-up [77].

The risk assessment charts were generated for high and low risk European regions, evaluating the 10-year risk of fatal CVD using two models: one based on total cholesterol and the other on the cholesterol/HDL ratio. These charts visually represent the 10-year risk of fatal CVD for individuals in different European regions. They allow clinicians to quickly estimate risk based on key factors like total cholesterol or the cholesterol/HDL ratio, facilitating informed decision-making for patient care. The risk charts were specifically designed for the primary prevention of CVD [37].

2.2.5 QRISK

QRISK is a prediction equation for CVD in UK, considering traditional risk factors such as age, ethnicity, smoking status and medical history. It was developed in 2007, employing existing risk variables such as age, smoking status, SBP and TC: HDL for CVD, as well as family history, antihypertensive treatment, and BMI. The QRISK algorithm was developed using a validated clinical research database, called QRESEARCH database, comprising of routinely acquired data from general practitioner clinical computer systems [78].

The modelling framework behind QRISK equation is derived from a cohort of 1.28 million patients, aged 35-74 years, registered at 318 practices between January 1, 1995, to April 1, 2007. A Cox PH model was employed to analyse the data and estimate the coefficients associated with the risk factor to predict CVD risk. In comparison to the previously used Framingham Equation, the current equation provided better predictions, as the Framingham equations overestimated the 10-year risk by 35% [70].

2.2.6 PREDICT-1 Equation

Due to the absence of a local cohort study, NZ's first national CVD risk factor management guidelines were developed using the FRS in 1990s. In 2002, a computerised decision support system was developed to assist general practitioners in implementing the national guidelines, while also generating a cohort study, also called a PREDICT cohort, to assess the applicability of the Framingham equation to NZ's ethnically and socioeconomically diverse 21st-century population[79]. The CVD risk equation for the NZ population was first developed in 2018, also known as the PREDICT-1 equation, using the PREDICT cohort [80]. PREDICT is an ongoing open cohort study in NZ and recruits participants post-completion of standardised cardiovascular disease risk assessments. The PREDICT-1 equations were adjusted further including BMI as one of the variables in 2018 and released for implementation in 2019 [81]. These equations are used to classify patients into three risk groups (<5%, 5-15% and >15%) for CVD risk assessments which is pivotal to guide decision making and management in primary care patients in NZ. The recommended interventions, goals and follow-up based on CVD risk assessment is explained in Table 2.

Table 2: Suggested interventions, goals, and monitoring strategies tailored to CVD risk levels [82].

Risk Category	**5-year CVD risk**	**Recommended Intervention & Goals**	**Follow-up**
Low Risk	<5%	• Medication management is generally not recommended. • The focus is on lifestyle modifications, including a healthier diet, more physical activity, and quitting smoking.	Reassessment in 5-10 years
Moderate Risk	5-15%	• Lipid-lowering and blood pressure-lowering medications are recommended for managing CVD risk in patients.	Reassessment every 2-5 years

		• Special consideration is given to prescribing these medications for individuals with higher risk levels.	
High Risk	>15%	• Treatment with lipid-lowering and blood pressure-lowering medications is strongly recommended. • This group is considered to have a risk level comparable to those with existing CVD.	Annual reassessment

2.2.6.1 Outcome Variables

An outcome variable, also referred to as the dependent or response variable, is the key variable that the model used for the development of PREDICT equation aims to predict. The outcome variable typically represents the event or condition that the model is examining, such as the occurrence of a disease, death, or other health-related event. The primary outcomes in the PREDICT study to measure CVD risk include the following [83]:

1. Myocardial infarction - A myocardial infarction (MI), commonly referred to as a heart attack, occurs when blood supply to the heart muscle decreases or stops, resulting in infarction i.e. the death of the tissue, of the heart muscles [84]. In 2015, around 15.9 million of MI cases were noted globally [85]. Over 3 million individuals experienced a ST elevation MI, whereas over 4 million suffered from a non-ST elevation MI [86]. This condition occurs approximately twice as frequently in men compared to women [87].
2. Unstable angina – Unstable angina is a condition characterised by unpredictable chest pain or discomfort which is triggered more easily than usual. It is characterised as a kind of acute coronary syndrome, which includes heart conditions resulting from reduced blood supply to

the heart [88]. Distinguishing unstable angina from non-ST elevation myocardial infarction (NSTEMI) can be difficult, as both conditions exhibit similar symptoms. The fundamental difference is whether the ischemia causes significant heart muscle damage to release detectable levels of biomarkers, such as troponin T or troponin I, into the bloodstream which, if present, categorises the condition as NSTEMI [89].

3. Other coronary heart disease – These include
    a. stable angina is characterised by chest discomfort linked to physical activities such as running or walking with minimal symptoms at rest. Symptoms subside few minutes post-activity and re-emerge when activity resumes [90].
    b. variant angina, like unstable angina, can occur at rest or during sleep due to a temporary spasm in the coronary arteries [91].
4. Ischaemic stroke – An ischaemic stroke occurs when blood flow to the brain is obstructed, resulting in insufficient oxygen and nutrition for brain cells. This results in a temporary loss of cognitive function. Insufficient blood supply translates over time into the death of brain tissue, resulting in irreversible damage [92].
5. Haemorrhagic stroke – The haemorrhagic stroke occurs when a cerebral blood vessel ruptures, resulting in bleeding either within the brain or on its exterior. The bleeding may arise from a ruptured aneurysm, a compromised segment of a blood artery that ruptures, or from arteriovenous malformations, which are abnormal convolutions of blood vessels that disrupt normal blood circulation and delivery of oxygen to the brain [93]. Haemorrhagic stroke accounts for 10-20% of all the new strokes that occur every year. They are less common than ischemic strokes but cause a significant number of deaths worldwide [94].
6. Transient ischemic attack – A transient ischemic attack (TIA), often referred to as a mini-stroke, occurs when there is a temporary disruption of blood flow to the brain, spinal cord, or retina. TIA doesn’t cause lasting damage or tissue injury, and symptoms typically resolve quickly. However, TIA serves as a significant warning indicator of a potential stroke, with the highest risk within the initial 48 hours [95].
7. Peripheral vascular disease – It refers to a condition impacting blood vessels outside the heart and brain, often those in the arms, legs, or other organs. This condition arises when

the arteries become narrowed, reducing blood flow, predominantly affecting the limbs [96]. It impacts approximately 200 million people globally [97].

8. Congestive heart failure – Heart failure, or congestive heart failure, occurs when the heart is unable to efficiently pump blood and oxygen throughout the body, which may result in organ damage and fluid retention in the tissues. The condition is primarily attributable to the weakening of the heart muscle, impairing its ability to effectively circulate blood. This condition is chronic and progressively deteriorates, being more prevalent among older adults [98].

2.2.6.2 Cox Proportional Hazard Model (Cox PH Model)

Survival analysis investigates the link between the occurrence of a particular event and the time projected for the event to take place. Cox PH regression analysis, widely regarded as the most commonly used survival analysis method, was employed in half of the CVD risk models [99]. The Cox PH models were also used to build the 5-year CVD risk prediction equation for the NZ population. The cohort included a gender-specific study of New Zealanders who were enrolled for the period until their first hospital admission or death due to CVD. The period of study was defined as the time between the initial evaluation and the first of: death due to CVD or hospital admission, the end of follow-up, or death due to other causes [83].

The Cox PH model [100] is employed to analyse 'survival time'. As a multivariate technique, it is frequently used in medical research to explain the relationship between the patient survival time with respect to multiple predictors. It also helps in identifying how specific [3]covariates influence the event incidence, expressed by the hazard function h(t).

The Cox PH model is expressed by the hazard function which is defined as the risk of having an event at time "t". In simple terms, it expresses the probability of experiencing an event in a period centered around t for an individual under observation.

[3] The predictor variables are called covariates in the Cox PH model

The relationship between hazard functions, which defines the proportional hazard assumption, can be expressed as follows:

$$h(\mathrm{t}) = \phi\, h_0(\mathrm{t}), \tag{2}$$

where,

- $h(\mathrm{t})$ represents an individual's hazard at a given time t,
- $h_0(\mathrm{t})$ is the baseline hazard, which is the hazard when all the covariates are zero at time t,
- ϕ is a scale parameter depending on the linear predictor $\beta\, \mathrm{Z_i}$. For a Cox PH model it is given by, $\phi = e^{\beta \mathrm{Z_i}}$, where $\mathrm{Z_i}$ is the vector of covariates for subject i.

The Cox model can be mathematically expressed as:

$$h(\mathrm{t}) = h_0(\mathrm{t}) e^{(\beta_1 z_1 + \beta_2 z_2 + \ldots + \beta_p z_p)}, \tag{3}$$

where,

- $h(\mathrm{t})$ is determined by p covariates $(z_1, z_2, \ldots, z_p)$,
- Each covariate's effect is captured by its corresponding coefficients $(\beta_1, \beta_2, \ldots, \beta_p)$,
- $h_0(\mathrm{t})$ is the baseline hazard when all the covariates are zero at time t,
- 't' in $h(\mathrm{t})$ indicates that the hazard will change over time.

An alternative form of the Cox PH Model is expressed in Equation 4. The term $e^{\beta_\mathrm{p} z_\mathrm{p}}$ represents the HR for p covariates.

$$\frac{h(\mathrm{t})}{h_0(\mathrm{t})} = e^{(\beta_1 z_1 + \beta_2 z_2 + \ldots + \beta_p z_p)} \tag{4}$$

The Cox PH model allows fitting regression models to censored survival data. A partial likelihood is used to estimate the regression coefficients of the Cox PH model, allowing for the baseline hazard to remain constant, thus focussing solely on the relationship between covariates and the hazard [101]. There are two differences between likelihood and partial likelihood. Firstly, the factors of the partial likelihood are conditional probabilities and secondly, it is a product of expressions for each

failure time. The censoring time does not contribute to the analysis as they are assumed to be non-informative.

Log-likelihood estimation

D.R. Cox proposed that the conditional probabilities at different failure times are independent of one another. By combining the likelihood for each individual across all failure times, a partial likelihood for β can be estimated [100].

Let t_j denote the j[th] observed failure time, ordered from smallest to largest. At time t_j, the risk set is defined as the set of all individuals who are still at risk of experiencing the event just prior to t_j. to The hazard function for subject i at failure time t_j is given by $h_i(t_j)$ and δ_i denotes the vector for censoring indicators which records whether each observed time for subject i is an event (1) or censored (0) . Under the Cox PH model, for subject i, we can express the hazard function as:

$$h_i(t_j) = \psi_i \, h_0(t_j), \tag{5}$$

where $h_0(t_j)$ is the baseline hazard function, and $\psi_i = e^{z_i \beta}$ is the relative risk for subject i, with z_i representing the covariate vector.

The expression for the partial likelihood for the Cox PH model is given as:

$$L(\beta) = \prod_{i=1}^{n} \left\{ \frac{e^{\beta z_i}}{\sum_{k \in R_i} e^{\beta z_k}} \right\}^{\delta_i}. \tag{6}$$

Where the denominator sums over all individuals k in the risk set corresponding to those still at risk when individual i experiences the event. The log partial likelihood can be obtained as:

$$l(\beta) = \log l(\beta) = \sum_{i=1}^{n} \delta_i (\beta z_i) - \sum_{i=1}^{n} \delta_i \ln \left\{ \sum_{k \in R_i} e^{\beta z_k} \right\} \tag{7}$$

Using the Equation 7, the maximum partial likelihood estimate (MPLE), denoted as $\hat{\beta}$ represents the value of β that maximizes the function $l(\beta)$. The variance of $\hat{\beta}$ is obtained by evaluating the information, $I(\hat{\beta}) = -l''(\beta)$, where $l''(\beta)$ is the second derivative of the log partial likelihood function

$l(\beta)$ with respect to β. The information is a measure of the curvature of the likelihood at $\hat{\beta}$. Higher curvature levels give more information with lower variance, whereas lower curvature values supply less information with higher variance. The standard error of $\hat{\beta}$ is $\frac{1}{\sqrt{I(\hat{\beta})}}$ and its variance is approximately $\frac{1}{I(\hat{\beta})}$.

The beta coefficients and the HRs obtained from the Cox PH model for all the risk factors applied to the NZ population in the PREDICT study are outlined in Table 3 [81]. The study also mentioned the significance of the SBP in predicting the CVD risk in the NZ population, along with its interaction with the age of the individual and their type of BP-lowering medication. Table 3 presents the beta coefficients and the HRs obtained from the data which is used to develop the PREDICT-1 equation.

Table 3: Beta Coefficients and Hazard Ratios for all the variables used to predict 5-year CVD risk [81].

Predictor variables	**Women**		**Men**	
	Beta Coefficient	**Hazard Ratio**	**Beta Coefficient**	**Hazard Ratio**
Age (centered)	0.0734393	1.076	0.0669484	1.069
Chinese or other Asian	-0.2680559	0.765	-0.4131973	0.662
Indian	0.2086713	1.232	0.3666816	1.443
Māori	0.4164622	1.517	0.3166164	1.372
Pacific	0.2268597	1.255	0.2217931	1.248
NZDep quantile	0.0957229	1.1	0.0631146	1.065
ExSmoking	0.1444243	1.155	0.0748648	1.078
Currently smoking	0.6768396	1.968	0.5317607	1.702
Family history of premature CVD	0.0645588	1.067	0.1275721	1.136
Atrial fibrillation	0.9293084	2.533	0.6250334	1.868
Diabetes	0.4967444	1.643	0.4107586	1.508
Systolic BP (centered)	0.0176523	1.018	0.0179827	1.018
TC:HDL-C (centered)	0.1361335	1.146	0.1296756	1.138
Underweight (<18.5)	0.6277962	1.873	0.5488212	1.731

Overweight (25.0–29.9)	0.0018215	1.002	-0.033177	0.967
Obesity class 1 (30.0–34.9)	-0.0169324	0.983	-0.0025986	0.997
Obesity class 2 (35.0–39.9)	0.0343351	1.035	0.1202739	1.128
Obesity class 3 (40.0+)	0.3196519	1.377	0.3799261	1.462
BMI unknown	0.0213595	1.022	-0.073928	0.929
On BP-lowering medication	0.3487781	1.417	0.2847596	1.329
On lipid-lowering medication	-0.0568366	0.945	-0.0256429	0.975
On either antiplatelet or anticoagulant medications	0.1393368	1.15	0.0701999	1.073
Age (centred) x diabetes	-0.0189779	0.981	-0.0124356	0.988
Age (centred) x systolic BP (centred)	-0.000471	1	-0.0004931	1
On BP-lowering medication x systolic BP (centred)	-0.0054002	0.995	-0.0049226	0.995

It is noted that the HR for SBP is slightly higher in men (1.20 per 10 mmHg increase) compared to women (1.19 per 10 mmHg increase), which means that for every 10 mmHg increase in SBP, the risk of CVD event increases by around 20% in men and 19% in women. BP lowering medication use is associated with increased HRs of 1.42 in women and 1.33 in men, suggesting higher baseline risk in treated individuals.

The study also compares the risk of CVD for different ethnic groups being used in the PREDICT cohort. Currently, Māori, Pacific, and Indian patients have a higher risk of CVD than Europeans, whereas Chinese or other Asian patients have a reduced risk of CVD. Furthermore, Māori, Indian, and Pacific patients with low deprivation scores had a CVD risk twice as high as European or Chinese patients with low deprivation scores [38].

To assess how well the model differentiates between the subjects' outcomes, discrimination, also known as 'separation', is evaluated. The discrimination ability can be quantified using several statistical indices, including the receiver operating characteristic (ROC) area, model χ^2 or

Spearman's ρ, etc. Among these, the ROC area is the most widely acknowledged measure of diagnostic discrimination [102]. In survival models, the area under the ROC curve is equivalent to concordance that can consider the censored observations. It is computed based on the Wilcoxon-Mann-Whitney two-sample rank test. It shows the overall evaluation of the model discrimination power, which is defined as the model's ability to produce an accurate ranking of survival times based on individual scores. Similarly, the concordance index (c-index) is useful to validate a survival model's predictive ability. It is the percentage of pairs in the data where your model predicts that the observation with the longer survival time will survive longer [103]. The c-index value of 1 suggests perfect prediction, whereas a value of 0.5 suggests that the prediction is randomly obtained [101]. The Royston's D statistic uses the concept of c-index but measures the separation between the predicted risks for patients who experience the event vs who do not experience the event, standardised by the variability of the predicted risk. The performance results of the Cox PH model used in the PREDICT study are mentioned in Table 4. Model performance was similar for both men and women with men having better separation as per Royston's D statistic value.

Table 4: Standard performance metrics for PREDICT-1 equations.

Gender	**Performance Metrics**	**Value (95% CI)**
Women	R^2	30 (29-31)
	Harrell's C-index	0.73 (0.72 – 0.73)
	Royston's D statistic	1.334 (1.291 – 1.377)
Men	R^2	29 (28 – 30)
	Harrell's C-index	0.73 (0.72 – 0.73)
	Royston's D statistic	1.318 (1.285 – 1.351)

2.3 Blood Pressure and CVD

BP is considered as one of the most common significant predictors in CVD risk prediction models [104]. It plays a crucial role in the circulatory system, as blood flows through the body due to pressure differences within the arterial system. BP is defined as the pressure exerted by circulating blood on

the walls of the arteries, driven by the heart's pumping action. Measured in millimetre of mercury (mmHg), it represents the force exerted per unit area within the artery system [105]. In most adults, the heart typically beats between 60 and 100 times per minute. Each beat creates a pressure wave in the arteries, which occurs as a regular, pulsating pattern. The highest pressure occurs during systole, that is, when the heart contracts and pumps blood into the arteries with the most force. The lowest pressure occurs during diastole, when the heart relaxes and refills with blood. This cycle of contraction and relaxation creates the rhythm of the heartbeat and the pressure changes in the arteries [106]. BP is usually expressed in terms of systolic pressure and diastolic pressure defined as follows:

1. Systolic Blood Pressure (SBP) - The maximum pressure exerted by the blood during ventricular contraction during a cardiac cycle.
2. Diastolic Blood Pressure (DBP). The lowest pressure exerted by blood on arteries during a cardiac cycle.

The American Diabetes Association (ADA) emphasises the importance of managing BP in diabetic patients to reduce cardiovascular risks. The Action to Control Cardiovascular Risk in Diabetes (ACCORD) BP trial, highlighted by the ADA, demonstrated that intensive management of BP significantly decreases stroke occurrence in diabetes patients [107]. A review in Cardiovascular Diabetology explores the impact of intensive BP lowering on cardiovascular outcomes in diabetic patients, including studies like ACCORD and the Systolic Blood Pressure Intervention Trial (SPRINT). This review discusses the long-term benefits and challenges associated with achieving lower SBP targets in this population [108]. Even within normal SBP ranges, maintaining lower levels is associated with a reduced risk of atherosclerotic CVD, highlighting the need of optimal SBP management [109].

The mortality rate from CVD varies significantly due to the changes in the absolute values of the primary risk factors like smoking, cholesterol, and BP [110]. A study found that every 1 mmHg increase in the SBP was associated with a 2-4% increase in the estimated relative risk of a CVD event [111]. The Ministry of Health guidelines predict a 20% reduction in CVD risk over five years for

every 10 mmHg decrease in SBP, based on NZ trials, based on a randomised controlled trial in NZ [53].

The Multiple Risk Factor Intervention Trial (MRFIT) has demonstrated a strong association between SBP and CHD mortality. MRFIT is a randomised, primary prevention trial conducted in 22 US clinical centres for men ranging from 35-57 years of age. Table 5 shows the relative risks pertaining to different SBP ranges, classified according to different stages of BP from normal (SBP<120 mmHg) to stage 3 hypertension (SBP≥180 mmHg) [112]. Different SBP ranges have an associated relative risk with respect to the reference SBP range (SBP<120 mmHg).

This association between SBP and CVD events emphasises the importance of accurately measuring BP. Given that relative risks are calculated based on specific SBP thresholds, even small deviations in measurement can result in significant misclassification of an individual's risk category. Accurate BP readings are essential not only for determining relative risk but also for guiding appropriate clinical interventions. Assuming there were measurement errors when measuring SBP, the individual's relative risk will be incorrectly estimated, which in turn causes insufficient or overcompensated treatment due to incorrect CVD predictions. Consequently, to ensure that the patients are getting the right medications and treatments based on the guidelines, it is crucial to measure BP accurately [113]. Therefore, it is vital to comprehend the significance of potential errors in the recorded SBP values, as they can categorise the patient as being at higher or lower risk, affecting the CVD risk prediction.

Table 5: Association of different SBP values with the relative risk of coronary heart disease (CHD) [112].

SBP (mmHg)	**Relative Risk**
<120	1.00 (Reference)
120-129	1.28 (1.19 – 1.36)
130-139	1.66 (1.56 -1.77)
140-159	2.45 (2.30 – 2.61)

160-179	3.42 (3.16-3.71)
180-209	5.26 (4.68-5.90)
>210	6.40 (4.74-8.65)

2.4 Sources of BP measurement errors

BP measurement errors may arise from various sources. This study focuses on two of the most common sources: errors arising from rounding the BP readings to the nearest zero end-digit and errors arising from inaccuracies from BP measurement devices and, as discussed in the section 2.4.1 and 2.4.2. while these are not the only contributors to inaccuracies, a few additional sources are also mentioned for context. Nonetheless, the primary analysis in this study focusses on the impact of these two key sources of BP errors, that correspond to distinct sources for manual and automatic BP measurements, highlighting their significance in the context of CVD risk assessment in NZ.

2.4.1 Terminal digit bias in BP measurement

A terminal digit bias, characterised by the observer's preference to round BP measurements to a specific end digit, typically zero, is referred to as zero end-digit preference [51]. This means that the observer rounds the end-digit towards a preferred level, either rounding up or rounding down [114].

Terminal digit bias is a common occurrence in both clinical and research settings [114], [115], [116], despite being one of the most frequently measured vital signs with the aim of ensuring accurate and reliable BP readings remains a challenge. The inherent variability of BP can increase when recommended measurement protocols are not consistently followed [117].

Without zero end-digit preference, 20% of BP readings from manual devices and 10% from automated devices would naturally end in zero[4]. However, studies from U.S. clinics reveal that

[4] Manual devices typically record even digits, whereas the automated devices can record any digit [269].

terminal digit preference occurs more frequently than these expected rates, with approximately 37.7% of SBP and 37.8% of DBP readings ending in zero in 2019 [115], [118], [119], [120].

Zero end-digit preference has an impact on patient care. Studies indicate that patients in clinics with higher rates of terminal digit preference are less likely to receive antihypertensive prescriptions and more likely to face adverse cardiovascular events. Terminal digit preference was found to increase at higher BP levels among patients with hypertension [121]. Some researchers might argue that rounding of BP above treatment thresholds is not a major issue, as these patients would still receive treatment. However, rounding often lowers the recorded value, which could influence the intensity of medication or other patient care strategies [118], [122].

2.4.2 BP measurement device errors

One prevalent source of error in BP measurement comes from the device itself [123]. Different techniques and algorithms used by measurement devices worldwide result in different ranges of BP deviations. Table 6 shows some of these techniques. As a result, it is necessary that the accuracy of the BP devices is validated with proper guidelines with consistent, quantifiable metrics.

Table 6: Some important methodologies and techniques used in the development of BP measurement devices.

Reference	**Methodology/Technique**
[124]	Evaluation of oscillometric algorithms focusing on cuff deflation speed and signal filtering.
[125]	Pulse Wave Analysis (PWA) using Fourier transform and waveform decomposition to estimate central BP.
[126]	Pulse Transit Time (PTT) methodology; challenges in calibration and environmental noise impact.
[127]	Photoplethysmography (PPG) signal processing with adaptive filtering and feature extraction.
[128]	Deep-learning-based blood pressure estimation using multi-channel photoplethysmogram and finger pressure with attention mechanism

To address this, internationally recognised standards establish protocols to measure the accuracy of BP measurement devices, as outlined in Table 7. In 2017, an international effort was initiated by the ANSI/AAMI SP10, ESH and ISO committees to establish a globally recognised standard to evaluate BP measurement device accuracy [129]. The resulting ISO 81060-2:2-2018 standard [49] consolidates prior evidence and has gradually replaced previous protocols employed worldwide. To ensure this initiative achieves a global impact in enhancing the availability of clinically validated BP devices, it is crucial to adhere to the universally recognised ISO standard and evaluate the accuracy of BP measurement devices prior to their market release for medical, community, and home applications [130].

Table 7: Overview of standards for assessing BP measurement device accuracy, including corresponding validation criteria.

Standard	**Organisation**	**Error Range (Mean ± SD)**	**Validation Sample Size (N)**	**Reference**
BHS Protocol	British Hypertension Society	A grading system (A, B, C, D) is applied based on the differences between paired readings within ≤ 5, 10, and 15 mmHg, assessed separately for each observer and for both SBP and DBP. Additionally, the criteria include mean differences of ≤ 5 mmHg and a standard deviation of ≤ 8 mmHg, as recommended by AAMI.	85	[131]

ESH	European Society of Hypertension	The criteria are based on the number of readings where the test-reference BP difference falls within ≤ 5, 10, and 15 mmHg. For individual BP measurements (Part 1), the test is considered a pass if at least 73.7% (73 out of 99) of differences are ≤ 5 mmHg, 87.9% (87 out of 99) are ≤ 10 mmHg, and 97.6% (96 out of 99) are ≤ 15 mmHg. For individual subjects (Part 2 - Accuracy), the evaluation is based on the number of subjects with 0, 2, or 3 absolute differences within ≤ 5 mmHg.	33	[132]
ANSI/AAMI SP10	Association for the Advancement of Medical Instrumentation (AAMI) / American National Standards Institute (ANSI)	Criterion 1 (Individual BP Readings): The mean BP difference must be ≤ 5 mmHg, with a standard deviation of ≤ 8 mmHg. Criterion 2 (Individual Subjects): The mean difference and standard deviation of BP readings must fall within the	≥ 85	[133]

		threshold determined by the mean value of Criterion 1.		
ISO 81060-2:2018	International Organization for Standardization	The ISO incorporates both criteria from the AAMI SP10 standard. Additionally, standardised Bland–Altman scatterplots to show agreement between test and reference measurements will be presented. The mean test-reference BP difference and standard deviation for each cuff subgroup must be reported, but no pass/fail criteria will be applied to the test device.	≥ 85	[49]

The ISO standard, like ANSI/AAMI SP10 states that BP errors, assumed to be normally distributed, must lie within a *tolerable error* range of -10 mmHg to +10 mmHg and the *probability of tolerable error* is fixed to at least 85%. BP errors are computed by taking the difference between the value obtained from the test device and the value obtained using a reference auscultatory method which is considered the gold standard for non-invasive BP (NIBP) methods. The ISO standard also incorporates the statistical principles from ANSI/AAMI SP10, specifying that the accuracy of BP measurement devices must be assessed according to two primary criteria with no fewer than 85 subjects. The first criteria specifies that the mean difference between the test device and the reference device is to be ± 5 mmHg or less, with a SD of 8 mmHg or less. The second criterion states that the mean difference remains within ± 5 mmHg and follow the upper limit threshold of SD for a given fixed value of the mean, also presented in Table 8. For example, if the mean error of the device

is 2 mmHg, to maintain 85% *probability of tolerable error*, its true SD must be no greater than 6.65 mmHg for the device to be regarded acceptable.

Table 8: Acceptable upper limit of SD for a specific value of mean BP error [133].

Sample mean error (mmHg)	0	± 0.5	± 1	± 1.5	± 2	± 2.5	± 3	± 3.5	± 4	± 4.5	± 5
SD, ≤ (mmHg)	6.95	6.93	6.87	6.78	6.65	6.47	6.25	5.97	5.64	5.24	4.81

The standard also evaluates the probability of accepting a device, which means the probability that for a true mean and true SD (σ) of BP errors, at least 85% of the errors lie within - 10 to 10 mmHg. The probability values for N= 85 are provided in Table 9. For example, if the true mean (μ) is 2 mmHg and standard deviation (σ) is 6, there is 89% probability of accepting the device, when 85 or more samples are gathered.

[5]*Table 9: Probability of accepting the device for different values of the true mean and SD of errors and a sample size of n=85 [133].*

μ	**σ = 5**	**σ = 6**	**σ = 7**	**σ = 8**
0	(0.95) 1.0	(0.90) 0.98	(0.85) 0.45	(0.79) 0.04
1	(0.95) 1.0	(0.90) 0.96	(0.84) 0.40	(0.79) 0.03
2	(0.94) 1.0	(0.89) 0.89	(0.83) 0.26	(0.77) 0.02
3	(0.91) 0.99	(0.86) 0.66	(0.81) 0.11	(0.76) 0.01
4	(0.88) 0.86	(0.83) 0.28	(0.78) 0.03	(0.73) 0.00
5	(0.84) 0.33	(0.79) 0.05	(0.75) 0.00	(0.70) 0.00
6	(.79) .02	(.74) .00	(.71) .00	(.67) .00
7	(.73) .00	(.69) .00	(.66) .00	(.63) .00

2.4.2.1 Review of BP Measurement Device Errors

Out of the standards outlined in Section 2.4.2 that are widely adopted globally, only approximately 20% of the devices currently available adhere to the specified requirements [134]. A study comparing

[5] In parentheses is the true probability of the tolerable error.

twelve various types of automated BP monitors found that only four of them had reasonable proof that the device's accuracy had been evaluated using the proposed worldwide criteria [135]. Multiple clinical studies have validated BP monitoring devices using the existing assessment methodologies for BP monitoring device standards. A study that tested 524 sphygmomanometers observed the degree of inaccuracy to be around 4-6 mmHg for 32% of the instances. In 19% of the cases, the degree of inaccuracy was 7-12 mmHg and > 13 mmHg in 7% of the case [136]. According to another survey done in the healthcare settings, more than 50% of the sphygmomanometers resulted in BP deviations of up to 12 mmHg, highlighting the presence of devices with SDs far exceeding the maximum acceptable limit of 6.95 mmHg [123].

Additionally, examining several clinical studies that provide different ranges of BP measurement device errors led to the discovery of several reported device inaccuracies using a sample size of less than 85. Even when the sample size criteria are followed, studies do not adhere to the minimum mean $\pm$ SD criteria. A list of studies providing device inaccuracy details using a sample size smaller than 85 have been provided in Table 10.

Table 10: BP device inaccuracies obtained from studies using different techniques.

Study	SBP/DBP	Method/Techniques	Sample Size	Device Inaccuracy ($\bar{x}$ ± SD)
[137]	SBP	Oscillometry	33	-0.7 ± 6.9
[138]		PTT	10	1.04 ± 6.88
[139]		PTT-PPG	33	1.17 ± 5.72
[140]		Standing	25	-0.462 ± 8
[141]		PAT	32	0.12 ± 6.15
[142]		PTT	33	-0.06 ± 6.63
[143]		PTT-linear	20	0 ± 6.73
		PTT-nonlinear		0 ± 5.56
[144]		ML	45	4.53±2.68
[138]		PTT	10	-2.16 ± 6.60

[139]	DBP	PTT-PPG	33	0.40 ± 7.11
[145]		PTT-IPG	15	-0.5 ± 5.07
[146]		PWV-	15	-0.06 ± 5.46
[137]		Oscillometry	33	-1.0 ± 5.1
[141]		PAT	32	1.31 ± 5.36
[142]		PTT	33	-0.25 ± 5.63

2.4.3 Other sources of BP errors

This section outlines other common sources of BP measurement errors, as detailed below:

2.4.3.1 Incorrect Positioning

Among the critical factors influencing BP measurement accuracy are the position of the cuff, body and arms.

Cuff Position

It is recommended that the cuff should be wrapped around the upper arm with the midline of the bladder placed over the brachial artery during BP measurement [147]. In recent years, wrist and finger monitors have gained popularity due to their ease of use, however BP readings taken from the wrist or fingers often differ from those taken at the upper arm. This variation occurs because BP naturally changes as it travels through the body, with differences influenced by vessel size and distance from the heart [106]. Table 11 illustrates the influence on BP measurement of positioning a cuff on the upper arm and wrist. For example, a study examined the average changes in BP readings, both SBP and DBP, when the cuff of the BP measurement device was positioned on the upper arm versus the wrist of the patients. The results showed that changing the cuff position from the upper arm to the wrist led to an increase in the average SBP reading, while the average DBP reading decreased.

Table 11: Changes in the SBP and DBP measurements due to different cuff position

N	Cuff Position	DBP (mmHg)	SBP (mmHg)	Reference
250	Upper Arm	80.7 ± 11.2	127.7 ± 15.7	[148]
	Wrist	75.7 ± 11.9	143 ± 22.2	
70 (Mixed hypertension group)	Upper Arm	83.4 ± 9.4	168.3 ± 18.4	[149]
	Wrist	83.2 ± 10.5	159.2 ± 18.5	
45 (Isolated systolic hypertension group)	Upper Arm	95.8 ± 11.5	174 ± 14.1	[149]
	Wrist	94.4 ± 11.5	163.8 ± 25.4	

Body Position

Recommendations stipulate that the patients are seated with proper back support and arms supported at the heart level [150]. Research shows that BP tends to be higher when a person is sitting compared to lying down [151], [152]. However, the differences between supine and sitting BP have been found to be relatively small [153]. For this reason, health professionals tend to not consider or underestimate the effect of position when interpreting the results of BP measurements. But even a mean difference of a few mmHg may have relevant implications [154]. Table 12 presents information on BP values for different positions of the subject during BP measurement.

Table 12: Changes in the SBP and DBP measurements due to different body position

N	Body Position	DBP (mmHg)	SBP (mmHg)	Reference
57	Sitting	79.5 ± 9.7	135.7 ± 24.8	[152]
	Supine	84.6 ± 10.5	141.3 ± 25.5	
229	Beach Chair	64.6 ± 11.2	114.6 ± 24.8	[155]
	Supine	72.5 ± 14.5	129.8 ± 27.5	
250	Sitting	83.0 ± 9.6	137.2 ± 13.7	[156]

	Supine	80.1 ± 9.1	139.3 ± 14.0	
245	Sitting	86.0 ± 14	136.7 ± 21.9	[157]
	Supine	83.5 ± 12.5	135.5 ± 20.3	

Arm/Leg Position

Proper positioning of the arm during BP measurement is essential for accurate readings. Ideally, the arm should be supported at heart level on a flat surface to avoid errors. Deviating from this position, such as moving the arm from horizontal to vertical, can increase BP readings by 5–6 mmHg due to hydrostatic pressure changes [50]. A study by Mariotti et al. examined the impact of arm positioning and postural hypotension during BP measurements. They observed that incorrect arm positioning while standing led to overestimated BP values [158].

Similarly, proper leg positioning is equally important during BP assessments. Guidelines emphasize that sitting with feet flat on the floor provides more accurate measurements. In contrast, sitting with crossed legs has been shown to result in higher BP readings[51] [159]. Table 13 presents studies that review the effect of arm and leg position on BP values.

Table 13: Changes in the SBP and DBP measurements due to different armf position

	N	**Position**	**DBP (mmHg)**	**SBP (mmHg)**	**Reference**
Arm	69	Arm low (on chair armrest)	88.6 ± 9.1	143.0 ± 19.9	[160]
		Arm high (at heart level)	77.7 ± 9.9	133.3 ± 20.7	
	57	Arm low (on the bed)	82.1 ± 13.4	142.1 ± 28.0	[152]
		Arm high (at heart level)	78.2 ± 14.4	137.4 ± 29.0	
Leg	238	Crossed	92.1 ± 11.2	153.6 ± 20.2	[161]
		Uncrossed	86.4 ± 10.8	145.3 ± 20.3	

	100	Crossed	84.9 ± 11.6	155.6 ± 19.3	[159]
		Uncrossed	80.9 ± 11.2	146.5 ± 18.6	

2.4.3.2 White coat effect

The white coat effect refers to elevated BP readings in a clinical setting, attributed to stress or anxiety induced by the presence of medical personnel or due to healthcare environment. In individuals exhibiting high BP in a clinical setting, approximately 15 to 30% experience white coat hypertension [162]. The white coat effect has been observed in 30-40% of children and adolescents referred for assessment of mild hypertension [163]. This results in inaccurate hypertension diagnosis and inappropriate medical care if overlooked, as the elevated BP readings frequently stem from situational stress rather than chronic hypertension [164].

2.4.3.3 Incorrect Cuff Size and placement

Using an arm bladder cuff with dimensions below the advised minimum ratio of 1:1.8 for width to length may result in an overestimation of blood pressure by 7 mmHg systolic and 4.3 mmHg diastolic [165]. The ideal cuff must have a bladder length of 80% and a width of at least 40% of the arm circumference, according to a length-to-width ratio of 2:1 [166]. A small sized cuff may lead to an increase in SBP by 3 mmHg when ideally a regular cuff size was required and by 9.6 mmHg when an extra-large BP cuff was appropriate. Studies show that only 74% of medical students select the correct cuff size for BP measurements [167]. Although BP device manufacturers attempt to mitigate this issue by including arm circumference range indicators on individual cuffs to assist in selecting the appropriate cuff size for each measurement, it remains unclear how accurately clinicians or patients follow these guidelines [168].

CHAPTER 3. ZERO END-DIGIT PREFERENCE IN BLOOD PRESSURE AND IMPLICATIONS FOR CARDIOVASCULAR DISEASE RISK PREDICTION—A STUDY IN NEW ZEALAND (Manuscript 1)

3.1 Prelude

Accurate blood pressure (BP) measurement is fundamental to cardiovascular disease (CVD) risk assessment and management. However, the phenomenon of end-digit preference, particularly the rounding of BP measurements to terminal digits, most commonly zero, has been documented in clinical practice. This rounding can introduce systematic biases, potentially affecting the precision of CVD risk prediction models. This chapter investigates the prevalence of zero end-digit preference in systolic blood pressure (SBP) readings and its impact on CVD risk classification within the New Zealand (NZ) population.

Utilising a substantial dataset of 427,299 individuals who underwent opportunistic CVD risk assessments in primary care settings across NZ, the researchers identified that 292,122 SBP readings had non-zero terminal digits. These readings were systematically rounded to the nearest zero end-digit to evaluate the effect of such rounding on CVD risk prediction. The analysis revealed that end-digit preference could lead to misclassification of individuals' risk categories, thereby influencing clinical decision-making process. The findings highlight the necessity for adherence to precise BP measurement guidelines and caution against arbitrary rounding practices.

In addition to the clinical implications, this study also includes a financial cost analysis to assess the economic impact of SBP measurement errors. Misclassification of CVD risk can lead to inappropriate treatment decisions, resulting in increased healthcare costs and resource allocation inefficiencies. By quantifying these financial consequences, the study provides a comprehensive perspective on the burden of measurement inaccuracies, emphasising the necessity for precise BP assessments in both clinical and economic contexts.

This study contributes to the broader discussion on the accuracy of clinical measurements and their implications for patient outcomes. By highlighting the potential inaccuracies introduced by end-digit

preference, it calls for enhanced training and awareness among healthcare professionals regarding BP measurement techniques. In conclusion, the investigation into zero end-digit preference in BP measurements within the NZ context provides critical insights into the subtle yet significant factors that can influence CVD risk assessment. Addressing these issues is essential for optimising patient care and ensuring the accuracy of predictive models used in cardiovascular risk stratification.

3.2 Introduction

For over eight decades, the significance of precise blood pressure (BP) measurement has been widely acknowledged [169]. BP is a critical metric for health monitoring and diagnosis in healthcare settings. High BP causes around 54% of strokes and 47% of coronary heart disease worldwide [170]. BP serves as a fundamental parameter for cardiovascular disease (CVD) risk assessment [171]. It is one of the significant variables in the models used to construct CVD risk prediction equations globally.

The most accurate method for BP measurement is arterial cannulation. It is invasive, time-consuming, and requires skilled personnel [55]. In routine practice, BP is measured non-invasively [172]. An occluding upper arm cuff is mostly employed for intermittent non-invasive monitoring. BP readings can be assessed either manually (by auscultation of Korotkoff sounds or palpation) or mechanically (for example, using oscillometric technique) [172]. In clinical and research settings, manual auscultatory sphygmomanometer is considered the traditional gold standard for non-invasive BP measurement [50]. However, these non-invasive methods introduce inaccuracies. Minor measurement errors result in misclassification of CVD risk of millions of individuals. This misclassification has clinical consequences [173]. Underestimation of BP by 5 mmHg would misclassify over 20 million Americans as pre-hypertensive instead of hypertensive. Conversely, overestimating BP by 5 mmHg misclassifies around 27 million people as hypertensive instead of pre-hypertensive [174]. Untreated hypertension significantly increases the risk of fatal strokes and myocardial infarctions [175]. On the other hand, overtreated hypertension can lead to adverse outcomes, such as increased risk of hypotension, and unnecessary healthcare costs [176].

To limit the consequences of inaccurate BP measurements, the International Organization for Standardization (ISO) has established globally accepted criteria that encode acceptable limits for

measurement errors, and protocols to test that devices meet these criteria. ISO 81060-2:2018, for instance, outlines the standards for clinical investigations involving automated, non-invasive sphygmomanometers, a guideline that has received whole or partial approval from numerous national regulatory bodies. It outlines the procedures for testing accuracy, performance, and safety, ensuring these devices meet international standards [49].

Despite training in standardized procedures, BP measurements may still be subject to limitations in accuracy [12][177]. In a prenatal clinic based in Canada, the redefinition of a treatment threshold by a single mmHg adjustment, shifting from systolic blood pressure (SBP) >140 mmHg to SBP ≥140 mmHg, resulted in a twofold increase in the percentage of patients identified as needing treatment, with the proportion escalating from 13% to 26% [178]. In a UK case-control study, the impact of terminal digit preference on disease outcomes was associated with increased mortality [179].

A New Zealand (NZ) study examined how rounding BP measurements to zero end-digits affects patient categorization for pharmacological management in primary care. It found that about 64.4% of SBP values ended in zero [180]. Measurements rounded to the nearest even number, as recommended, should have around 20% zero end-digit [181]. The study employed three distinct risk prediction algorithms sourced from different countries, each incorporating a range of risk factors. One of these algorithms is based on the US population, which was developed from data obtained from the Framingham Heart Study [67]. The remaining two algorithms were the 2004 British Hypertension Society (BHS) guidelines (BHS-IV) [182], and the 2005 Joint British Societies' (JBS) guidelines (JBS2)[183]. The primary objective of this NZ based study [180] was to assess the consequences of rounding BP measurements to zero end-digits when applying diverse risk prediction models from different countries within primary care in the context of NZ. The results showed misclassification of 1 in 41 patients, potentially altering treatment decisions. Under the JBS2 guidelines, 1 in 19 would be misclassified, and under the BHS-IV guidelines, 1 in 12, primarily leading to increased treatment. At the time of this study, NZ did not have a locally developed CVD risk prediction equation.

Initially NZ was using the Framingham Risk Score for their CVD risk management guidelines. However, these equations were developed targeting the US population, which might result in

incorrect estimation of risk for other populations with diverse ethnic backgrounds. This led to the development of a region-specific CVD risk equation [184]. In 2018, NZ introduced its CVD risk prediction equation known as the PREDICT-1 equation [38]. It estimates an individual's 5-year CVD risk and is specifically tailored to the NZ population from a study cohort also known as PREDICT. The five-year CVD risk was used to guide decision-making for primary prevention. The New Zealand CVD risk management guidelines recommend the selection of a 5-year risk assessment [185] instead of the more commonly used 10-year risk, as most trials of CVD risk reduction have about 5 years follow-up [186], [187]. Table 2 in Section 2.2.6 summarises the recommended interventions, goals and follow-up based on CVD risk assessment for clinicians based on the three risk categories. These risk categories are used in current CVD management guidelines in NZ: <5% (low risk group), 5-15% (moderate risk group) and >15% (high risk group).

The patients are categorised into three risk groups based on the data collected from the PREDICT study cohort. However, the potential impact of rounding BP measurements on decisions related to pharmacological treatment, as guided by the PREDICT-1 equation, has not been fully explored. This paper investigates the misclassification arising from rounding practices of SBP and evaluates its effect on determining eligibility for CVD risk management through medication.

Our primary focus is on the potential consequences of misclassifying moderate-risk patients as low-risk. Misclassification in this context leads to significant undertreatment, as low-risk patients typically receive recommendations for lifestyle modifications rather than pharmacological interventions. While high-risk patients receive comprehensive medical treatment, many moderate-risk patients are also prescribed medications based on clinical evaluation. However, those misclassified in the low-risk category are at greater risk of not receiving necessary medical intervention. Additionally, the paper also examines how these misclassifications due to rounding impact overall healthcare costs within the NZ healthcare system.

3.3 Methods

3.3.1 Study Population

PREDICT is an ongoing, open cohort study for the NZ population. Cardiovascular risk is composed of various events, including ischemic or haemorrhagic cerebrovascular events, peripheral vascular

disease, and congestive heart failure. PREDICT is embedded into practice management software systems facilitating healthcare needs of at least a third of the country's residents [38]. The cohort under examination is gender specific. The study period spans from the initial assessment until the earliest of the following events: hospital admission, the conclusion of the follow-up period, death due to CVD or other causes.

In this paper, the risk profile was gathered for every person whose first risk assessment was conducted between October 2004 and December 2018 within routine healthcare settings. The 5-year CVD risk score, estimated as the percentage risk of an individual having a CVD event in 5 years, is evaluated based on the risk factors outlined in Table 2. Specifically, for each patient, the 5-year absolute risk of a CVD event was obtained by applying 'PREDICT-1' [38], given by:

$$Risk\ (\%) = \left(1 - S_0^{e^{(\beta_1 z_1 + \beta_2 z_2\ + \ldots + \beta_p z_p)}}\right) \times 100, \tag{8}$$

where S_0 is the baseline survival function, β_i coefficients are the beta estimates obtained for each variable mentioned in Table 14 and z_i are corresponding risk factors.

The coefficients are estimated by applying the Cox Proportional Hazard Model (Cox PH). The Cox PH model is [188] expressed as:

$$h(\mathrm{t}) = \mathrm{h}_0(\mathrm{t}) e^{(\beta_1 z_1 + \beta_2 z_2\ + \ldots + \beta_p z_p)}, \tag{9}$$

where:

h(t) is determined by p covariates ($z_1, z_2, \ldots., z_p$). The impact of each covariate is measured by their respective coefficients ($\beta_1, \beta_2, \ldots\ldots, \beta_p$).

$\mathrm{h}_0(\mathrm{t})$ is the baseline hazard when all the covariates are zero at time t.

and ‘t’ in h(t) indicates that the hazard will change over time.

Table 14: Risk factors obtained from PREDICT study

Risk Factors	**Type**	**Categories**
Age (centered),	Numeric	
Ethnicity	Categorical	European Māori Pacific Indian Chinese or other Asian
New Zealand Deprivation Index (NZDep)	Categorical	1(least deprived) 2 3 4 5(most deprived)
Ex-smoker	Categorical	0 = No 1 = Yes
Current smoker	Categorical	0 = No 1 = Yes
Family history of premature CVD	Categorical	0 = No 1 = Yes
Atrial fibrillation	Categorical	0 = No 1 = Yes
Diabetes	Categorical	0 = No 1 = Yes
Systolic blood pressure (SBP, centered)	Numeric	

Total Cholesterol to High-Density Lipoprotein Cholesterol (TC:HDL) ratio (centered)	Numeric	
BMI	Categorical	normal underweight overweight obesity class 1 obesity class 2 obesity class 3 bmi unknown
On BP-lowering medication	Categorical	0 = No 1 = Yes
On lipid-lowering medication	Categorical	0 = No 1 = Yes
On either antiplatelet or anticoagulant medications	Categorical	0 = No 1 = Yes

3.3.2 Study Design

For patients whose SBP did not end in zero, a second SBP value was assigned by rounding the original measurement to the nearest zero end-digit. End digits one to four were rounded down, and six to nine were rounded up. The end digits of five were randomly rounded up or down with equal probability. This method replicates common rounding practices observed in routine healthcare for manually recorded measurements [180].

The impact of rounding SBP values is assessed by examining how rounding to the nearest zero-end digit affects CVD risk prediction and alters the model. Figure 5 illustrates the overall design to study the misclassification rates when a new survival model is fitted on the existing dataset with rounded SBP values. The detailed explanation of the design and the simulation study conducted is as follows:

- Simulations were conducted for men and women separately. A CoxPH model (Model 1) was fitted using a subset of patients whose original BP ($SBP_{original}$) was measured without a zero-end digit. The beta coefficients derived from Model 1 were used to calculate the 5-year CVD risk (Original Risk) for each individual, categorizing patients into three distinct CVD risk categories.
- This subset of non-zero SBP values were rounded to the nearest zero-end digit ($SBP_{rounded}$). A new model (Model 2) was then fitted on this dataset. Coefficients of Model 2 were used to estimate the 5-year CVD risk (Rounded Risk). This process was repeated 10,000 times to account for variability and ensure robustness in the findings. At each simulation, individuals were categorized into one of three CVD risk categories based on their calculated 5-year risk. The total number of individuals in each risk category was recorded. The average number of individuals in each risk category across all 10,000 simulations was then computed, along with the 95% confidence intervals (CIs) to quantify the uncertainty around the risk classifications.
- The model estimates in terms of Hazard Ratio (HR) were compared for both the models, i.e. Model 1 and Model 2, by assessing the relative difference between the two models, evaluated using equation (10). The classification results obtained from $SBP_{original}$ and $SBP_{rounded}$ were compared. This comparison allowed for an assessment of the extent of misclassification and the potential impact of rounding SBP values on the accuracy of CVD risk prediction.

$$Relative\ Difference\ = \frac{\mathrm{HR}_{Model\,2} - \mathrm{HR}_{Model\,1}}{\mathrm{HR}_{Model\,1}} \times 100 \qquad (10)$$

To account for device calibration to the nearest 2 mmHg, we estimate the proportion of rounding in the whole cohort by dividing the data it into three groups: (1) non-zero end-digit SBP values, (2) true zero-end digits (20% prevalence), and (3) zero-end digits likely due to rounding. Using this estimate, we apply stratified sampling to select the same percentage of patients from the non-zero group and

round their SBP values. Stratified sampling ensures proportional representation across quantiles based on key CVD factors minimizing bias in risk profiles. This method accounts for variations across different risk profiles, thereby reducing the risk of over- or under-representation. After adjusting the data through rounding, we refit the Cox PH model using the updated dataset. The same simulation approach, as previously outlined in this paper, is then applied to assess the impact of rounding on CVD risk predictions and model performance.

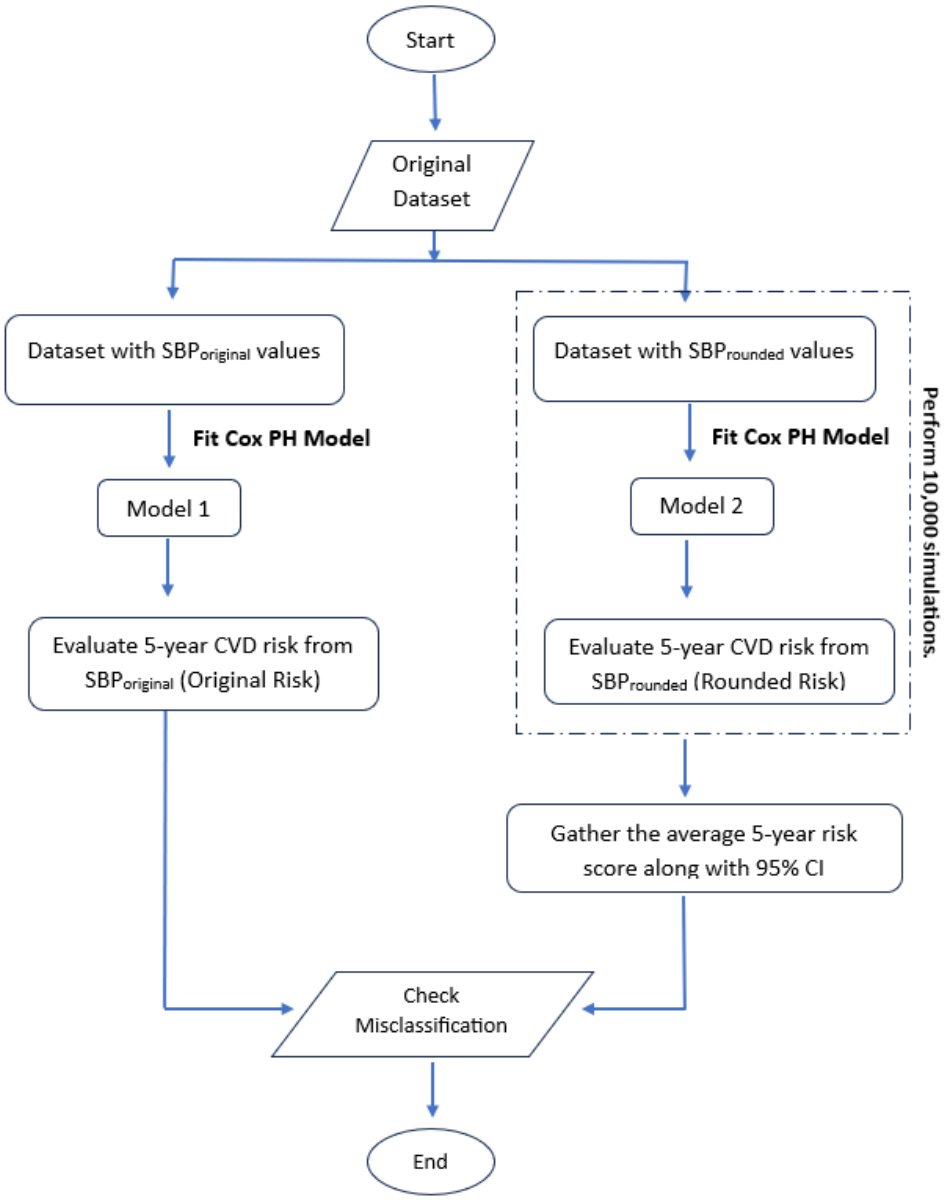


Figure 5: Process to evaluate misclassification rate by introducing a new model with rounded SBP.

To study the financial impact, a comprehensive cost assessment is conducted for key resources used for CVD risk assessment. The cost associated with these resources were gathered from

multiple official sources that provide the latest cost schedules and medication prices relevant to CVD risk assessment and are presented below:

1) Data Sources:
    a) Government Databases - Cost Resource manual and Community Pharmaceutical Schedule provided by Pharmaceutical Management Agency (PHARMAC) [189], [190].
    b) Ministry of Health - Cardiovascular Disease Risk Assessment and Management for Primary Care document highlighting key variables for CVD risk assessment [82].
2) Cost Variables:
    a) Consultation Fees
    b) Diagnostic Procedures
    c) Medications

Based on these resources, the average cost in 5 years, associated with CVD risk management has been estimated. Additionally, a minimum and maximum cost range was estimated to account for variations.

3.4 Results

Our dataset includes BP measurements from 427,299 individuals in the PREDICT cohort. Among these, men (n = 241,036) reported an average age of 51.29 years (SD = 10.07 years), while women (n = 186,263) had an average age of 55.85 years (SD = 8.87 years). For men, 32.16% of the SBP values ended in a zero, while for women, this figure was 30.96%. Overall, approximately 32% of the total dataset consisted of SBP values ending in zero, with the remaining having non-zero end digits.

Figure 6 shows the distribution of people across three CVD risk categories. The largest group of people falls into the lowest risk category, that is, having less than 5% risk of developing CVD within the next five years. A smaller proportion of people fall into the 5-15% risk group, with the smallest proportion in the high-risk category, which is >15%.

Table 15 gives the demographics and risk factors pertaining to SBP end-digit groups for men and women. After excluding individuals with a history of CVD or those with SBP values ending in zero, the study resulted in a final cohort of 163,528 men and 128,594 women.

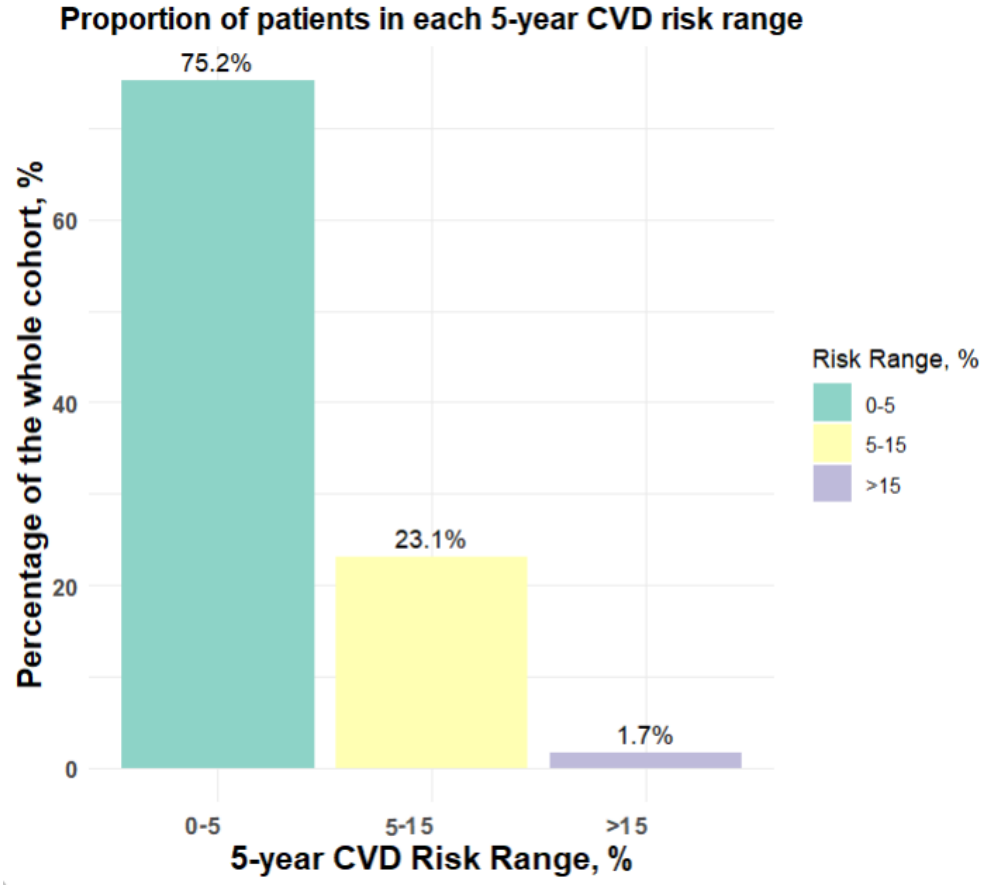


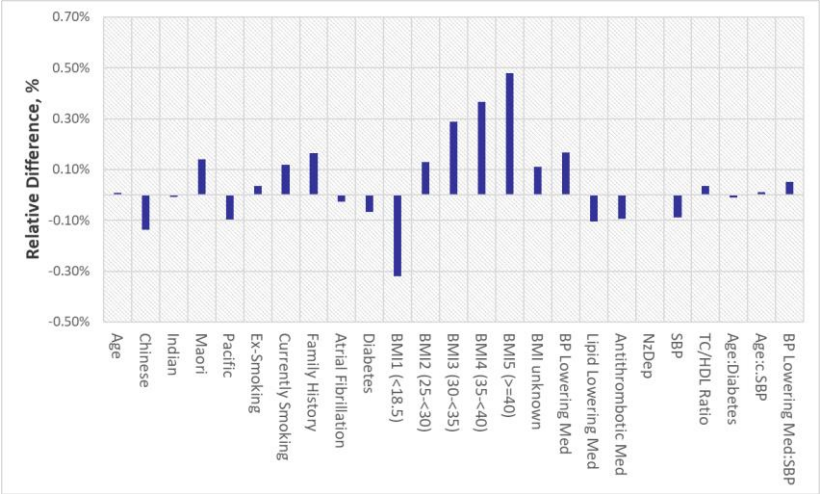
Figure 6: Proportion of patients in each 5-year CVD risk range

Table 15: Overview of demographics and risk factors about SBP end-digit groups among men and women

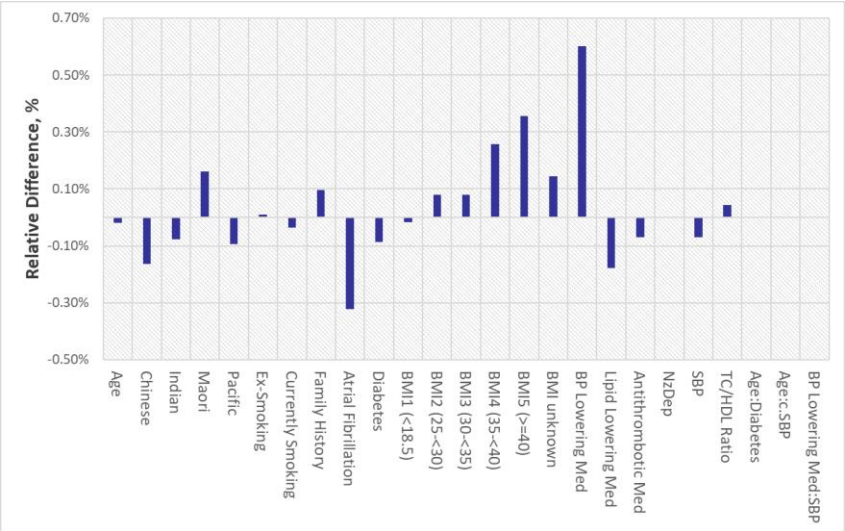
Risk Factors	**MEN (N= 163,528)**			**WOMEN (N= 128,594)**		
	SBP End-Digit (%)			**SBP End-Digit (%)**		
	Zero	**Other**	**Total**	**Zero**	**Other**	**Total**
Self-Identified Ethnicity						
European	29.30%	70.70%	138,195	28.50%	71.50%	103,922
Māori	32.30%	67.70%	31,473	30.30%	69.70%	27,141
Pacific	37.30%	62.70%	35,083	35.20%	64.80%	27,660
Indian	35.70%	64.30%	21,500	34.90%	65.10%	14,100
Chinese or other Asian	41.60%	58.40%	14,785	38.40%	61.60%	13,440
NZ Dep quintile						
1 (least deprived)	31.00%	69.00%	53,260	30.60%	69.40%	41,210
2	31.30%	68.70%	47,818	30.00%	70.00%	36,545
3	31.30%	68.70%	42,676	30.20%	69.80%	33,131
4	32.70%	67.30%	44,152	31.70%	68.30%	34,078
5 (most deprived)	34.40%	65.60%	53,130	32.20%	67.80%	41,299
Ex-smoker						
Yes	29.60%	70.40%	44,540	27.70%	72.30%	28,929
No	32.70%	67.30%	196,496	31.60%	68.40%	157,334
Current smoker						
Yes	33.70%	66.30%	40,287	30.90%	69.10%	23,798
No	31.90%	68.10%	200,749	31.00%	69.00%	162,465

Family history of premature cardiovascular disease						
Yes	28.10%	71.90%	23,452	27.40%	72.60%	21,969
No	32.60%	67.40%	217,584	31.40%	68.60%	164,294
Atrial fibrillation						
Yes	28.00%	72.00%	4060	26.70%	73.30%	1982
No	32.20%	67.80%	236,976	31.00%	69.00%	184,281
Diabetes						
Yes	31.80%	68.20%	24,063	32.00%	68.00%	22,475
No	32.20%	67.80%	216,973	30.80%	69.20%	163,788
Blood-pressure-lowering medication						
Yes	28.20%	71.80%	49,261	28.40%	71.60%	50,728
No	33.20%	66.80%	191,775	31.90%	68.10%	135,535
Lipid-lowering medication						
Yes	30.10%	69.90%	40,013	30.40%	69.60%	31,478
No	32.60%	67.40%	201,023	31.10%	68.90%	154,785
Antithrombotic medication						
Yes	29.80%	70.20%	24,384	30.20%	69.80%	19,186
No	32.40%	67.60%	216,652	31.10%	68.90%	167,077
BMI						
normal	33.20%	66.80%	41,862	31.40%	68.60%	44,654
underweight	35.90%	64.10%	781	31.50%	68.50%	1869
overweight	30.80%	69.20%	82,268	30.00%	70.00%	46,082
obesity class 1	30.90%	69.10%	46,854	30.20%	69.80%	29,691
obesity class 2	30.90%	69.10%	17,787	30.10%	69.90%	16,621
obesity class 3	32.20%	67.80%	10,070	30.40%	69.60%	13,918
BMI unknown	35.60%	64.40%	41,414	33.10%	66.90%	33,428

While comparing the estimates for Model 1 and Model 2 for both men and women, the highest relative difference for men and women was about 0.60%. This relative difference is illustrated in Figure 7. Further details regarding the HRs of the models are presented in Table 36 and Table 37 in Appendix A.



(a)



(b)

Figure 7: Relative change in the HR due to rounding for (a) men and (b) women respectively.

To evaluate the stability and reliability of the Cox PH model in categorising individuals into distinct risk categories, we performed a series of simulations with varying sample sizes (ranging from 10 to 15,000). Analysis for women presented in Figure 8 indicated that within roughly 5000 iterations, the variability in classification accuracy stabilized, as demonstrated by minimal fluctuations in the percentage of correct classifications and overlapping 95% confidence intervals. Same pattern was observed for men as well. The results suggest that beyond 5000 simulations, additional runs contribute little to overall accuracy improvement. We increased the simulation count to 10,000 for the analysis, to enhance the robustness of our estimates, mitigate potential random fluctuations associated with smaller counts, and strengthen the robustness of our conclusions.

Figure 9 shows the plots for misclassification in three risk groups due to rounding SBP for men and women. More under classification is noted than over classification with higher proportion of misclassification noted for men compared to women. The plots show that approximately 4.24% of high-risk men and 3.21% of high-risk women are misclassified into lower-risk categories. 1.19% of men are overclassified into a moderate-risk group, and 0.47% into a high-risk group. For women, 0.62% are overclassified into the moderate-risk group, and 0.20% into the high-risk group.

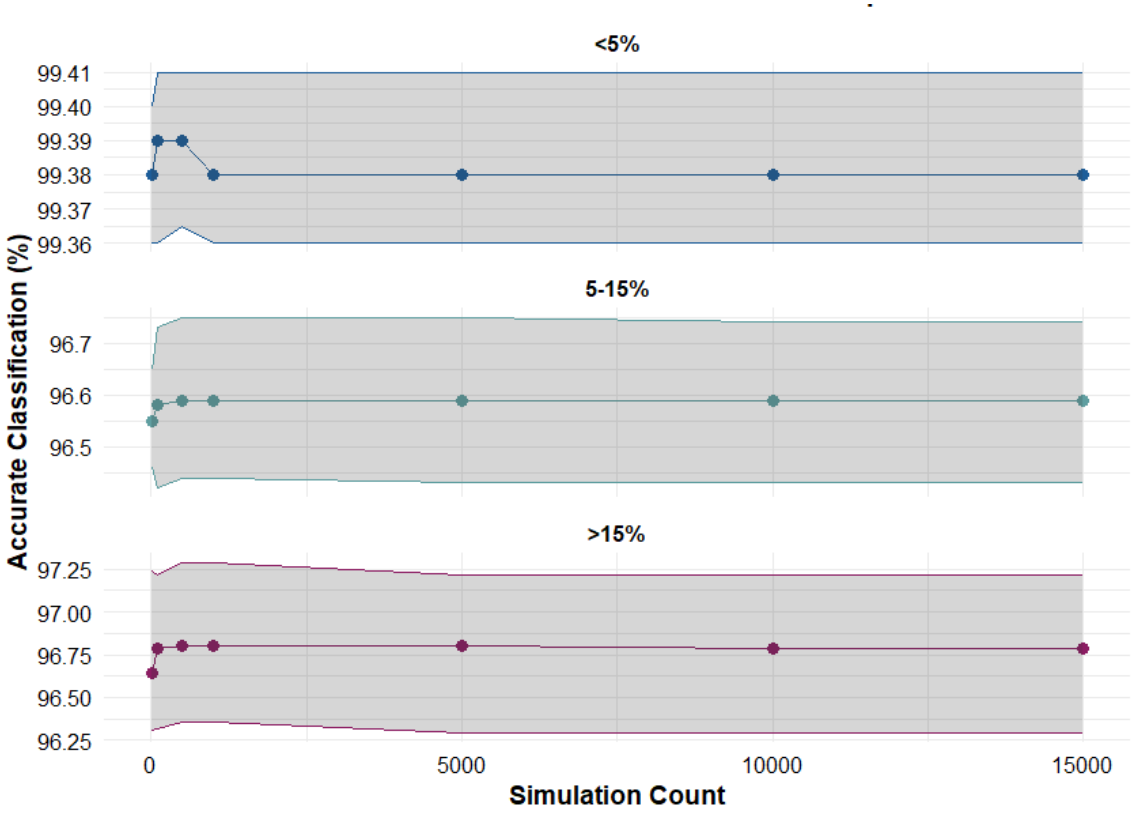


Figure 8: Accurate classification across risk groups by simulation count for women

Table 38 in Appendix A provides the 95% CI values for the misclassification rates for each risk group. While there is minimal deviation in the HRs for all the predictors used in Equation (1), there is still a notable misclassification in cardiovascular risk groups when rounding the SBP values to the nearest zero-end digit. The average number of people misclassified into different risk groups are mentioned in Table 16. For instance, among women in our sample, 670 moderate risk patients are misclassified as low risk, leading to undertreatment, whilst 654 low risk patents are misclassified as moderate risk, which could lead to overtreatment. Similarly, we found that 1,291 men previously in moderate risk group were misclassified to low-risk group, and 1,354 men initially in the low-risk category are incorrectly classified as moderate risk.

The results in Table 16 are generated by rounding all non-zero SBP values. As discussed earlier, around 32% of the data has zero end-digit. Assuming device calibration to the nearest 2 mmHg, there is around 12% data that was rounded to nearest zero end-digit. After adjusting for entire cohort, on average, 304 patients (193 men and 111 women) were undertreated, while 308 patients (201 men and 107 women) were overtreated due to misclassification.

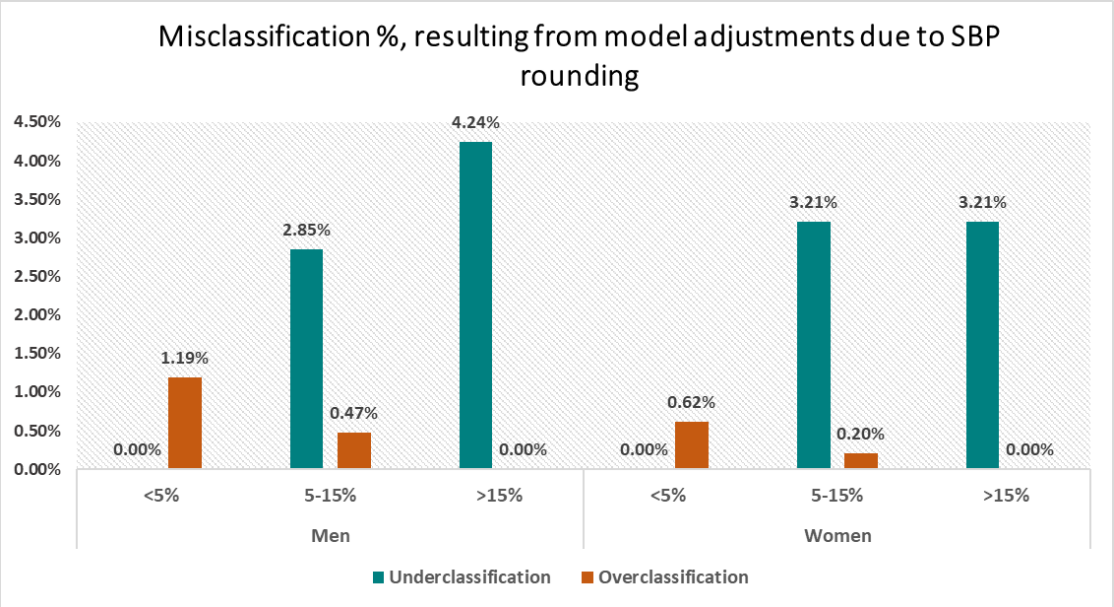


Figure 9: Misclassification plots due to rounding per risk group for men and women

Table 16: Misclassification count for men and women along with 95% CI's

		Rounded Risk			
	Original Risk	**<5%**	**5–15%**	**>15%**	**Total**
Women	**<5%**	105,629 (105,599–105,656)	654 (627–684)	0	106,283
	5–15%	670 (639–703)	20,196 (20,162–20,227)	43 (35–51)	20,909
	>15%	0	45 (39–52)	1357 (1350–1363)	1402
Men	**<5%**	112,257 (112,214–112,296)	1354 (1315–1397)	0	113,611
	5–15%	1291 (1250–1334)	43,839 (43,784–43,888)	215 (193–241)	45,345
	>15%	0	194 (181–208)	4378 (4364–4391)	4572

Results from stratified sampling across 100 iterations showed minimal deviation in the misclassification rates, with a maximum variation of only 0.001 and overlapping confidence intervals. The average estimates of the misclassification rate are presented in Table 39 in the Appendix A. No significant changes were observed in the overall survival model when generalized to the whole

cohort. Further details on the hazard ratios from the model can be found in Table 36 and Table 37 in the Appendix A.

3.5 Financial Impact

The cost data was sourced from the PHARMAC Cost Resource Manual [189]. This manual provides key pricing information relevant to the funding of pharmaceuticals, including consultation and diagnostic costs. The gathered data, along with minimum and maximum cost ranges, are presented in Table 17. This table outlines the total consultation and diagnostic costs (in NZD) over a 5-year period for the three cardiovascular risk groups. The estimates account for the differences in follow-up assessments specific to each risk group. They also reflect the variations in the frequencies of healthcare interventions required for effective risk management across different risk levels.

Additionally, medication cost estimates were derived from various reports on type of medications recommended in New Zealand. These estimates were based on the community pharmaceutical schedule [190], which lists all medications funded for New Zealand residents. Table 18 details the daily per-dose cost of three commonly prescribed medications for managing CVD: statins, antihypertensives, and aspirin[30], [191], [192].

The medications are typically prescribed for long-term CVD prevention[193], [194], [195]. For individuals with a CVD risk of less than 5%, lifestyle changes are primarily recommended, and medication is not generally prescribed. In contrast, for those with a risk greater than 15%, the total estimated cost in 5 years per person ranges from NZD $1,198 to NZD $5,908, reflecting variations in the type and dosage of medications prescribed.

For moderate risk patients, there is evidence that statins and antihypertensives are recommended, as they have been shown to provide significant benefits in reducing cardiovascular events [196], [197]. Addition of antithrombotics is considered beneficial only for secondary prevention in those with higher CVD risk (≥15%), where the benefits generally outweigh the risks despite the potential for side effects like gastrointestinal bleeding and ulceration [193]. However, the decision to prescribe additional medications ultimately depends on the physician's assessment of the patient’s overall

health profile, including comorbidities and specific risk factors [82]. For cost estimation of moderate risk group, prices of statins and antihypertensives are used, as they are commonly prescribed medications. The average estimated total cost in 5 years is NZD $1255.38 per person ranging from NZD $766 to NZD $4669. This approach conservatively estimates cost as the actual cost could vary based on medication type and the number of medications prescribed.

Table 17: Estimated Cost breakdown for consultation and diagnostic resources used in CVD risk assessment.

	Frequency of Assessment (in 5 Years)	**Minimum Cost**	**Cost**	**Maximum Cost**
GP Visit		NZD 80.00	NZD 80.00	NZD 80.00
Nurse Visit		NZD 40.00	NZD 40.00	NZD 40.00
Lipid Profile Test		NZD 11.64	NZD 22.89	NZD 46.50
Blood Glucose Test		NZD 15.12	NZD 24.27	NZD 39.50
Electrocardiogram (ECG)		NZD 60.00	NZD 68.80	NZD 85.00
Cost (Annual)		NZD 206.76	NZD 235.96	NZD 291.00
Cost (in 5 years)				
Risk Category: <5%	2	NZD 413.52	NZD 471.92	NZD 582.00
Risk Category: 5–15%	3	NZD 620.28	NZD 707.88	NZD 873.00
Risk Category: >15%	5	NZD 1033.80	NZD 1179.80	NZD 1455.00

Table 18: Estimated Cost breakdown for medication prescribed for CVD management.

	Minimum Cost	**Cost**	**Maximum Cost**
Statins	NZD 0.05	NZD 0.05	NZD 0.05
Antihypertensives	NZD 0.03	NZD 0.25	NZD 2.03
Antithrombotics	NZD 0.01	NZD 0.14	NZD 0.36
Cost (Annual)	NZD 32.85	NZD 160.60	NZD 890.60
Cost (in 5 years)			
Risk Category: <5%	NA	NA	NA
Risk Category: 5–15%	NZD 146.00	NZD 547.50	NZD 3796.00
Risk Category: >15%	NZD 164.25	NZD 803.00	NZD 4453.00

For around 2008 overclassifications into a moderate-risk category for men and women, as shown in Table 16, expenses of approximately NZD $1.57 million will be incurred, with expenses ranging up to NZD $8.2 million. Additionally, the misclassification of 1961 men and women into the low-risk category leads to undertreatment, which significantly impacts hypertension management and increases the likelihood of adverse outcomes, such as cardiovascular events, renal dysfunction, and all-cause mortality[198]. Undertreated patients are at a higher risk of experiencing these adverse events, contributing to increased healthcare costs. The inpatient cost for managing cardiovascular events ranges from NZD $1,200 for a hospital medical ward to NZD $5,500 per patient for an ICU, expenses that could largely be avoided through accurate risk assessment and timely intervention[189].

3.6 Discussion and Conclusion

This research evaluated the impact of rounding by applying it to the entire subset of non-zero SBP records and then generalizing the results to the entire cohort. Although the overall model showed only slight variations in both cases, notable misclassification was still observed. This misclassification contributes to an increased healthcare burden, as it affects the accuracy of risk classification, potentially leading to inappropriate medical interventions. As a result, patients may not receive the necessary treatment, ultimately compromising their health outcomes and increasing the likelihood of adverse cardiovascular events. Potential solutions to reduce rounding errors emphasise the need for education programs highlighting the impact of rounding for healthcare practitioners with periodic retraining in BP measurement. Implementing these strategies can improve the reliability of CVD risk assessment tools, ultimately supporting better clinical outcomes and more informed decision-making for both healthcare providers and policymakers.

The practice of recording BP with a zero end-digit, known as "zero end-digit preference," is common in primary care, particularly during routine measurements [199]. This poses significant implications when healthcare professionals employ equations like the PREDICT-1 equation to estimate a patient's 5-year CVD risk [118]. It was assumed that the measurement device being used were standard manual auscultatory sphygmomanometers as they remain the gold standard for non-

invasive BP measurement due to its accuracy compared to other devices. The manual sphygmomanometers continue to serve as the reference standard for validating other devices in development and clinical studies as per almost all the international guidelines [49], [200]. Thus, it becomes crucial to understand the impact of rounding, a prevalent practice in clinical settings. [180]. However, there has been no information gathered since then, regarding the type of measurement device used, emphasizing the need for improved data collection in the future.

This work used a substantial patient cohort sourced from over 400 general practitioners (GPs) and analysed data from standard primary care settings, where healthcare providers measured SBP during routine visits [180]. It was assumed that individuals with non-zero SBP values had their BP recorded more accurately, given the lack of evidence for rounding to other end digits [180]. The analysis estimates the upper-bound impact of rounding errors by rounding non-zero SBP values and also includes a generalized version for the entire cohort.

The prevalence and implications of zero end-digit preference have been addressed in previous studies. Research involving 85,000 BP measurements in patients with ischaemic heart disease in England discovered that zero end-digit preference accounted for 64% of SBP and 59% of DBP measurements [114]. A study demonstrated that error variance of the associations between BP and both BMI and age was higher when zero end-digit preference for BP readings were employed in comparison to unbiased readings [118].

The observed effect of BP rounding on CVD risk misclassification highlights the importance of establishing measures to minimise rounding inaccuracies in clinical assessments. Standardising BP measurement techniques is an essential initial step, prioritising the training of clinical personnel to accurately document readings. Many large surveys including National Health and Nutrition Examination Survey have advocated appropriate certification and periodic re-certification as a means to mitigate these errors [201]. Transitioning to digital sphygmomanometers, where feasible, can reduce human error and provide consistency across providers. The use of automated devices has demonstrated efficacy in reducing zero end-digit preference in studies like Hypertension Optimal Treatment and Anglo-Scandinavian Cardiac Outcomes Trial [118]. It is recommended to revise clinical guidelines to consider the effects of BP rounding on CVD risk assessment, in addition to

conducting regular educational sessions that highlight the importance of precise measurement in risk classification.

Evidence shows higher proportion of misclassifications in lower risk categories for both men and women, a trend that may be influenced by the greater representation of individuals within the lower-risk group among NZ's population. This study demonstrated that BP measurements were more frequently rounded down than up at various stages of hypertension. The propensity to round down may result in a higher frequency of individuals being under classified instead of overclassified into lower-risk categories, as BP demonstrates a positive correlation with CVD risk [202]. This influence can potentially lead to either mistreatment of an individual when considering other relevant factors [203]. If undertreated, the individual might not be prescribed necessary medications or lifestyle changes to manage their actual higher risk, leading to delay in necessary treatment which might increase the likelihood of adverse CVD events. Conversely, overtreatment may subject individuals to unnecessary interventions and frequent monitoring, leading to inefficiencies in healthcare resource allocation [204] [205]. While estimating cost, it was assumed that all individuals in the moderate risk group are only prescribed statins and antihypertensives. Despite being a conservative estimate, it amounted to be a large financial burden. With additional information, the estimate might be much higher. Hence, accurate classification is foremost in CVD risk management.

Research highlights the importance of achieving treatment targets to prevent cardiovascular events. Maintaining optimal BP control, as per the 2017 American College of Cardiology / American Heart Association guidelines could prevent 71.9 cardiovascular events per 1,000 treated individuals over 10 years [206]. Failure to initiate appropriate interventions or adhere to recommended treatments increases the risk of cardiovascular events by 20% and mortality by 35%. This leads to extended hospital stays and a heavier long-term healthcare burden. Accurate risk classification and early management are therefore essential in reducing both clinical and financial strain[207].

The likelihood of BP rounding may be dependent on certain confounding factors, such as socioeconomic status, e.g. high NZDep scores, or access to better healthcare facilities. However, due to the lack of data or research on these specific factors, it was not feasible to account for them in the study. This limitation emphasises the need to conduct future research to investigate the

potential impact of these variables on BP rounding practices, thereby enabling more precise generalization and adjustment for potential confounders.

We acknowledge that there are various other sources of BP measurement errors, such as device calibration issues, repeated measurements and patient factors like white coat hypertension or improper cuff size, which can all influence treatment decisions [176], [208]. However, this study specifically focuses on the impact of rounding errors on treatment thresholds established by the Ministry of Health in NZ [82]. The results provide additional insights into the existing body of knowledge, highlighting that while they are not the sole determinant, they are an important factor in clinical decision-making. Future research should also consider the additional sources of error to provide a more comprehensive understanding of the challenges in accurate BP measurement and improve clinical outcomes.

The interpretation of our study findings is particular to the New Zealand context, as we are investigating the impact of rounding BP on CVD risk prediction using New Zealand based guidelines and data. These outcomes may differ in various nations due to variations in attributes, such as ethnicity, socioeconomic level, and healthcare practices. Furthermore, our healthcare cost estimate is based on cost schedules from PHARMAC, a NZ government agency, which may vary from cost structures and insurance procedures in other nations. Future research in other regions may yield comparative insights and evaluate the relevance of these findings across varied healthcare systems.

CHAPTER 4. IMPACT OF DEVICE-RELATED INACCURACIES IN BLOOD PRESSURE MEASUREMENTS ON CARDIOVASCULAR RISK CLASSIFICATION IN NEW ZEALAND (Manuscript 2).

4.1 Prelude

Chapter 2 reviewed research on device inaccuracies in clinical practice. Building on this, Chapter 4 examines the impact of blood pressure (BP) measurement errors arising from inaccuracies in automated cuff-based sphygmomanometers. Even when within internationally allowable error limits, these inaccuracies can distort BP readings, leading to potential misclassification of patients into incorrect cardiovascular risk categories. This misclassification may, in turn, influence clinical decision-making and patient outcomes. The aim of this study is to quantify how such device errors affect risk classification and evaluate their potential clinical implications.

A 5-year CVD risk was assessed for 427,299 individuals aged between 30-74 years. Baseline SBP values were used to evaluate a 5-year CVD risk, subsequently adjusted after considering device-related inaccuracies in Systolic Blood Pressure (SBP) measurements. Survival models were used to estimate risk, with hazard ratios as the foundation for model comparisons. The misclassification rates across three CVD risk categories (<5%, 5–15%, >15%) resulting from device inaccuracies were examined, together with the financial implications of these misclassifications.

The highest recorded misclassification rates indicated that as much as 7.50% of men were inaccurately classified into higher-risk groups, while roughly 1.97% were misclassified into lower-risk categories. Among women, around 5.65% were overclassified, and approximately 1.02% were under-classified.

Automated device-related inaccuracies in SBP measurements affect CVD risk classification and this could have significant financial implications. Instructing healthcare workers on these effects is crucial to advocate for more stringent standards for blood pressure devices, hence improving accuracy and clinical results.

4.2 Introduction

Blood Pressure (BP) measurement is crucial in assessing cardiovascular health. High BP contributes to heart attacks, heart failure and other cardiovascular conditions. In particular, systolic blood pressure (SBP) is a significant predictor of cardiovascular events [209],[210]. Even within normal SBP ranges, maintaining lower levels is linked to a reduced risk of cardiovascular diseases (CVD), emphasising the importance of optimal SBP management [109]. According to the World Health Organization (WHO), CVD accounts for around 32% of global mortality [211].

Manual measurement using a manual auscultatory devices, and a stethoscope is considered the gold standard of BP measurement. However, this method is technically demanding and often leads to faulty values. To address these challenges, automated BP devices have been developed that simplify the measurement process and reduce technical difficulties [212]. Despite advancements, BP measurement is still prone to errors. Automated devices may be more reliable in reducing human error compared to manual sphygmomanometers, but these devices are not immune to inaccuracies [213]. Even minor errors in BP measurement can have significant consequences. For instance, an underestimation of BP by merely 5 mmHg can result in the misclassification of millions of individuals [174], which poses substantial risk for fatal strokes and myocardial infarctions [175]. Hence, it becomes crucial to understand the impact of errors in BP measurement on CVD risk prediction when measurements were taken using automated BP devices that have a certain level of inaccuracy.

Different types of errors can affect BP measurements, including systematic and random errors. A study in the US investigated the impact of systematic and random errors in measuring BP on the prevalence of high BP among adults [214]. Systematic errors, arising from number of technician and device related factors, were added and subtracted from the original measurements. These adjustments caused the proportion of high BP to increase to 44.4% with higher BP measurements and dropped to 21.9% with lower measurements, compared to 32.0% when standardized measurements were used. Similarly with random errors, 21.4% of US adults not taking antihypertensive medication and 20.5% of those taking antihypertensive medication had their high BP classification changed. These systematic and random errors were selected based on acceptable

error limits as per international standards, prior studies of the differences in SBP and DBP between routine care and research-grade measurements [215][216][177][217].

To address the challenges posed by biases in BP measurements, national regulators have made substantial efforts to standardise global regulations for medical devices. Manufacturers must adhere to standardised protocols when designing BP devices to ensure that inaccuracies remain within an acceptable range, generally represented as the mean error ± standard deviation (SD) of BP errors for non-invasive methods. ISO 81060-2:2018 specifies the criteria for the clinical evaluation of automated, non-invasive sphygmomanometers [49]. This standard has been adopted and acknowledged by many national regulators, superseding region-specific standards [132].

ISO standard establishes standards for assessing the acceptable accuracy of sphygmomanometers. Its origin trace back to the establishment of the American Standard for manual, electronic, or automated sphygmomanometers known as SP10 [200]. ISO, similar to SP10, mandates a minimum sample size (N) of 85 participants for assessing the inaccuracy of the BP device. Considering data collection protocols, such as using appropriately sized cuffs, ensuring patients are seated for 5 minutes and other similar measures, the standard defines acceptable error limits. For a particular sample mean error, the SD of the error must not exceed a threshold. Table 8 in Section 2.4.2 shows the upper limit of SD for a given mean as stipulated by the standard. The standards set error limits but do not absolutely prevent measurement inaccuracies. The impact of these error margins on patient's health, particularly regarding misdiagnoses of conditions like CVD, remains uncertain and invites further analysis.

Therefore, this study aims to investigate the impact of BP measurement inaccuracies on the misclassification of patients into CVD risk groups within the New Zealand (NZ) population. NZ has its own CVD risk prediction equation known as the PREDICT-1 equation, which is currently used to classify patients into three risk categories (<5%, 5-15%, and >15%) as established by the Ministry of Health [53]. These risk categories guide clinicians in recommending appropriate interventions, setting goals, and planning follow-up care based on the assessed cardiovascular risk. Hence, this research focuses on how these BP measurement inaccuracies could influence clinical decision-

making and patient outcomes. The research further investigates the effects of these device-related misclassifications on the total cost of healthcare in the NZ healthcare system.

4.3 Methods

The PREDICT-1 equation is a CVD risk prediction equation developed using the data gathered for 427,299 people in NZ. These equations were derived using the Cox proportional hazard (Cox PH) model, which allows for the assessment of time-to-event data [188]. The cohort study used to develop these equations is the PREDICT study, an ongoing, open cohort study in NZ that automatically enrols participants when primary health-care practitioners conduct standardised CVD risk assessments using PREDICT decision support software [218].

The software automatically fills PREDICT risk factor templates using patient records. Clinicians must complete any missing data prior to calculating CVD risk and finalising recruitment. The participant risk factor profiles recorded by the software are consistently linked to national databases that document drug dispensing, ICD-coded hospitalisations, and CVD related fatalities.

4.3.1 Study Design

This study examined primary care patients who underwent CVD risk assessments at primary healthcare organisations using the PREDICT software. The cohort comprised 427,299 individuals, both men and women, aged 30-74 years at the time of their initial risk assessment. Individuals having a history of CVD, congestive heart failure, or renal disease were excluded based on diagnoses from general practitioners, hospital discharge records, and the dispensing of anti-anginal medications and loop diuretics. Data for 5 ethnic groups, namely European, Māori, Pacific, Indian and Chinese has been recorded. For each individual in the cohort, two BP measurements were taken, and their average was used in the development of the PREDICT-1 equation. All analyses were conducted over three risk categories, namely: low (<5%), moderate (5-15%) and high (>15%).

Recognizing that SBP measurements can be subject to errors that typically follow a normal distribution [49], we simulated these errors according to international standards such as ISO. Hence, normally distributed random values were generated using specified means and their respective SD

that align with acceptable limits outlined in these standards[49]. The maximum allowed mean error for SBP measurements is 5 mmHg. Therefore, we considered three mean values for our simulation:

1. Mean = 0 mmHg, SD = 6.95 mmHg.
2. Mean = 2.5 mmHg, SD = 6.47 mmHg.
3. Mean = 5 mmHg, SD = 4.81 mmHg.

These three scenarios encompass a range of potential errors, including the minimum, intermediate, and maximum mean errors. We focus on the potential problems that arise when patients who are at moderate risk are misclassified as low risk. This misclassification can lead to significant undertreatment because low-risk patients are usually given lifestyle recommendations rather than medication. While moderate-risk patients typically receive medication based on clinical evaluations, high-risk patients get more comprehensive medical care, similar to those with a prior CVD.

4.3.2 Simulation and Analysis

The simulation is designed to assess the impact of device errors on CVD risk prediction by analysing how errors in SBP measurements would potentially influence risk classification. The study assumes that the SBP values originally recorded in the dataset are error-prone, referred as SBP_{error}, meaning they were measured using a device with inherent inaccuracies. The device inaccuracies are removed in the form of normally distributed errors with above specified means and standard deviations, referred as $SBP_{adjusted}$. The design of the study is illustrated in Figure 10, and explained in the following steps:

1. Stratification by Risk Categories

All analyses are stratified into three risk groups. This stratification allows us to evaluate the impact of device errors across different levels of risk, providing a comprehensive understanding of how measurement errors influence risk prediction outcomes.

2. Calculation of 5-Year CVD Risk with SBP_{error}

In the initial stage of the analysis, the 5-year CVD risk is calculated using the original SBP_{error} values. The 5-year CVD risk for each individual is estimated by applying the Cox PH model (original model)

on the dataset, where SBP_{error} values are used as risk factors as expressed in Equation (11). This 5-year risk score is referred as the original risk. The original risk represents the predicted probability of a CVD event occurring within 5 years.

$$\frac{h(t)}{h_0(t)} = e^{(\beta_1 SBP_{error} + \beta_2 z_2 + \ldots\ldots + \beta_p z_p)} \quad (11)$$

In equation (11), $\frac{h(t)}{h_0(t)}$ is the hazard ratio (HR) that describes how the hazard changes as a function of the covariates.

h(t) is the hazard at time t

h0(t) is the baseline hazard when all covariates are zero

Covariates SBP_{error}, z_2 ,…, z_p are individual risk factors and $\beta_1, \ldots, \beta_p$ are their corresponding beta coefficients.

3. Adjustment of SBP and Recalculation of 5-Year CVD Risk

Recognizing that SBP measurements can have device-related errors, the next step is to adjust these measurements and obtain $SBP_{adjusted}$, in which simulated errors are removed. This adjustment allows for recalculating the CVD risk with adjusted SBP values, providing an approximation to a simulated actual SBP. Once the SBP_{error} values are replaced with $SBP_{adjusted}$, a new 5-year CVD risk is calculated for each individual referred to as adjusted risk. The adjusted risk represents the risk estimation after accounting for device inaccuracies, using the $SBP_{adjusted}$. All other variables in the dataset are unaltered. An updated CoxPH model (adjusted model) as expressed in Equation (12) is used to calculate the adjusted risk.

$$\frac{h(t)}{h_0(t)} = e^{(\beta_1 SBP_{adjusted} + \beta_2 z_2 + \ldots\ldots + \beta_p z_p)} \quad (12)$$

4. Simulation Study for $SBP_{adjusted}$

To reduce the risk of variability while generating errors in the above step, the process is repeated 10,000 times. This simulation is useful to evaluate the practical impact of SBPadjusted measurements on risk prediction. It also highlights the variability in risk classification.

For each simulation, the 5-year CVD risk is recalculated using the $SBP_{adjusted}$ values. The impact of these corrections is analysed by comparing the Adjusted Risk to the Original Risk, specifically focusing on how individuals are classified into different risk groups. The average misclassification rate across the simulations is calculated, along with its 95% confidence interval, to assess how measurement errors affect CVD prediction. To assess comparisons between the model estimates of the two models being tested, the relative difference between the hazard ratios is examined, which is defined in Equation (13) as,

$$Relative\ Difference,\% = \frac{\mathrm{HR}_{Adjusted\ Model} - \mathrm{HR}_{Original\ Model}}{\mathrm{HR}_{Original\ Model}} * 100 \quad (13)$$

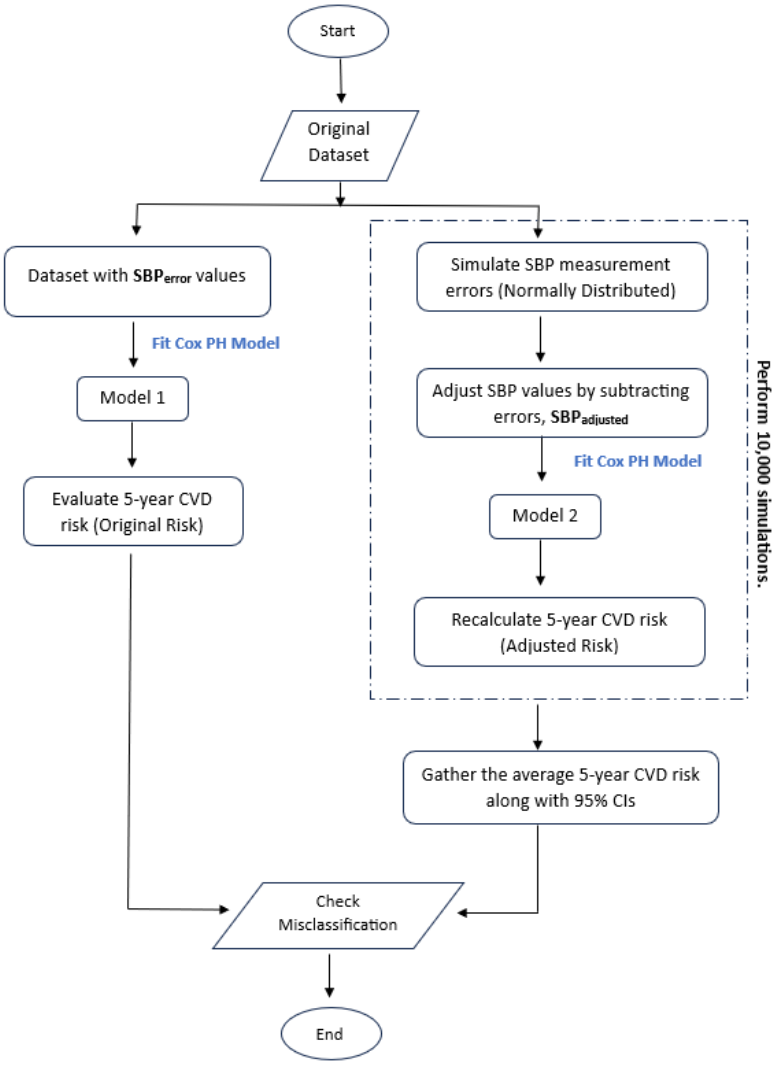


Figure 10: Process flow diagram to evaluate misclassification rate by introducing a new model with adjusted SBP

4.4 Results

The dataset used in the analysis contains 241,036 men with an average age of 51.29 years (Interquartile Range, IQR = 14 years, and SD = 10.07 years) and 186,263 women with an average age of 55.85 years (IQR = 13 years and SD = 8.87 years). The HRs for adjusted models represent the average of 10,000 simulations conducted for Model 2, with a maximum variation of only 0.007 between the HRs across all simulations. More detail is provided in Table 40 and Table 41 in Appendix B regarding the HRs for men and women. Upon examining the variations in HRs for both genders at SD=6.95 mmHg (the maximum permissible SD), we observed no substantial discrepancies between the original dataset and the dataset that accounts for the errors. The highest relative difference among the HRs for all the variables was observed 3.5% and 1.6% for men and women; also presented in Figure 31 in Appendix B.

Even though there was no notable difference observed in the models, misclassification of risk categories due to device errors persisted. Figure 11 and Figure 12 provide plots showing misclassification patterns for different combinations of mean and SD, highlighting the impact of device errors. While the degree of misclassification varied, the maximum observed misclassification rates showed that up to 7.50% of men were overclassified into higher-risk categories, and approximately 1.97% were under-classified into lower-risk categories. Among women, a maximum of 5.65% were overclassified, while around 1.02% were under-classified. The detailed version of the misclassification rates with their corresponding 95% CIs are provided in Table 42, Table 43 and Table 44 in Appendix B.

According to the established guidelines, the maximum allowable SD is 6.95 mmHg, resulting in about 4896 patients being undertreated and 4650 being overtreated due to misclassification also presented in Table 19. These results indicate an unnecessary expenditure of roughly NZD 5.84 million with the potential to increase to NZD 21.71 million in 5 years. This estimate is based on cost data presented in Chapter 3, that provides a detailed breakdown of expenses associated with CVD risk assessment and treatment [219]. Table 20 presents comprehensive average misclassification counts, for the three risk groups. Furthermore, undertreated patients do not receive adequate treatment, which

increases their risk of unfavourable CVD events and all-cause death, resulting in increased healthcare costs. The inpatient cost of managing cardiovascular events ranges from $1,200 for a hospital medical ward to $5,500 per patient for an ICU, charges that may be significantly reduced with precise risk assessment and appropriate management [189].

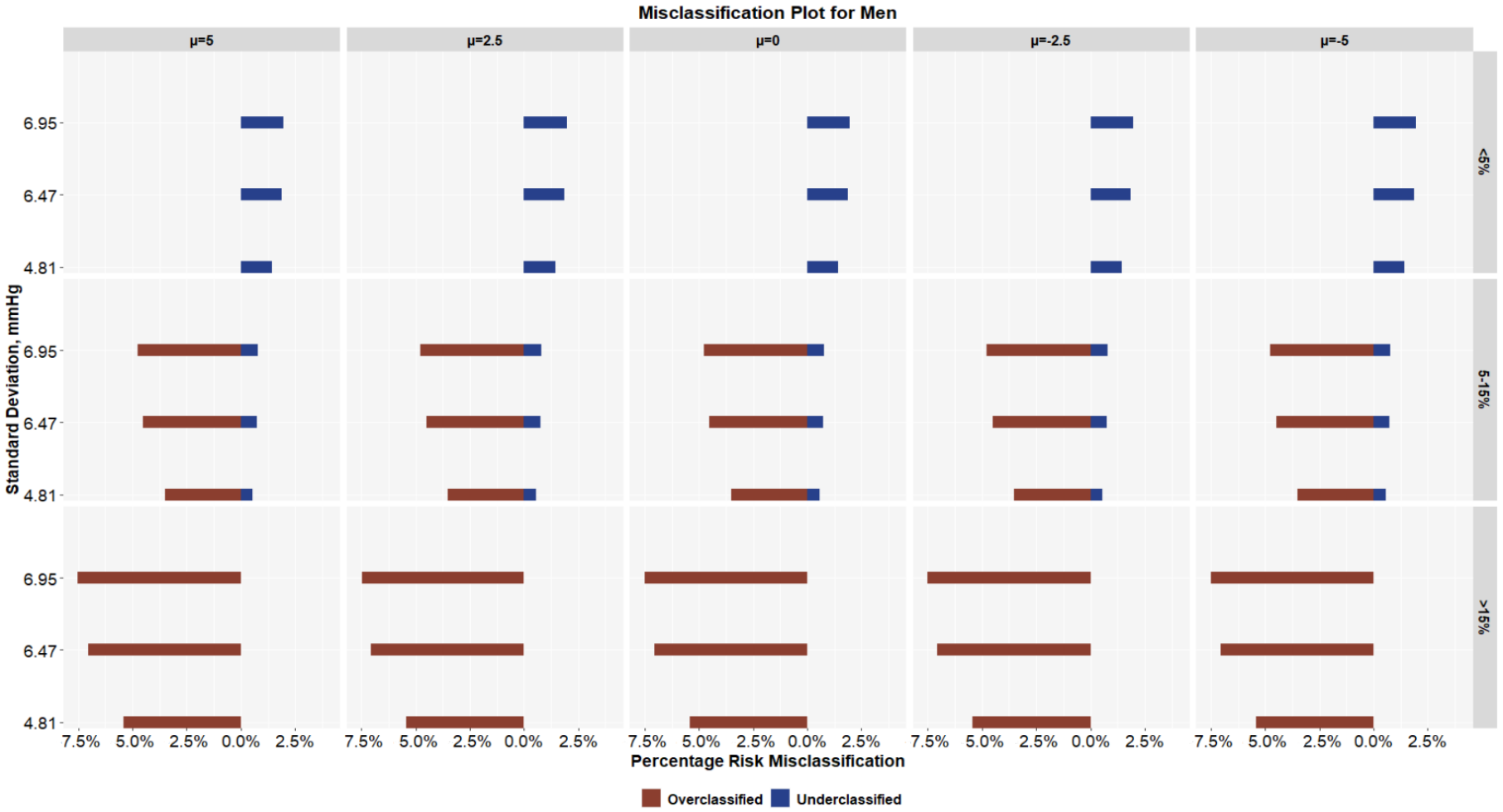


Figure 11: Misclassification plot for different combination of acceptable device errors for men.

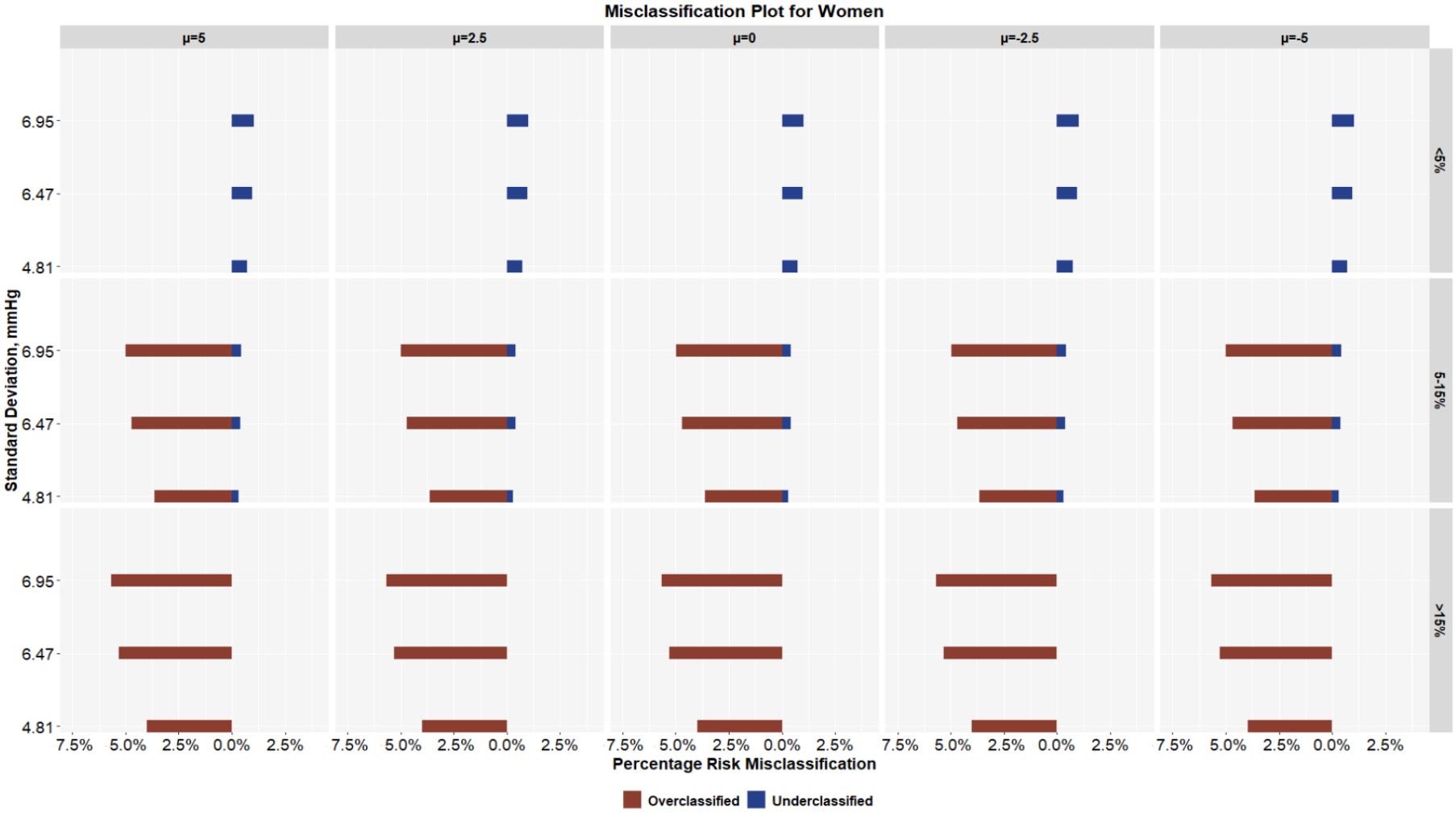


Figure 12: Misclassification plot for different combination of acceptable device errors for women.

Table 19: Misclassification count after accounting for (a) error ~ N (0,6.95) (b) error ~ N (2.5,6.47) (c) error ~ N (5,4.81)

(a)

	Model 1 with SBP_{error}	**Model 2 with SBP_{actual}**			
		<5%	**5-15%**	**>15%**	**Total**
Men Results	**<5%**	164890	3321	0	168211
	5-15%	3153	62477	521	66151
	>15%	0	501	6173	6674
Women Results	**<5%**	152667	1575	0	154242
	5-15%	1497	28440	128	30065
	>15%	0	111	1845	1956

(b)

	Model 1 with SBP_{error}	**Model 2 with SBP_{actual}**			
		<5%	**5-15%**	**>15%**	**Total**
Men Results	**<5%**	165082	3129	0	168211
	5-15%	2977	62686	488	66151
	>15%	0	471	6203	6674
Women Results	**<5%**	152762	1480	0	154242
	5-15%	1411	28535	119	30065
	>15%	0	103	1853	1956

(c)

	Model 1 with SBP_{error}	**Model 2 with SBP_{actual}**			
		<5%	**5-15%**	**>15%**	**Total**
Men Results	**<5%**	165804	2407	0	168211
	5-15%	2320	63465	366	66151
	>15%	0	363	6311	6674
Women Results	**<5%**	153113	1129	0	154242
	5-15%	1091	28885	89	30065
	>15%	0	78	1878	1956

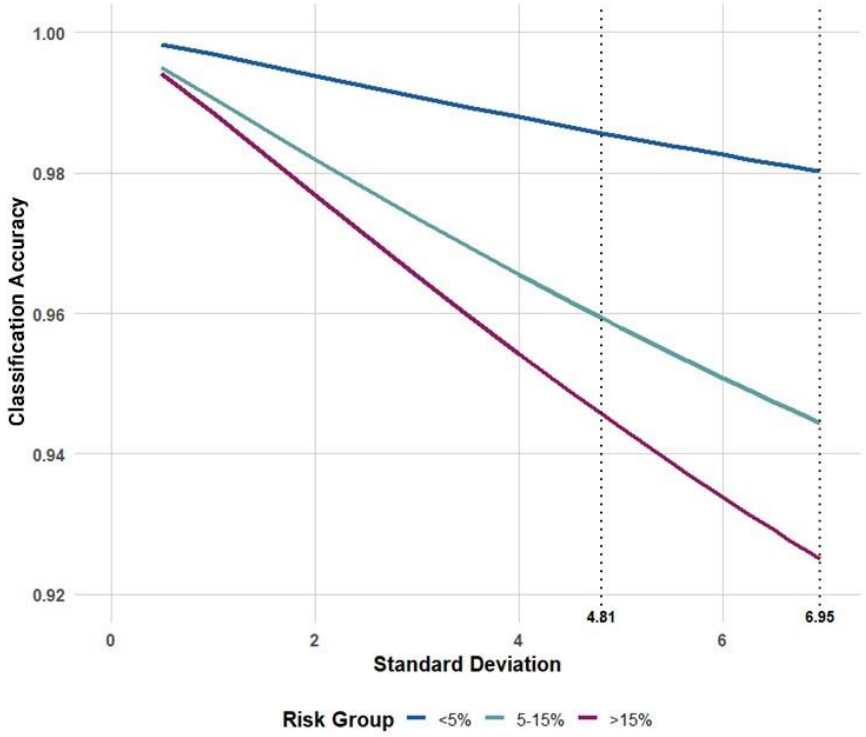
Table 20: 5-year cost estimates for the three risk groups.

Risk Group	Cost	Minimum Cost	Maximum Cost
<5%	NZD 471.92	NZD 413.52	NZD 582.00
5-15%	NZD 1255.38	NZD 766.28	NZD 4669.00
>15%	NZD 1982.80	NZD 1198.05	NZD 5908.00

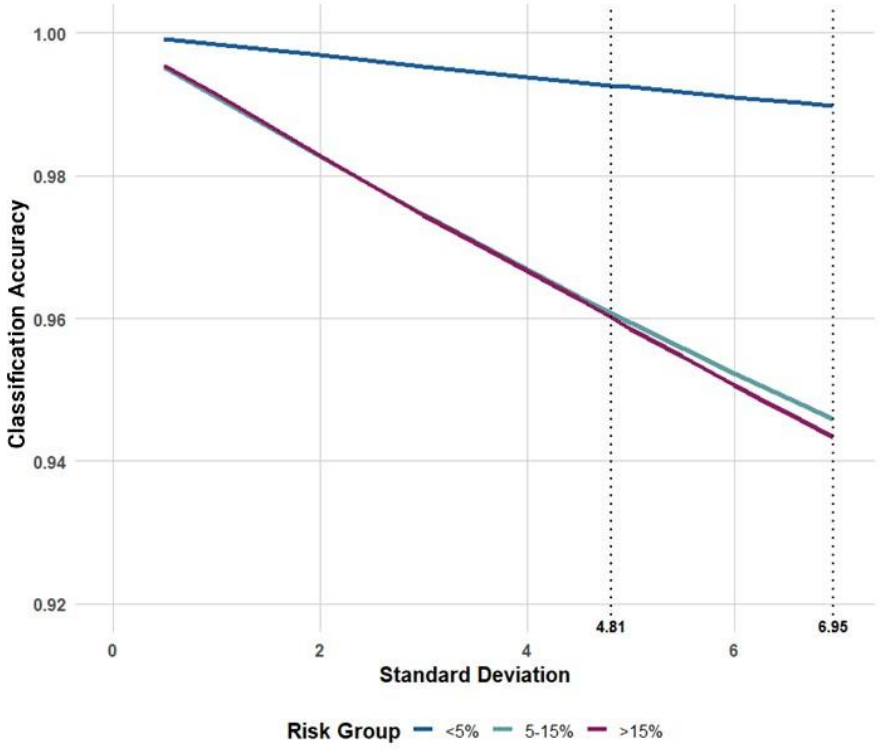
Figure 13 illustrates the classification accuracy for different SD in device errors. This highlights the effect of varying SD on classification accuracy for (a) men and (b) women. The two vertical dotted lines represent the permissible range of SD values as per ISO. The figure shows the percentage of correct classifications across three CVD risk categories. As the SD increases from 0 to 6.95 mmHg, a general decline in classification accuracy is observed for all risk groups. Even within the allowable SD range, a noticeable decrease in classification accuracy is evident. At the upper limit of the permissible SD (6.95 mmHg), the classification accuracy reaches approximately 92.5% for men, while for women, the accuracy is around 94.5%. At the SD value of 4.81 mmHg, which represents the maximum permissible SD for the minimum mean error (0 mmHg), the classification accuracy for men is slightly higher than at 6.95 mmHg, though still showing a notable decline compared to lower SD values. At an SD of 4.81 mmHg, the classification accuracy for men is roughly 94.5%, while for women it is around 96%, suggesting that despite this acceptable level of device error, a certain degree of misclassification remains. It is also noteworthy that the higher risk categories (5-15% and >15%) are more susceptible to incorrect classification despite having a relatively smaller population being classified into these categories.

4.5 Discussion and Conclusion

The results suggest that the estimation of five-year CVD risk in NZ patients may either be overestimated or underestimated even when BP measurements are obtained using devices that comply with allowable inaccuracies as defined by international standards. Despite adherence to these standards, such devices can still introduce misclassification across different CVD risk categories.



(a)



(b)

Figure 13: Accurate Risk classification after accounting for the normal errors with a range of standard deviation for (a) men and (b) women.

The findings indicate a stronger propensity toward overestimation of risk for both men and women compared to underestimation, a trend potentially influenced by a higher proportion of patients in the lower-risk categories within NZ population. Overestimation may result in unnecessary treatments, contributing to a significant healthcare burden. Conversely, underestimation increases the likelihood of adverse events over time, resulting in increased long-term healthcare costs and negatively affecting patient outcomes by delaying preventive interventions.

Several studies have highlighted higher inaccuracies particularly when employing different techniques or algorithms that deviate from standard practices [141], [220]. However, in this study, we adopt a conservative approach by examining the impact of device errors strictly within the allowable range defined by international standards. This approach ensures that our analysis focuses on understanding the potential implications of inaccuracies that are considered acceptable in clinical practice.

BP measurements may be affected by multiple factors associated with the patient, the technician, or the device, resulting in systematic or random deviations from the true BP value [176], [208]. BP overestimation may arise if measurement protocols are not properly executed, including neglecting to adhere to guidelines for repeated readings over many visits or relying solely on a single BP measurement instead of averaging multiple readings as recommended in clinical guidelines [51], [221]. Conversely, underestimation may occur due to factors such as the use of an oversized cuff, incorrect deflation rate, or observer errors such as hearing issues[51]. Random errors can also arise from inconsistencies in measurement techniques or terminal digit preference [222], [223]. While systematic and random errors can be reduced by following standardized protocols such as, employing trained professionals, using validated equipment, and averaging multiple readings, this analysis aimed to assess the influence of errors that persist within the accepted clinical range, on risk classification and patient outcomes.

Various studies have demonstrated that BP measured in routine clinical practice frequently differs from readings taken by following a standardised protocol[167]. For instance, BP readings in outpatient settings are significantly elevated compared to those recorded under controlled settings. Research comparing clinical and standardised measurements revealed that SBP was, on average

12.7 mmHg higher, and DBP was 12.0 mmHg higher when a specified protocol, such as that employed in the Systolic Blood Pressure Intervention Trial (SPRINT), was not adhered to [224]. These findings highlight the importance of adhering to precise measurement protocols, as errors may result in significant misclassification of high BP status, particularly around critical thresholds. This reinforces the idea that both systematic and random errors—originating from the patient, the clinician, or the device—can significantly affect BP readings and, consequently, risk classification and patient management. Future research should also examine the additional sources of error to provide a comprehensive understanding of the challenges in accurate BP measurement and improve clinical outcomes. This study's findings, focussing on device related errors, contribute significant insights to the current knowledge base. Although device accuracy is not the sole determinant in clinical decision making, it is a significant factor in ensuring reliable CVD risk assessments.

This research evaluated a new Cox PH model after adjusting BP measurement errors to evaluate not only the changes in risk classification but also the impact of errors on the overall model accuracy and estimates. The original model was built using BP data that was assumed to include measurement errors. The coefficient from the original model inherently accommodates those errors. Therefore, if the original model estimates are used for the adjusted BP measurements, it could still reflect the bias from the original model, leading to unreliable predictions. Hence, development of a new model allows for a nuanced analysis of how adjusting for measurement errors affects the predictive accuracy and validity of the risk model, a critical consideration in real-world clinical scenarios where precision is paramount.

The most accurate method for measuring BP is the invasive technique, which involves direct arterial measurement. However, due to the associated risks and invasiveness, it is not commonly used in clinical practice. The manual auscultatory method, a non-invasive technique, is regarded as the gold standard for BP monitoring and serves as the benchmark for evaluating the accuracy of other devices. Aneroid manometers, also extensively used, require frequent calibration to maintain precision [50]. Although these methods are commonly employed, it is not entirely devoid of error, indicating that all non-invasive BP measurement devices entail a certain level of inaccuracy. These inaccuracies are often presumed to conform to a normal distribution, as specified by international

standards. In this study, these inherent inaccuracies have been accounted for by assuming that all BP readings in the dataset contain a degree of measurement error. However, due to the lack of detailed data on the specific devices used to record BP, this assumption has been uniformly applied across all readings. This approach enables us to assess the overall effect of device-related errors on CVD risk classification, providing a broader perspective on how device inaccuracies may influence clinical outcomes.

Figure 4 in this research illustrates a distinct inverse relationship between the increasing SD of device errors and the decreasing classification accuracy. Notably, even within the internationally specified acceptable range of SD, a degree of misclassification persists. This observation prompts an important discussion regarding whether the existing allowable range for device inaccuracies is sufficiently stringent for clinical studies, especially those involving CVD risk prediction. It may be necessary to reevaluate or restrict these limits, ensuring that a smaller SD is required to get a classification accuracy level considered suitable for clinical decision-making. Moreover, healthcare professionals and researchers working with such cohorts should engage in discussions about what constitutes an acceptable level of classification accuracy, weighing the trade-off between device error tolerance and clinical outcomes.

CHAPTER 5. IMPACT OF DEVICE RELATED INACCURACIES ON CVD RISK PREDICTION ACROSS EUROPEAN, MĀORI AND PACIFIC GROUPS (Manuscript 3).

5.1 Prelude

Chapter 2 highlights review of the two primary sources of Blood Pressure (BP) measurement error: rounding to the nearest zero-end digit and inaccuracies inherent in BP measurement devices.

Chapters 4 and 5 analysed the impact of these errors on cardiovascular disease (CVD) risk classification. Although model estimates did not show significant differences, notable misclassification was observed, leading to potential healthcare burden, including increased costs and inappropriate treatment decisions. While rounding errors occur in only a subset of measurements, device inaccuracies affect all BP readings, even within internationally accepted limits. The findings indicated that device inaccuracies resulted in higher misclassification rates compared to rounding errors, highlighting the importance of ensuring BP measurement reliability.

Previous analysis focused on the overall cohort; however, CVD risk varies across ethnic groups. Māori and Pacific populations in New Zealand face a disproportionate burden of CVD due to higher prevalence rates of risk factors. This chapter investigates how BP measurement inaccuracies influence CVD risk classification for these populations compared to Europeans. The analysis begins with an exploration of key risk factors with the highest hazard ratios (HR), followed by an assessment of how misclassification occurs after error adjustment. This approach highlights potential disparities in risk categorisation and the role of measurement inaccuracies in increasing health inequities.

5.2 Introduction

Cardiovascular diseases (CVD) and diabetes constitute almost 17% of health loss in New Zealand (NZ), a globally recognised measure of reduced healthy life due to premature death, illness, or disability [225] where the economic burden tends to heavily affect specific demographic groups due to differences in socioeconomic, behavioural, and cultural risk factors [226], [227], [228]. Even so, a significant decrease in the incidence and mortality of CVD has been observed in NZ over the past

two decades owing to effective prevention and access to treatment [225], [229]. However, despite this consistent reduction in the overall incidence and case-fatality rates of CVD since the 1980s [230], some population groups, such as Pacific and Māori in NZ, continue to face an unevenly high CVD burden [231]. The prevalence of CVD is disproportionately distributed across ethnic groups, with Māori and Pacific individuals experiencing a greater burden [232], [233], [234], [235]. Compared to NZ Europeans, who make up about 70% of the population, Māori and Pacific individuals are more likely to exhibit CVD risk factors such as smoking, high blood pressure, diabetes, history of atrial fibrillation and obesity [28], [236], [237], [238]. Furthermore, Māori and Pacific groups have reported significantly higher use of antihypertensives than other ethnic groups [239]. These factors have importantly contributed to a higher likelihood of developing and dying from CVD in these groups [240], [241].

An experimental investigation using general practice data revealed higher rates of obesity among Pacific and Māori people in comparison to people of European/other and Asian ethnicities [242]. Research assessing the impact of smoking-related fatalities revealed that approximately 22.6% of deaths among Māori and 13.8% among Pacific people were linked to smoking [243] that was found related to differences in socioeconomic deprivation [244]. Further research indicates that the average systolic blood pressure (SBP) among individuals in NZ is rising due to various causes, including elevated obesity rates and lifestyle choices [239]. Diabetes is also an important concern, the rates of which are anticipated to increase substantially in NZ, placing strain on the healthcare system [245]. A study indicated that Pacific individuals across all age demographics demonstrated a higher likelihood of diabetes diagnosis compared to their European counterparts [246], [247]. A recent primary care study highlighted a higher prevalence of atrial fibrillation among Māori and Pacific people aged ≥ 45 years (4.2%) compared to Europeans of the same age group (3.8%) [248]. Additionally, there was a significant association between ethnicity and the use of antihypertensive medications ($p<0.001$), with Māori and Pacific peoples being the most frequent users [239]. This likely reflects that individuals on antihypertensive medications generally have a higher CVD risk profile.

Precise diagnosis and control of such risk factors plays an important role towards mitigating CVD risk and enhancing health outcomes [249]. Risk prediction equations, such as the PREDICT-1 equation developed from the NZ-based PREDICT cohort, assess an individual's likelihood of developing CVD [81]. The cohort involved a web-based CVD risk assessment and management programme developed to support CVD risk assessment and management in routine primary care practice [250]. The PREDICT-1 equations assign hazard ratios (HRs) to quantify the relative impact of each factor on CVD risk. The study findings indicate that Māori men and Pacific men had a 34% and 19% higher chance of getting a CVD respectively compared to Europeans. Similarly, Māori and Pacific women had a 48% and 22% higher chance of getting a CVD respectively compared to Europeans. Table 21 summarises the risk factors with their corresponding HRs to highlight those with the highest influence.

Among the remaining risk factors, SBP serves as a fundamental parameter for CVD risk assessment [171]. It has been invariably found to be a significant predictor in CVD risk prediction equations globally. Even so, errors in BP measurements due to inaccuracies in BP measurement devices or other factors such as observer bias and improper cuff size [180] have the ability to introduce bias into risk prediction models.

Table 21: Hazard Ratios (HRs) for key CVD risk factors stratified by gender [81].

Risk Factor	**Men**	**Women**
Diabetes	1.51	1.64
Currently Smoking	1.70	1.97
History of Atrial Fibrillation	1.87	2.53
On BP lowering medication	1.33	1.42
Body Mass Index (BMI)		
Underweight (<18.5 kg/m2)	1.73	1.87
Obesity class 3 (≥40.0 kg/m2)	1.46	1.38

The immediate effect is the misclassification of individuals into incorrect risk categories, potentially and negatively influencing treatment decisions and health outcomes. It's been predicted that misclassifications due to errors in BP measurement may result in a healthcare burden of around 8.2 million NZD [219]. Previous chapters have examined the impact of both rounding and device errors on CVD risk prediction, demonstrating that measurement inaccuracies can influence risk classification. The studies used the PREDICT-1 equation to classify patients into three risk categories (<5%, 5-15%, and >15%) as established by the NZ Ministry of Health [53], serving as guidelines to clinicians in recommending appropriate interventions, setting goals, and planning follow-up care based on the assessed cardiovascular risk. While both types of errors contribute to misclassification, device errors introduce greater variability due to the allowable inaccuracy range, which can reach up to 6.95 mmHg SD, exceeding the maximum deviation of ±5 mmHg seen with rounding. This suggests that device errors, particularly at their upper bound, have the most significant impact on risk misclassification.

Building on this, the current study aims to understand how these device errors affect risk prediction across different ethnic groups in NZ. This gap in knowledge is particularly critical for Māori and Pacific populations in NZ, who already bear a disproportionate burden of CVD attributed to elevated rates of obesity, smoking, and other social determinants of health. Furthermore, inaccuracies in BP measurements, even within internationally acceptable ranges, may aggravate existing health inequities by affecting risk assessments and undermining the validity of risk prediction models for these vulnerable groups.

The primary aim of this study is to examine the trends from significant, yet meaningful, risk factors identified by the PREDICT-1 equation for Māori, Pacific, and European. The analysis particularly considers men and women separately. In addition, the study evaluates the impact of BP device inaccuracies on the classification of these ethnic groups into CVD risk categories. Specifically, the analysis will focus on investigating the extent to which these inaccuracies contribute to the misclassification of individuals across different risk groups. By examining the overall shifts in risk classification for Māori, Pacific, and European populations, this research seeks to provide an

understanding of how device inaccuracies influence CVD risk predictions, potentially for ethnic groups that are at a higher risk.

5.3 Method

5.3.1 Data and Data Sources

In this study, participants were enrolled in the cohort during their initial CVD risk assessment, performed by primary care professionals. The health professionals documented health data in the PREDICT software, widely used to compute CVD risk profiles using risk factors including age, gender, smoking status, diabetes, blood pressure, BMI, and cholesterol levels, along with medical history (including previous cardiovascular disease and atrial fibrillation). This information is securely maintained in the practice management system and in an anonymised central database. With appropriate permissions, risk profiles from the central database were linked to a unique National Health Index number, facilitating anonymised integration with national and regional health data for additional research.

Approximately 95% of the country's population is registered with primary health organisations that provide most primary health care services across the country. About one-third of the population gets care from clinics employing the PREDICT software, predominantly situated in the Auckland and Northland regions, which include both, extensive urban and rural populations, offering a diverse representation of NZ's varying socioeconomic and ethnic demographics [38]. The participating practices encompass the largest Māori, Pacific, Indian, Chinese, and other Asian demographics in the country, ensuring a broad and representative study sample.

It's important to note that ethnicity was self-reported in the primary health management system and corroborated with public health organisation enrolment and hospitalisation data. When individuals identified with various ethnicities, a prioritisation process was employed to designate them to a singular ethnic group. This method adheres to national ethnicity data regulations and prioritised ethnic groupings in the subsequent order, as follows: Māori, Pacific, Indian (including Fijian Indian), Other Asian (including Chinese), European, Middle Eastern/Latin American/African (MELAA), Other, and Unknown. Ethnic groups comprising fewer than 1,000 individuals were omitted from the analysis. All participants had comprehensive data on the essential variables necessary for CVD risk

assessment. Participants without data on the New Zealand Index of Socioeconomic Deprivation (NZDep) were omitted.

5.3.2 Study Design

This study included 427,299 men and women aged 30-74 years at the time of their first PREDICT risk assessment. Simulation was performed by generating random errors for SBP based on a normal distribution with specified means and their respective SD, as presented in Table 8 in Section 2.4.2. These random errors reflect the acceptable error limits outlined by the standard [49].

The population is divided into three CVD risk categories: low (<5%), moderate (5-15%), and high (>15%) according to the Ministry of Health [82]. The original SBP data recorded in the dataset, termed SBP_{error}, is presumed to be affected by device inaccuracies. Subsequently, a 5-year CVD risk is calculated using the SBP_{error} values. A Cox PH model is applied to estimate the risk for everyone, using SBP_{error} as one of the key risk factors. This risk score is referred to as the original risk and represents the predicted likelihood of a CVD event within five years, evaluated using Equation (14):

$$\frac{h(t)}{h_0(t)} = e^{(\beta_1 SBP_{error} + \beta_2 z_2 + \ldots + \beta_p z_p).} \quad (14)$$

Here, $h(t)$ represents the hazard at time t, $h_0(t)$ is the baseline hazard, SBP_{error}, z_2 ,…, z_p are individual risk factors and $\beta_1, \ldots, \beta_p$ are their corresponding Cox regression coefficients.

The SBP measurements are then adjusted by removing the simulated errors, resulting in $SBP_{adjusted}$ values. This adjustment is implemented to better determine the true SBP, facilitating the recalculation of the 5-year CVD risk using the adjusted SBP values. The new risk score, adjusted risk, reflects the updated estimate of CVD risk, while other factors remain unchanged. A new Cox PH model (adjusted model) is applied to compute this adjusted risk as expressed in Equation (15):

$$\frac{h(t)}{h_0(t)} = e^{(\beta_1 SBP_{adjusted} + \beta_2 z_2 + \ldots + \beta_p z_p)} \quad (15)$$

The process is performed 10,000 times to minimize the impact of random fluctuations. By comparing the adjusted risk and original risk for each individual, the study specifically examines how errors in SBP measurements impact the classification of individuals from different ethnic groups into various

CVD risk categories, focusing on misclassification rates within each group. The misclassification rates are evaluated by comparing the adjusted and original risk classifications, with the average misclassification rate calculated across all simulations.

This study also examines how risk prediction changes in response to the highest permissible device error, with an SD of 6.95 mmHg. These values represent the maximum allowable error in BP measurements according to international standards. By focusing on this threshold, the study assesses the upper-bound impact of device inaccuracies within the acceptable limits outlined by the standards. The study specifically examines how these errors influence the risk classification across different risk factors. These factors include smoking status, diabetic status, obesity levels (as measured by BMI), history of atrial fibrillation and their status of taking antihypertensives.

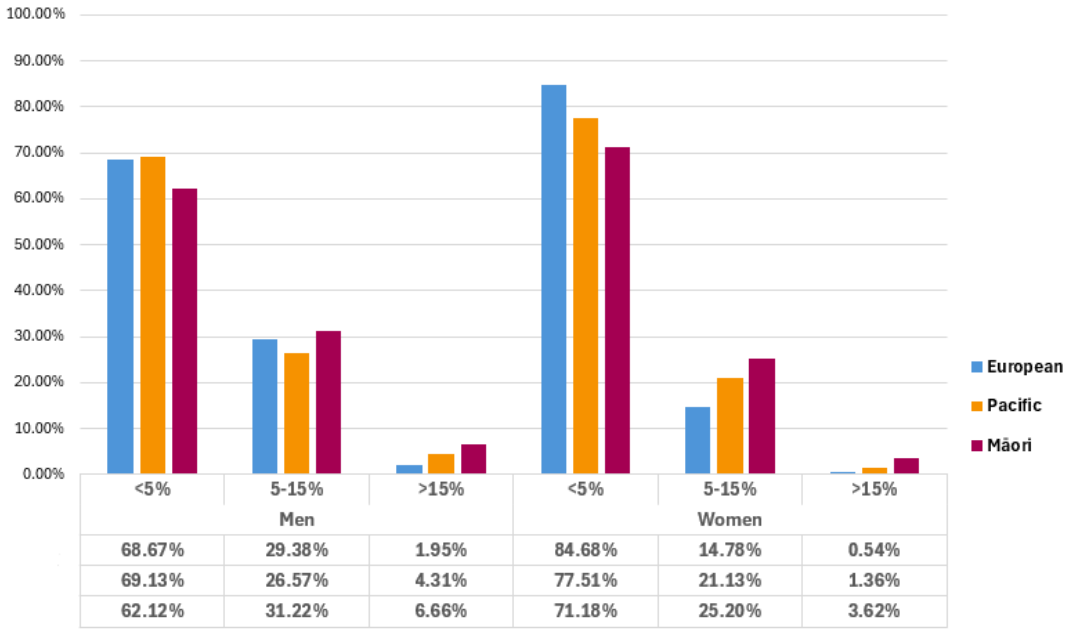
5.4 Results

5.4.1 Exploratory Analysis

The PREDICT cohort dataset comprises 241,036 men with a mean age of 51.29 years (SD = 10.07 years) and 186,263 women with a mean age of 55.85 years (SD = 8.87 years). Table 22 presents the ethnic composition of the cohort, focusing on Māori, Pacific, and European populations. Figure 14 illustrates the distribution of individuals across the three risk groups, stratified by ethnicity. The plot indicates that a higher proportion of Europeans fall into the <5% risk category, whereas Māori and Pacific people are more frequently classified into risk categories above 5%.

Table 22: Ethnic representation of Māori, Pacific, and European populations within the cohort

	European	**Māori**	**Pacific**
Men	138195 (57.08%	31473 (53.70%)	35083 (55.92%)
Women	103922 (42.92%	27141 (46.30%)	27660 (44.08%)



	<5%	5-15%	>15%	<5%	5-15%	>15%
	Men			Women		
European	68.67%	29.38%	1.95%	84.68%	14.78%	0.54%
Pacific	69.13%	26.57%	4.31%	77.51%	21.13%	1.36%
Māori	62.12%	31.22%	6.66%	71.18%	25.20%	3.62%

Figure 14: Proportion of population in three risk groups per ethnicity.

The distribution of SBP by ethnicity (European, Māori, and Pacific) and gender (men and women) is distributed roughly in a bell shape, indicating a normal-like distribution across the populations, also presented in Figure 32 in Appendix C. Table 23 provides the mean SBP values, indicating that Pacific people (both men and women) have a lower mean SBP compared to Māori and European populations.

Table 23: Mean SBP values for European, Māori, and Pacific populations, stratified by men and women.

Ethnicity	**Gender**	**SBP, mmHg (Average)**
European	Men	129.86
European	Women	129.49
Māori	Men	130.14
Māori	Women	128.86
Pacific	Men	127.57
Pacific	Women	128.58

Table 24: Description of Age, BMI range, Diabetes Status and Smoking status in the PREDICT cohort for European, Māori and Pacific, stratified by gender.

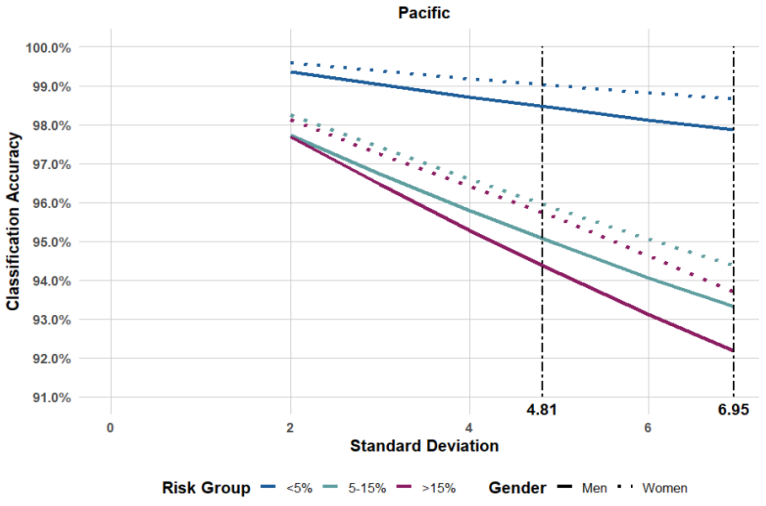
		Men			Women		
		European	**Māori**	**Pacific**	**European**	**Māori**	**Pacific**
Age	**30-34**	1037 (0.75%)	665 (2.11%)	829 (2.36%)	474 (0.46%)	464 (1.71%)	613 (2.22%)
	35-44	10939 (7.92%)	13382 (42.52%)	16989 (48.43%)	4236 (4.08%)	2955 (10.89%)	3426 (12.39%)
	45-54	64215 (46.47%)	10539 (33.29%)	10072 (28.71%)	18667 (17.96%)	15358 (56.59%)	15202 (54.96%)
	55-64	39866 (28.85%)	5175 (16.44%)	5193 (14.80%)	55276 (53.19%)	6260 (23.06%)	5863 (21.20%)
	65-74	22138 (16.02%)	1712 (5.44%)	2000 (5.70%)	25269 (24.32%)	2104 (7.75%)	2556 (9.24%)
	Mean (SD)	54.15 (9.04)	47.07 (9.39)	46.11 (9.76)	58.61 (7.94)	51.14 (8.19)	50.97 (8.75)
BMI	**Less than 18.5**	367 (0.27%)	65 (0.21%)	51 (0.15%)	1111 (1.07%)	189 (0.70%)	50 (0.18%)
	18.5 - <25	24110 (17.45%)	2979 (9.47%)	2335 (6.66%)	28156 (27.09%)	4069 (14.99%)	1731 (6.26%)
	25 - <30	52156 (37.74%)	8419 (26.75%)	7833 (22.33%)	27218 (26.19%)	6319 (23.28%)	4640 (16.78%)
	30 - <35	24507 (17.73%)	8300 (26.37%)	10629 (30.30%)	14399 (13.86%)	5759 (21.22%)	6563 (23.73%)
	35 - <40	6498 (4.70%)	4240 (13.47%)	6400 (18.24%)	6296 (6.06%)	3694 (13.61%)	5791 (20.94%)
	40 or more	2333 (1.69%)	3127 (9.94%)	4395 (12.53%)	3665 (3.53%)	3578 (13.18%)	6352 (22.96%)
	NA	28224 (20.42%)	4343 (13.80%)	3440 (9.81%)	23077 (22.21%)	3533 (13.02%)	2533 (9.16%)
	Mean (SD)	28.36 (4.74)	32.07 (6.89)	33.34 (6.87)	27.88 (6.16)	32.11 (8.02)	35.53 (8.15)
Diabetes	**Non-diabetic**	129328 (93.58%)	27644 (87.83%)	29238 (83.34%)	97364 (93.69%)	23297 (85.84%)	20235 (73.16%)
	Diabetic	8867 (6.42%)	3829 (12.17%)	5845 (16.66%)	6558 (6.31%)	3844 (14.16%)	7425 (26.84%)
Smoking Status	**Non-smoker**	121787 (88.13%)	20647 (65.60%)	25883 (73.78%)	94088 (90.54%)	17481 (64.41%)	23601 (85.33%)
	Smoker	16408 (11.87%)	10826 (34.40%)	9200 (26.22%)	9834 (9.46%)	9660 (35.59%)	4059 (14.67%)
History of Atrial Fibrillation	**No**	135124 (97.78%)	30968 (98.4%)	34766 (99.1%)	102577 (98.71%)	26829 (98.85%)	27466 (99.3%)
	Yes	3071 (2.22%)	505 (1.6%)	317 (0.9%)	1345 (1.29%)	312 (1.15%)	194 (0.7%)
On BP lowering medication	**No**	109495 (79.23%)	25256 (80.25%)	28069 (80%)	75945 (73.08%)	19997 (73.68%)	18884 (68.27%)
	Yes	28700 (20.77%)	6217 (19.75%)	7014 (20%)	27977 (26.92%)	7144 (26.32%)	8776 (31.73%)

Table 24 presents the description of all the risk factors considered in this study, stratified as follows: Europeans, Māori and Pacific. The Māori and Pacific population have a larger proportion of younger people. The average BMI is relatively higher in Māori and Pacific groups. The highest prevalence of diabetes is observed among Pacific women, with 26.84% of the population diagnosed with diabetes. This is followed by Pacific men at 16.66%, that is, almost three times more likely compared to European men. Within each ethnic group, individuals diagnosed with diabetes is generally higher for women, especially in the Pacific demographic. The Māori population has a notably higher diabetes prevalence compared to the European population. 12.17% of Māori men and 14.16% of Māori women are diagnosed with diabetes, which is higher than the European male (6.42%) and female (6.31%) population. A greater proportion of the European demographic have a history of Atrial Fibrillation compared to Māori and Pacific. The higher prevalence of Atrial Fibrillation among NZ Europeans likely reflects their older age distribution compared to Māori and Pacific populations, as the condition is more common in older individuals [251]. A higher proportion of women are on BP-lowering medication compared to men, with Pacific women exhibiting the highest prevalence at 31.73%. Māori people have a highest smoking rate followed by Pacific. Compared to Europeans, Māori individuals are almost three times more likely to be smokers. Māori women have a slightly higher rate of smoking compared to Māori men. On the contrary, Pacific men have a higher prevalence of being currently smoking than Pacific women.

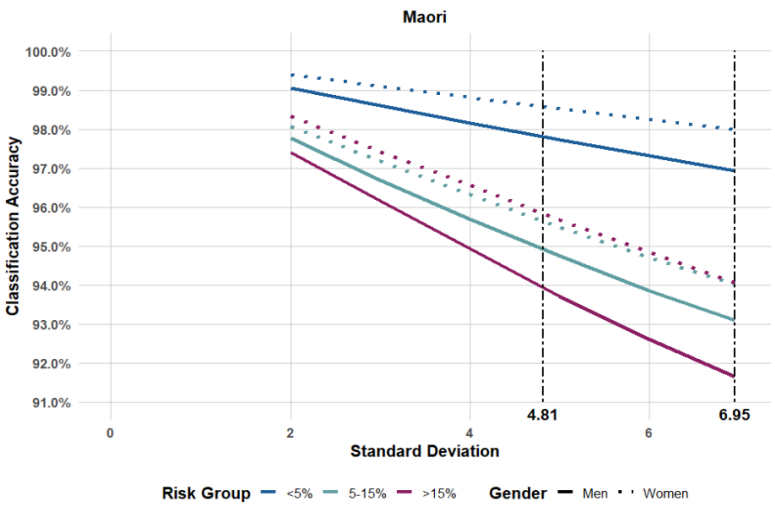
5.4.2 Comparative Modelling

Figure 15 shows the percentage of accurate classifications for various standard deviations from up to the acceptable range as mentioned by the standards for a BP measurement device. This is used to compare the impact on risk group classification accuracy for each ethnicity, stratified by risk group (<5%, 5-15%, >15%) and gender. As SD increases, classification accuracy decreases for both men and women, with women having slightly higher accuracy across all SD ranges. Ethnic differences in classification accuracies are evident, with Pacific and Māori populations having lower baseline accuracy, especially in the >15% risk group, compared to Europeans. Europeans show higher accuracy and more gradual declines. The lowest accuracy is observed in Māori's which is around

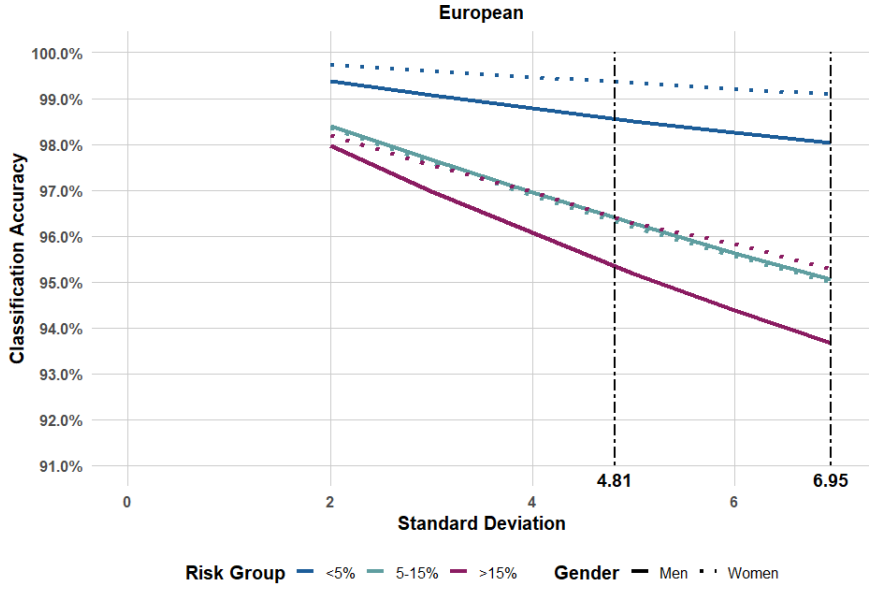
91.5%. The vertical dotted lines mark the allowable SD range for device errors, within which a clear reduction in classification accuracy is still observed.



(a)



(b)



(c)

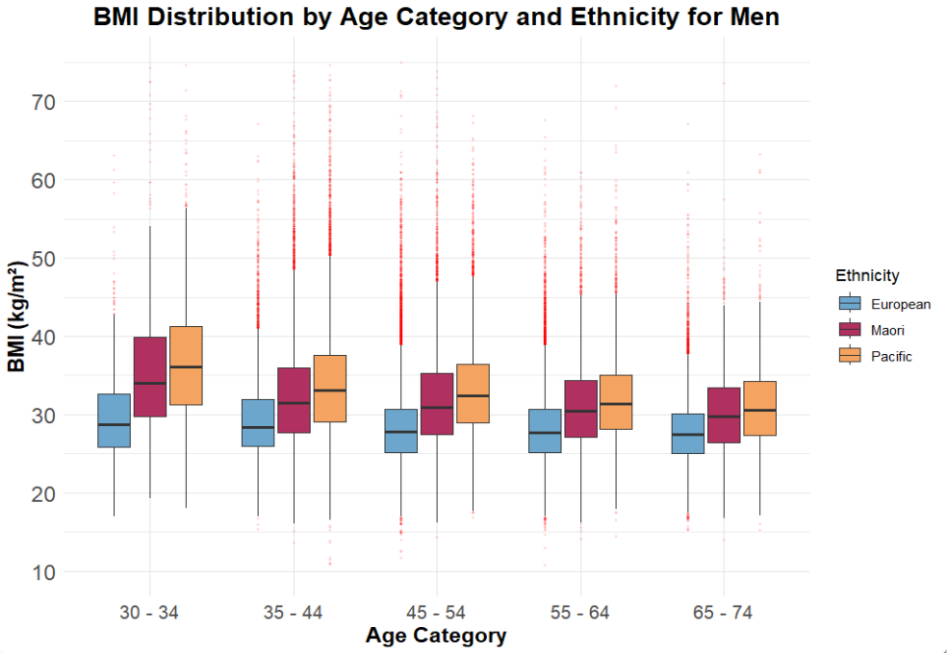
Figure 15: Accurate Risk classification after accounting for the normal errors with a range of standard deviation for (a) Māori, (b) Pacific and (c) European.

5.4.3 Risk Adjustment by Key Indicators

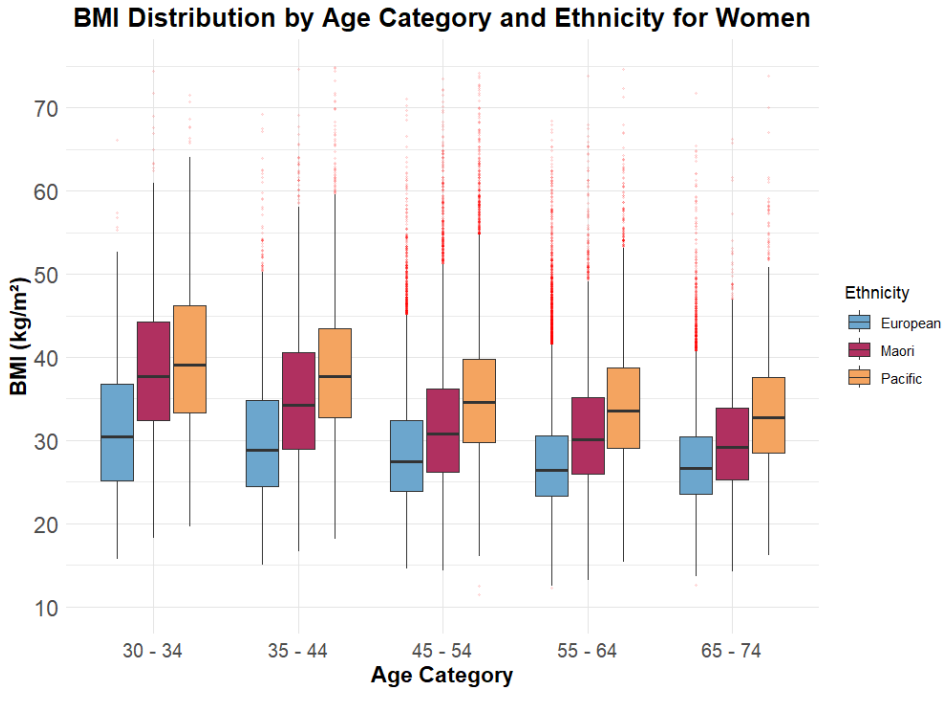
The PREDICT-1 equation uses multiple attributes for estimating CVD risk. In this research, the attributes or risk factors with highest HRs are observed for changes in risk category assignment when the SBP is adjusted for simulated errors.

BMI

BMI is distributed across age groups in the cohort for European, Māori and Pacific population as depicted in Figure 16. Pacific men tend to exhibit the highest BMI across all age categories, followed by Māori and then European men. A similar pattern is observed for women, where Pacific women have the highest BMI. The median BMI for women is slightly higher than men for all ethnic groups. Even the mean BMI is higher for Pacific men (33.34 kg/m2) and Pacific women (35.53 kg/m2).



(a)



(b)

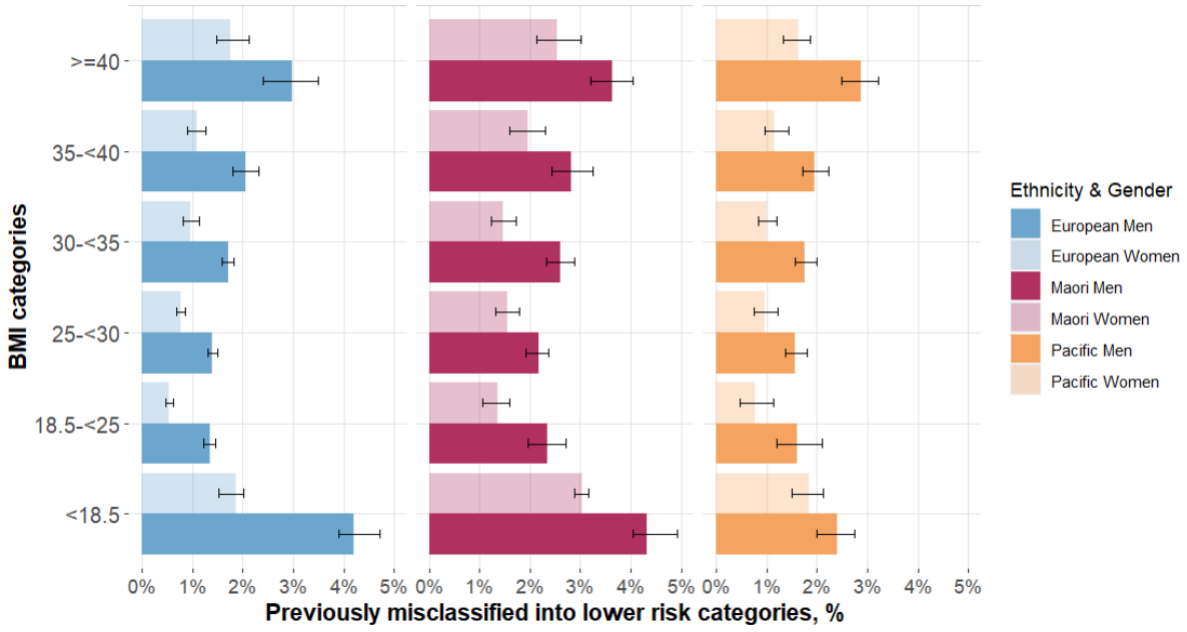
Figure 16: Variation of BMI by age category and Ethnicity for (a) Men and (b) Women

The correlation of BMI with age for each combination of ethnicity and gender entails that the relationships are not strong, with the weakest correlation for European men (-0.0718) and the strongest for Pacific women (-0.172), indicating a modest inverse association across groups, also presented in Table 25.

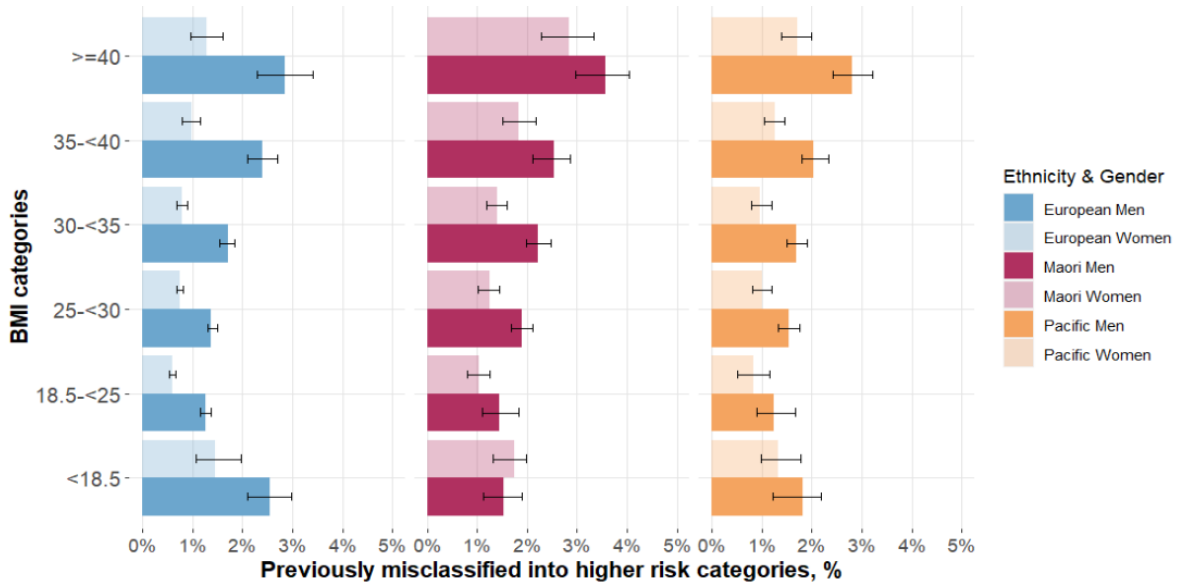
Table 25: Correlation of BMI with age for ethnicity and gender

Ethnicity	Gender	Correlation (Age, BMI)
European	Women	-0.102
European	Men	-0.0718
Māori	Women	-0.165
Māori	Men	-0.111
Pacific	Women	-0.172
Pacific	Men	-0.146

Analysis for BMI categories, as illustrated in Figure 17, reveals that the error adjustment of the recorded BP results in a higher prevalence of people being misclassified into lower risk categories. This effect is particularly notable in individuals categorized as underweight (BMI <18.5 kg/m²) and those with a BMI ≥ 40 kg/m², across both genders. However, no statistically significant gender differences were observed in the patterns of changes in 5-year risk predictions for any of the assessed risk factors.



(a)



(b)

Figure 17: Impact of error adjustment (SD = 6.95) on risk classification across ethnic groups and gender, highlighting changes in risk group assignment based on BMI categories.

Diabetes

Pacific men and women consistently exhibit the highest proportions of diabetes in all age categories compared to Māori and European populations, as shown in Figure 18. Among Māori, a distinct trend is observed where younger women (particularly those aged 30–44 years) have a notably higher proportion of diabetes compared to their male counterparts. However, this trend reverses in older age groups, with Māori men aged 65–74 years showing a higher proportion of diabetes. For European men and women, the proportion of diabetes is the lowest across all age groups.

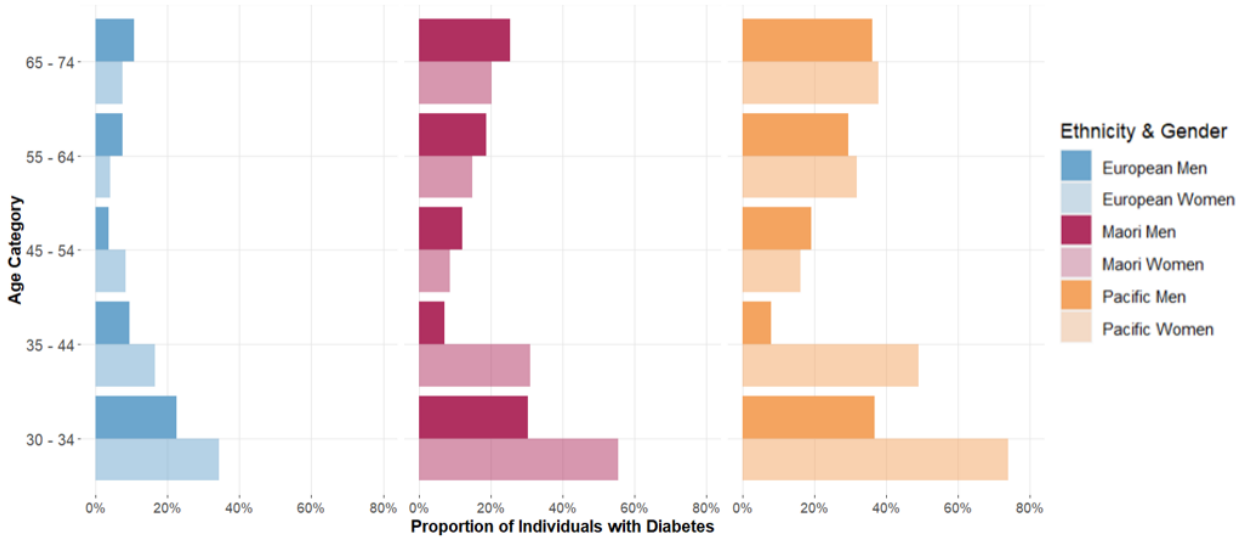


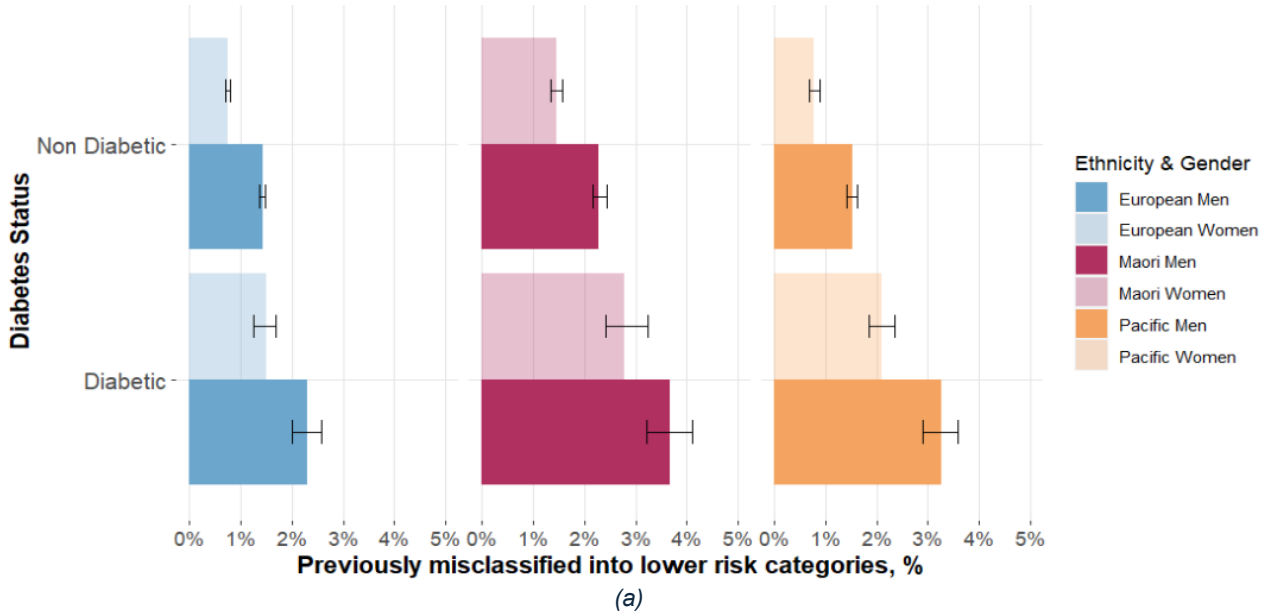
Figure 18: Proportion of individuals with diabetes for various ethnicity, age groups and gender

Overall, a general pattern is seen across ethnicities where diabetes prevalence is relatively higher among younger populations (30–44 years), decreases slightly in middle and then increases in older age groups for all ethnicities. There is no general trend for diabetes with age. Pacific and Māori males have a weak positive correlation with age as shown in Table 26. For remaining combinations of ethnicities and genders negative correlation can be observed.

Table 26: Correlation of diabetes with age for ethnicity and gender

Ethnicity	Gender	Correlation (Age, Diabetes)
European	Women	-0.778
European	Men	-0.345
Māori	Women	-0.631
Māori	Men	0.333
Pacific	Women	-0.529
Pacific	Men	0.441

Figure 19 illustrates the changes in risk classification following the adjustment of BP error (SD = 6.95), stratified by diabetic status (diabetic vs. non-diabetic) across ethnic groups and gender. The results indicate that the 5-year risk prediction changes more substantially, both increasing and decreasing, for diabetic individuals compared to non-diabetic individuals across all ethnic groups. Among all three ethnic groups, Māori men experience the highest risk reclassification, with around



(a)

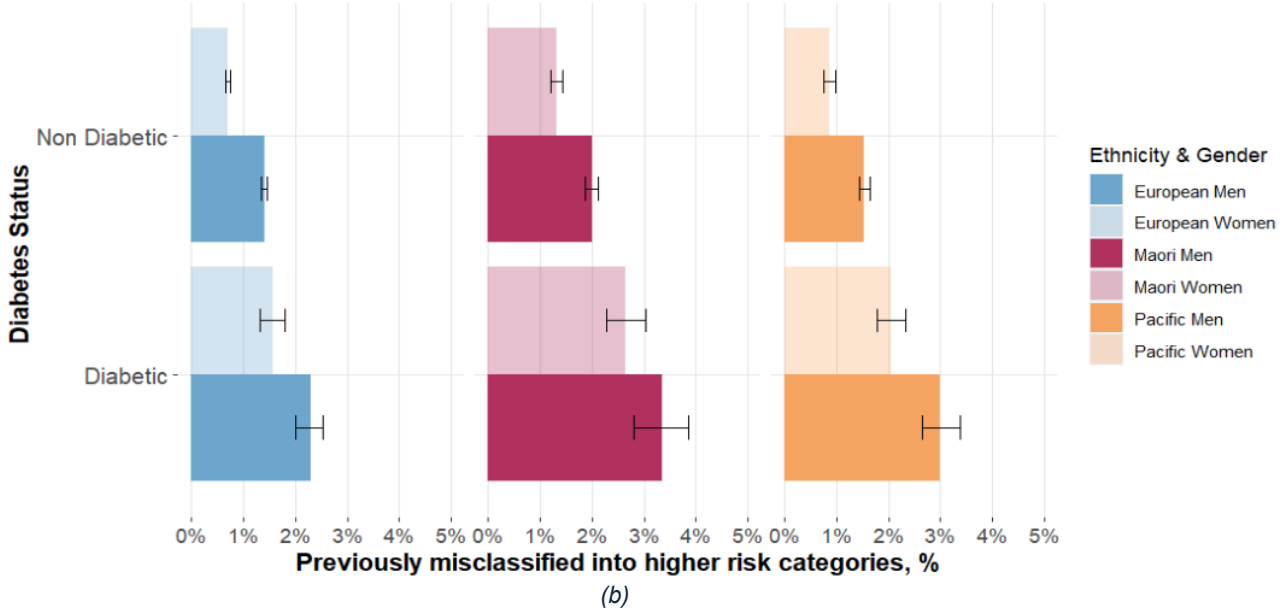


Figure 19: Impact of error adjustment (SD = 6.95) on risk classification across ethnic groups and gender, highlighting changes in risk group assignment based on diabetic status.

3.67% (3.21%-4.10%) of them moving to a higher risk group and 3.35% (2.81%-3.85%) moving to a lower risk group after error adjustment.

Smoking Status

Figure 20 highlights ethnic and gender disparities in smoking prevalence across age groups. Māori have the highest proportion of smokers, with women smoking more than men, although the difference between genders is not significant across age groups. In contrast, Pacific men have a higher proportion of smokers compared to Pacific women. Europeans have the lowest proportion of smokers among both men and women. Overall, smoking prevalence tends to decrease with age, with the highest proportions observed in younger age groups across all ethnicities.

The decrease in smoking habits with age is enumerated in Table 27. There exists a strong negative correlation with age for all ethnicities and genders indicating inverse proportionality.

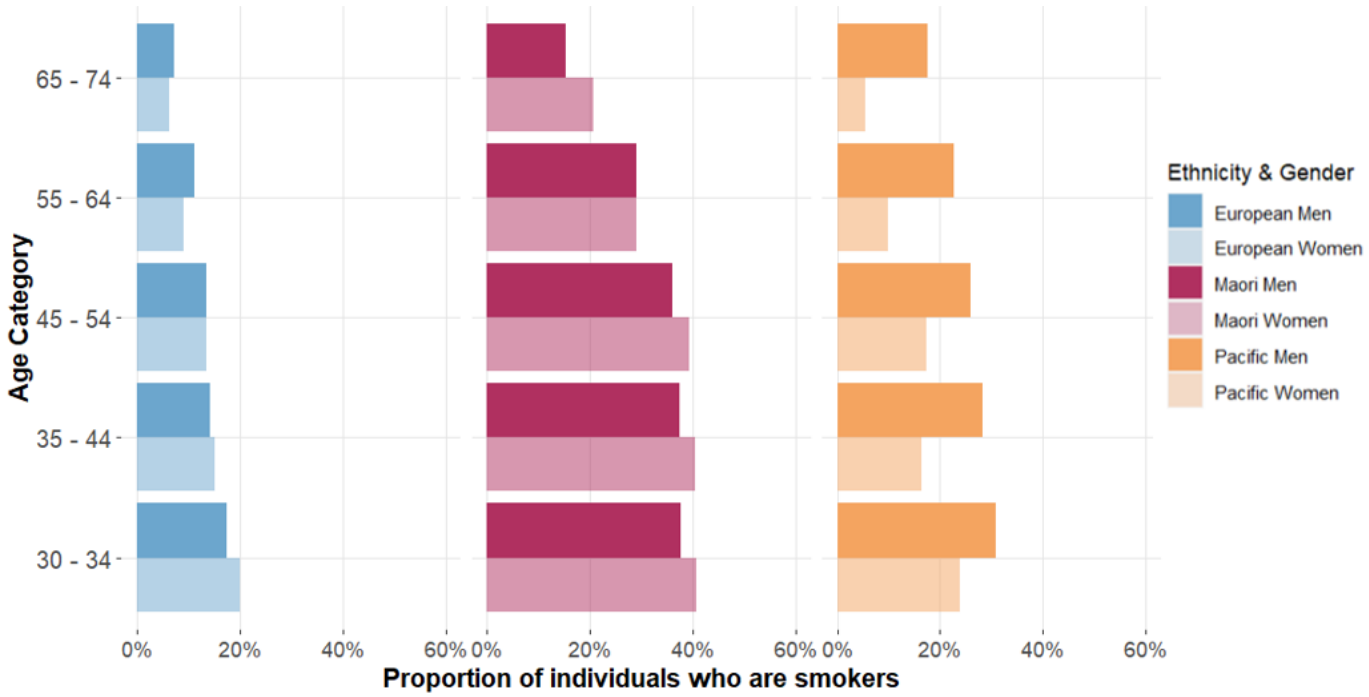
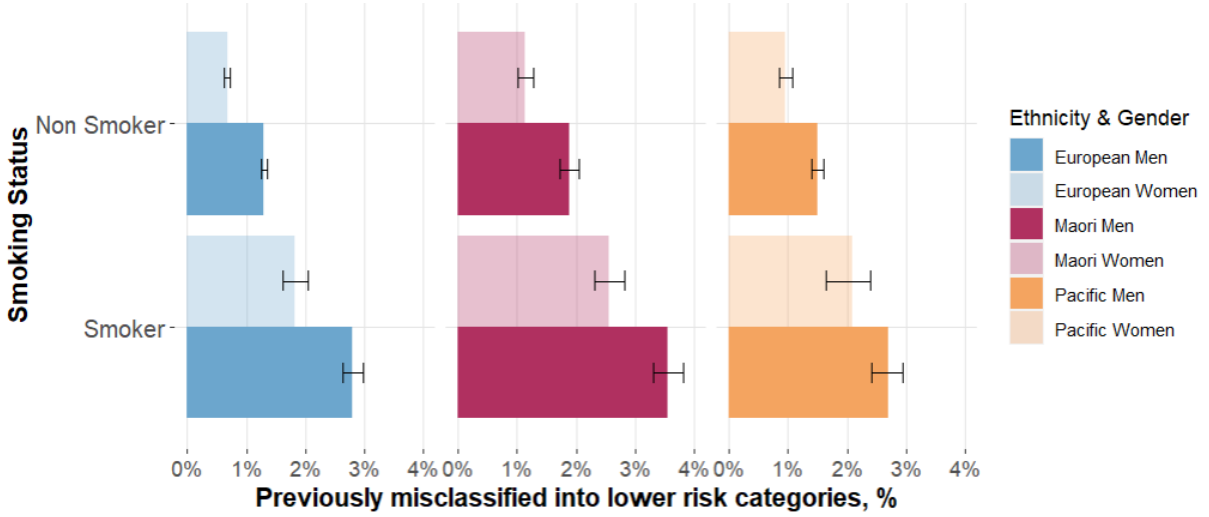


Figure 20: Proportion of individuals who are smokers for various ethnicity, age groups and gender

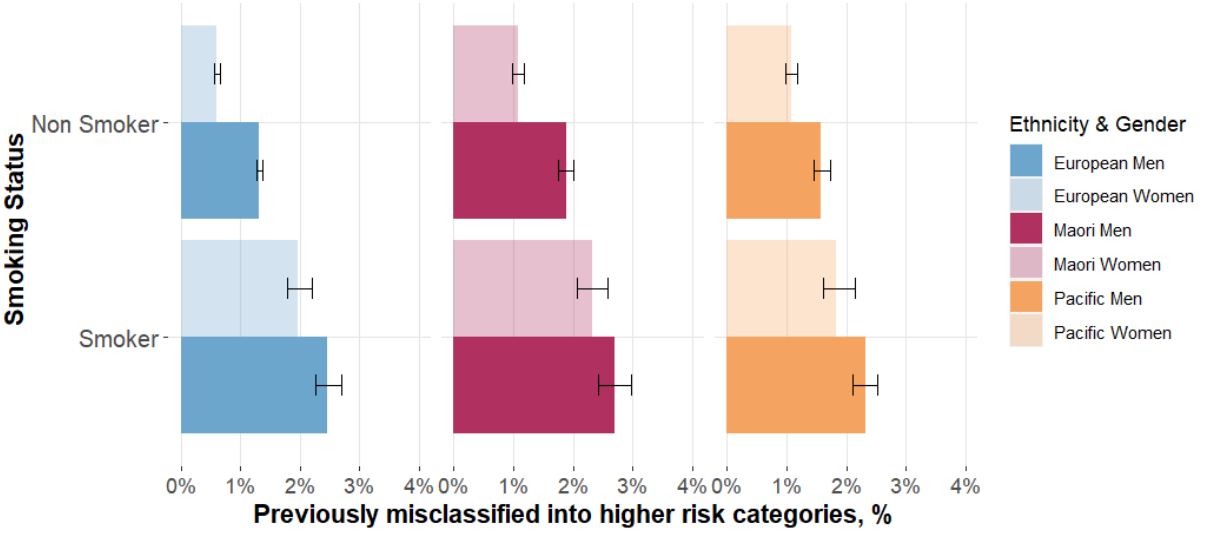
Table 27: Correlation of smoking status (currently) with age for ethnicity and gender.

Ethnicity	Gender	Correlation (Age, Smoking)
European	Women	-0.936
European	Men	-0.943
Māori	Women	-0.919
Māori	Men	-0.883
Pacific	Women	-0.941
Pacific	Men	-0.922

Figure 21 shows the estimated impact of error adjustment on the risk group classification based on smoking status. Among ethnic groups, Māori individuals exhibit the most notable change in risk classification, with around 3.55% (3.31%-3.82%) previously misclassified into lower risk group and 2.70% (2.42%-2.97%) misclassified into higher risk group, followed by Pacific and European groups. These changes on the effects are consistently more pronounced in men than women across all ethnicities.



(a)



(b)

Figure 21: Impact of error adjustment (SD = 6.95) on risk classification across ethnic groups and gender, highlighting changes in risk group classification based on smoking status.

Usage of BP lowering medication

Figure 22 illustrates the trends in the use of anti-hypertensive medications by ethnicity and gender across age groups. The use of antihypertensive medications notably increases with age, with women exhibiting a higher proportion of usage compared to men. Among the three ethnicities, Pacific men and women have the highest proportion of individuals on antihypertensive medication.

BP usually increases with age [252], hence the use of medication is also expected to increase. Table 28 shows a strong positive correlation between usage and age with European females being the highest at 0.91.

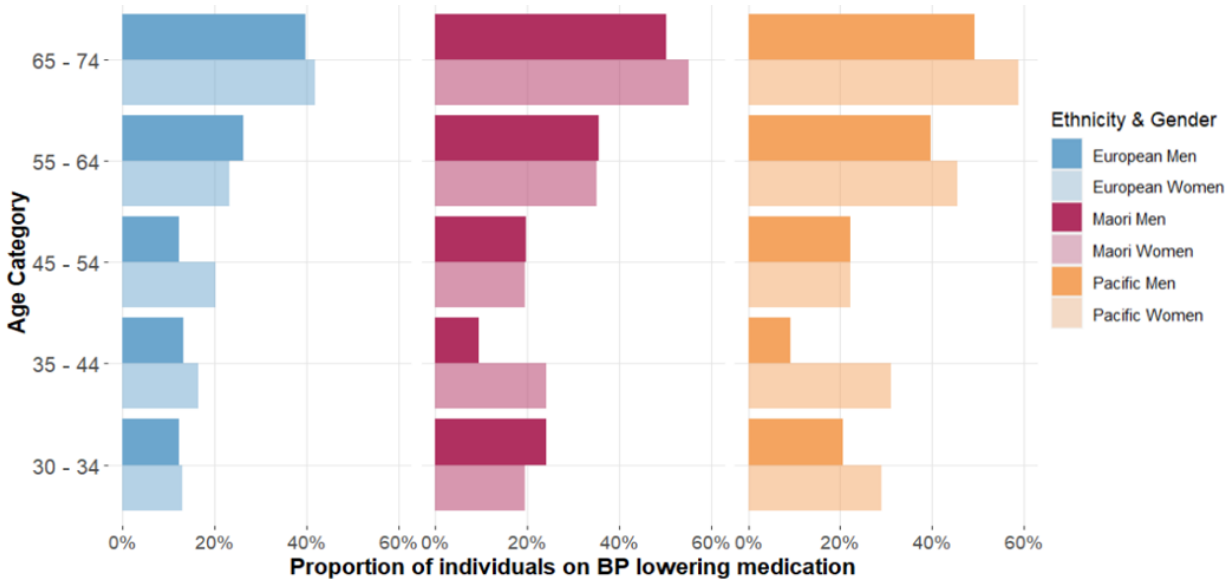
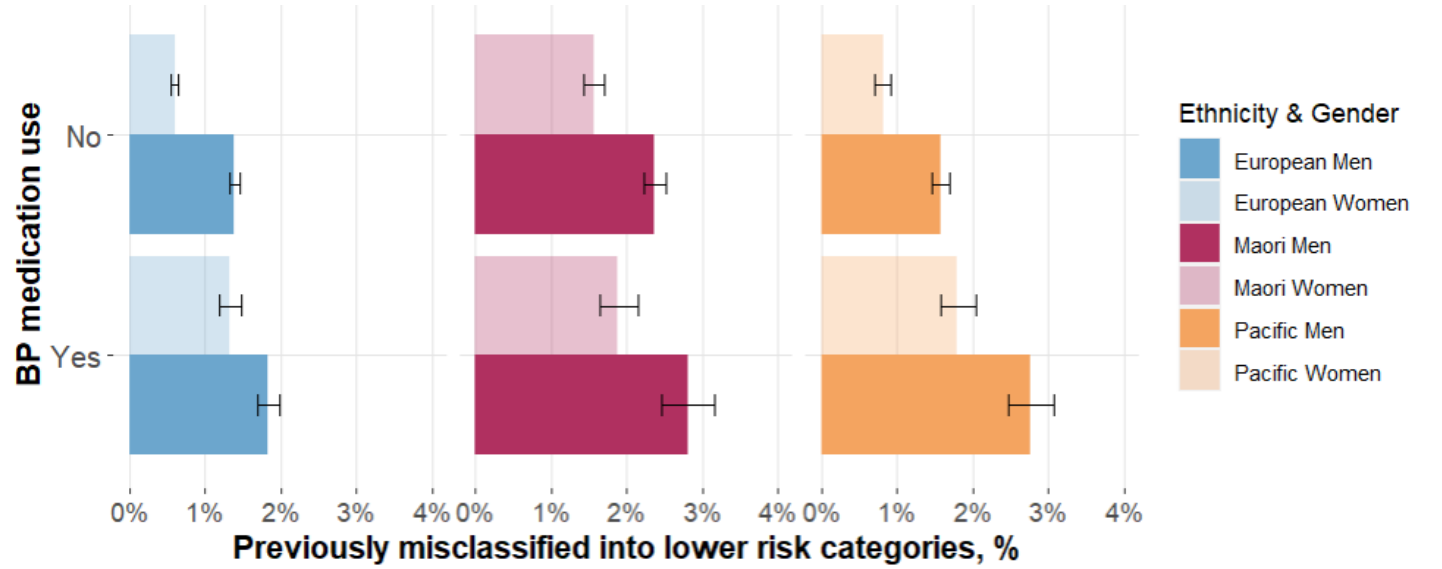


Figure 22: Proportion of individuals with on BP lowering medication for various ethnicity, age groups and gender

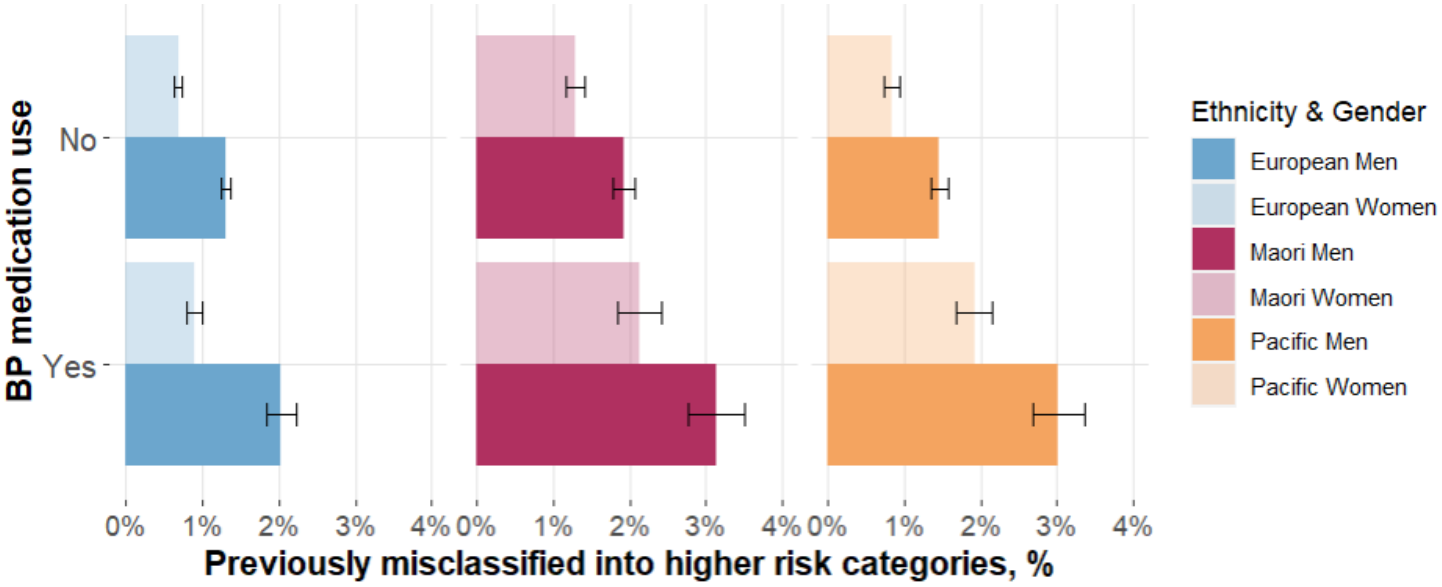
Table 28: Correlation of usage of BP lowering medication with age for ethnicity and gender.

Ethnicity	Gender	Correlation (Age, Uses BP Med)
European	Women	0.917
European	Men	0.908
Māori	Women	0.895
Māori	Men	0.862
Pacific	Women	0.838
Pacific	Men	0.906

Figure 23 illustrates the impact of error adjustments on the changes in the risk group classification based on the status of using anti-hypertensives. The analysis reveals greater shifts in risk classification among Māori and Pacific populations compared to Europeans. However, contrary to the shift observed in the previous risk factors, there is a higher proportion of shift in the risk reclassification to a lower risk category. Māori men on BP-lowering medications have the highest misclassification rates, with 2.82% placed in lower risk categories and 3.13% in higher risk categories.



(a)



(b)

Figure 23: Impact of error adjustment (SD = 6.95) on risk classification across ethnic groups and gender, highlighting changes in risk group assignment based on BP lowering medication status.

Atrial Fibrillation

Figure 24 highlights the ethnic and gender disparities in the history (incidence) of atrial fibrillation across the age groups. The proportion of individuals with a history of atrial fibrillation increases with age with the highest proportion observed in Māori individuals and then Europeans. Atrial fibrillation increases with age across all ethnicities and genders.

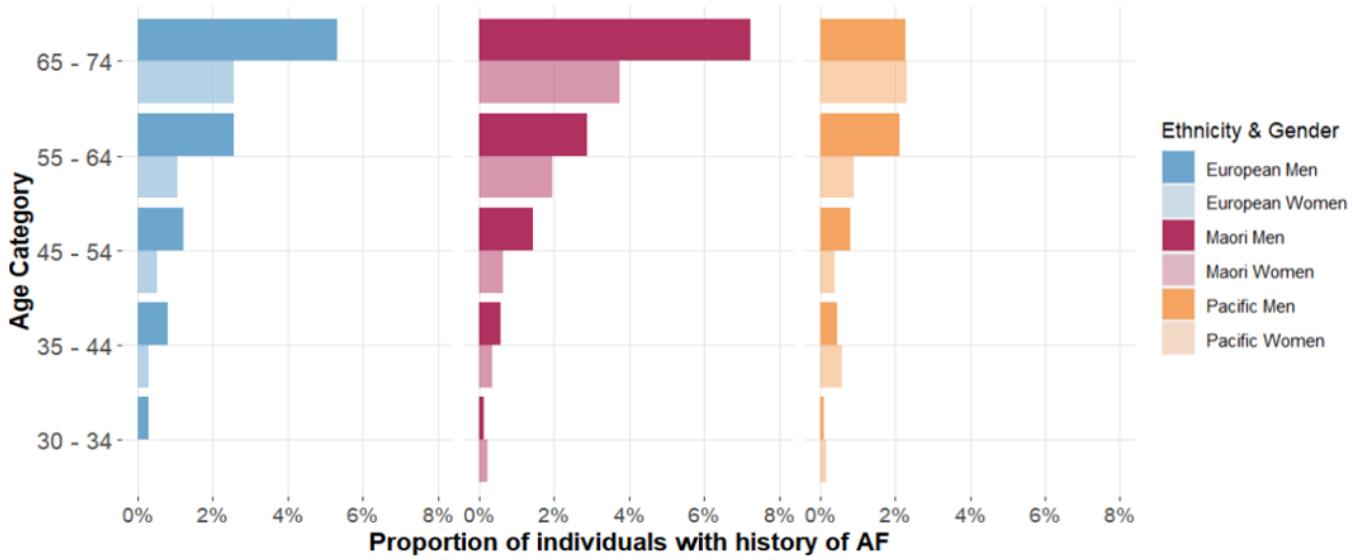


Figure 24: Proportion of individuals with history of Atrial Fibrillation for various ethnicity, age groups and gender

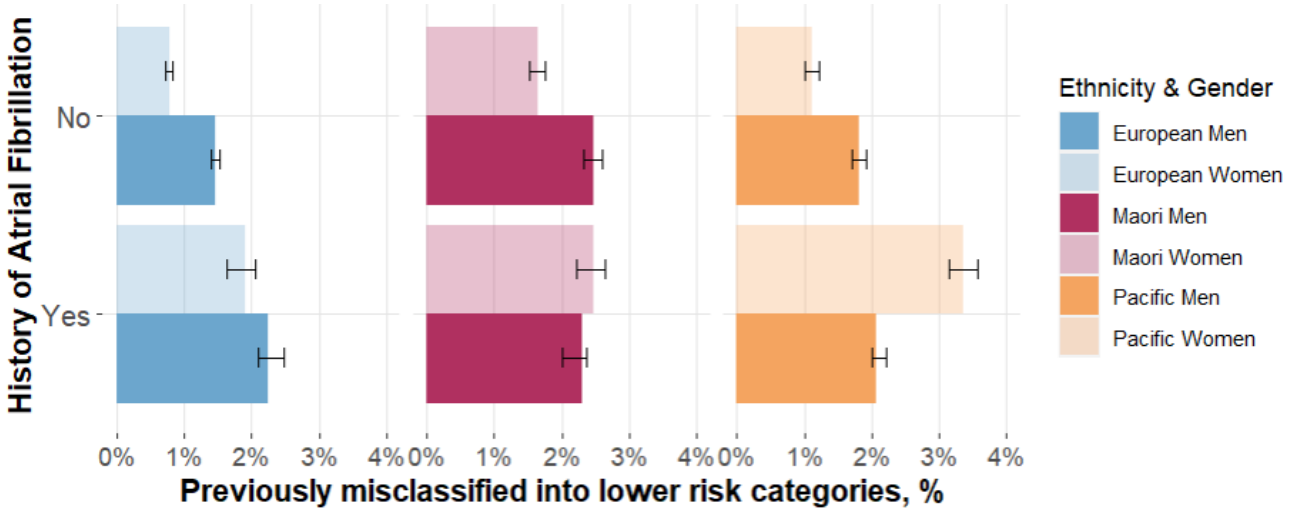
There is generally an increasing trend of history of atrial fibrillation with a strong positive correlation with age. Correlation is highest for European men at 0.896 and lowest for Pacific females at 0.633 as shown in Table 29.

Table 29: Correlation of having history of Atrial Fibrillation with age for ethnicity and gender.

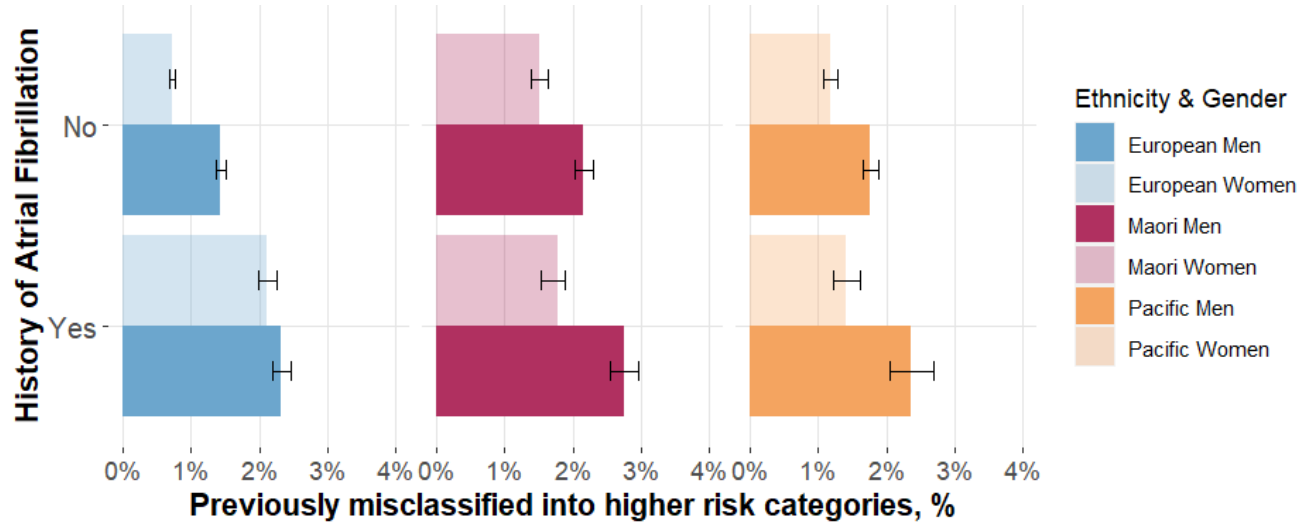
Ethnicity	Gender	Correlation (Age, have Atrial Fibrillation)
European	Women	0.895
European	Men	0.896
Māori	Women	0.811
Māori	Men	0.823
Pacific	Women	0.633
Pacific	Men	0.852

As depicted in Figure 25, there is no clear distinction between increased and decreased risks. Pacific women with a history of atrial fibrillation exhibit the highest shift to a higher risk category (3.36%),

followed by Māori women with approximately a 2.46% shift. On the other hand, around 2.76% of Māori men with a history of atrial fibrillation have their risk reclassified into a lower risk category after adjustment of the errors. While these differences are notable, no clear pattern emerges in risk shifts across ethnicities for individuals with a history of atrial fibrillation.



(a)



(b)

Figure 25: Impact of error adjustment (SD = 6.95) on risk classification across ethnic groups and gender, highlighting changes in risk group assignment based on the history of Atrial Fibrillation.

5.5 Discussion and Conclusion

This article demonstrates ethnicity-based analysis to study the impact of BP errors arising from the inaccurate BP measurement devices, on the 5-year CVD risk prediction for Europeans, Māori and Pacific population in NZ. The analysis includes one of the largest cohorts ever recruited in NZ. Even

though the errors induced by the device show insignificant changes to the overall model estimates, some notable patterns in the misclassification of people into different risk groups for the three ethnicities under test emerge clearly. The results highlight the level of classification accuracy decreasing even within the acceptable range of BP errors, especially for Māori and Pacific populations compared to Europeans.

The data used for analysis is from a study undertaken by primary care clinicians. The data is also combined with information from national databases. About a third of the country's population is served by clinics that use PREDICT software. Consequently, the cohort is large with ethnically diverse populations. No factor other than BMI has missing values. The cohort is likely to be representative of the population of NZ in which CVD risk is recommended as around 90% of eligible patients. It also contains about 35% of all primary care practices in NZ were included in the study [38].

It is unlikely that any significant population subgroup has been excluded from this study [28]. While national surveys provide self-reported data on CVD risk factor levels by ethnicity, the sample sizes for non-European ethnic groups are typically limited. However, since the PREDICT-1 equation uses these ethnic groups with similar numbers for CVD risk prediction, our study considers similar data for our analysis to study the overall impact with the existing established equation. The decision to participate in PREDICT cohort is determined at the primary health organisation level, rather than the individual practice level. This minimises the likelihood of selection biases at the practice level. The PREDICT cohort includes large rural and urban regions, incorporating the largest Māori, Pacific and other ethnic demographics in NZ.

The results show higher prevalence of smoking among Māori people and higher prevalence of being diabetic among Pacific people. Despite a reduction in smoking rate, Māori and Pacific are still the highest consumers[253]. There is also a three times higher prevalence of diabetes among Pacific peoples compared to Europeans; and this finding has been well documented since the 1990s [254]. The results for average age of Māori and Pacific people being more than Europeans was expected due to the guidelines that recommend their CVD risk to be assessed 10 years earlier than Europeans [255].

For all the risk factors considered, Māori people are mostly affected due to adjustment of BP errors, which is consistent with the higher HR of Māori compared to other ethnic groups in the PREDICT-1 equation [81]. Research also indicates that at a given age, Māori have an increased prevalence of CVD compared to non-Māori. Therefore, risk assessment is recommended a decade earlier for Māori patients [237].

There is a growing body of evidence that differences in access to healthcare between Māori, Pacific and other ethnic groups are an important contributor to evidence-practice gaps that may in turn lead to poorer CVD outcomes. If these populations are already at higher risk and get affected the most by errors, they are likely to also suffer from increased medical costs due to over treatment or undertreatment. There is evidence suggesting increased healthcare burden when a patient is undertreated, leading to more adverse CVD outcomes, which in turn leads to increased healthcare costs, potentially in the millions of dollars in NZ [219].

The study assumes device errors to be normally distributed. The inaccuracies of the device tend to follow a normal distribution, as outlined by international standards. This analysis addresses inherent inaccuracies by assuming that all BP readings in the dataset possess a certain level of measurement error. Nevertheless, owing to the absence of comprehensive data regarding the specific devices employed to measure BP, this presumption has been consistently applied to all readings. This method allows us to comprehensibly evaluate the impact of device-related errors on CVD risk classification, offering a wider view of how device inaccuracies affect clinical outcomes. Various other sources of error in BP measurement should also be considered in future research to evaluate their impact on different ethnic groups in NZ. However, analysis for Māori, Pacific and Europeans assessed in primary care adds to what has been a limited body of knowledge about the impact of allowable device errors on the CVD risk prediction. As the PREDICT dataset becomes larger and more complete, it will be possible to undertake more in-depth research on the CVD risk profiles of these ethnic groups. Targeting these modifiable risk factors must be a priority to reduce the widening health inequalities between these population groups.

Although there are certain steps taken to mitigate ethnicity related disparities in CVD management such as starting the risk assessment for Māori, Pacific and South Asian population 15 years earlier

than other population groups [82]. However, the results suggest more ethnicity focussed CVD risk predictions and more targeted interventions may help to address these health inequities. A ‘one-size fits all’ approach may not work to achieve equity of CVD and its risk factors in NZ. The results prompt an important discussion regarding whether the existing allowable range for device inaccuracies is sufficiently stringent for clinical studies, especially those involving CVD risk prediction.

Crucially, while all variables in the CVD risk prediction equation contribute to overall risk, this study prioritizes those with higher HRs. These variables exhibit a particularly strong association with CVD risk, highlighting their clinical significance. By focusing on these high-impact variables, the study evaluates how adjustments for device inaccuracies influence risk scores, while acknowledging the relevance of other predictors in the equation. Although this analysis emphasises the most influential predictors, future research should also investigate other variables with comparatively lower HRs, to provide a more comprehensive understanding of CVD risk prediction.

Our findings emphasise the significant influence of inaccurate BP measurements on CVD risk prediction in NZ’s ethnically diverse population, emphasising the need for more awareness among health professionals. Educational initiatives should be prioritised to inform practitioners about the implications of measurement inaccuracies and to stress the importance of proper device selection and usage. Additionally, collecting detailed data on the specific devices used for BP measurement is crucial for refining analyses and achieving more robust insights into the impact of measurement inaccuracies on CVD risk prediction.

CHAPTER 6. BLOOD PRESSURE MEASUREMENT DEVICE ACCURACY EVALUATION: STATISTICAL CONSIDERATIONS WITH AN IMPLEMENTATION IN R (Manuscript 4)

6.1 Prelude

Previous chapters have highlighted significant inconsistencies in blood pressure (BP) measurement accuracy due to device inaccuracies. A review of existing studies revealed that many reported BP measurement errors exceeded the internationally accepted accuracy criteria, and many do not adhere to the use of minimum sample size threshold, raising concerns about the reliability of these devices in both research and clinical practice. Some studies report measurement errors well within allowable limits, whereas others reveal deviations exceeding acceptable thresholds. Yet both are often used interchangeably in risk assessments, due to the lack of research regarding the implications of not adhering to the criteria (both the acceptable limits and also the sample size criteria).

Building on this, Chapter 6 systematically examines the statistical considerations in evaluating BP measurement device accuracy. It aims to provide a structured framework for assessing device adherence to accuracy criteria and quantifying the implications of deviations from these standards. The study aims to establish a more reliable approach for assessing the impact of measurement inaccuracies on clinical outcomes.

Additionally, this study presents an implementation in R, offering a practical tool for researchers and clinicians to assess BP device accuracy. This implementation facilitates a comprehensive evaluation of measurement errors, allowing for improved decision-making regarding device selection. By addressing the consequences of non-adherence to accuracy standards, this study contributes to a more precise understanding of BP measurement reliability and its broader implications for clinical management.

6.2 Introduction

Blood pressure (BP) is extensively used to assist health monitoring and diagnosis in healthcare settings. However, inaccuracies in BP measurement can result in misjudgements, potentially leading to severe consequences[171]. The clinical gold standard for BP measurement is performed using arterial cannulation[55], however, arterial cannulation is invasive, time-consuming, and can only be performed by skilled personnel. It is also linked with cases of ischemia, lesion of nerves or vessels, embolism, and other complications[256]. In regular cases, BP is measured non-invasively[172] which yields measurement inaccuracies. Even slight measurement inaccuracies can result in misclassifying millions of individuals [257]. Hence, a precise measurement of blood pressure holds significant importance in public health. Underestimating true BP by merely 5 mmHg or less can have significant clinical consequences as several studies have inferred incorrect tagging of more than 20 million Americans as pre-hypertensive when, in fact, they are suffering from hypertension. Untreated hypertension can lead to a 25% increased risk of fatal strokes and fatal myocardial infarctions[171]. Conversely, if there is an overestimation of true BP by 5 mmHg, nearly 30 million Americans may receive inappropriate treatment with antihypertensive medications. This could result in exposure to potential side effects of the drugs, psychological distress due to misdiagnosis, and unnecessary financial liability[257]. In healthcare domains such as intensive care, accurate BP measurement is even more crucial. As a result, regulating BP measurement devices is a critical matter and suitable processes must be used for clinical investigations to validate BP devices.

National regulators have made significant efforts towards global harmonization of the standards for medical devices. When designing a blood pressure (BP) measurement device, manufacturers must adhere to standardized protocols, ensuring that the device's inaccuracy falls within an acceptable range, typically expressed as mean error ± SD of BP errors for non-invasive techniques. Even when within acceptable limits, continuous efforts are made to improve the accuracy using improved methods by adding parameters associated with blood pressure [137], [139], [258]. This pursuit aims to provide healthcare professionals with more reliable BP readings, reducing the likelihood of errors and supporting informed decision-making. The International Organization for Standardization (ISO) established in 1947 defines standards that are accepted worldwide. It comprises representatives

from various national standards organizations. The ISO 81060-2:2018 standard defines the criterion for the clinical investigation of automated, non-invasive sphygmomanometers[49] and has been approved for use currently and recognized in whole or part by many national regulators. It supersedes region-specific standards such as EN 1060-4:2004[259] and has been adopted in law, in contrast to validation protocols such as those recommended by the British Hypertension Society [260] and the European Society of Hypertension[261].

ISO 81060-2:2018 stipulates criteria for determining the acceptable accuracy of sphygmomanometers that originated from the initial work of the Committee of US Association for the Advancement of Medical Instrumentation (AAMI) in creating the American Standard for manual, electronic, or automated sphygmomanometers known as SP10[262]. The standard also specifies safety, labelling, and performance requirements designed to ensure the safety and effectiveness of the device. ISO 81060-2:2018, like SP10, mandates a minimum sample size (N) of 85 participants to be used to evaluate the BP device inaccuracy [49]. In addition to N>=85, the standard requires the BP errors to be within -10 mmHg to 10 mmHg, also known as the tolerable error limit, and the estimated probability of tolerable error ($\hat{p}$) to be at least 85%. In practice, it is found that accuracy requirements are difficult to achieve, and process requirements are costly. Manufacturers attempt to adhere to this standard. However, only a small fraction of manufacturers can do so [263]. A study reports that less than 20% of the devices accessible today conform to an established guideline [55].

While compliance with this standard is appropriate for devices that are to be marketed, there are purposes other than regulation of medical devices for which studies involving fewer participants can still yield useful information. For instance, early evaluation of experimental devices would benefit from an earlier checkpoint, as it is often difficult for clinicians to gather 85 participants [141], [144], [264]. Currently, to our knowledge, there is no official method for evaluating studies with fewer participants. As a result, various research work in this field adopted potentially incorrect pass/fail criteria of the standard apparently without recognizing the difference between their research methods and those assumed by the standard. This paper aims to inform researchers and BP device manufacturers about the potential effects of employing different sample sizes for the validation of a BP measurement device.

We also offer recommendations to adjust the appropriate acceptance range (upper limit of acceptable SD for a certain mean error) required for any study to adhere to criteria similar to the SP10 requirements. In addition to the different acceptance limits for different sample sizes, this paper provides a brief comparison of previous studies that investigated novel BP measurement methods with different sample sizes and also assesses their adherence to the current standard.

6.3 SP10 Statistical Considerations

6.3.1 SP10 Acceptance Criteria

Multiple techniques are used for automated, non-invasive BP measurement. Most of the researchers/clinicians use an inflatable cuff to hinder the flow of blood in the upper arm. As the cuff is deflated, various methods can be employed to estimate the systolic and diastolic blood pressure (SBP; DBP)[265]. The error in estimation is the difference between the values obtained from the test device and the value obtained using a reference method, which is normally specified as auscultation by trained observers[266]. Acceptance criteria evolved from SP10's inception in 1987 to reflect a more defined statistical treatment and is currently adopted in the ISO 81060-2:2018.

Initially, the standard required that manufacturers should maintain the mean of errors within ±5 mmHg with a SD no greater than 8 mmHg[133]. However, these static values did not consider the relation between the mean and SD of errors. For the same SD, $\hat{p}$ will be different if the sample mean is 0 mmHg and if the sample mean is 5 mmHg. Hence, SP10's Table F.1 (reproduced as in Table 8 in Section 2.4.2) was introduced to span different values of the sample mean and the upper limit of SD such that 85% of the errors are within the tolerable error range, where $\hat{p} = 0.85$. These values of acceptable SD for a given sample mean represent the acceptance limit. For instance, for a mean error of 2 mmHg, the SD must be less than or equal to 6.65 mmHg to accept the device. But this estimated probability of tolerable error ($\hat{p}$) is itself an estimate.

How far off it is from the true value depends on the sample size. As per SP10, a sample size of N=85 yields a 90% chance or confidence that $\hat{p}$ will not differ by more than about 0.07 from its true probability of tolerable error (p) [133], given by:

$$\hat{p} - p = 1.645 \cdot K, \quad (16)$$

Where, 1.645 is the critical value from the standard normal distribution, $K = \sqrt{\frac{1}{2\pi(N-1)}}$ and N is the sample size. We will refer to this difference as the "90% confidence between p and $\hat{p}$". In Equation (16), K is the SD of the distribution of probability of tolerable error which is assumed to be asymptotically normal according to the SP10 standard, where the mean of the distribution is $\hat{p}$. Thus, for the device to be acceptable, $\hat{p}$ must be at least 85% for N=85, because then one can be confident that p is at least 78%, as per the standard.

6.3.2 Brief review of the problem

6.3.2.1 Acceptance Criteria

According to SP10, Table 8 can be used with any number of participants, but they only consider the acceptance limit that is suitable for N ≥ 85. However, studies with fewer participants than the minimum of 85 specified by SP10 are not uncommon. For instance, one study proposes a novel BP estimation method based on Pulse Arrival Time (PAT) to estimate SBP and DBP [141]. Using 32 subjects, they report the BP error limit, mean error ± SD. These limits are 0.12 ± 6.15 (SBP, mmHg) and 1.31 ± 5.36 (DBP, mmHg). Another study validates a wireless BP monitor using 33 participants[137]. The estimated BP errors were -0.7 ± 6.9 mmHg for SBP and -1 ± 5.1 mmHg for DBP. A new calibration procedure that accounts for Sympathetic Nervous System (SNS) on BP-PTT (Pulse Transit Time) was also proposed to estimate BP values using 10 subjects[138]. All these studies attain the $\hat{p} = 0.85$ criteria mentioned in SP10, but the sample size is less than 85. For smaller sample sizes, there is little guidance on how the acceptance limit should change such that one can be 95% certain that the true probability is at least 78%, which is recognised as the threshold for acceptability by SP10. While these studies are potentially valuable, it would be inappropriate to interpret results by making comparison to the criteria in the standard which is just fixed for N=85.

At present, the 90% confidence between $\hat{p}$ and p which is evaluated using Equation (16), only considers the sample size of a specific study. Using this 90% difference value, the standard makes some assumptions about the $\hat{p}$ such that $p \geq 0.78$. However, variations on the 90% difference are

not only due to changes in sample size but also with the value of $\hat{p}$, obtained from the reported sample mean and Sd from the BP device. To tackle this issue, we propose a methodology that provides a more flexible approach to evaluating the 90% confidence between $\hat{p}$ and p with respect to the sample size, sample mean, and SD from a statistical point of view. With this approach, the value of $\hat{p}$ can be evaluated for different sample sizes, which we can use to study the changes in the acceptance limit of different sample sizes such that the devices under test adhere to the SP10 criteria.

6.3.2.2 Probability of Acceptance (P_A)

As a result of the acceptance limit varying with sample size, the probability of acceptance, which essentially gives the probability of meeting the SP10 criterion for a particular sample mean, SD, and sample size, will fluctuate. In this regard, this research also provides a mechanism for a more relevant comparison of the mean and SD of the data from studies with varying sample sizes. To effectively compare studies with smaller sample sizes and distinguish the methodology (e.g., techniques and mathematical methods) being used to develop the BP measurement devices, a more robust statistical treatment is required to re-evaluate the literature less subjectively as per the international standards. We present aspects behind the computation of the probability of acceptance, denoted by P_A.

There are two key outcomes from this work. First, we study the changes in the acceptance limit for different sample sizes such that they adhere to the standards when the sample size is less than N = 85. Secondly, we provide a methodology to evaluate the probability of acceptance P_A, allowing comparison of different studies with varying sample sizes, assessing the accuracy of different methods and techniques being tested to build BP measurement devices. This work has a companion R package called ‘bpAcc’ which implements the methodology introduced in this paper. This enables manufacturers and researchers to better judge their compliance with the accuracy criteria of ISO 81060-2:2018 using a smaller sample size and more appropriately compare studies performed using different sample sizes.

6.4 Methodology

This section outlines the theoretical details of this research starting from the protocols currently in use by the SP10 standard. The parameters utilised in this section have also been outlined in Table 45 within Appendix D.

6.4.1 Brief review of the statistical components of SP10

6.4.1.1 Average error and tolerable error

For each of the N participants, k=3 pairs of blood pressure measurements are obtained: one measurement, $\delta_{k_1}^j$, produced by the usual auscultatory reference method, and the other, $\delta_{k_2}^j$, produced by the device being assessed. The difference, $\epsilon_k^j = \delta_{k_2}^j - \delta_{k_1}^j$, is called an error, and the average error for the j^{th} participant is:

$$\delta^j = \sum_{k=1}^{3} \frac{\epsilon_k^j}{3}, \quad j = 1, \dots, N \tag{17}$$

Statistically, we assume the average errors δ^j produced by the device D follows a θ–parameterized distribution $\mathbb{F} = \mathbb{F}(\theta)$. The maximum average error accepted, also known as tolerable error, is denoted by Δ. The tolerable error is set to Δ = 10 mmHg in this work, following the SP10 standard. Hence, the probability of tolerable error is given by:

$$\mathbb{P}(|\delta j| \leq 10; \theta) = \mathbb{F}(10; \theta) - \mathbb{F}(-10; \theta). \tag{18}$$

The errors produced by any device are deemed acceptable if p, which is the true probability of tolerable error, is a minimum of γ_p, i.e.,

$$p = \mathbb{P}\left(|\delta^j| \leq 10; \theta\right) \geq \gamma_p, 0 \leq \gamma_p \leq 1 \tag{19}$$

Fundamentally, we assume the errors δ^j follow a normal distribution with parameters $\theta = (\mu_p, \sigma_p)^T$, $\mu_p \in \mathbb{R}$, $\sigma_p > 0$. μ_p and σ_p are the mean and SD of the errors produced by the device readings and will be referred to as true mean error or true bias, and true SD, respectively.

6.4.1.2 σ_{γ_P} -- Acceptance Curve

Our interest focuses on ${\sigma_{\gamma_p}}^{MAX}$, a bivariate function of (μ_p, γ_p) given by

$$ {\sigma_{\gamma_p}}^{MAX} = \sigma_{\gamma_p}(\mu_p, \gamma_P) = max\left\{\sigma_p;\ \mathbb{P}\left(|\delta^j| \leq 10;\ \theta = (\mu_p, \sigma_p)^T\right) \geq \gamma_P\right\}, \qquad (20) $$

for $\mu_p \in (-10, 10)$, and fixed $\gamma_P \in (0, 1)$. The curve is called the σ_{γ_P}- acceptance curve, or simply acceptance curve. For every μ_p, ${\sigma_{\gamma_p}}^{MAX}$ is given by the maximum sample SD producing a probability of tolerable error of at least γ_P. Figure 26 shows σ_{γ_P}- acceptance curves represented by σ^{MAX} for $\gamma_P \in \{0.75, 0.80, 0.85, 0.90, 0.95\}$, and μ_p in (10,10).

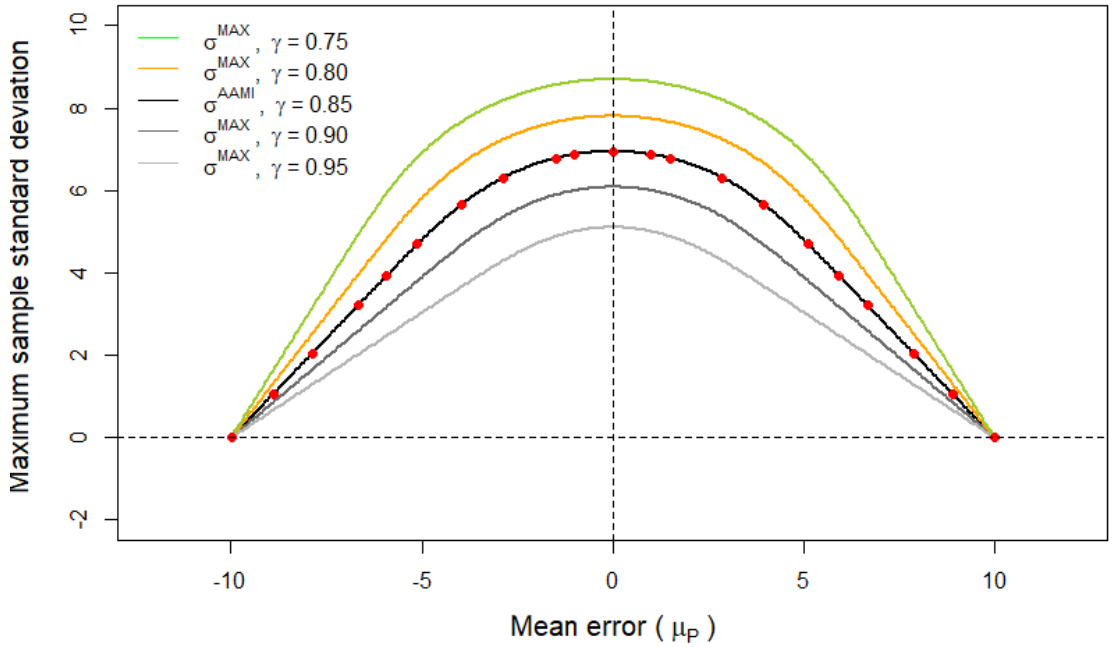


Figure 26: Acceptance curves obtained from Equation (20) for $\gamma_p \in \{0.75, 0.80, 0.85, 0.90, 0.95\}$; μ_p in (−10, 10).

6.4.1.3 ANSI/AAMI- SP10 acceptance criterion

As per the SP10 standard[133], a device D will be deemed acceptable if the estimated probability of tolerable error $\hat{p}$ is at least γ_p = 0.85, and the sample size is 85 subjects. From Equation (20), we define the σ^{AAMI}- acceptance curve as

$$ \sigma^{AAMI} = {\sigma_{0.85}}^{MAX}\left\{\sigma_p;\ \mathbb{P}\left(\left|\delta^j\right| \leq 10;;\ \theta = (\mu_p, \sigma_p)^T\right) \geq 0.85\right\}, \qquad (21) $$

with $\mu_p \in$ (-10, 10), $\theta = (\mu_p, \sigma_p)^T$. Since σ_{γ_P}- acceptance curves narrow down as shown in Figure 26 as γ_P decreases, without loss of generality, we assume that the σ^{AAMI}- acceptance curve is obtained at γ_p = 0.85, which is as follows:

$$\sigma^{AAMI} = max\left\{\sigma_p;\ \mathbb{P}\left(\left|\delta^j\right| \leq 10; \theta\right) \geq\ 0.85\right\} = \left\{\sigma_p;\ \mathbb{P}\left(\left|\delta^j\right| \leq 10; \theta\right) =\ 0.85\right\}. \quad (22)$$

The σ^{AAMI}- acceptance curve is the thick black line and is given by the solution to:

$$\mathbb{P}\left(|\delta^j| \leq \Delta;\ \theta = (\mu_p, \sigma_p)^T\right) =\ 0.85. \quad (23)$$

with Δ = 10, and $\mu_0 \in (-10{,}10)$.

However, in practice, only size-limited samples of BP measurements are available to test the device D. In the following, we will introduce the statistical assumptions required in this work. Let S_ι = $\{\delta_\iota^1, \ldots, \delta_\iota^n\}$ be a size-n sample of BP average errors δ_ι with $\delta_\iota \sim N(\mu_p, \sigma_p)$. The S_ι- sample mean and S_ι- sample SD are denoted by $\bar{x}_\iota$ and s_ι respectively. We will remove the superscripts 1, …, n for simplicity.

Crucially, we replace μ_p with $\bar{x}_\iota$ in Equation (23). Then, the highest permissible value for s_ι rendering the device $\sigma_{0.85}$- acceptable, denoted by $\sigma_{0.85}^{AAMI}$, is the function of $\bar{x}_\iota$ given by:

$$\sigma_{0.85}^{AAMI} =\ \sigma_{0.85}^{AAMI}(\bar{x}_\iota) = \left\{\sigma_p;\ \mathbb{P}(|\delta_\iota| \leq\ 10;\ \bar{x}_\iota) =\ 0.85\right\}. \quad (24)$$

As a result, the device D is deemed acceptable under the SP10 acceptance criterion if and only if $s_\iota \leq \sigma_{0.85}^{AAMI}(\bar{x}_\iota)$.

The values $\sigma_{0.85}^{AAMI}$ for fixed $\bar{x}_\iota$ are obtained by directly applying the bisection method to Equation (22) as a function of σ_p. Table 30 gives $\sigma_{0.85}^{AAMI}$ to selected values of $\mu_p = \bar{x}_\iota$. The pairs $(\bar{x}_\iota, \sigma_{0.85}^{AAMI})$ mentioned in the table are the values displayed in Figure 26 as red dots.

Table 30: The upper limit of SD for selected values of $\bar{x}_\iota$.

$\bar{x}_\iota$	0	±1	±1.5	±2.87	±3.96	±5.12	±6.67	±7.15	±7.88	±8.9	±9.99
$\sigma_{0.85}^{AAMI}$	6.947	6.874	6.782	6.311	5.664	4.696	3.213	2.75	2.045	1.061	0.01

6.4.2 Sampling Distribution of Sampling Proportions

SP10's confidence limits for p, or true probability of tolerable error, rely on approximations of the Gaussian density using Taylor expansions around the mean and SD [133], providing a biased, standard error depending just on N of the form $\frac{1}{2\pi(N-1)}$ and 95% confidence limits given by $\hat{p} \pm 1.645 * \sqrt{\frac{1}{2\pi(N-1)}}$.

In this paper, as opposed to [133], we adopt a statistical standpoint to address the uncertainty attached to $\hat{p}$. Consider the binomial random variable, say Y, given by the number of errors falling in the interval [-10 ,10]. The probability of 'occurrence' or errors falling in [-10, 10], denoted with p, is central to this paper. Essentially, we estimate p via maximum likelihood estimation (MLE) using the sampling distribution of proportions which results from the theoretical probability distribution of random-sampled proportions of fixed-size N from the population of errors. The MLE of p is given by $\frac{Y}{N}$. This method represents our main modelling framework allowing us to estimate p and compute probabilities associated with any sample. This framework has been implemented in the package `bpAcc' in R software [267].

The distribution of p, or proportion's sampling distribution, is asymptotically normal, based on the Central Limit Theorem, requiring a reasonably large sample size for estimation accuracy. Specifically, it requires N * $\hat{p}$ ≥ 5 and N*(1 – $\hat{p}$) ≥ 5. Under such conditions, the distribution of p is approximately normal with mean $\hat{p}$, and SD:

$$sd = \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{N}} \tag{25}$$

With the proposed approach the 90% confidence between $\hat{p}$ and p is given by:

$$p - \hat{p} = 1.645 * \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{N}} \tag{26}$$

To comply with the SP10 standards, the 90% confidence between $\hat{p}$ and p should be such that p ≥ 0.78. In this way, one can evaluate an updated value of $\hat{p}$ for any sample size (N) using Equation

(27). This results in changes in the acceptance limits for different sample sizes which will be discussed in Section 5.

$$\hat{p} \; - \; 1.645 * \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{N}} \; = \; 0.78 \tag{27}$$

6.4.3 Evaluation of the Probability of Acceptance (P_A)

6.4.3.1. Evaluating the probability

According to the standard, there is 95% certainty that p ≥ 0.78 with a sample size of N = 85, where 95% is the threshold for the probability of acceptance. This will serve as the benchmark for our proposed methodology, given by:

$$P_A \; = \; 1 - \; \emptyset \left(0.78, \hat{p}, \sqrt{\frac{\hat{p}(1 - \hat{p})}{N}} \right), \tag{28}$$

where ∅ is the cumulative density function of the Normal Distribution. Fundamentally, Equation (28) compares the probability of acceptance for previously published studies with reported mean and SD for the BP errors under different sample sizes. For cases where P_A ≥ 0.95, the device meets the SP10 standard. Currently, the acceptance region provided for N=85 is used to validate devices that have used smaller sample sizes, however with the proposed approach, we can now provide more insights on whether those devices are complying to the SP10 standards with fewer sample sizes or not.

For inference purposes, the proposed framework relies on reasonably large sample sizes i.e. N ≥ 39 such that N*$\hat{p}$ ≥ 5 and, N*(1 – $\hat{p}$) ≥ 5. However, the results provided by the simulation study described in Section 3.3.2 have shown closer approximations even for small sample sizes (N < 39), as shown in Table 31. For instance, μ_P= 2 and σ_P= 5.5 is used to check for cases of samples that are less than 39 to compare the value of P_A. The selection of sample size for comparing the values of P_A in Table

31 is informed by some of the previous studies that have utilized smaller sample sizes to assess device inaccuracy through various evaluation methods [138], [140], [143], [145].

Table 31: Simulated P_A and P_A obtained from Normal approximation using proposed method vs the method currently in use in the SP10 standard, for small sample sizes, with $\mu_P = 2$ and $\sigma_P = 5.5$

	N=10	**N=15**	**N=20**	**N=25**
Simulated P_A	0.95	0.964	0.974	0.982
P_A using proposed framework	0.931	0.965	0.982	0.99
P_A using SP10 method	0.84	0.893	0.926	0.948

6.4.3.2. Simulation Study

The simulation study conducted to evaluate the probability of acceptance compares the results obtained with the proposed framework. We investigated a simple situation in which N random numbers from a normal distribution with a known mean and SD were generated. The proportion of errors that fall within the tolerable error range is calculated, yielding the estimated probability of tolerable error, $\hat{p}$. The simulations are conducted for sim.count = 20000 errors and the proportion of $\hat{p} \geq 0.78$ are evaluated to determine the value of P_A.

To get an estimate of the proportion of instances that have $\hat{p} \geq 0.78$, this process was repeated 50 times. The proportion obtained in each of these repetitions is comparable with a maximum difference of 0.004. For mean = 2, SD = 5.5 and sample size = 25, the median of these repetitions was 9.982. Future simulations would yield in similar medians of proportions with only minute differences. We can demonstrate through simulations that our modelling framework is a better approximation than the present technique for N < 39.

6.5 Software Implementation

The concept of acceptance region and the probability of acceptance have been implemented in the package “bpAcc” for the R statistical software. The function to evaluate the acceptance region for different sample sizes is `AcceptR()` which directly computes Equation (21). Here, $\gamma_P = \hat{p}$ is evaluated for a given sample size, N, using Equation (27). The function `PAccept()` gives the probability of acceptance for a study that has reported a sample mean error and SD for a sample size to validate a BP measurement device. This function directly evaluates Equation (28). Arguments for both the functions from the package are provided in Table 32 and Table 33. The Comprehensive R Archive Network contains concise documentation on user guidance, providing detailed descriptions of package functions and examples. Users can access this documentation when downloading the package in R

Table 32: Arguments for AcceptR() from the R package bpAcc.

Argument	**Comments**
N	$S\iota$–Sample size.
distribution	Distribution the errors are pulled from. Default is “normal”, i.e., normally distributed δ_ι^k-errors
criteria	The underlying standard criteria for testing and data analysis. The default is "SP10:2006".

Table 33: Arguments for Paccept() from the R package bpAcc.

Argument	**Comments**
N	$S\iota$–Sample size.
Xbar, sd	Sample mean and sample SD of $\delta_\iota^{\mathrm{k}}$-errors distribution
distribution	Distribution the errors are pulled from. Default is “normal”, i.e., normally distributed δ_ι^{k}-errors
criteria	The underlying standard criteria for testing and data analysis. The default is “SP10:2006”.

6.5.1 `AcceptR()` *function*

Figure 27 provides an upper limit on the sample SD to make sure that $\hat{p}$ is at least 87.47% for N=33. If the sample mean error is between two values in the Table 8, linear interpolation is implemented. As an example, if sample mean is -0.7 mmHg. This is (- 0.7+0.5)/ (-1+0.5) = 0.40 = 40% of the distance between -1.0 and -0.5, so one uses 0.40*6.45 + (1-0.40) * 6.50 = 6.48. The sample SD would have to be 6.48 or less to accept the device.

```
> ## n = 33, xbar = -0.7, sd = 6.9

> AcceptR(n = 33)

-----------------------------------------
-------------

 For 33 samples, 87.47% of errors must be within -10
mmHg to 10 mmHg.

-----------------------------------------------------
----------                  xbar       sd

    0.0     6.522419

    0.5     6.503214

    1.0     6.445114

    1.5     6.346654

    2.0     6.205400

    2.5     6.018054

    3.0     5.780903

    3.5     5.491068

    4.0     5.149097

    4.5     4.762372
```

Figure 27: Sample output from AcceptR() for sample size N = 33.

6.5.2 `PAccept()` *function*

During the initial research and development phase of a BP measurement device, different methods can be compared by evaluating P_A which gives the probability of a device meeting the standards using the "`PAccept()`" function. Usually, different sample sizes are used to evaluate the device. This function can be directly used to determine how far existing studies or devices are from the acceptable standard.

For instance, when two methods to develop a device are compared where Method 1 provides a device inaccuracy with sample mean error $\pm$ SD = 4 $\pm$ 5.1 and Method 2 provides sample mean error $\pm$ SD = 3 $\pm$ 6.2 for N = 33, to validate which method provides better accuracy and is acceptable as per the standards is evaluated using the R code chunk provided in Figure 28.

.

```
> ## n = 33, xbar = 4, sd = 5.1

> PAccept(xbar=4,sd=5.1,N=33)

-----------------------------------------------------

The probability of acceptance as per SP10 is
0.9557126

The device is meeting the SP10 criteria.

-----------------------------------------------------



> ## n = 33, xbar = 3, sd = 6.2

> PAccept(xbar=3,sd=6.2,N=33)



-----------------------------------------------------

The probability of acceptance as per SP10 is
0.8801012

 The device is not meeting the SP10 criteria.

-----------------------------------------------------
```

Figure 28: Sample output from PAccept() for two different cases with different device inaccuracy.

6.6 Applications

6.6.1 Acceptance Region for different sample size

Using the proposed methodology, we can obtain the value of $\hat{p}$ such that the standards criteria is also met using Equation (27). This adjusted value of $\hat{p}$ for a given N is termed as the revised estimated probability of tolerable error or revised $\hat{p}$ as illustrated in Figure 29. The figure indicates the changes in revised $\hat{p}$ for different sample sizes. This implies that there will be different acceptance regions per sample size as opposed to a single acceptance region in the SP10 standard [133], which is illustrated in Figure 30. The figure shows the acceptance region for a range of sample sizes between 5 to 85 each showing the upper limit of SD for a given mean error that must be

followed such that 95% of the time, the true probability is at least 0.78 to adhere to the SP10 standards.

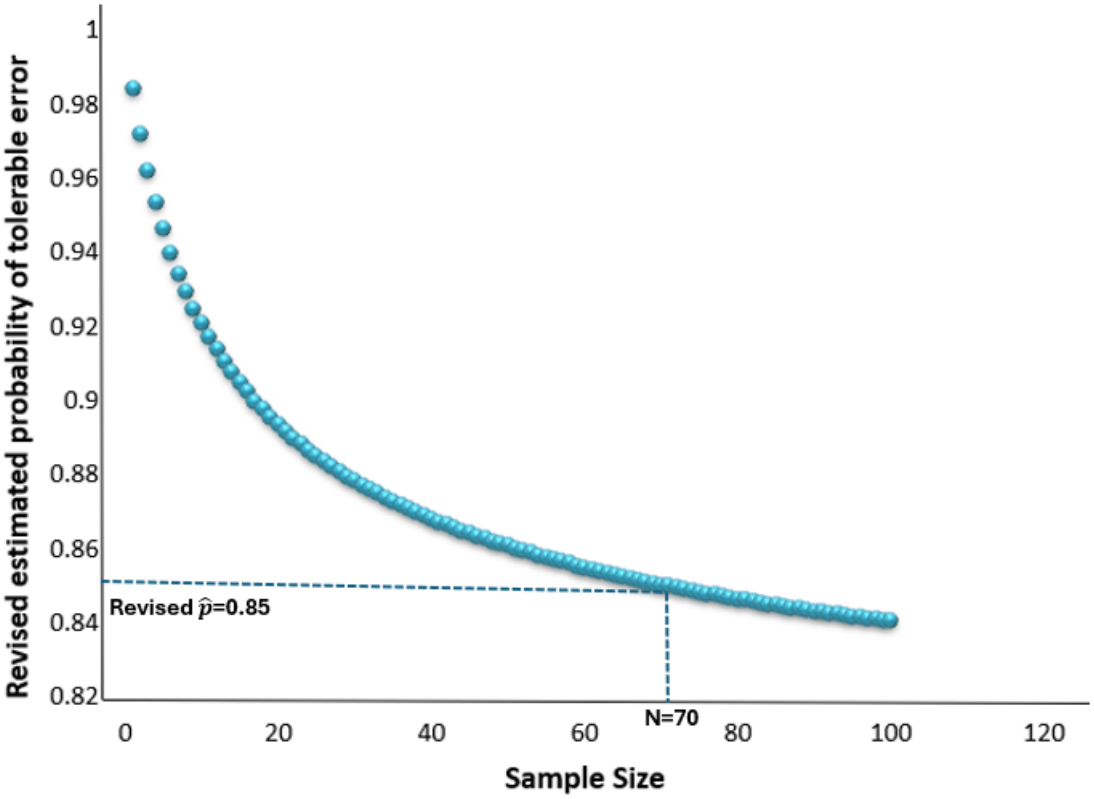


Figure 29: Relationship between the revised estimated probability of tolerable error for different sample sizes (N).

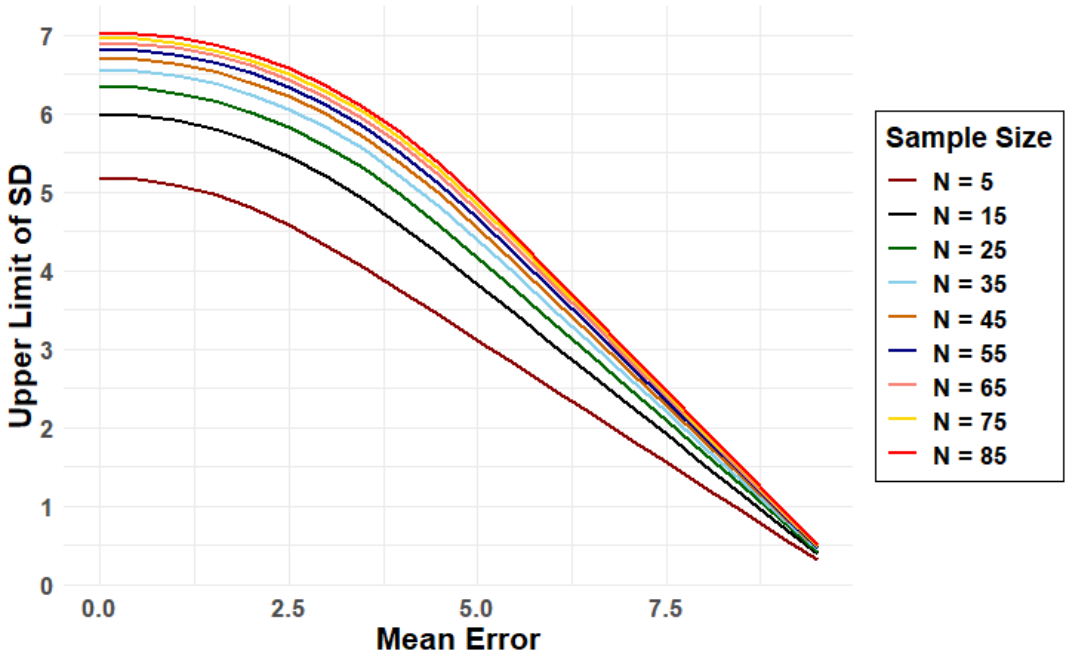


Figure 30: Changes in the Acceptance Region (Upper Limit of SD for a given sample mean, $\bar{x}$) as per SP10 for different sample sizes.

6.6.2 BP Technologies: Comparison of different methods

Since the acceptance region varies for different sample size based on the revised $\hat{p}$, the value of PA will also vary for any reported mean error and SD. The P_A values can be directly evaluated using Equation (28). Table 34 and Table 35 provide a list of different studies that have reported device inaccuracy based on their development techniques or research methods. The techniques or methods outlined in the tables represent only a subset of the diverse range of technologies employed in the development of blood pressure (BP) measurement devices. While the table highlights specific studies utilizing various methods, it is important to recognize that numerous other technologies and approaches are also being explored within the field of device development. The device inaccuracy in both the tables signifies the error ($\bar{x}$ ± SD) associated with the BP device. These BP estimation errors are measured in mmHg.

The proposed methods allow us to evaluate the Probability of Acceptance of devices reported by these studies and hence also provide comparison between different BP development techniques/methods. For instance, a study conducted to develop a BP device using oscillometric method has reported mean of BP errors $\bar{x}$ as -0.7 and SD as 6.9 for a sample size 33 [23]. The SP10 standard states that 85 samples should be used and for $\bar{x}$ = -0.7 mean BP error, the SD should not be more than 6.95. Even though the SD reported by the study with 33 samples is less than 6.95, it would be incorrect to interpret this as compliance to the SP10 criteria because the smaller sample size will also influence the acceptability. By evaluating P_A, we can analyse that effect. For this study P_A ~ 0.87 which is less than the threshold of acceptability upon which SP10 is based i.e. P_A ≥0.95. Hence, the device is not meeting the criteria of acceptability.

Table 34: Comparison statistics of the previous clinical studies that have reported device inaccuracy based on SBP values.

Study	**Method/Techniques**	**Sample Size**	**Device Inaccuracy ($\bar{x}$ ± SD)**	**Probability, P_A ($p \geq 0.78$)**
[137]	Oscillometry	33	-0.7 ± 6.9	0.873

[138]	PTT	10	1.04 ± 6.88	0.730
[139]	PTT-PPG	33	1.17 ± 5.72	0.997
[140]	Standing	25	-0.462 ± 8	0.539
[141]	PAT	32	0.12 ± 6.15	0.984
[142]	PTT	33	-0.06 ± 6.63	0.934
[143]	PTT-linear	20	0 ± 6.73	0.859
	PTT-nonlinear		0 ± 5.56	0.995
[144]	ML	45	4.53±2.68	0.999

Table 35: Comparison statistics of the previous clinical studies that have reported device inaccuracy based on DBP values.

Study	**Method/Techniques**	**Sample Size**	**Device Inaccuracy ($\bar{x}$ ± SD)**	**Probability, P_A ($p \geq 0.78$)**
[138]	PTT	10	-2.16 ± 6.60	0.732
[139]	PTT-PPG	33	0.40 ± 7.11	0.825
[145]	PTT-IPG	15	-0.5 ± 5.07	0.999
[146]	PWV-	15	-0.06 ± 5.46	0.991
[137]	Oscillometry	33	-1.0 ± 5.1	0.999
[141]	PAT	32	1.31 ± 5.36	0.999
[142]	PTT	33	-0.25 ± 5.63	0.999

6.7 Discussion and Conclusion

International standards such as ISO 86010-2 serve an important purpose in providing clarity to consumers, manufacturers, and regulators that medical devices (at least with respect to the scope of the standard) are safe and effective. With this purpose, standards provide clear pass/fail criteria, which reflects the level of a device's performance and acceptability. In this regard, the pass/fail criteria set out in ISO 86010-2:2018, inherited from SP10, is broadly recognized and represents an implicit definition of what constitutes acceptable errors in blood pressure measurement. In SP10, the mathematical translation of this definition into pass/fail criteria utilises an approximate approach that results in formulas for confidence intervals that are functions only on the sample size, disregarding the sampling errors in the form of estimated probability of tolerable error. In this work, we have proposed a method using a solid statistical theory to determine confidence intervals. The proportion's sampling distribution is a more accurate statistical approach to study the random errors producing p since it additionally takes the mean and the SD of measurement errors into consideration.

By detailing the expected changes in device acceptability, the paper contributes valuable knowledge to the existing research in this field. This work also provides an adjusted acceptance limit of BP errors based on the same definition of acceptable performance underlying SP10 standard for studies that use a sample size less than 85. The adjusted limits are expected to be useful in initial validation of BP technologies. Device manufacturers can use these adjusted acceptance limits to estimate compliance with the standards using smaller sample sizes, reducing cost of development and/or allowing faster iterative development.

An important use-case for this research is the ability to compare reported results, for example, when there are several technologies/methods/algorithms being used for estimating BP, and the individual reports utilize different sample sizes, as shown in Table 34. In each case, the reported device inaccuracy ($\bar{x}$ and SD) appear to be within the SP10 criteria. However, the sample sizes are significantly smaller. Using the methods presented here, calculation of probability of acceptance, P_A, allows a quantitative comparison of the existing literature and also to the SP10 criteria. For most of the presented studies, it is apparent that the reported results do not reach an equivalent level of confidence as SP10. P_A allows direct comparison of results with different means and standard

deviations, for example a study with high ± mean and low SD[144] and a study with much lower ± mean and higher SD [142]. Many hypertension societies now offer clinicians a comprehensive list of blood pressure measurement devices, facilitating informed decision-making for clinical trials. Currently the list can only assess whether the device is recommended or not based on their performance, utilizing the acceptance range specified in the standards [268]. With the introduction of the proposed method, clinicians can now more appropriately compare the reported inaccuracies across varying sample sizes. This functionality empowers clinicians when evaluating devices for their specific research needs.

We demonstrate that no more than 70 samples are required to maintain the 85% estimated probability of tolerable error, as opposed to the N=85 stated in the standards, as illustrated in Figure 29. There are cases where being able to correctly interpret results of a study with a smaller sample would be beneficial, for example, with population subset with only infants. Our framework still makes a statistical assumption of reasonably large sample sizes, ideally N ≥ 39, Although, the proposed method is optimized for at least 39 samples, it is instructive to see how the framework performs for fewer samples. For this, we performed a simulation study with varying the mean and the SD of the error distribution. We experimented with high variance but not exceeding the SP10 mandate for the SD, that is, $\sigma \leq 6.9$. The approximations with the proposed framework demonstrated closer results compared to the approximations with the existing framework. Table 31 presents one such result, while additional scenarios are elaborated in detail in Table 46, Table 47 and Table 48 in Appendix D. The results get more distant for both frameworks as the sample size decreases. We witness this relation because with smaller samples sizes, the tendency of the sampling distribution to approximate normal distribution decreases. These results are further confirmation that caution should be applied when using smaller sample sizes, and particularly when N < 39. To extend the proposed framework for less than 39 samples, further research is required.

This study presents a formal statistical evaluation of the device's conformity with international standards, primarily through the evaluation of the probability of acceptance P_A depending on the mean error, SD of the error, and the sample size. While evaluating a device's inaccuracy, international standards also mandate that device follows guidelines in regard to selecting the cuff

size, providing subject with a resting period prior to measuring etc. Any deviation from these protocols has the potential to introduce bias, though the specific impact remains inconclusive as most studies do not explicitly state whether these protocols were adhered to. Additionally, to ensure enough samples in varied categories, including, to cover high and low blood pressure groups, ranges and distributions of arm sizes, large sample size is necessary. In such cases, a smaller sample size will not be representative of the population.

With significant modifications for the standard SP10, we introduce a mathematical framework to accommodate different underlying definitions of acceptable error and confidence. This is of relevance to those developing new BP measurement technologies which are often tested initially in smaller samples, such as cuffless, wearable BP measurement devices that perform continuous readings to gather trends for a long period of time. These technologies help in assisting real time fluctuations which might be useful for clinical trials that aim to gather longitudinal data on blood pressure trends and responses to interventions.

Finally, to assist in calculations presented, this paper also introduces the companion R package “bpAcc”, an implementation of this methodology involving functions to directly compute the acceptance limit and P_A without having to deal with the mathematical complexities. At present, our framework has infrastructure to afford normally distributed errors, as stated by the argument “distribution” from “PAccept()” and “AcceptR()”. Future work includes upgrading both functions to handle errors other than normal. Initial steps have been taken in this direction with both functions being currently trained and tested using the one-parameter (λ, or degrees of freedom) Student-t distribution. We are focused on selecting real-valued distributions with practical benefits for clinicians and manufacturers, rather than a theory–based selection of choices. Essentially, more data is useful, but experiments are often expensive. We aim to provide choices spanning various sample sizes, by providing statistical infrastructure to maximize the user’s ability to identify faulty devices (e.g., Type I error) for BP measurement. Over time, both functions will be enhanced with further arguments to handle, e.g., criteria other than “SP10:2006”.

CHAPTER 7. DISCUSSION AND CONCLUSION

7.1 Conclusions

The primary goal of this thesis is to investigate the impact of BP measurement errors on CVD risk prediction in the NZ population and their potential effect on treatment decisions. We particularly focus on two critical sources of BP measurement errors, namely, rounding errors and the errors arising due to device inaccuracies.

The results on the impact of rounding BP measurements to the nearest zero end-digit is presented in Chapter 3. The methodology involved a simulation study where a second reading was generated for a sample of 292,122 non-zero SBP records by rounding to the nearest zero end-digit. Cox PH models were subsequently employed on both, the original and rounded datasets, to evaluate changes in the risk classification and on the overall model predictability, measured by HR. This approach was then generalised to the full cohort, accounting for device calibration to the nearest 2 mmHg, to represent real-world rounding situations in the current cohort [219]. The results show that, even though there were slight variations in the overall model predictability, there was notable changes in the risk classification. Approximately 4.24% of high-risk men and 3.21% of high-risk women are misclassified into lower-risk categories. Additionally, 1.19% of men and 0.62% of women are overclassified into the moderate-risk group, while 0.47% of men and 0.20% of women are overclassified into the high-risk group. To examine the implications of the changes in these risk classifications due to rounding on the overall healthcare costs within the NZ healthcare system, a detailed healthcare cost analysis was conducted, and presented in Table 17 and Table 18 in Chapter 3.

Chapter 4 presents the results of the analysis on the impact of device related inaccuracies in BP measurement on CVD risk classification into different risk groups in NZ. A simulation study was performed using the entire dataset, by adjusting the original SBP measurements after accounting for the device-related inaccuracies, in contrast with the approach adopted in Chapter 3, as all BP measurement devices have a certain form of device inaccuracy. Likewise, Cox PH models were

applied to the original dataset with baseline SBP values and subsequently to the dataset with the adjusted SBP values. Our key focus was on the changes in HRs between both models, as well as on the misclassification rates across three CVD risk groups and the economic consequences of these misclassifications. The changes in the HR of both models were analysed along with the rate of misclassification across three CVD risk categories resulting from device inaccuracy, together with the financial implications of these misclassifications.

Chapter 5 presents the impact of BP device inaccuracies on the classification of European, Māori and Pacific individuals into different CVD risk categories. While the study design was similar to that adopted in our analyses presented in Chapter 4, it extended the analysis by investigating re-classification as it interrelates with significant risk factors with higher HRs, such as BMI, diabetes status, smoking status, usage of BP lowering medications and history of atrial fibrillation. This study examined the trends among different ethnic groups through comprehensive exploratory analysis along with comparative analysis to assess the accuracy of risk classification within these groups. This research also evaluated the rate of misclassification among CVD risk categories when the SBP is adjusted for the highest permissible device error (SD=6.95 mmHg). Moreover, we discovered multiple instances in previous research where the range of acceptable errors for BP measurement devices was not adhered to, typically due to non-compliance with sample size requirements or variations from established acceptance limits.

A significant shortcoming identified in the existing methodology was its sole dependence on the sample size, neglecting the potential variability of acceptance limits in relation to alterations in the mean and SD of BP errors. To address this gap, Chapter 6 provides a novel methodology that incorporates both the mean and SD of BP errors to provide a comprehensive framework for evaluating device accuracy. Here, a simulation was performed to validate the proposed approach, demonstrating a better approximation to estimate the probability of device acceptance compared to the methodology outlined in the ISO standard [49].

Additionally, Chapter 6 also presented the concept of evaluating the probability of acceptance (P_A), which offers a practical means of comparing studies employing different device development

techniques. At this stage, we reviewed previous research, evaluating P_A under various conditions. Finally, the study presents an R package “bpAcc”, which facilitates calculations by providing functions, `AcceptR()` and `PAccept()`, to directly compute the acceptance limit and P_A, thereby reducing mathematical complexities.

7.2 Contributions

Theory advancements

The ISO standard defines acceptance criteria based on an approximation that depends only on sample size for establishing confidence intervals, while overlooking the impact of sampling errors via mean and SD. The work presents a more rigorous statistical approach for establishing appropriate confidence intervals. The sampling distribution of sampling proportions provides a more precise approach for analysing random errors, as it contains both the mean and SD of measurement errors. This enhancement more effectively accounts for the variability generated by such errors, yielding a more accurate and dependable framework for evaluating BP measurement device acceptance as opposed to the methodology presented by the standard which does not account for this variability.

Real-world applications

This research also presents revised BP error acceptance limits for sample sizes below 85, in accordance with the SP10 standard's criteria for acceptable performance. The revised limitations can aid in the preliminary validation of BP technologies, allowing manufacturers to assess compliance with standards using fewer samples, hence lowering development costs and expediting iterative development. Another application of this research is the ability to compare reported results from diverse technologies, methodologies, or algorithms employed for measuring BP, particularly when they use differing sample sizes for reporting BP device inaccuracy. The methodology proposed allows for the evaluation of the probability of acceptance, enabling a quantitative comparison of the current literature to identify best methods to develop a BP measurement device. A key takeaway from this research is that acceptance limits are dependent upon sample size; thus, criteria established for sample sizes of N ≥ 85 are inapplicable to smaller sample sizes.

R package

The concept of “acceptance region” and “probability of device acceptance” have been implemented in an R package “bpAcc”. The package consists of two main functions, `AcceptR()` and `PAccept()` as illustrated in Table 32 and Table 33 in Chapter 6. `AcceptR()` provides the acceptance limits for a particular sample size to ensure adherence to the standards. `PAccept()` directly computes the probability of acceptance which is useful to determine how far existing studies or devices are from the acceptable standard.

Practical implications of BP errors

This study makes contributions towards the current body of literature on the impact of BP measurement inaccuracies on CVD risk prediction using the PREDICT-1 equation. This study assesses the implications of BP errors on patient classification within NZ's national cohort and the model's predictive accuracy, in contrast to prior studies that mainly examined the influence of rounding errors using international risk equations. The results quantify the extent of this influence, providing data for healthcare practitioners and regulators to reconsider BP measurement practices. The findings also suggest disparities among ethnic groups, with Māori exhibiting the highest rates of misclassification, hence increasing the risk of both undertreatment and overtreatment. The study further investigates the behaviour of risk factors in the context of BP errors in European, Māori, and Pacific populations, contributing to discussions on possible benefits of ethnicity-specific risk prediction models.

Healthcare costs estimate

This work also makes a significant contribution by performing a comprehensive analysis of healthcare costs, evaluating the expenses associated with CVD risk assessment and medication over a five-year period for each individual. Integrating clinical risk estimations with economic evaluations provides significant insights into the financial burden of CVD management. The results indicate an unnecessary expenditure of around NZD 8.2 million from rounding errors ranging up to NZD 21.71 million from acceptable device inaccuracy, both attributable to overtreatment. Additionally, the results emphasise the increased likelihood of adverse events due to undertreatment,

further straining the healthcare system. This analysis illustrates the long-term financial repercussions of misclassification resulting from BP measurement errors, emphasising the necessity for precise risk evaluation to optimise resource allocation.

Implications for Clinical Practice and Policy

This study suggests benefits that could be derived from specialised educational programs for healthcare professionals regarding the effects of rounding errors and the significance of precise BP recording. Regular retraining helps solidify best practices, maintaining uniformity and minimising errors. The findings also motivate investigations into the adequacy of the currently allowable range for the inaccuracy of BP measurement devices in clinical research and whether more rigorous limits could enhance risk classification. Furthermore, the insights assist policymakers in enhancing protocols that improve BP measurement accuracy and outcomes for patients.

7.3 Recommendations for Future Work

We also identify new directions that broaden the work further. The following sections describe basic ideas for future work:

Better data collection

The probability of BP rounding may be influenced by specific confounding variables, such as socioeconomic status, for instance, elevated NZDep scores, or access to superior healthcare facilities. Nevertheless, owing to the absence of data or research regarding these criteria, it was impracticable to incorporate them into our study. Future research should explore the association between BP rounding tendencies and socioeconomic disparities, potentially integrating linked healthcare data to assess whether individuals in lower socioeconomic groups experience greater measurement inaccuracies.

The BP measurement device errors are assumed to adhere to a normal distribution, as outlined by international standards and have been accounted for in the study by assuming that all BP readings in the dataset contain a certain level of measurement error. However, owing to the absence of comprehensive data regarding the specific devices employed for BP recording, this assumption has

been consistently applied to all readings. It allows us to evaluate the holistic impact of device-related errors on CVD risk classification, offering a wider understanding of how device inaccuracies affect clinical outcomes. Future work should focus on gathering detailed device-level data to evaluate whether measurement errors vary across different BP monitors. Conducting validation studies comparing device-specific accuracy would help refine error estimates and improve risk classification adjustments in clinical settings.

Standardisation of Ethnic Representation

This study is limited by the unequal representation of ethnic groups in the dataset, with more records of Europeans compared to other ethnic groups. The dataset was not modified to balance the number of records among ethnicities, as the study aimed to assess the influence of BP measurement inaccuracies on CVD risk prediction utilising the PREDICT-1 equation. Since this equation was developed using the same dataset and is currently being used in the NZ healthcare system, retaining the original proportions ensures its applicability to the real-world settings. However, future research may investigate the standardisation of the dataset by balancing the number of records among ethnic groups. This would provide a more thorough assessment of the influence of BP measurement errors on CVD risk prediction, guaranteeing equitable representation of all ethnic groups.

Other sources of BP errors

This research focusses on analysing the impact of rounding errors and device errors on the CVD risk prediction in NZ, while acknowledging the other sources of BP errors, such as incorrect posture, white coat effect and incorrect cuff size and placement. The results offer insights into the current body of information, emphasising that although they are not the sole determinant, they constitute a significant element in clinical decision-making. However, further research is necessary to investigate the impact of additional sources of error and the inter-relation between these errors, to expand understanding of BP measurement challenges and increase clinical outcomes.

Consideration of small sample size

The simulation study conducted in Chapter 6 to validate the approximations from the proposed methodology, as presented in Table 31 in Chapter 6, highlight that the probability of acceptance values get more distant as the sample size decreases. This occurs due to a smaller sample size that reduces the propensity of the sampling distribution to approximate a normal distribution. These findings additionally highlight the need for careful consideration when employing smaller sample sizes, especially when $N < 39$. Further research must be conducted to expand the proposed framework for smaller sample sizes.

REFERENCES

[1] T. Gaziano, K. S. Reddy, F. Paccaud, S. Horton, and V. Chaturvedi, "Cardiovascular Disease," in *Disease Control Priorities in Developing Countries.*, 2nd ed., New York: Oxford University Press, 2006, ch. 33.

[2] P. Libby *et al.*, "Atherosclerosis," *Nat Rev Dis Primers*, vol. 5, no. 1, Dec. 2019, doi: 10.1038/s41572-019-0106-z.

[3] Heart Research Institute, "Cardiovascular disease: impacts and risks," NZ. Accessed: Oct. 15, 2024. [Online]. Available: https://www.hri.org.nz/health/learn/cardiovascular-disease/cardiovascular-disease-impacts-and-risks

[4] Institute for Health Metrics and Evaluation, "Causes of death, World, 2021." Accessed: Oct. 15, 2024. [Online]. Available: https://ourworldindata.org/grahttps://ourworldindata.org/grapher/annual-number-of-deaths-by-causepher/annual-number-of-deaths-by-cause

[5] F. Akter, N. Masood, S. Shakil Malik, S. Mubarik, and Y. Wang, "Prediction of cardiovascular diseases mortality- and disability-adjusted life-years attributed to modifiable dietary risk factors from 1990 to 2030 among East Asian countries and the world." [Online]. Available: http://ghdx.healthdata.org/gbd-results-tool

[6] C. Kreatsoulas and S. S. Anand, "The impact of social determinants on cardiovascular disease," *Can J Cardiol*, vol. 26, p. 8, 2009.

[7] WHO, "Cardiovascular diseases (CVDs)," Jun. 2021. Accessed: Oct. 15, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[8] WHO Team, "Monitoring health for the SDGs, Sustainable Development Goals." Accessed: Oct. 15, 2024. [Online]. Available: https://www.who.int/publications/i/item/9789240094703

[9] I. Ferreira-González, "The Epidemiology of Coronary Heart Disease," *Revista Española de Cardiología (English Edition)*, vol. 67, no. 2, pp. 139–144, Feb. 2014, doi: 10.1016/j.rec.2013.10.002.

[10] World Stroke Organization, "World Stroke Organization (WSO): Global Stroke Fact Sheet 2022." [Online]. Available: http://ghdx.healthdata.org/gbd-results-tool

[11] S. S. Martin *et al.*, "2024 Heart Disease and Stroke Statistics: A Report of US and Global Data from the American Heart Association," *Circulation*, vol. 149, no. 8, pp. E347–E913, Feb. 2024, doi: 10.1161/CIR.0000000000001209.

[12] World Economic Forum and Harvard School of Public Health, "The Global Economic Burden of Non-communicable Diseases," Geneva, 2011. [Online]. Available: www.weforum.org/EconomicsOfNCD

[13] S. Cohen, D. Janicki-Deverts, and G. E. Miller, "Psychological Stress and Disease," *JAMA*, vol. 298, no. 14, pp. 1685–1687, 2007.

[14] A. Narendrula, K. Ajani, J. Lang, E. Brinza, and C. T. Longenecker, “Psychological distress and health perception in patients with a previous myocardial infarction or stroke: a national cross-sectional study,” *BMC Cardiovasc Disord*, vol. 23, no. 1, Dec. 2023, doi: 10.1186/s12872-023-03422-5.

[15] New Zealand Ministry of Health, “Mortality 2016 data tables,” Wellington, Apr. 2020. Accessed: Oct. 16, 2024. [Online]. Available: https://www.health.govt.nz/publications/mortality-2016-data-tables

[16] New Zealand Ministry of Health, “Annual Update of Key Results 2017/18: New Zealand Health Survey,” Apr. 2019. Accessed: Oct. 16, 2024. [Online]. Available: https://www.health.govt.nz/publications/annual-update-of-key-results-201718-new-zealand-health-survey

[17] New Zealand Ministry of Health, “Annual Data Explorer 2022/2023: New Zealand Health Survey,” 2024. Accessed: Oct. 16, 2024. [Online]. Available: https://minhealthnz.shinyapps.io/nz-health-survey-2022-23-annual-data-explorer/_w_5a990bce/#!/home

[18] New Zealand Ministry of Health, “Mortality 2017 data tables,” Wellington, Dec. 2019. Accessed: Oct. 16, 2024. [Online]. Available: https://www.health.govt.nz/publications/mortality-2017-data-tables

[19] D. Adhikary, S. Barman, R. Ranjan, and H. Stone, “A Systematic Review of Major Cardiovascular Risk Factors: A Growing Global Health Concern,” *Cureus*, Oct. 2022, doi: 10.7759/cureus.30119.

[20] R. H. Eckel, S. M. Grundy, and P. Z. Zimmet, “The metabolic syndrome,” in *Lancet*, Elsevier Limited, Apr. 2005, pp. 1415–1428. doi: 10.1016/S0140-6736(05)66378-7.

[21] World Health Organization, “Obesity and Overweight [Fact Sheet],” Mar. 2024. Accessed: Oct. 16, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

[22] PwC New Zealand, “The Economic and Social Cost of Type 2 Diabetes,” Wellington, 2021.

[23] New Zealand Ministry of Health, “Living well with diabetes : a plan for people at high risk of or living with diabetes, 2015-2020,” Ministry of Health, Oct. 2015.

[24] Best Practice Advocacy Centre, “A rising tide of type 2 diabetes in younger people: what can primary care do?,” 2021. [Online]. Available: www.bpac.org.nz

[25] J. Hoek, E. Ozarka, and P. Gendall, “The 5000 annual deaths caused by smoking also matter,” 2024. Accessed: Oct. 16, 2024. [Online]. Available: https://www.phcc.org.nz/briefing/5000-annual-deaths-caused-smoking-also-matter

[26] G. Gallucci, A. Tartarone, R. Lerose, A. V. Lalinga, and A. M. Capobianco, “Cardiovascular risk of smoking and benefits of smoking cessation,” Jul. 01, 2020, *AME Publishing Company*. doi: 10.21037/jtd.2020.02.47.

[27] C. Grey *et al.*, “Trends in ischaemic heart disease: patterns of hospitalisation and mortality rates differ by ethnicity (ANZACS-QI 21),” *NZMJ*, vol. 13, 2018, [Online]. Available: www.nzma.org.nz/journal

[28] V. Selak *et al.*, “Ethnic differences in cardiovascular risk profiles among 475,241 adults in primary care in Aotearoa, New Zealand,” *NZMJ*, vol. 133, p. 1521, 2020, [Online]. Available: www.nzma.org.nz/journal

[29] T. Blakely, G. Disney, J. Atkinson, A. Teng, and J. P. Mackenbach, “A Typology for Charting Socioeconomic Mortality Gradients: ‘go Southwest,’” *Epidemiology*, vol. 28, no. 4, pp. 594–603, Jul. 2017, doi: 10.1097/EDE.0000000000000671.

[30] Best Practice Advocacy Centre, “Hypertension in adults: the silent killer,” May 2024. Accessed: Oct. 21, 2024. [Online]. Available: https://bpac.org.nz/2023/hypertension.aspx

[31] N. Wilson, C. Cleghorn, N. Nghiem, and T. Blakely, “Prioritization of intervention domains to prevent cardiovascular disease: a country-level case study using global burden of disease and local data,” *Popul Health Metr*, vol. 21, no. 1, Dec. 2023, doi: 10.1186/s12963-023-00301-1.

[32] T. Blakely *et al.*, “Disease-related income and economic productivity loss in New Zealand: A longitudinal analysis of linked individual-level data,” *PLoS Med*, vol. 18, no. 11, Nov. 2021, doi: 10.1371/journal.pmed.1003848.

[33] G. K. Ghodeshwar, A. Dube, and D. Khobragade, “Impact of Lifestyle Modifications on Cardiovascular Health: A Narrative Review,” *Cureus*, Jul. 2023, doi: 10.7759/cureus.42616.

[34] M. J. Korhonen *et al.*, “Lifestyle Changes in Relation to Initiation of Antihypertensive and Lipid-Lowering Medication: A Cohort Study,” *J Am Heart Assoc*, vol. 9, no. 4, Feb. 2020, doi: 10.1161/JAHA.119.014168.

[35] P. Wilson, R. D’Agostino, D. Levy, A. Belanger, H. Silbershatz, and W. Kannel, “Prediction of coronary heart disease using risk factor categories,” *Circulation*, vol. 97, no. 18, pp. 1837–1847, 1998, doi: 10.1161/01.CIR.97.18.1837.

[36] D. C. Goff *et al.*, “2013 ACC/AHA guideline on the assessment of cardiovascular risk: A report of the American college of cardiology/American heart association task force on practice guidelines,” *J Am Coll Cardiol*, vol. 63, no. 25 PART B, pp. 2935–2959, 2014, doi: 10.1016/j.jacc.2013.11.005.

[37] R. M. Conroy *et al.*, “Estimation of ten-year risk of fatal cardiovascular disease in Europe: The SCORE project,” *Eur Heart J*, vol. 24, no. 11, pp. 987–1003, Jun. 2003, doi: 10.1016/S0195-668X(03)00114-3.

[38] R. Pylypchuk *et al.*, “Cardiovascular disease risk prediction equations in 400 000 primary care patients in New Zealand: a derivation and validation study,” *The Lancet*, vol. 391, no. 10133, pp. 1897–1907, 2018, doi: 10.1016/S0140-6736(18)30664-0.

[39] A. P. Kengne *et al.*, “Blood pressure variables and cardiovascular risk: New findings from advance,” *Hypertension*, vol. 54, no. 2, pp. 399–404, 2009, doi: 10.1161/HYPERTENSIONAHA.109.133041.

[40] D. Tran, N. Lekhak, K. Gutierrez, and S. Moonie, “Risk factors associated with cardiovascular disease among adult Nevadans,” *PLoS One*, vol. 16, no. 2 February, pp. 1–11, 2021, doi: 10.1371/journal.pone.0247105.

[41] D. Gu *et al.*, “Blood pressure and risk of cardiovascular disease in chinese men and women,” *Am J Hypertens*, vol. 21, no. 3, pp. 265–272, 2008, doi: 10.1038/ajh.2007.59.

[42] J. J. Mourad, “The evolution of systolic blood pressure as a strong predictor of cardiovascular risk and the effectiveness of fixed-dose ARB/CCB combinations in lowering levels of this

preferential target," *Vasc Health Risk Manag*, vol. 4, no. 6, pp. 1315–1325, 2008, doi: 10.2147/vhrm.s4073.

[43] B. Kaur, S. Kaur, L. N. Yaddanapudi, and N. V. Singh, "Comparison between invasive and noninvasive blood pressure measurements in critically ill patients receiving inotropes," *Blood Press Monit*, vol. 24, no. 1, pp. 24–29, Feb. 2019, doi: 10.1097/MBP.0000000000000358.

[44] B. H. McGhee and E. J. Bridges, "Monitoring arterial blood pressure: what you may not know," *Crit Care Nurse*, vol. 22, no. 2, pp. 60–64, 66, Apr. 2002, doi: 10.4037/ccn2002.22.2.60.

[45] E. P. Mccutxheon and R. F. Rushmer, "Korotkoff Sounds," *Circ Res*, no. 2, 1967, [Online]. Available: http://ahajournals.org

[46] M. Sharma *et al.*, "Cuff-Less and Continuous Blood Pressure Monitoring: A Methodological Review," Jun. 01, 2017, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/technologies5020021.

[47] Y. Perera, J. Raitt, K. Poole, D. Metcalfe, and A. Lewinsohn, "Non-invasive versus arterial pressure monitoring in the pre-hospital critical care environment: a paired comparison of concurrently recorded measurements," *Scand J Trauma Resusc Emerg Med*, vol. 32, no. 1, Dec. 2024, doi: 10.1186/s13049-024-01240-y.

[48] H. Zheng, "Hypertension in Obesity: Impact of Miscuffing in Blood Pressure Measurement," Jan. 01, 2024, *Elsevier Inc.* doi: 10.1016/j.amjmed.2024.08.028.

[49] International Organization for Standardization (ISO), "ISO 81060-2:2018 – Non-invasive sphygmomanometers – Part 2: Clinical investigation of automated measurement type," Geneva, Switzerland, 2018.

[50] G. Ogedegbe and T. Pickering, "Principles and Techniques of Blood Pressure Measurement," *Cardiol Clin*, vol. 28, no. 4, pp. 571–586, Nov. 2010, doi: 10.1016/j.ccl.2010.07.006.

[51] N. Kallioinen, A. Hill, M. S. Horswill, H. E. Ward, and M. O. Watson, "Sources of inaccuracy in the measurement of adult patients' resting blood pressure in clinical settings: A systematic review," 2017, *Lippincott Williams and Wilkins*. doi: 10.1097/HJH.0000000000001197.

[52] K. E. Foti, L. J. Appel, K. Matsushita, J. Coresh, G. Caleb Alexander, and E. Selvin, "Digit Preference in Office Blood Pressure Measurements, United States 2015–2019," *Am J Hypertens*, vol. 34, no. 5, pp. 521–530, May 2021, doi: 10.1093/ajh/hpaa196.

[53] Ministry of Health, *Cardiovascular Disease Risk Assessment and Management for Primary Care.*, vol. 8, no. 1. 2018.

[54] E. O'Brien and G. S. Stergiou, "The pursuit of accurate blood pressure measurement: A 35-year travail," Aug. 01, 2017, *Blackwell Publishing Inc.* doi: 10.1111/jch.13005.

[55] G. S. Stergiou, B. S. Alpert, S. Mieke, J. Wang, and E. O'Brien, "Validation protocols for blood pressure measuring devices in the 21st century," *J Clin Hypertens*, vol. 20, no. 7, pp. 1096–1099, 2018, doi: 10.1111/jch.13294.

[56] S. A. Noh, H. S. Kim, S. H. Kang, C. H. Yoon, T. J. Youn, and I. H. Chae, "History and evolution of blood pressure measurement," Dec. 01, 2024, *BioMed Central Ltd*. doi: 10.1186/s40885-024-00268-7.

[57] J. E. Sharman *et al.*, “The urgency to regulate validation of automated blood pressure measuring devices: a policy statement and call to action from the world hypertension league,” Feb. 01, 2023, *Springer Nature*. doi: 10.1038/s41371-022-00747-0.

[58] S. W. Hee *et al.*, “Does the low prevalence affect the sample size of interventional clinical trials of rare diseases? An analysis of data from the aggregate analysis of clinicaltrials.gov,” *Orphanet J Rare Dis*, vol. 12, no. 1, Mar. 2017, doi: 10.1186/s13023-017-0597-1.

[59] S. Eloranta and M. Boman, “Predictive models for clinical decision making: Deep dives in practical machine learning,” Aug. 01, 2022, *John Wiley and Sons Inc*. doi: 10.1111/joim.13483.

[60] A. Laupacis, N. Sekar, and I. G. Stiell, “Clinical prediction rules: A review and suggested modifications of methodological standards,” *J Am Med Assoc*, vol. 277, no. 6, pp. 488–494, 1997, doi: 10.1001/jama.277.6.488.

[61] New Zealand Ministry of Health, “The assessment and management of cardiovascular risk,” Wellington, 2003. [Online]. Available: www.nzgg.org.nz

[62] K. G. M. Moons *et al.*, “Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker,” *Heart*, vol. 98, no. 9, pp. 683–690, 2012, doi: 10.1136/heartjnl-2011-301246.

[63] L. Ruilope, L. Queveo-Aguado, C. Fernández-Labandera, P. Martínez-Muñoz, C. Catalina-Romero, and Fernández-Meseguer, “Application of an expert system to improve cardiovascular risk in workplace,” *J Hypertens*, p. e292, 2018, doi: 10.1097/01.hjh.0000539853.20878.ce.

[64] David G. Kleinbaum, L. L. Kupper, A. Nizam, and E. S. Rosenberg, *Applied Regression Analysis and Other Multivariable Methods*. Cengage Learning, 2014.

[65] J. Kaur and B. S. Khehra, “Fuzzy Logic and Hybrid based Approaches for the Risk of Heart Disease Detection: State-of-the-Art Review,” *Journal of The Institution of Engineers (India): Series B*, vol. 103, no. 2, pp. 681–697, 2022, doi: 10.1007/s40031-021-00644-z.

[66] B. Martins, D. Ferreira, C. Neto, A. Abelha, and J. Machado, “Data Mining for Cardiovascular Disease Prediction,” *J Med Syst*, vol. 45, no. 1, pp. 1–8, 2021, doi: 10.1007/s10916-020-01682-8.

[67] K. Anderson, P. Odell, P. F Wilson, and W. Kannel, “Cardiovascular disease risk profiles,” *Am Heart J*, vol. 121, no. 1, pp. 293–298, 1991.

[68] ACC/AHA, “2013 ACC/AHA Guideline on the Assessment of Cardiovascular Risk,” *Circulation*, pp. S49–S73, 2014, doi: 10.1161/01.cir.0000437741.48606.98.

[69] S. S. Khan *et al.*, “Development and Validation of the American Heart Association’s PREVENT Equations,” *Circulation*, vol. 149, no. 6, pp. 430–449, Feb. 2024, doi: 10.1161/CIRCULATIONAHA.123.067626.

[70] J. Hippisley-Cox, C. Coupland, Y. Vinogradova, J. Robson, M. May, and P. Brindle, “Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: Prospective open cohort study,” *Br Med J*, vol. 335, no. 7611, pp. 136–141, Jul. 2007, doi: 10.1136/bmj.39261.471806.55.

[71] K. N. Karmali and D. M. Lloyd-Jones, “Implementing Cardiovascular Risk Prediction in Clinical Practice: The Future Is Now,” *J Am Heart Assoc*, vol. 6, no. 4, pp. 1–3, 2017, doi: 10.1161/JAHA.117.006019.

[72] R. B. D'Agostino *et al.*, "General cardiovascular risk profile for use in primary care: The Framingham heart study," *Circulation*, vol. 117, no. 6, pp. 743–753, 2008, doi: 10.1161/CIRCULATIONAHA.107.699579.

[73] N. R. Cook, P. M. Ridker, and W. Hospital, "Under-Ascertainment in the Women ' s Health Study," *JAMA Intern Med*, vol. 174, no. 12, pp. 1964–1971, 2015, doi: 10.1001/jamainternmed.2014.5336.Further.

[74] Y. Huang and C.-Y. Wang, "Cox Regression with Dependent Error in Covariates," *Biometrics*, vol. 74, no. 1, pp. 118–126, 2018, doi: 10.1111/biom.12741.Cox.

[75] A. C. Razavi *et al.*, "PREVENT Equations: A New Era in Cardiovascular Disease Risk Assessment," *Circ Cardiovasc Qual Outcomes*, vol. 17, no. 4, p. E010763, Apr. 2024, doi: 10.1161/CIRCOUTCOMES.123.010763.

[76] M. Kavousi *et al.*, "Comparison of application of the ACC/AHA guidelines, Adult Treatment Panel III guidelines, and European Society of Cardiology guidelines for cardiovascular disease prevention in a European cohort," *JAMA - Journal of the American Medical Association*, vol. 311, no. 14, pp. 1416–1423, 2014, doi: 10.1001/jama.2014.2632.

[77] I. M. Graham *et al.*, "Systematic Coronary Risk Evaluation (SCORE): JACC Focus Seminar 4/8," Jun. 22, 2021, *Elsevier Inc.* doi: 10.1016/j.jacc.2021.04.052.

[78] Cox Julia Hippisley, Vingradova Yana, Coupland Carol, and Pringle Mike, "Comparison of key practice characteristics between general practices in England and Wales and general practices in the QRESEARCH database.," *QRESEARCH*, Nov. 2005.

[79] S. Wells *et al.*, "Cohort Profile: The PREDICT cardiovascular disease cohort in New Zealand primary care (PREDICT-CVD 19)," *Int J Epidemiol*, vol. 46, no. 1, 2017, doi: 10.1093/ije/dyv312.

[80] C. Schilling *et al.*, "Australian general practitioners initiate statin therapy primarily on the basis of lipid levels; New Zealand general practitioners use absolute risk," *Health Policy (New York)*, vol. 121, no. 12, pp. 1233–1239, 2017, doi: 10.1016/j.healthpol.2017.09.022.

[81] G. Devlin *et al.*, *HISO 10071:2019 Cardiovascular Disease Risk Assessment Data Standard*. *Wellington: Ministry of Health*, vol. 5013. 2019.

[82] New Zealand Ministry of Health, "Cardiovascular Disease Risk Assessment and Management for Primary Care.," Wellington, 2018.

[83] R. Pylypchuk *et al.*, "Cardiovascular disease risk prediction equations in 400 000 primary care patients in New Zealand: a derivation and validation study," *The Lancet*, vol. 391, no. 10133, pp. 1897–1907, 2018, doi: 10.1016/S0140-6736(18)30664-0.

[84] National Heart Lung and Blood Institute, "What Are the Signs and Symptoms of Coronary Heart Disease?," Sep. 2015, Accessed: Dec. 23, 2024. [Online]. Available: https://web.archive.org/web/20150224034615/http://www.nhlbi.nih.gov/health/health-topics/topics/cad/signs

[85] GBD 2015 Disease and Injury Incidence and Prevalence Collaborators, "Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015," *The Lancet*, vol. 388, pp. 1545–1602, Oct. 2016.

[86] H. D. White and D. P. Chew, "Acute myocardial infarction," *Lancet*, vol. 389, no. 10065, pp. 197–210, 2008, doi: 10.1016/S0140-6736(16)30677-8.

[87] J. L. Anderson, “2013 ACCF/AHA guideline for the management of ST-elevation myocardial infarction: A report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines,” Jan. 29, 2013. doi: 10.1161/CIR.0b013e3182742cf6.

[88] E. A. Amsterdam *et al.*, “2014 AHA/ACC guideline for the management of patients with non-st-elevation acute coronary syndromes: A report of the American college of cardiology/American heart association task force on practice guidelines,” Dec. 23, 2014, *Lippincott Williams and Wilkins*. doi: 10.1161/CIR.0000000000000134.

[89] K. Thygesen *et al.*, “Third Universal Definition of Myocardial Infarction,” *Eur Heart J*, vol. 126, pp. 2020–2035, 2012, doi: 10.1161/CIR.0b013e31826e1058/-/DC1.

[90] T. A. Rousan and U. Thadani, “Stable angina medical therapy management guidelines: A critical review of guidelines from the European Society of Cardiology and National Institute for Health and Care Excellence,” 2019, *Radcliffe Cardiology*. doi: 10.15420/ecr.2018.26.1.

[91] Yoshiki Kusama, Eitaro Kodani, Akihiro Nakagomi, and Toshiaki Otsuka, “Variant Angina and Coronary Artery Spasm: The Clinical Spectrum, Pathophysiology, and Management,” *Journal of Nippon Medical School*, vol. 78, no. 1, 2011.

[92] S. K. Feske, “Ischemic Stroke,” Dec. 01, 2021, *Elsevier Inc.* doi: 10.1016/j.amjmed.2021.07.027.

[93] S. Chen, L. Zeng, and Z. Hu, “Progressing haemorrhagic stroke: categories, causes, mechanisms and managements,” Nov. 01, 2014, *Dr. Dietrich Steinkopff Verlag GmbH and Co. KG*. doi: 10.1007/s00415-014-7291-1.

[94] M. A. Ikram, R. G. Wieberdink, and P. J. Koudstaal, “International epidemiology of intracerebral hemorrhage,” *Curr Atheroscler Rep*, vol. 14, no. 4, pp. 300–306, Aug. 2012, doi: 10.1007/s11883-012-0252-1.

[95] J. D. Easton *et al.*, “Definition and evaluation of transient ischemic attack: A scientific statement for healthcare professionals from the American heart association/American stroke association stroke council; council on cardiovascular surgery and anesthesia; council on cardiovascular radiology and intervention; council on cardiovascular nursing; and the interdisciplinary council on peripheral vascular disease,” *Stroke*, vol. 40, no. 6, pp. 2276–2293, Jun. 2009, doi: 10.1161/STROKEAHA.108.192218.

[96] J. Nordanstig *et al.*, “Peripheral arterial disease (PAD) – A challenging manifestation of atherosclerosis,” Jun. 01, 2023, *Academic Press Inc.* doi: 10.1016/j.ypmed.2023.107489.

[97] M. A. Allison *et al.*, “Health Disparities in Peripheral Artery Disease: A Scientific Statement from the American Heart Association,” Jul. 18, 2023, *Lippincott Williams and Wilkins*. doi: 10.1161/CIR.0000000000001153.

[98] M. Walli-Attaei *et al.*, “Characteristics, management, and outcomes in women and men with congestive heart failure in 40 countries at different economic levels: an analysis from the Global Congestive Heart Failure (G-CHF) registry,” *Lancet Glob Health*, vol. 12, no. 3, pp. e396–e405, Mar. 2024, doi: 10.1016/S2214-109X(23)00557-0.

[99] J. Damen *et al.*, “Prediction models for cardiovascular disease risk in the general population: Systematic review,” *BMJ (Online)*, vol. 353, 2016, doi: 10.1136/bmj.i2416.

[100] D. R. Cox, “Regression Models and Life-Tables,” *Journal of the Royal Statistical Society.*, vol. 34, no. 2, pp. 187–220, 1972.

[101] Dirk F. Moore, *Applied Survival Analysis Using R*. Cham, Switzerland: Springer, 2016.

[102] B. Van Calster, V. Van Belle, Y. Vergouwe, and E. W. Steyerberg, "Discrimination ability of prediction models for ordinal outcomes: Relationships between existing measures and a new measure," *Biometrical Journal*, vol. 54, no. 5, pp. 674–685, 2012, doi: 10.1002/bimj.201200026.

[103] Z. Zhang, "Semi-parametric regression model for survival data: Graphical visualization with R," *Ann Transl Med*, vol. 4, no. 23, pp. 1–8, 2016, doi: 10.21037/atm.2016.08.61.

[104] J. B. Echouffo-Tcheugui, G. D. Batty, M. Kivimäki, and A. P. Kengne, "Risk Models to Predict Hypertension: A Systematic Review," *PLoS One*, vol. 8, no. 7, Jul. 2013, doi: 10.1371/journal.pone.0067370.

[105] Heart Research Institute NZ, "High blood pressure: definition, causes and symptoms." Accessed: Dec. 26, 2024. [Online]. Available: https://www.hri.org.nz/health/learn/risk-factors/high-blood-pressure

[106] S. S. Mousavi, M. A. Reyna, G. D. Clifford, and R. Sameni, "A Survey on Blood Pressure Measurement Technologies: Addressing Potential Sources of Bias," *Sensors*, vol. 24, no. 6, pp. 1–33, 2024, doi: 10.3390/s24061730.

[107] G. Pasquale Passarella, Tatiana A. Kiseleva, Farida V. Valeeva, "Hypertension Management in Diabetes: 2018 Update," *Diabetes Spectrum*, vol. 31, no. 3, pp. 218–224, 2018.

[108] A. Grossman and E. Grossman, "Blood pressure control in type 2 diabetic patients," *Cardiovasc Diabetol*, vol. 16, no. 1, pp. 1–15, 2017, doi: 10.1186/s12933-016-0485-3.

[109] Y. Xu, J. Moller, R. Wang, and Y. Liang, "Life-course blood pressure trajectories and cardiovascular diseases: A population-based cohort study in China," *PLoS One*, vol. 15, no. 10 October, pp. 1–9, 2020, doi: 10.1371/journal.pone.0240804.

[110] G. A. Mensah *et al.*, "Decline in Cardiovascular Mortality: Possible Causes and Implications," Jan. 20, 2017, *Lippincott Williams and Wilkins*. doi: 10.1161/CIRCRESAHA.116.309115.

[111] P. Froom and U. Goldbourt, "Secular decrease in blood pressure and reduction in mortality from cardiovascular disease in Israeli workers," *J Hum Hypertens*, vol. 18, no. 2, pp. 113–118, 2004, doi: 10.1038/sj.jhh.1001645.

[112] D. M. Lloyd-Jones, J. C. Evans, M. G. Larson, C. J. O'Donnell, and D. Levy, "Differential impact of systolic and diastolic blood pressure level on JNC-VI staging," *Hypertension*, vol. 34, no. 3, pp. 381–385, 1999, doi: 10.1161/01.HYP.34.3.381.

[113] E. J. Wallis, L. E. Ramsay, and P. R. Jackson, "Cardiovascular and coronary risk estimation in hypertension management," *Heart*, vol. 88, no. 3, pp. 306–312, 2002, doi: 10.1136/heart.88.3.306.

[114] S. de Lusignan, J. Belsey, N. Hague, and B. Dzregah, "End-digit preference in blood pressure recordings of patients with ischaemic heart disease in primary care," *J Hum Hypertens*, vol. 18, no. 4, pp. 261–265, Apr. 2004, doi: 10.1038/sj.jhh.1001663.

[115] S. Thavarajah, W. B. White, and G. A. Mansoor, "Terminal digit bias in a specialty hypertension faculty practice," *J Hum Hypertens*, vol. 17, no. 12, pp. 819–822, 2003, doi: 10.1038/sj.jhh.1001625.

[116] Prasad N, Wheeldon N M, and MacDonald T M, “Evaluating the use of a semiautomated cuff-oscillometric sphygmomanometer in the hypertension clinic,” *British Journal of Clinical Practice*, vol. 48, no. 6, pp. 307–309, 1994.

[117] P. Muntner *et al.*, “Blood Pressure Assessment in Adults in Clinical Practice and Clinic-Based Research: JACC Scientific Expert Panel,” Jan. 29, 2019, *Elsevier USA*. doi: 10.1016/j.jacc.2018.10.069.

[118] P. J. Nietert, A. M. Wessell, C. Feifer, and S. M. Ornstein, “Effect of terminal digit preference on blood pressure measurement and treatment in primary care,” *Am J Hypertens*, vol. 19, no. 2, pp. 147–152, Feb. 2006, doi: 10.1016/j.amjhyper.2005.08.016.

[119] E. Kim *et al.*, “End-digit preference and the quality of blood pressure monitoring in diabetic adults,” *Diabetes Care*, vol. 30, no. 8, pp. 1959–1963, Aug. 2007, doi: 10.2337/dc07-0020.

[120] J. W. Graves *et al.*, “The Impact of Observer and Patient Factors on the Occurrence of Digit Preference for Zero in Blood Pressure Measurement in a Hypertension Specialty Clinic: Evidence for the Need of Continued Observation,” *Am J Hypertens*, vol. 19, no. 6, pp. 567–572, Jun. 2006, doi: 10.1016/j.amjhyper.2005.04.004.

[121] Alsanjari O N *et al.*, “Trends and transient change in end-digit preference in blood pressure recording: studies of sequential and longitudinal collected primary care data,” *Int J Clin Pract*, vol. 66, no. 1, pp. 37–43, 2012.

[122] M. Greiver *et al.*, “Trends in end digit preference for blood pressure and associations with cardiovascular outcomes in Canadian and UK primary care: A retrospective observational study,” *BMJ Open*, vol. 9, no. 1, Jan. 2019, doi: 10.1136/bmjopen-2018-024970.

[123] B. Mishra, N. D. Sinha, H. Gidwani, S. K. Shukla, A. Kawatra, and S. C. Mehta, “Equipment errors: A prevalent cause for fallacy in blood pressure recording-A point prevalence estimate from an indian health university,” *Indian Journal of Community Medicine*, vol. 38, no. 1, pp. 15–21, 2013, doi: 10.4103/0970-0218.106622.

[124] M. Forouzanfar, H. R. Dajani, V. Z. Groza, M. Bolic, S. Rajan, and I. Batkin, “Oscillometric blood pressure estimation: Past, present, and future,” *IEEE Rev Biomed Eng*, vol. 8, pp. 44–63, 2015, doi: 10.1109/RBME.2015.2434215.

[125] Martin Proença *et al.*, “Pulse Wave Analysis Techniques,” in *The Handbook of Cuffless Blood Pressure Monitoring*, 2019, pp. 107–137.

[126] R. Mukkamala *et al.*, “Toward Ubiquitous Blood Pressure Monitoring via Pulse Transit Time: Theory and Practice,” *IEEE Trans Biomed Eng*, vol. 62, no. 8, pp. 1879–1901, Aug. 2015, doi: 10.1109/TBME.2015.2441951.

[127] Q. Liu, C. Yang, S. Yang, C. F. Kwong, J. Wang, and N. Zhou, “Photoplethysmography-based non-invasive blood pressure monitoring via ensemble model and imbalanced dataset processing,” *Phys Eng Sci Med*, 2024, doi: 10.1007/s13246-024-01445-6.

[128] J. Kyung *et al.*, “Deep-learning-based blood pressure estimation using multi channel photoplethysmogram and finger pressure with attention mechanism,” *Sci Rep*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-36068-6.

[129] G. S. Stergiou *et al.*, “A universal standard for the validation of blood pressure measuring devices: Association for the Advancement of Medical Instrumentation/European Society of Hypertension/International Organization for Standardization (AAMI/ESH/ISO) Collaboration

Statement," *J Hypertens*, vol. 36, no. 3, pp. 472–478, 2018, doi: 10.1097/HJH.0000000000001634.

[130] J. E. Sharman *et al.*, "Lancet Commission on Hypertension group position statement on the global improvement of accuracy standards for devices that measure blood pressure," *J Hypertens*, vol. 38, no. 1, pp. 21–29, Jan. 2020, doi: 10.1097/HJH.0000000000002246.

[131] B. Williams *et al.*, "Guidelines for management of hypertension: Report of the fourth working party of the Bristish Hypertension Society, 2004 - BHS IV," Mar. 2004. doi: 10.1038/sj.jhh.1001683.

[132] E. O'Brien *et al.*, "European Society of Hypertension International Protocol revision 2010 for the validation of blood pressure measuring devices in adults," *Blood Press Monit*, vol. 15, no. 1, pp. 23–38, 2010, doi: 10.1097/MBP.0b013e3283360e98.

[133] ANSI/AAMI SP10, "American National Standard for Manual, electronic, or automated sphygmomanometers," *Ansi/Aami Sp10:2002/(R)2008*, 2008.

[134] G. S. Stergiou, B. S. Alpert, S. Mieke, J. Wang, and E. O'Brien, "Validation protocols for blood pressure measuring devices in the 21st century," *J Clin Hypertens*, vol. 20, no. 7, pp. 1096–1099, 2018, doi: 10.1111/jch.13294.

[135] A. de Greeff, I. Lorde, A. Wilton, P. Seed, A. J. Coleman, and A. H. Shennan, "Calibration accuracy of hospital-based non-invasive blood pressure measuring devices," *J Hum Hypertens*, vol. 24, no. 1, pp. 58–63, 2010, doi: 10.1038/jhh.2009.29.

[136] D. Mion and A. M. G. Pierin, "How accurate are sphygmomanometers?," *J Hum Hypertens*, vol. 12, no. 4, pp. 245–248, 1998, doi: 10.1038/sj.jhh.1000589.

[137] Q. Wang, H. Zhao, W. Chen, N. Li, and Y. Wan, "Validation of the iHealth BP7 wrist blood pressure monitor, for self-measurement, according to the European Society of Hypertension International Protocol revision 2010," *Blood Press Monit*, vol. 19, no. 1, pp. 54–57, 2014, doi: 10.1097/MBP.0000000000000017.

[138] H. W. Lui and K. L. Chow, "A Novel Calibration Procedure of Pulse Transit Time based Blood Pressure measurement with Heart Rate and Respiratory Rate," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2018-July, pp. 4318–4322, 2018, doi: 10.1109/EMBC.2018.8513364.

[139] X. Ding, B. P. Yan, Y. T. Zhang, J. Liu, N. Zhao, and H. K. Tsang, "Pulse Transit Time Based Continuous Cuffless Blood Pressure Estimation: A New Extension and A Comprehensive Evaluation," *Sci Rep*, vol. 7, no. 1, pp. 1–11, 2017, doi: 10.1038/s41598-017-11507-3.

[140] P. S. Pandian *et al.*, "Smart Vest: Wearable multi-parameter remote physiological monitoring system," *Med Eng Phys*, vol. 30, no. 4, pp. 466–477, 2008, doi: 10.1016/j.medengphy.2007.05.014.

[141] A. Esmaili, M. Kachuee, and M. Shabany, "Nonlinear Cuffless Blood Pressure Estimation of Healthy Subjects Using Pulse Transit Time and Arrival Time," *IEEE Trans Instrum Meas*, vol. 66, no. 12, pp. 3299–3308, Dec. 2017, doi: 10.1109/TIM.2017.2745081.

[142] M. Masè, W. Mattei, R. Cucino, L. Faes, and G. Nollo, "Feasibility of cuff-free measurement of systolic and diastolic arterial blood pressure," *J Electrocardiol*, vol. 44, no. 2, pp. 201–207, 2011, doi: 10.1016/j.jelectrocard.2010.11.019.

[143] T. Wibmer *et al.*, “Pulse transit time and blood pressure during cardiopulmonary exercise tests,” *Physiol Res*, vol. 63, no. 3, pp. 287–296, 2014, doi: 10.33549/physiolres.932581.

[144] Y. K. Jung, H. C. Baek, M. I. Soo, J. J. Myoung, Y. K. In, and I. K. Sun, “Comparative study on artificial neural network with multiple regressions for continuous estimation of blood pressure,” *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, vol. 7 VOLS, pp. 6942–6945, 2005, doi: 10.1109/iembs.2005.1616102.

[145] T. H. Huynh, R. Jafari, and W. Y. Chung, “Noninvasive cuffless blood pressure estimation using pulse transit time and impedance plethysmography,” *IEEE Trans Biomed Eng*, vol. 66, no. 4, pp. 967–976, 2019, doi: 10.1109/TBME.2018.2865751.

[146] T. H. Huynh, R. Jafari, and W. Y. Chung, “A robust bioimpedance structure for smartwatch-based blood pressure monitoring,” *Sensors (Switzerland)*, vol. 18, no. 7, 2018, doi: 10.3390/s18072095.

[147] Y. Li *et al.*, “Effect of cuff positioning on the accuracy of blood pressure measurement with automated electronic blood pressure monitors,” *J Clin Hypertens*, vol. 22, no. 7, pp. 1163–1172, Jul. 2020, doi: 10.1111/jch.13902.

[148] P. Sareen, K. Saxena, B. Sareen, and B. Taneja, “Comparison of arm and calf blood pressure,” *Journal of Anaesthesia*, vol. 56, no. 1, pp. 83–85, Jan. 2012.

[149] M. Kayrak *et al.*, “A comparison of blood pressure and pulse pressure values obtained by oscillometric and central measurements in hypertensive patients,” *Blood Press*, vol. 19, no. 2, pp. 98–103, 2010, doi: 10.3109/08037050903516318.

[150] G. Mancia *et al.*, “2013 Practice guidelines for the management of arterial hypertension of the European Society ofHypertension (ESH) and the European Society of Cardiology (ESC): ESH/ESCTask Force for the Management of Arterial Hypertension,” 2013, *Lippincott Williams and Wilkins*. doi: 10.1097/HJH.0b013e328364ca4c.

[151] E. Privšek, M. Hellgren, L. Råstam, U. Lindblad, and B. Daka, “Epidemiological and clinical implications of blood pressure measured in seated versus supine position,” *Medicine (United States)*, vol. 97, no. 31, Aug. 2018, doi: 10.1097/MD.0000000000011603.

[152] R. T. Netea, J. W. M. Lenders, P. Smits, and T. Thien, “Both body and arm position significantly influence blood pressure measurement,” *J Hum Hypertens*, vol. 17, no. 7, pp. 459–462, Jul. 2003, doi: 10.1038/sj.jhh.1001573.

[153] T. G. Pickering *et al.*, “Recommendations for blood pressure measurement in humans and experimental animals: Part 1: Blood pressure measurement in humans - A statement for professionals from the Subcommittee of Professional and Public Education of the American Heart Association Council on high blood pressure research,” *Circulation*, vol. 111, no. 5, pp. 697–716, Feb. 2005, doi: 10.1161/01.CIR.0000154900.76284.F6.

[154] S. D. Pierdomenico *et al.*, “Prognostic value of different indices of blood pressure variability in hypertensive patients,” *Am J Hypertens*, vol. 22, no. 8, pp. 842–847, Aug. 2009, doi: 10.1038/ajh.2009.103.

[155] K. Chachula, F. Lieb, F. Hess, J. Welter, N. Graf, and A. Dullenkopf, “Non-invasive continuous blood pressure monitoring (ClearSight™ system) during shoulder surgery in the beach chair position: a prospective self-controlled study,” *BMC Anesthesiol*, vol. 20, no. 1, Dec. 2020, doi: 10.1186/s12871-020-01185-6.

[156] G. Cicolini *et al.*, "Differences in blood pressure by body position (supine, fowler's, and sitting) in hypertensive subjects," *Am J Hypertens*, vol. 24, no. 10, pp. 1073–1079, Oct. 2011, doi: 10.1038/ajh.2011.106.

[157] Romana T. Netea, Paul Smits, Jacques W.M. Lenders, and Theo Thien, "Does it matter whether blood pressure measurements are taken with subjects sitting or supine?," *J Hypertens*, vol. 16, pp. 263–268, 1998.

[158] G. Mariotti *et al.*, "Arm position as a source of error in blood pressure measurement," *Clin Cardiol*, vol. 10, no. 10, pp. 591–593, 1987, doi: 10.1002/clc.4960101016.

[159] Foster-Fitzpatrick Lucille;, A. Ortiz, H. Sibilano, R. Marcantonio, and L. T. Braun, "The Effects of Crossed Leg on Blood Pressure Measurement," *Nurs Res*, vol. 48, no. 2, pp. 105–108, Mar. 1999.

[160] R. Netea, J. Lenders, P. Smits, and T. Thien, "Arm position is important for blood pressure measurement," *J Hum Hypertens*, vol. 13, pp. 105–109, 1999, [Online]. Available: http://www.stockton-press.co.uk/jhh

[161] R. Pinar, N. Sabuncu, and A. Oksay, "Effects of crossed leg on blood pressure," *Blood Press*, vol. 13, no. 4, pp. 252–254, 2004, doi: 10.1080/08037050410000903.

[162] S. S. Franklin, L. Thijs, T. W. Hansen, E. O'Brien, and J. A. Staessen, "White-coat hypertension new insights from recent studies," *Hypertension*, vol. 62, no. 6, pp. 982–987, 2013, doi: 10.1161/HYPERTENSIONAHA.113.01275.

[163] H. Badeli and F. Assadi, "Strategies to Reduce Pitfalls in Measuring Blood Pressure," 2014. [Online]. Available: www.ijpm.ir

[164] D. Xu *et al.*, "The Influence of Factors such as Anxiety on the White Coat Effect during the Treatment of Patients with Hypertension," *Rev Cardiovasc Med*, vol. 23, no. 11, Nov. 2022, doi: 10.31083/j.rcm2311359.

[165] N. D. Markandu, F. Whitcher, A. Arnold, and C. Carney, "The mercury sphygmomanometer should be abandoned before it is proscribed," *J Hum Hypertens*, vol. 14, no. 1, pp. 31–36, 2000, doi: 10.1038/sj.jhh.1000932.

[166] T. Pickering, H. E. John, A. J. Lawrence, and Falkner Bonita E, "Recommendations for Blood Pressure Measurement in Humans: An AHA Scientific Statement from the Council on High Blood Pressure Research Professional and Public Education Subcommittee," *The Journal of Clinical Hypertension*, vol. 7, no. 2, 2005.

[167] M. K. Rakotz *et al.*, "Medical students and measuring blood pressure: Results from the American Medical Association Blood Pressure Check Challenge," *J Clin Hypertens*, vol. 19, no. 6, pp. 614–619, Jun. 2017, doi: 10.1111/jch.13018.

[168] J. Ishigami, J. Charleston, E. R. Miller, K. Matsushita, L. J. Appel, and T. M. Brady, "Effects of Cuff Size on the Accuracy of Blood Pressure Readings: The Cuff(SZ) Randomized Crossover Trial," *JAMA Intern Med*, vol. 183, no. 10, pp. 1061–1068, Oct. 2023, doi: 10.1001/jamainternmed.2023.3264.

[169] S. Wright, R. F. Schneider, H. E. Ungerleider, and N. Y. New York, "Factors of Errors in Blood Pressure Readings.", *The American Heart Journal*, 1988.

[170] C. Y. Wu, H. Y. Hu, Y. J. Chou, N. Huang, Y. C. Chou, and C. P. Li, "High blood pressure and all-cause and cardiovascular disease mortalities in community-dwelling older adults," *Medicine (United States)*, vol. 94, no. 47, 2015, doi: 10.1097/MD.0000000000002160.

[171] J. Handler, “The Importance of Accurate Blood Pressure Measurement,” *Perm J*, vol. 13, no. 3, pp. 51–54, 2009, doi: 10.7812/tpp/09-054.

[172] A. S. Meidert and B. Saugel, “Techniques for non-invasive monitoring of arterial blood pressure,” *Front Med (Lausanne)*, vol. 4, pp. 1–6, 2017, doi: 10.3389/fmed.2017.00231.

[173] M. Gulati, L. A. Peterson, and A. Mihailidou, “Assessment of blood pressure skills and belief in clinical readings,” *Am J Prev Cardiol*, vol. 8, Dec. 2021, doi: 10.1016/j.ajpc.2021.100280.

[174] D. W. Jones, L. J. Appel, S. G. Sheps, E. J. Roccella, and C. Lenfant, “Measuring Blood Pressure Accurately New and Persistent Challenges.”, *JAMA*, vol. 289, no. 8, 2022.

[175] D. Zhou, B. Xi, M. Zhao, L. Wang, and S. P. Veeranki, “Uncontrolled hypertension increases risk of all-cause and cardiovascular disease mortality in US adults: The NHANES III Linked Mortality Study,” *Sci Rep*, vol. 8, no. 1, pp. 1–7, 2018, doi: 10.1038/s41598-018-27377-2.

[176] P. Muntner *et al.*, “Measurement of blood pressure in humans: A scientific statement from the american heart association,” *Hypertension*, vol. 73, no. 5, pp. E35–E66, May 2019, doi: 10.1161/HYP.0000000000000087.

[177] O. Tang *et al.*, “Comparison of automated clinical and research blood pressure measurements: Implications for clinical practice and trial design,” *J Clin Hypertens*, vol. 20, no. 12, pp. 1676–1682, Dec. 2018, doi: 10.1111/jch.13412.

[178] S. W. Wen, M. S. Kemmer, J. Hoey, J. A. Hanley, and R. H. Usher, “Terminal Digit Preference, Random Error and Bias in Routine Clinical Measurement of Blood Pressure”, 1993.

[179] D. Wingfield, G. K. Freeman, and C. J. Bulpitt, “Selective recording in blood pressure readings may increase subsequent mortality.”, *QJM: An International Journal of Medicine,* vol. 95, 2002.

[180] J. Broad, S. Wells, R. Marshall, and R. Jackson, “Zero end-digit preference in recorded blood pressure and its impact on classification of patients for pharmacologic management in primary care - PREDICT-CVD-6,” *British Journal of General Practice*, vol. 57, no. 544, pp. 897–903, 2007, doi: 10.3399/096016407782317964.

[181] O. E. Ayodele, E. O. Sanya, O. O. Okunola, and A. A. Akintunde, “End digit preference in blood pressure measurement in a hypertension specialty clinic in Southwest Nigeria,” *Cardiovasc J Afr*, vol. 23, no. 2, pp. 85–89, Mar. 2012, doi: 10.5830/CVJA-2011-045.

[182] B. Williams *et al.*, “Guidelines for management of hypertension: Report of the fourth working party of the Bristish Hypertension Society, 2004 - BHS IV,” Mar. 2004. doi: 10.1038/sj.jhh.1001683.

[183] P. J. Wood D, Wray R, Poulter N, Williams B, Kirby M, Patel V, Durrington P, Reckless J, Davis M, Sivers F, “Joint British Societies’ guidelines on prevention of cardiovascular disease in clinical practice.,” *Heart*, vol. 1–52, no. 91, 2005.

[184] P. Brindle, A. McConnachie, M. Upton, C. Hart, G. Smith, and G. Watt, “The accuracy of the Framingham risk-score in different socioeconomic groups: a prospective study,” *British Journal of General Practice*, vol. 55, pp. 838–845, 2005.

[185] Te whatu ora, “The assessment and management of cardiovascular risk,” New Zealand Guidelines Group, 2003. Accessed: Nov. 07, 2024. [Online]. Available: https://www.tewhatuora.govt.nz/assets/Publications/Cardiovascular-Publications/assessment-and-management-of-cardiovascular-risk-full.pdf

[186] B. Mihaylova *et al.*, “The effects of lowering LDL cholesterol with statin therapy in people at low risk of vascular disease: Meta-analysis of individual data from 27 randomised trials,” *The Lancet*, vol. 380, no. 9841, pp. 581–590, Aug. 2012, doi: 10.1016/S0140-6736(12)60367-5.

[187] The Blood Pressure Lowering Treatment Trialists’ Collaboration, “Blood pressure-lowering treatment based on cardiovascular risk: a meta-analysis of individual patient data,” *The Lancet*, pp. 591–598, 2014, Accessed: Nov. 14, 2024. [Online]. Available: www.thelancet.com

[188] J. E. Helmreich, *Regression Modeling Strategies with Applications to Linear Models, Logistic and Ordinal Regression and Survival Analysis (2nd Edition)*, vol. 70, no. Book Review 2. 2016. doi: 10.18637/jss.v070.b02.

[189] PHARMAC, “Cost Resource Manual,” Jan. 2020. Accessed: Oct. 21, 2024. [Online]. Available: https://pharmac.govt.nz/medicine-funding-and-supply/the-funding-process/policies-manuals-and-processes/economic-analysis/cost-resource-manual#s12

[190] PHARMAC, “Community Schedule,” Oct. 2024. Accessed: Oct. 21, 2024. [Online]. Available: https://schedule.pharmac.govt.nz/2024/10/01/Schedule.pdf

[191] Best Practice Advocacy Centre, “Oral anticoagulant selection in primary care,” Nov. 2023. Accessed: Oct. 21, 2024. [Online]. Available: https://bpac.org.nz/2023/docs/anticoagulants.pdf

[192] Best Practice Advocacy Centre, “Prescribing statins to reduce cardiovascular risk,” May 2021. Accessed: Oct. 21, 2024. [Online]. Available: https://bpac.org.nz/2021/statins.aspx

[193] Best Practice Advocacy Centre, “The use of ANTITHROMBOTICS MEDICINES in general practice,” 2011. Accessed: Oct. 21, 2024. [Online]. Available: https://bpac.org.nz/bpj/2011/october/docs/bpj_39_antithrombotic_pages_10-21.pdf

[194] D. Canoy *et al.*, “Antihypertensive drug effects on long-term blood pressure: an individual-level data meta-analysis of randomised clinical trials,” *Heart*, vol. 108, no. 16, pp. 1281–1289, 2022, doi: 10.1136/heartjnl-2021-320171.

[195] H. Lu, D. M. Jin, M. Liu, Y. M. Liu, J. F. Wang, and D. F. Geng, “Long-term efficacy and safety of statin treatment beyond six years: A meta-analysis of randomized controlled trials with extended follow-up,” *Pharmacol Res*, vol. 81, pp. 64–73, 2014, doi: 10.1016/j.phrs.2014.02.006.

[196] Best Practice Advocacy Centre, “An update on statins,” 2010. Accessed: Oct. 21, 2024. [Online]. Available: https://bpac.org.nz/bpj/2010/august/statins.aspx

[197] Best Practice Advocacy Centre, “Cardiovascular disease risk assessment in primary care: managing lipids,” Oct. 2018. Accessed: Oct. 21, 2024. [Online]. Available: https://bpac.org.nz/2018/docs/lipids.pdf

[198] S. M. Hamrahian, O. H. Maarouf, and T. Fülöp, “A Critical Review of Medication Adherence in Hypertension: Barriers and Facilitators Clinicians Should Consider,” *Patient Prefer Adherence*, vol. 16, no. September, pp. 2749–2757, 2022, doi: 10.2147/PPA.S368784.

[199] G. Beevers, G. Y. H. Lip, and E. O’brien, “Blood pressure measurement: Part II—Conventional sphygmomanometry: Technique of auscultatory blood pressure measurement,” *BMJ*, vol. 322, no. 7293, p. 1043, Apr. 2001, doi: 10.1136/bmj.322.7293.1043.

[200] ANSI/AAMI SP10, “American National Standard for Manual, electronic, or automated sphygmomanometers,” *Ansi/Aami Sp10:2002/(R)2008*, 2008.

[201] Y. Ostchega, R. J. Prineas, R. Paulose-Ram, C. M. Grim, G. Willard, and D. Collins, “National Health and Nutrition Examination Survey 1999-2000: Effect of observer training and protocol standardization on reducing blood pressure measurement error,” *J Clin Epidemiol*, vol. 56, no. 8, pp. 768–774, Aug. 2003, doi: 10.1016/S0895-4356(03)00085-4.

[202] K. N. Karmali and D. M. Lloyd-Jones, “Global risk assessment to guide blood pressure management in cardiovascular disease prevention,” *Hypertension*, vol. 69, no. 3, pp. e2–e9, Mar. 2017, doi: 10.1161/HYPERTENSIONAHA.116.08249.

[203] R. S. Armstrong, “Nurses’ knowledge of error in blood pressure measurement technique,” *Armstrong RS. International Journal of Nursing Practice*, vol. 8, pp. 118–126, 2002, doi: 10.1046/j.1322-7114.2002.00348.x.

[204] M. A. E. F. M. D. Badawy *et al.*, “Evaluation of cardiovascular diseases risk calculators for CVDs prevention and management: scoping review,” *BMC Public Health*, vol. 22, no. 1, pp. 1–11, 2022, doi: 10.1186/s12889-022-13944-w.

[205] I. García-Fernández-Bravo, A. Torres-Do-Rego, A. López-Farré, F. Galeano-Valle, P. Demelo-Rodriguez, and L. A. Alvarez-Sala-Walther, “Undertreatment or Overtreatment With Statins: Where Are We?,” *Front Cardiovasc Med*, vol. 9, no. April, pp. 1–16, 2022, doi: 10.3389/fcvm.2022.808712.

[206] A. P. Bress *et al.*, “Potential Cardiovascular Disease Events Prevented with Adoption of the 2017 American College of Cardiology/American Heart Association Blood Pressure Guideline,” *Circulation*, vol. 139, no. 1, pp. 24–36, Jan. 2019, doi: 10.1161/CIRCULATIONAHA.118.035640.

[207] C. Chen, X. Li, Y. Su, Z. You, R. Wan, and K. Hong, “Adherence with cardiovascular medications and the outcomes in patients with coronary arterial disease: ‘Real-world’ evidence,” *Clin Cardiol*, vol. 45, no. 12, pp. 1220–1228, 2022, doi: 10.1002/clc.23898.

[208] T. G. Pickering *et al.*, “Recommendations for blood pressure measurement in humans: an AHA scientific statement from the Council on High Blood Pressure Research Professional and Public Education Subcommittee.,” *The Journal of Clinical Hypertension*, vol. 7, no. 2, pp. 102–109, 2005, doi: 10.1111/j.1524-6175.2005.04377.x.

[209] J. He and P. K. Whelton, “Elevated systolic blood pressure and risk of cardiovascular and renal disease: Overview of evidence from observational epidemiologic studies and randomized controlled trials,” *Am Heart J*, vol. 138, no. 3 II, pp. 12–17, 1999, doi: 10.1016/s0002-8703(99)70312-1.

[210] F. U. S. Mattace-Raso *et al.*, “Blood pressure components and cardiovascular events in older adults: The Rotterdam Study,” *J Am Geriatr Soc*, vol. 52, no. 9, pp. 1538–1542, 2004, doi: 10.1111/j.1532-5415.2004.52419.x.

[211] S. S. Mousavi, M. A. Reyna, G. D. Clifford, and R. Sameni, “A Survey on Blood Pressure Measurement Technologies: Addressing Potential Sources of Bias,” *Sensors*, vol. 24, no. 6, pp. 1–33, 2024, doi: 10.3390/s24061730.

[212] A. Visher and T Burkard, “Principles of Blood Pressure Measurement–Current Techniques, Office vs Ambulatory Blood Pressure Measurement.,” in *Hypertension: from basic research to clinical practice*, 2017, pp. 85–96.

[213] R. N. Morcos, K. J. Carter, F. Castro, S. Koirala, D. Sharma, and H. Syed, “Sources of error in office blood pressure measurement,” *Journal of the American Board of Family Medicine*, vol. 32, no. 5, pp. 732–738, 2019, doi: 10.3122/jabfm.2019.05.190085.

[214] S. Sakhuja *et al.*, “Potential impact of systematic and random errors in blood pressure measurement on the prevalence of high office blood pressure in the United States,” *J Clin Hypertens*, vol. 24, no. 3, pp. 263–270, Mar. 2022, doi: 10.1111/jch.14418.

[215] R. Agarwal, “Implications of blood pressure measurement technique for implementation of Systolic Blood Pressure Intervention Trial (SPRINT),” *J Am Heart Assoc*, vol. 6, no. 2, 2017, doi: 10.1161/JAHA.116.004536.

[216] P. E. Drawz *et al.*, “Concordance between Blood Pressure in the Systolic Blood Pressure Intervention Trial and in Routine Clinical Practice,” *JAMA Intern Med*, vol. 180, no. 12, pp. 1655–1663, Dec. 2020, doi: 10.1001/jamainternmed.2020.5028.

[217] F. S. Ahmad, C. Chan, M. B. Rosenman, and W. S. Post, “Validity of Cardiovascular Data From Electronic Sources: The Multi-Ethnic Study of Atherosclerosis and HealthLNK,” *Circulation*, vol. 136, no. 13, pp. 1207–1216, 2017, doi: 10.1161/CIRCULATIONAHA.117.027436.Validity.

[218] S. Wells *et al.*, “Cohort Profile: The PREDICT cardiovascular disease cohort in New Zealand primary care (PREDICT-CVD 19),” *Int J Epidemiol*, vol. 46, no. 1, 2017, doi: 10.1093/ije/dyv312.

[219] T. Chandel, V. Miranda, A. Lowe, and T. C. Lee, “Zero End-Digit Preference in Blood Pressure and Implications for Cardiovascular Disease Risk Prediction—A Study in New Zealand,” *J Clin Med*, vol. 13, no. 22, p. 6846, Nov. 2024, doi: 10.3390/jcm13226846.

[220] E. O’Brien and N. Atkins, “Accuracy of the Dinamap Portable Monitor, model 8100: A review of the evidence for accuracy,” *Blood Press Monit*, vol. 2, no. 1, pp. 31–33, 1997.

[221] J. E. Schwartz *et al.*, “Reliability of Office, Home, and Ambulatory Blood Pressure Measurements and Correlation With Left Ventricular Mass,” *J Am Coll Cardiol*, vol. 76, no. 25, pp. 2911–2922, Dec. 2020, doi: 10.1016/j.jacc.2020.10.039.

[222] G. Whitlock *et al.*, “Random errors in the measurement of 10 cardiovascular risk factors,” *Eur J Epidemiol*, vol. 17, pp. 907–909, 2001.

[223] T. Niyonsenga, A. Vanasse, J. Courteau, and L. Cloutier, “Impact of Terminal Digit Preference by Family Physicians and Sphygmomanometer Calibration Errors on Blood Pressure Value: Implication for Hypertension Screening,” *J Clin Hypertens*, vol. 10, no. 5, p. 341, 2008, [Online]. Available: www.lejacq.com

[224] R. Agarwal, “Implications of blood pressure measurement technique for implementation of Systolic Blood Pressure Intervention Trial (SPRINT),” *J Am Heart Assoc*, vol. 6, no. 2, 2017, doi: 10.1161/JAHA.116.004536.

[225] New Zealand Ministry of Health, “Health loss in New Zealand 1990–2013:A report from the New Zealand Burden of Diseases, Injuries and Risk Factors Study.,” 2016.

[226] E. Craig *et al.*, “The Determinants of Health for Māori Children and Young People in New Zealand,” *New Zealand Child and Youth Epidemiology Service, University of Otago*, 2013.

[227] Debbie. Ryan, Corina. Grey, and B. Mischewski, “Tofa Saili : a review of evidence about health equity for Pacific people in New Zealand,” *Pacific Perspectives Ltd.*, 2019.

[228] T. Tane, V. Selak, K. Eggleton, and M. Harwood, “Understanding the barriers and facilitators that influence access to quality cardiovascular care for rural Indigenous peoples: protocol for a scoping review,” *BMJ Open*, vol. 12, no. 12, p. e065685, Dec. 2022, doi: 10.1136/bmjopen-2022-065685.

[229] C. Grey *et al.*, "First and recurrent ischaemic heart disease events continue to decline in New Zealand, 2005-2015," *Heart*, vol. 104, no. 1, pp. 51–57, Jan. 2018, doi: 10.1136/heartjnl-2017-311613.

[230] T. Blakely, M. Tobias, J. Atkinson, and Y. Huang, "Tracking disparity : trends in ethnic and socioeconomic inequalities in mortality," Ministry of Health, Wellington, 2007.

[231] T. Riddell and S. M. Wells, "Ethnic and socioeconomic disparities in the prevalence of cardiovascular disease in New Zealand," *New Zealand Medical Journal*, vol. 121, no. 1285, 2018.

[232] L.-C. Yeh, R. Timutimu, P. Himona, N. Talamaivao, P. Searle, and D. Weerasekera, "Tatau kahukura : Māori health chart book 2015," *New Zealand Ministry of Health*, 2015.

[233] L. McCormack, L. C. Yeh, B. Braybrook, and L. Clyne, "Tupu ola moui: Pacific Health Chart Book," *New Zealand Ministry of Health*, 2012.

[234] C. Grey *et al.*, "Trends in ischaemic heart disease: patterns of hospitalisation and mortality rates differ by ethnicity (ANZACS-QI 21)," *NZMJ*, vol. 13, 2018, [Online]. Available: www.nzma.org.nz/journal

[235] C. Grey, R. Jackson, S. Wells, R. Marshall, T. Riddell, and A. J. Kerr, "Twenty-eight day and one-year case fatality after hospitalisation with an acute coronary syndrome: A nationwide data linkage study," *Aust N Z J Public Health*, vol. 38, no. 3, pp. 216–220, 2014, doi: 10.1111/1753-6405.12241.

[236] V. A. Cameron *et al.*, "A cohort study comparing cardiovascular risk factors in rural Māori, urban Māori and non-Māori communities in New Zealand," *BMJ Open*, vol. 2, no. 3, 2012, doi: 10.1136/bmjopen-2011-000799.

[237] T. Riddell, R. Jackson, S. Wells, J. Broad, and L. Bannink, "Assessing Māori/non-Māori differences in cardiovascular disease risk and risk management in routine primary care practice using web-based clinical decision support: (PREDICT CVD-2)," *N Z Med J*, vol. 120, no. 1250, 2007.

[238] C. Grey *et al.*, "A comparative analysis of the cardiovascular disease risk factor profiles of Pacific peoples and Europeans living in New Zealand assessed in routine primary care: PREDICT CVD-11," *Journal of the New Zealand Medical Association NZMJ*, vol. 123, p. 8716, 2010.

[239] R. M. Mclean, S. Williams, J. I. Mann, J. C. Miller, and W. R. Parnell, "Blood pressure and hypertension in New Zealand: results from the 2008/09 Adult Nutrition Survey," *Journal of the New Zealand Medical Association NZMJ*, vol. 126, p. 8716, 2013.

[240] New Zealand Ministry of Health, "Mortality and Demographic Data," Wellington, Nov. 2012. Accessed: Nov. 29, 2024. [Online]. Available: https://www.health.govt.nz/publications/mortality-and-demographic-data-2012

[241] T. Riddell, "Heart failure hospitalisations and deaths in New Zealand: patterns by deprivation and ethnicity," *N Z Med J*, vol. 118, no. 1208, 2005.

[242] C. B. Silva, C. Lynch, M. Turley, and S. Johnston, "Obesity in 2021/22: An experimental analysis using data from general practices. Wellington: Ministry of Health," *New Zealand Ministry of Health*, Nov. 2022.

[243] M. Walsh and K. Wright, “Ethnic inequities in life expectancy attributable to smoking,” *NZMJ*, vol. 133, p. 1509, 2020, [Online]. Available: www.nzma.org.nz/journal

[244] K. Autret and T. A. Bekelman, “Socioeconomic Status and Obesity,” *J Endocr Soc*, vol. 8, no. 11, Sep. 2024, doi: 10.1210/jendso/bvae176.

[245] L. Holder-Pearson and J. G. Chase, “Socio-Economic Inequity: Diabetes in New Zealand,” *Front Med (Lausanne)*, vol. 9, May 2022, doi: 10.3389/fmed.2022.756223.

[246] J. Schmidt-Busby, J. Wiles, D. Exeter, and T. Kenealy, “Understandings of disease among Pacific peoples with diabetes and end-stage renal disease in New Zealand,” *Health Expectations*, vol. 22, no. 5, pp. 1122–1131, Oct. 2019, doi: 10.1111/hex.12946.

[247] W. Cheuk Chan, M. Lee, and D. Papaconstantinou, *Understanding the heterogeneity of the diabetes population in Metro Auckland in 2018*. Auckland: Counties Manukau Health, 2020.

[248] A. M. Tomlin, H. S. Lloyd, and M. W. Tilyard, “Atrial fibrillation in New Zealand primary care: Prevalence, risk factors for stroke and the management of thromboembolic risk,” *Eur J Prev Cardiol*, vol. 24, no. 3, pp. 311–319, Feb. 2017, doi: 10.1177/2047487316674830.

[249] C. M. Mangione *et al.*, “Behavioral Counseling Interventions to Promote a Healthy Diet and Physical Activity for Cardiovascular Disease Prevention in Adults Without Cardiovascular Disease Risk Factors: US Preventive Services Task Force Recommendation Statement,” *JAMA*, vol. 328, no. 4, pp. 367–374, Jul. 2022, doi: 10.1001/jama.2022.10951.

[250] L. Bannink, S. Wells, J. Broad, T. Riddell, and R. Jackson, “Web-based assessment of cardiovascular disease risk in routine primary care practice in New Zealand: the first 18,000 patients (PREDICT CVD-1),” *N Z Med J*, vol. 119, no. 1245, 2006.

[251] R. Teh *et al.*, “Atrial fibrillation incidence and outcomes in two cohorts of octogenarians: LiLACS NZ,” *BMC Geriatr*, vol. 23, no. 1, Dec. 2023, doi: 10.1186/s12877-023-03902-5.

[252] W. Cheng *et al.*, “Age-related changes in the risk of high blood pressure,” *Front Cardiovasc Med*, vol. 9, Sep. 2022, doi: 10.3389/fcvm.2022.939103.

[253] Janine Nip, Richard Edwards, Jude Ball, Janet Hoek, and Andrew Waa, “Smoking prevalence and trends: key findings in the 2022/23 NZ Health Survey,” New Zealand Ministry of Health, Dec. 2023.

[254] G. Sundborn *et al.*, “Ethnic differences in the prevalence of new and known diabetes mellitus, impaired glucose tolerance, and impaired fasting glucose. Diabetes Heart and Health Survey (DHAH) 2002-2003, Auckland New Zealand,” *N Z Med J*, vol. 120, no. 1257, 2007.

[255] D. A. Taylor *et al.*, “Cardiovascular disease risk prediction in older people: A qualitative study,” *British Journal of General Practice*, vol. 71, no. 711, pp. E772–E779, Oct. 2021, doi: 10.3399/BJGP.2020.1038.

[256] B. V. Scheer, A. Perel, and U. J. Pfeiffer, “Clinical review: Complications and risk factors of peripheral arterial catheters used for haemodynamic monitoring in anaesthesia and intensive care medicine,” *Crit Care*, vol. 6, no. 3, pp. 198–204, 2002.

[257] D. W. Jones, L. J. Appel, S. G. Sheps, E. J. Roccella, and C. Lenfant, “Measuring Blood Pressure Accurately New and Persistent Challenges,” *JAMA*, vol. 289, no. 8, pp. 1027–1030, 2003.

[258] Y. Imai *et al.*, “The accuracy and the performance of the A&D TM 2421, a new ambulatory blood pressure monitoring device based on the cuff-oscillometric method and the Korotkoff sound technique,” *Am J Hypertens*, pp. 719–26, 1992.

[259] British Standards Institution., *BS EN ISO 81060-4:2023:Non-invasive sphygmomanometers - Part 4: Test procedures to determine the overall system accuracy of automated non-invasive sphygmomanometers.* London, UK: British Standards Institution., 2004.

[260] N. A. E O’Brien, J Petrie, W Littler, M de Swiet, P L Padfield, K O’Malley, M Jamieson, D Altman, M Bland, “The British Hypertension Society protocol for the evaluation of automated and semi-automated blood pressure measuring devices with special reference to ambulatory systems,” *J Hypertens*, vol. 8, no. 7, pp. 607–619, 1990.

[261] E. O’Brien *et al.*, “Working Group on Blood Pressure Monitoring of the European Society of Hypertension International Protocol for validation of blood pressure measuring devices in adults,” *Blood Press Monit*, vol. 7, no. 1, pp. 3–17, 2002, doi: 10.1097/00126097-200202000-00002.

[262] G. S. Stergiou *et al.*, “A universal standard for the validation of blood pressure measuring devices: Association for the Advancement of Medical Instrumentation/European Society of Hypertension/International Organization for Standardization (AAMI/ESH/ISO) Collaboration Statement,” *Hypertension*, vol. 71, no. 3, pp. 368–374, 2018, doi: 10.1161/HYPERTENSIONAHA.117.10237.

[263] B. Alpert, B. Friedman, and D. Osborn, “AAMI blood pressure device standard targets home use issues.,” *Biomedical instrumentation & technology / Association for the Advancement of Medical Instrumentation*, vol. Suppl Home, pp. 69–72, 2010.

[264] Y. Dong, J. Kang, Y. Yu, K. Zhang, Z. Li, and Y. Zhai, “A Novel Model for Continuous Cuff-less Blood Pressure Estimation,” in *11th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP)*, Budapest, Hungary, 2018.

[265] N. Kallioinen, A. Hill, M. S. Horswill, H. E. Ward, and M. O. Watson, “Sources of inaccuracy in the measurement of adult patients’ resting blood pressure in clinical settings: A systematic review,” *J Hypertens*, vol. 35, no. 3, pp. 421–441, 2017, doi: 10.1097/HJH.0000000000001197.

[266] P. Bonnafoux, “Auscultatory and oscillometric methods of ambulatory blood pressure monitoring, advantages and limits: a technical point of view.,” *Blood Press Monit*, vol. 1, no. 3, pp. 181–185, 1996.

[267] Miranda Victor, Chandel Tanvi, Lowe Andrew, and Lee Chuan Tet, “_bpAcc: Blood Pressure Device Accuracy Evaluation: Statistical Considerations_. R package version 0.0-2,” 2024, *Auckland*.

[268] R. Padwal *et al.*, “The Hypertension Canada blood pressure device recommendation listing: Empowering use of clinically validated devices in Canada,” *J Clin Hypertens*, vol. 22, no. 5, pp. 933–936, May 2020, doi: 10.1111/jch.13868.

[269] A. A. Nargesi *et al.*, “Manual or automated sphygmomanometer? A historical cohort to quantify measurement bias in blood pressure recording,” *J Clin Hypertens*, vol. 16, no. 10, pp. 716–721, Oct. 2014, doi: 10.1111/jch.12400.

APPENDIX

APPENDIX A

Table 36: Hazard ratios for the original model and the rounded model for men.

Variables	Original Dataset	Full Rounding	Generalised (12% Rounding)
Age	1.0715	1.0716	1.0715
Chinese	0.6508	0.6499	0.6509
Indian	1.1849	1.1848	1.1848
Māori	1.3663	1.3682	1.3658
Pacific	1.2497	1.2485	1.2495
Ex-Smoking	1.1167	1.1171	1.1167
Currently Smoking	1.7809	1.783	1.7811
Family History	1.1598	1.1617	1.1606
Atrial Fibrillation	1.7895	1.789	1.7915
Diabetes	1.6338	1.6327	1.6344
BMI1 (<18.5)	1.756	1.7504	1.7563
BMI2 (25–<30)	0.9345	0.9357	0.9345
BMI3 (30–<35)	0.9754	0.9782	0.9753
BMI4 (35–<40)	1.0955	1.0995	1.0957
BMI5 (>=40)	1.4224	1.4292	1.4226
BMI unknown	0.9053	0.9063	0.9054
BP Lowering Med	1.309	1.3112	1.3081
Lipid Lowering Med	0.9425	0.9415	0.9426
Antithrombotic Med	1.0589	1.0579	1.0588
NZDep	1.0597	1.0597	1.0597
SBP	1.0188	1.0179	1.0187
TC/HDL Ratio	1.1422	1.1426	1.1423
Age:Diabetes	0.9845	0.9844	0.9844
Age:c.SBP	0.9994	0.9995	0.9994
BP Lowering Med:SBP	0.9963	0.9968	0.9964

Table 37: Hazard ratios for the original model and the rounded model for women.

Variables	Original Dataset	Full Rounding	Generalised (12% Rounding)
Age	1.0761	1.0759	1.0761
Chinese	0.6712	0.6701	0.6707
Indian	1.166	1.1651	1.1655
Māori	1.6181	1.6207	1.6185
Pacific	1.2874	1.2862	1.2875
Ex-Smoking	1.1409	1.141	1.1412
Currently Smoking	1.9554	1.9547	1.956
Family History	1.0323	1.0333	1.0324
Atrial Fibrillation	2.5997	2.5913	2.5995
Diabetes	1.5993	1.5979	1.5993
BMI1 (<18.5)	1.6933	1.693	1.6945
BMI2 (25–<30)	0.9985	0.9993	0.9992
BMI3 (30–<35)	0.9977	0.9985	0.9983
BMI4 (35–<40)	1.0474	1.0501	1.0487
BMI5 (>=40)	1.377	1.3819	1.3789
BMI unknown	1.0377	1.0392	1.038
BP Lowering Med	1.414	1.4225	1.415
Lipid Lowering Med	0.9578	0.9561	0.9575
Antithrombotic Med	1.1394	1.1386	1.1388
NZDep	1.1035	1.1035	1.1036
SBP	1.0189	1.0182	1.0187
TC/HDL Ratio	1.1317	1.1322	1.1318
Age:Diabetes	0.9856	0.9856	0.9856
Age:c.SBP	0.9995	0.9995	0.9995
BP Lowering Med:SBP	0.9925	0.9925	0.9926

Table 38: Misclassification rate for men and women along with the 95% CI (Full rounding)

		Rounded Risk		
	Original Risk	**<5%**	**5-15%**	**>15%**
Women	**<5%**	99.38% (99.36–99.41%)	0.62% (0.59–0.64%)	0.00%
	5–15%	3.21% (3.06–3.36%)	96.59% (96.43–96.74%)	0.20% (0.17–0.24%)
	>15%	0.00%	3.21% (2.78–3.71%)	96.79% (96.29–97.22%)
Men	**<5%**	98.81% (98.77–98.84%)	1.19% (1.16–1.23%)	0.00%
	5–15%	2.85% (2.76–2.94%)	96.68% (96.56–96.79%)	0.47% (0.43–0.53%)
	>15%	0.00%	4.24% (3.96–4.55%)	95.76% (95.45–96.04%)

Table 39: Misclassification rate for men and women along with the 95% CI (12% rounding)

		Rounded Risk		
	Original Risk	**<5%**	**5-15%**	**>15%**
Women	**<5%**	99.90% (98.88–99.91%)	0.10% (0.09–0.12%)	0.00%
	5–15%	0.53% (0.44–0.63%)	99.44% (99.33–99.53%)	0.03% (0.02–0.05%)
	>15%	0.00%	0.74% (0.49–1.06%)	99.26% (98.94–99.51%)
Men	**<5%**	99.82% (99.80–99.84%)	0.18% (0.16–0.20%)	0.00%
	5–15%	0.43% (0.43–0.48%)	99.50% (99.44–99.55%)	0.08% (0.06–0.10%)
	>15%	0.00%	0.71% (0.70–0.93%)	99.29% (99.07–99.46%)

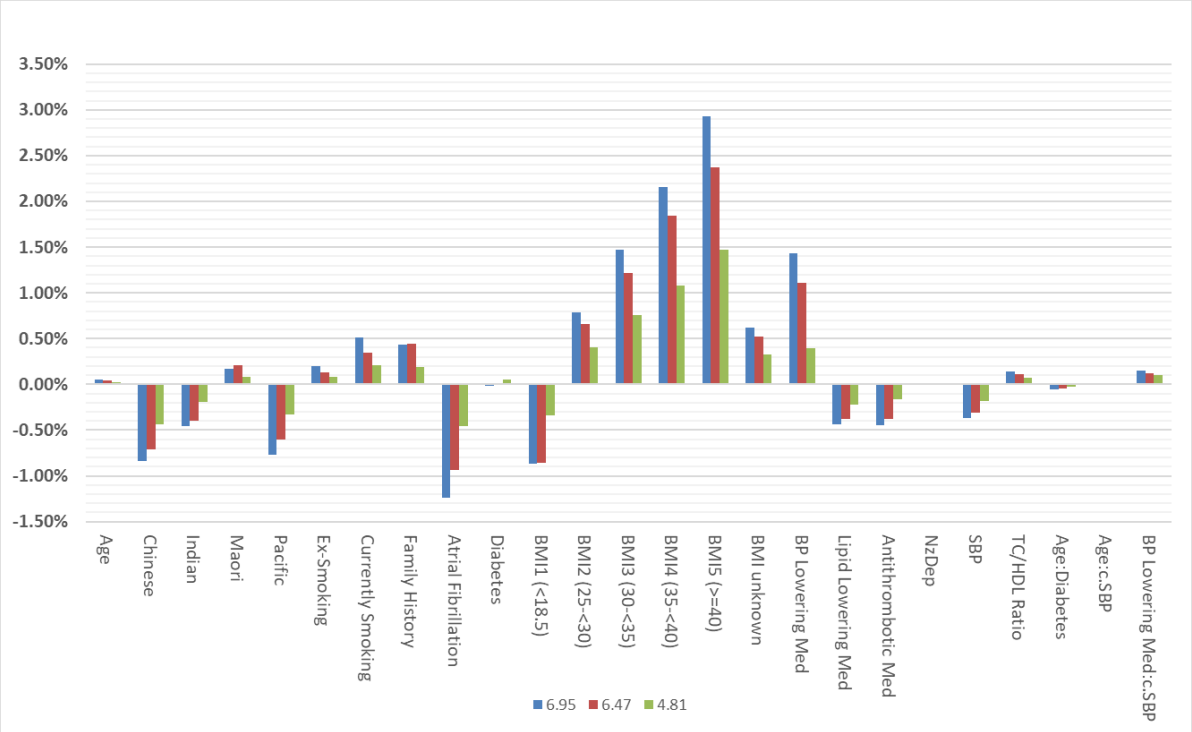
APPENDIX B

Table 40: HR from Cox PH model with original and adjusted data for measurement errors for men.

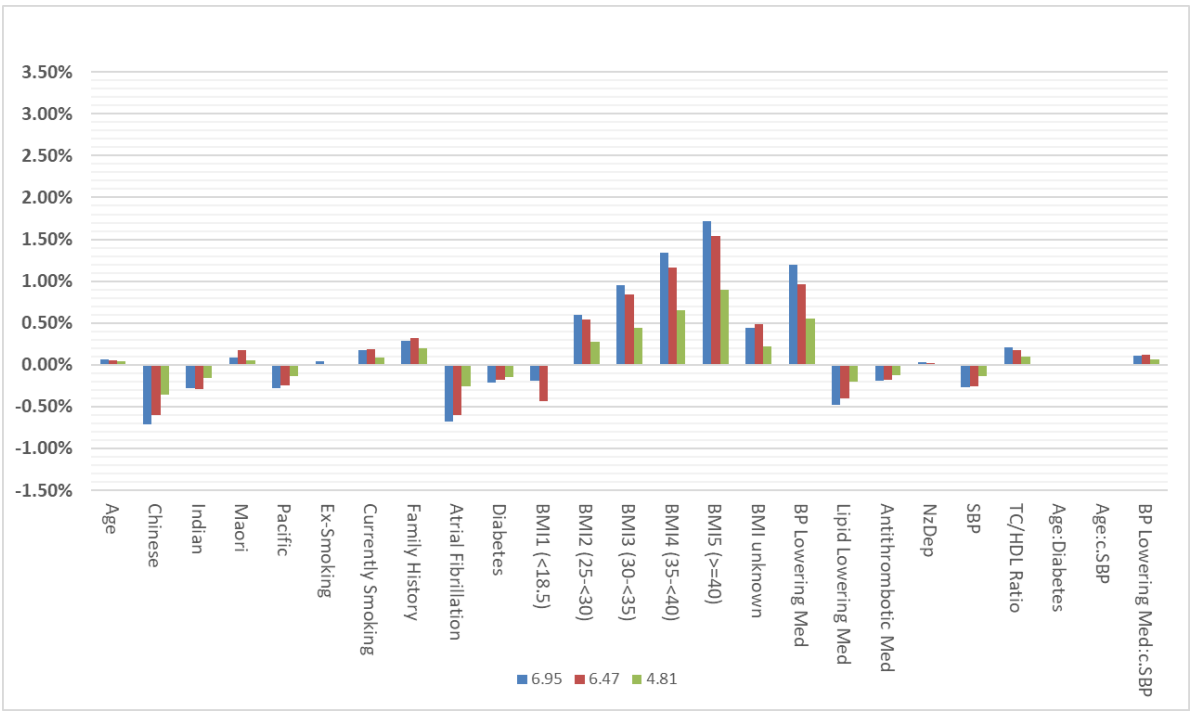
Variables	**Original Dataset**	$\in \sim \boldsymbol{N(0, 6.95)}$	$\in \sim \boldsymbol{N(2.5, 6.47)}$	$\in \sim \boldsymbol{N(5, 4.81)}$
Age	1.0707	1.0713	1.0711	1.0710
Chinese	0.6632	0.6576	0.6585	0.6603
Indian	1.2056	1.2001	1.2008	1.2033
Māori	1.3706	1.3729	1.3734	1.3718
Pacific	1.2648	1.2550	1.2572	1.2607
Ex-Smoking	1.0989	1.1011	1.1003	1.0998
Currently Smoking	1.695	1.7037	1.7008	1.6985
Family History	1.122	1.1268	1.1270	1.1242
Atrial Fibrillation	1.783	1.7610	1.7663	1.7749
Diabetes	1.5967	1.5964	1.5969	1.5975
BMI1 (<18.5)	1.7499	1.7348	1.7349	1.7439
BMI2 (25-<30)	0.9382	0.9456	0.9443	0.9420
BMI3 (30-<35)	0.976	0.9904	0.9879	0.9834
BMI4 (35-<40)	1.0721	1.0952	1.0918	1.0837
BMI5 (>=40)	1.4233	1.4650	1.4571	1.4443
BMI unknown	0.9322	0.9379	0.9371	0.9352
BP Lowering Med	1.3216	1.3405	1.3362	1.3268
Lipid Lowering Med	0.9735	0.9692	0.9699	0.9714
Antithrombotic Med	1.0499	1.0452	1.0460	1.0482
NZDep	1.064	1.0641	1.0641	1.0641
SBP	1.0186	1.0149	1.0155	1.0167
TC/HDL Ratio	1.1458	1.1474	1.1470	1.1466
Age:Diabetes	0.9858	0.9853	0.9854	0.9855
Age:c.SBP	0.9995	0.9996	0.9996	0.9995
BP Lowering Med:SBP	0.9958	0.9973	0.9970	0.9968

Table 41: HR from Cox PH model with original and adjusted data for measurement errors for women.

Variables	**Original Dataset**	$\in \sim \boldsymbol{N(0, 6.95)}$	$\in \sim \boldsymbol{N(2.5, 6.47)}$	$\in \sim \boldsymbol{N(5, 4.81)}$
Age	1.0756	1.0763	1.0761	1.0760
Chinese	0.7031	0.6981	0.6989	0.7006
Indian	1.1149	1.1118	1.1116	1.1132
Māori	1.5498	1.5512	1.5525	1.5506
Pacific	1.2762	1.2726	1.2731	1.2744
Ex-Smoking	1.1305	1.1310	1.1303	1.1304
Currently Smoking	1.9547	1.9580	1.9583	1.9564
Family History	1.0534	1.0564	1.0567	1.0555
Atrial Fibrillation	2.465	2.4482	2.4501	2.4587
Diabetes	1.6248	1.6213	1.6219	1.6225
BMI1 (<18.5)	1.7584	1.7550	1.7508	1.7581
BMI2 (25-<30)	0.99	0.9959	0.9953	0.9927
BMI3 (30-<35)	0.9765	0.9858	0.9847	0.9808
BMI4 (35-<40)	1.0312	1.0450	1.0432	1.0379
BMI5 (>=40)	1.3748	1.3984	1.3960	1.3871
BMI unknown	1.0158	1.0202	1.0208	1.0180
BP Lowering Med	1.4665	1.4840	1.4805	1.4747
Lipid Lowering Med	0.9333	0.9288	0.9296	0.9315
Antithrombotic Med	1.1342	1.1321	1.1322	1.1328
NZDep	1.1096	1.1099	1.1098	1.1097
SBP	1.0181	1.0154	1.0155	1.0167
TC/HDL Ratio	1.1414	1.1438	1.1434	1.1426
Age:Diabetes	0.9835	0.9833	0.9834	0.9834
Age:c.SBP	0.9995	0.9996	0.9996	0.9995
BP Lowering Med:SBP	0.9923	0.9934	0.9935	0.9929



(a)



(b)

Figure 31: Relative Change in Hazard Ratios After Accounting for Acceptable Device Errors for (a) Men and (b) Women, with Standard Deviation of 6.95 mmHg, 6.47 mmHg, and 4.81 mmHg.

Table 42: Misclassification after adjustment for Error ~ N (0,6.95)

	Model 1 with SBP_{error}	**Model 2 with SBP_{actual}**		
		<5%	**5-15%**	**>15%**
Men Results	**<5%**	98.03% (97.97%-98.09%)	1.97% (1.91% - 2.03%)	0.00%
	5-15%	4.77% (4.62%-4.92%)	94.45% (94.26% - 94.62%)	0.79% (0.71% - 0.88%)
	>15%	0.00%	7.50% (6.89% - 8.15%)	92.50% (91.85% -93.11%)
Women Results	**<5%**	98.98% (98.93%-99.03%)	1.02% (0.97%-1.07%)	0.00%
	5-15%	4.98% (4.76%-5.21%)	94.59% (94.37%-94.81%)	0.43% (0.36%-0.51%)
	>15%	0.00%	5.65% (4.65%-6.80%)	94.35% (93.20%-95.35%)

Table 43: Misclassification after adjustment for Error ~ N (2.5,6.47)

	Model 1 with SBP_{error}	**Model 2 with SBP_{actual}**		
		<5%	**5-15%**	**>15%**
Men Results	**<5%**	98.14% (98.08%-98.20%	1.86% (1.80%-1.92%)	0.00%
	5-15%	4.50% (4.36%-4.65%)	94.76% (94.58%-94.93%)	0.74% (0.66%-0.82%)
	>15%	0.00%	7.06% (6.46%-7.73%)	92.94% (92.27%-93.54%)
Women Results	**<5%**	99.04% (98.99%-99.09%)	0.96% (0.91%-1.01%)	0.00%
	5-15%	4.70% (4.49%-4.92%)	94.91% (94.69%-95.12%)	0.40% (0.33%-0.48%))
	>15%	0.00%	5.30% (4.35%-6.39%)	94.70% (93.61%-95.65%)

Table 44: Misclassification after adjustment for Error ~ N (5,4.81)

	Model 1 with SBP_{error}	**Model 2 with SBP_{actual}**		
		<5%	**5-15%**	**>15%**
Men Results	**<5%**	98.57% (98.52%-98.62%)	1.43% (1.38%-1.48%)	0.00%
	5-15%	3.51% (3.38%-3.64%)	95.94% (95.78%-96.08%)	0.55% (0.49%-0.63%)
	>15%	0.00%	5.43% (4.90%-6.01%)	94.57% (93.99%-95.10%)
Women Results	**<5%**	99.27% (99.22%-99.31%)	0.73% (0.69%-0.78%)	0.00%
	5-15%	3.63% (3.45%-3.83%)	96.08% (95.88%-96.26%)	0.30% (0.24%-0.36%)
	>15%	0.00%	3.99% (3.17%-4.96%)	96.01% (95.04%-96.83%)

APPENDIX C

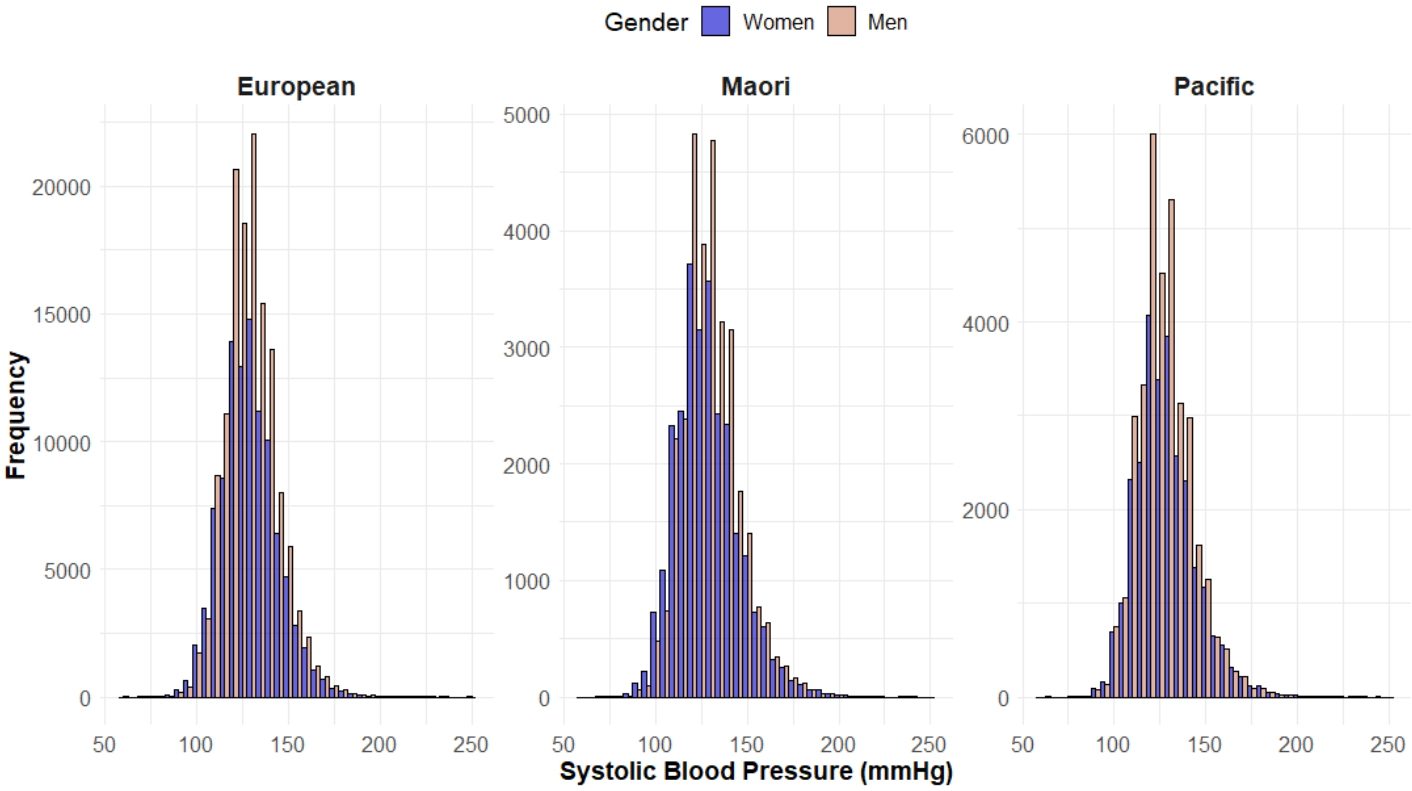


Figure 32: Distribution of SBP by Ethnicity and Gender

APPENDIX D

Table 45: List of parameters used in Section 3 with their description.

Parameter	Description
ϵ_k^j	BP error for the j[th] participant, which is the difference between the test device measurement, $\delta_{k_2}^j$ and the reference device, $\delta_{k_1}^j$.
δ^j	Average of the errors for three measurements for the j[th] participant.
Δ	Tolerable error i.e. errors within -10 mmHg and 10 mmHg
$\hat{p}$	Estimated probability of tolerable error
p	True probability of tolerable error
γ_p	Maximum value of the probability of a tolerable error
θ	Parameterised Distribution for BP errors
μ_p, σ_p	True mean and true SD respectively
σ^{MAX}	Maximum SD for a certain value of mean and γ_P
σ^{AAMI}	Maximum SD for a certain value of mean as per SP10 where γ_P= 0.85.
$\bar{x}_\iota$, s_ι	Sample mean and sample SD of BP errors respectively.
$\emptyset$	Cumulative Density Function of the standard normal distribution.

Table 46: Simulated P_A and P_A obtained from Normal approximation using proposed method vs the method currently in use in the SP10 standard, for small sample sizes, with $\mu_P = 0$ and $\sigma_P = 6.9$.

	N=10	N=15	N=20	N=25
Simulated P_A	0.83	0.83	0.84	0.85
P_A using proposed framework	0.74	0.79	0.82	0.85
P_A using SP10 method	0.71	0.75	0.79	0.81

Table 47: Simulated P_A and P_A obtained from Normal approximation using proposed method vs the method currently in use in the SP10 standard, for small sample sizes, with $\mu_P = 2.5$ and $\sigma_P = 6.9$.

	N=10	N=15	N=20	N=25
Simulated P_A	0.75	0.74	0.74	0.74
P_A using proposed framework	0.65	0.68	0.71	0.73
P_A using SP10 method	0.64	0.67	0.69	0.72

Table 48: Simulated P_A and P_A obtained from Normal approximation using proposed method vs the method currently in use in the SP10 standard, for small sample sizes, with μ_P= 5 and σ_P= 6.9.

	N=10	N=15	N=20	N=25
Simulated P_A	0.52	0.46	0.42	0.38
P_A using proposed framework	0.42	0.40	0.38	0.37
P_A using SP10 method	0.41	0.39	0.37	0.36