TU WIEN Informatics

Geschlechtsspezifische Aspekte in der Risikostratifikation nach transienten ischämischen Attacken

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieurin

im Rahmen des Studiums

Data Science

eingereicht von

Claudia Kößldorfer, BSc.
Matrikelnummer 11825357

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Dr.in techn. Alexandra Posekany

Wien, 29. April 2025

Claudia Kößldorfer	Alexandra Posekany

Technische Universität Wien
A-1040 Wien ▪ Karlsplatz 13 ▪ Tel. +43-1-58801-0 ▪ www.tuwien.at

TU WIEN Informatics

Gender-specific aspects in risk stratification after transient ischaemic attacks

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieurin

in

Data Science

by

Claudia Kößldorfer, BSc.
Registration Number 11825357

to the Faculty of Informatics

at the TU Wien

Advisor: Dr.in techn. Alexandra Posekany

Vienna, April 29, 2025

Claudia Kößldorfer | Alexandra Posekany

Technische Universität Wien
A-1040 Wien ▪ Karlsplatz 13 ▪ Tel. +43-1-58801-0 ▪ www.tuwien.at

Erklärung zur Verfassung der Arbeit

Claudia Kößldorfer, BSc.

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.
Ich erkläre weiters, dass ich mich generativer KI-Tools lediglich als Hilfsmittel bedient habe und in der vorliegenden Arbeit mein gestalterischer Einfluss überwiegt. Im Anhang „Übersicht verwendeter Hilfsmittel“ habe ich alle generativen KI-Tools gelistet, die verwendet wurden, und angegeben, wo und wie sie verwendet wurden. Für Textpassagen, die ohne substantielle Änderungen übernommen wurden, haben ich jeweils die von mir formulierten Eingaben (Prompts) und die verwendete IT- Anwendung mit ihrem Produktnamen und Versionsnummer/Datum angegeben.

Wien, 29. April 2025

Claudia Kößldorfer

Danksagung

Zu Beginn möchte ich mich bei meiner Betreuerin, Dr.in techn. Alexandra Posekany, für ihre Unterstützung und wertvollen Ratschläge bedanken, die mich durch das Erstellen dieser Arbeit geleitet haben.

Sehr dankbar bin ich auch Dr. Peter Sommer und Dr. Simon Fandler-Höfler dafür, dass sie mich mit ihrem medizinischen Wissen und durch essenzielles Feedback unterstützt haben. Dies ermöglichte mir, besser mit der Materie vertraut zu werden.

Ein besonderer Dank gilt meiner Schwester dafür, dass sie meine Arbeit gelesen und mir geholfen hat, Unklarheiten auszubessern. Bei meiner Mutter und meinem Partner muss ich mich für ihre immerwährende Unterstützung bedanken, welche es mir ermöglichte, dieses Masterstudium zu machen.

Acknowledgements

I would like to express my gratitude to my supervisor, Dr.in techn. Alexandra Posekany, for her continuous support and guidance throughout the process of creating this thesis.

I am also grateful to Dr Peter Sommer and Dr Simon Fandler-Höfler for providing valuable feedback and relevant information about the medical context, which helped me gain a deeper understanding of the subject matter.

A special thank you goes to my sister for carefully reading through my thesis and helping me improve its clarity. I am also deeply thankful to my mother and my partner, who have both supported and encouraged me throughout my studies.

Kurzfassung

Diese Arbeit untersucht geschlechtsspezifische Unterschiede in den Gesundheitsergebnissen, der Risikostratifizierung, den diagnostischen und sekundären Präventionspraktiken nach transitorischen ischämischen Attacken (TIA), basierend auf Daten aus dem Österreichischen Stroke Unit Register.

Zwei verschiedene Patientengruppen wurden definiert und analysiert. Einerseits ein Subregister, dessen Datenerhebung 2015 endete und sich mit der Erfassung von TIA-spezifischen klinischen Variablen befasste, und andererseits aktuelle Daten aus dem allgemeinen Register. In beiden Populationen wurden Geschlechterunterschiede in den Gesundheitsergebnissen und Pflegestrategien analysiert und verglichen. Zusätzlich wurden in dem TIA-Subregister die ABCD2- und ABCD3-I-Risikoscores auf Unterschiede in der prädiktiven Genauigkeit basierend auf Geschlecht untersucht.

Die Ergebnisse zeigten, dass Frauen, selbst nach Anpassung durch Alter und klinische Risikofaktoren, mit höherer Wahrscheinlichkeit eine schlechtere funktionelle Genesung nach einer TIA hatten. Allerdings wurde kein Geschlechtsunterschied für die Sterblichkeit beobachtet. Die ABCD2- und ABCD3-I-Scores hatten für Männern und Frauen weitgehend die gleiche prädiktive Fähigkeit, wobei die berechneten Score-Versionen besser funktionierten als die von den Klinikern eingetragenen. Für die Sekundärpräventionsmethoden wurden stärkere geschlechtsspezifische Unterschiede in den aktuelleren Daten nachgewiesen. Dabei erhielten Frauen selterner Karotisdurchblutungsinterventionen, echokardiografische Verfahren, duale Thrombozytenaggregationshemmer, hochdosierte Heparintherapie und Rehabilitationsdienste.

Während Risikostratifikationstools bei beiden Geschlechtern nach einer TIA ähnlich gut funktionierten, waren Unterschiede in den diagnostischen und sekundären Behandlungspraktiken eindeutig. Weitere Forschung ist erforderlich, um Bedenken hinsichtlich dieser Unterschiede anzusprechen, insbesondere da sie in den neueren klinischen Daten ausgeprägter sind.

Abstract

This thesis investigates gender-specific differences in health outcomes, risk stratification, diagnostic and secondary prevention practices following transient ischaemic attacks (TIA), using data from the Austrian Stroke Unit Registry.

Two distinct patient populations were defined and analysed. On the one hand, a sub-registry that focused on and collected TIA-specific clinical variables, but stopped gathering data in 2015. And on the other hand, recent data from the general registry. Across both populations, gender differences in health outcomes and care strategies were analysed and compared. Further, the ABCD2 and ABCD3-I risk scores were evaluated for differences in predictive accuracy based on gender in the TIA-subset.

The results showed that women were more likely to experience worse functional recovery following a TIA, even after adjusting for age and clinical risk factors, although no gender difference was observed for mortality. The predictive performance of the ABCD2 and ABCD3-I scores was broadly comparable between men and women, with calculated score versions outperforming clinician-entered ones. In the recent registry data, more pronounced gender-specific differences in secondary prevention practices were observed. Women were less likely to receive carotid interventions, echocardiographic procedures, dual antiplatelet therapy, high-dose heparin, and rehabilitation services.

While risk stratification tools performed similarly for both genders after a TIA, differences in diagnostic and secondary treatment practices were evident. Further research is needed to address concerns regarding these disparities, especially as they are more pronounced in recent clinical data.

Contents

Kurzfassung **xi**

Abstract **xiii**

1 Introduction **1**
1.1 Research Objectives . . . 2
1.2 Scope and Limitations of the Thesis . . . 4
1.3 Structure of the Thesis . . . 5

2 Literature Review **7**
2.1 Transient Ischaemic Attacks (TIA) . . . 7
2.2 Gender-Specific Research on Strokes and TIAs . . . 10
2.3 Austrian Stroke Unit Registry . . . 12
2.4 Risk Stratification Research on the Austrian Stroke Unit Registry . . . 12
2.5 Research Gap . . . 13

3 Methodology **15**
3.1 Research Design . . . 15
3.2 Statistical Methods . . . 17

4 Data **25**
4.1 Data Filtering . . . 25
4.2 ABCD2 & ABCD3-I Scores . . . 27
4.3 Additional Variables . . . 33
4.4 Data Structure . . . 38

5 Results **57**
5.1 Gender Differences in Health Outcomes . . . 57
5.2 Predictive Accuracy of the ABCD2 and ABCD3-I Scores . . . 68
5.3 Gender Differences in Diagnostics and Treatments . . . 92

6 Discussion **119**
6.1 Research Question 1: Gender Differences in Health Outcomes . . . 119

6.2 Research Question 2: Predictive Accuracy of the ABCD2 and ABCD3-I Scores 121
6.3 Research Question 3: Gender Differences in Diagnostics and Treatment 122

7 Conclusion **125**

Appendix **127**
Patient Characteristics and Risk Factors 127
Gender Differences in Health Outcomes 132
Predictive Accuracy of the ABCD2 and ABCD3-I Scores 136
Gender Differences in Diagnostics and Treatments 203

Overview of Generative AI Tools Used **215**

List of Figures **217**

List of Tables **219**

Bibliography **223**

CHAPTER 1

Introduction

Strokes occur when the brain's blood supply is interrupted, preventing sufficient oxygen and nutrients from reaching the brain tissue. Within minutes, affected brain cells can begin to die, potentially causing permanent damage. Depending on the extent of the impairment and the parts of the brain that are impacted, a person may experience a number of different limitations, such as issues with speech, mobility, or memory. Transient ischaemic attacks (TIA) are different from strokes, despite presenting with similar symptoms, as they are brief and do not cause lasting brain damage. Although sometimes referred to as 'mini-strokes' and easily underestimated due to their temporary nature, TIAs should not be taken lightly. [Johc]

TIAs are acute neurological events and serve as critical warning signs, indicating a raised short-term risk of a future stroke - particularly within the first two days [Eas+09]. After a TIA, early diagnosis and treatment are important to avoid severe health consequences. To improve patient outcomes and prioritise care, clinicians use risk stratification tools developed specifically for patients after a TIA. Two of the most commonly used are the ABCD2 and ABCD3-I scores, which combine clinical factors such as age, blood pressure, symptom duration, and the presence of diabetes into a single value [Eas+09]. The ABCD3-I score additionally includes imaging findings and whether a second TIA has recently occurred [Mer+10]. Their primary purpose is to estimate short-term stroke risk and identify high-risk patients who may benefit from urgent intervention or hospitalisation.

However, one key factor is absent from these scores: gender. Existing research suggests that men and women often differ in their risk factors, clinical presentation, prognosis, and even the treatments they receive following a TIA. Yet, these gender-specific variations are not accounted for in commonly used risk assessment tools. This raises an important question: do the ABCD2 and ABCD3-I scores perform equally well for both men and women, or might they lead to systematic misclassification? If omitting gender causes these scores to consistently under- or overestimate risk for some patients, the scores may unintentionally misguide treatment decisions.

This thesis addresses that gap by investigating the role gender may play in the predictive performance of the ABCD2 and ABCD3-I scores. To do so, data from the Austrian Stroke Unit Registry [Öst] — a nationwide database collecting standardised clinical information from stroke units across Austria - is used. The analysis focuses on patients who experienced a transient ischaemic attack or minor stroke. The registry covers patient characteristics, treatments received, diagnostic procedures, and follow-up outcomes, enabling not only a gender-based evaluation of the scores but also broader analyses.

In addition to evaluating the predictive performance of the scores, this thesis also explores potential gender differences in clinical outcomes, diagnostic methods, and secondary prevention treatments following a TIA. While such gender-specific differences have been examined in international studies, they have not yet been analysed using Austrian data. The analyses are performed across two distinct populations from the Austrian Stroke Unit Registry - one from a TIA-specific sub-registry with broader variable coverage, and another from more recent entries to ensure findings reflect current clinical practice. All medical information was reviewed and verified in consultation with two neurologists, Dr. Peter Sommer and Dr. Simon Fandler-Höfler, to ensure clinical accuracy and contextual relevance throughout the thesis.

1.1 Research Objectives

This thesis investigates gender-specific differences for patients who experienced a transient ischaemic attack (TIA). Three core aspects are explored: clinical health outcomes, risk stratification, and the use of diagnostic and treatment methods. In prior research, gender has been linked to differences in stroke and TIA characteristics, but it is unknown how much of this transfers to post-TIA risk assessment. Moreover, such differences in care and prognosis have not yet been explored in the Austrian clinical context.

To address these issues, three distinct research questions are formulated. Each is focused on a separate topic and being addressed using a different analytical approach. The aim is to gain a more complete understanding of whether current standards of care adequately account for gender-specific factors after a TIA.

1.1.1 Research Question 1

Are there gender-specific differences in health outcomes among patients after experiencing a transient ischaemic attack (TIA)?

TIAs are often caused by the same underlying vascular problems can result in strokes, even though they themselves resolve quickly and without lasting damage. Therefore, even after the symptoms of the current event have passed, a risk for a future stroke remains. By comparing the health outcomes of men and women after a TIA, it can be determined whether this risk presents differently for each gender.

The clinical outcomes like stroke recurrence (both early and within 90 days), functional recovery (measured by the modified Rankin Scale at follow-up), and the occurrence of

complications such as epileptic seizures or pneumonia can be analysed. To distinguish the effect of gender from other factors, adjustments for age and comorbidities will also be made.

Examining these outcomes can provide information on how gender may affect a patient's prognosis after a TIA. Both long- and short-term differences can be explored and any patterns identified that may otherwise go unnoticed.

1.1.2 Research Question 2

Are there gender-specific differences in the predictive accuracy of the ABCD2 and ABCD3-I scores for recurrence of ischaemic events after a transient ischaemic attack (TIA)?

Following a TIA, identifying which patients are most at risk of a subsequent stroke is a time-critical matter. Clinicians use risk stratification tools like the ABCD2 and ABCD3-I scores to quickly and systematically perform these assessments. However, neither score includes gender as a factor, despite evidence of differences in many aspects of TIAs and strokes between men and women.

If the performance of these scores varies by gender, it may result in systematic over- or under-classification of stroke risk. This would affect decisions about hospital admission, imaging, or preventive treatment and influence how limited healthcare resources are allocated. It's possible that some patients will receive needless interventions while others who truly require them won't.

To investigate this, the predictive accuracy of the ABCD2 and ABCD3-I scores is evaluated separately for men and women. The analysis also takes into account the impact of controlling for other patient characteristics and risk factors on the scores' predictive ability. This helps determine whether the current models are robust or if they may benefit from adjustments.

1.1.3 Research Question 3

Are there gender-specific differences in diagnostic methods and secondary prophylactic treatments after a transient ischaemic attack (TIA)?

In order to prevent future strokes and other complications, precise diagnosis and appropriate follow-up care are important. Once the TIA has passed, the next steps have to be determined, whether that be imaging and cardiac assessments or the prescription of medication and rehabilitation. These should be applied consistently and equitably across genders, since they are important in keeping patients healthy and stable.

However, research has revealed that in many medical specialities men and women are not always treated equally. In stroke care, previous studies have observed gender disparities in diagnosis and treatments. It is unknown if there are similar differences in the Austrian clinical setting after a TIA.

This thesis examines whether men and women differ in their likelihood of receiving diagnostic procedures (such as MRI or echocardiography) and secondary preventive treatments (such as antiplatelet agents, anticoagulants, or access to rehabilitation). Additionally, treatment rates for some illnesses, like atrial fibrillation, will be compared between genders.

Identifying any systematic differences can contribute to a better understanding of how gender may influence which care is provided after a TIA. The findings may also help guide more consistent, fair treatment strategies.

1.2 Scope and Limitations of the Thesis

This thesis uses two distinct patient populations from the Austrian Stroke Unit Registry. Each population offers specific advantages, but also presents certain limitations that affect the scope of the analysis.

Population 1 is based on a TIA sub-registry, which includes many variables not usually recorded in the broader registry. This allows for a more detailed characterisation of patients and enables analyses involving these clinical variables — specifically the evaluation of the ABCD2 and ABCD3-I risk scores. The general predictive accuracy of both scores in this population was previously evaluated in a 2016 study by Knoflach et al. [Kno+16]. Elements of that study's approach will be adapted in parts of this thesis. This population, however, only provides data from 2010 to 2015 and may not reflect current clinical practices.

In contrast, Population 2, covers the years 2018 to 2024 and provides more recent information on patients' health outcomes and care practices in Austria. While this strengthens the relevance of the findings, the dataset is limited by a reduced set of variables, and TIA-specific information is unavailable. This makes it impossible to calculate the risk scores or assess their predictive accuracy.

Additionally, not all variables are consistently recorded and available for every patient in the registry. Missing data were handled in different ways depending on the context of the variable and of the missingness. This affects the sample sizes for certain analyses and may introduce a small amount of bias in some comparisons.

The data of both populations was limited to information documented during routine patient care. As is common for registry-based studies, this means certain details are not captured — such as a patient's socio-economic background or personal health beliefs. The gender of patients' is only recorded as either male or female, so no analysis beyond this binary can be performed.

This registry collects data from stroke units across Austria, reflecting real clinical practices rather than a controlled research setting. As a result, differences in clinical decision-making are likely to occur, such as how diagnoses are assigned, how outcomes are evaluated, or how consistently diagnostic and treatment procedures are performed. Furthermore,

because the data in this thesis is observational in nature, it is only possible to identify associations rather than causal links. Lastly, the findings and conclusions may not be fully generalisable, since healthcare practices, patient demographics, and access to care may differ across countries.

1.3 Structure of the Thesis

The thesis is organised into six chapters following the Introduction, as outlined below:

- **Chapter 2: Literature Review**
  Reviews existing literature on gender-specific differences in TIAs and strokes. It also introduces the Austrian Stroke Unit Registry and the study by Knoflach et al. [Kno+16], which informs parts of the analysis.
- **Chapter 3: Methodology**
  Outlines the research design and presents the statistical methods used to analyse each of the three research questions.
- **Chapter 4: Data**
  Describes the two populations constructed from the Austrian Stroke Unit Registry, explains the filtering process, and outlines key preprocessing steps. It also defines new variables used in the analysis, including the ABCD2 and ABCD3-I risk scores.
- **Chapters 5: Results**
  Presents the findings for each of the three research questions using both populations.
- **Chapter 6: Discussion**
  Interprets the results in the context of previous research and discusses potential explanations for observed differences.
- **Chapter 7: Conclusion**
  Summarises the key findings and highlights the thesis's contributions as well as limitations.

CHAPTER 2

Literature Review

This chapter provides an overview of transient ischaemic attacks (TIA). Their clinical definition, risk assessment approaches, and the implications for stroke prevention are examined. Specifically discussed are the development and application of the risk assessment tools ABCD2 and ABCD3-I scores.

Studies exploring gender-related differences in stroke and TIA presentation, outcomes and treatments are reviewed, highlighting the need for more personalised approaches.

The Austrian Stroke Unit Registry is introduced, as well as the study on whose approach this paper is based.

2.1 Transient Ischaemic Attacks (TIA)

Brief episodes of neurological dysfunction referred to as transient ischaemic attacks (TIA) are important indicators for potential future stroke risk. Their definition has shifted from being based on the duration of symptoms to focusing more on whether brain tissue damage occurred. This chapter will discuss the current understanding of TIAs and focus on risk assessment for preventing subsequent strokes.

2.1.1 Definition

The clinical definition of transient ischaemic attacks (TIA) has been subject to change. The historical and classic definition describes them as a "sudden, focal neurologic deficit that lasts for less than 24 hours, is presumed to be of vascular origin, and is confined to an area of the brain or eye perfused by a specific artery." [Alb+02]

The fact that this definition is based on the duration of the symptoms has been criticised [Eas+09]. Specifically, as the cut-off is mostly arbitrary and can be misleading, since many patients who experience transient events lasting less than 24 hours also have cerebral

infarction linked to these. This then constitutes a misclassification of the patients. In practice this 24-hour threshold can also lead to delayed interventions from care providers, rather than administering immediate treatment. In these cases the false expectation could be for the symptoms to resolve themselves, which would only happen in the event of a TIA. This is extremely dangerous, as deficits lasting longer than an hour without effective therapy are likely to result in permanent deficits for the patient.

The symptoms of most TIAs do not exceed the one-hour mark, which further shows the arbitrary nature of the 24-hour cut-off. However, this does not mean that this threshold should be updated or replaced altogether. This is due to the fact that there is no time mark that reliably allows a proper distinction between events with or without tissue infarction. The latter was found to be the more relevant classification, as the focus is shifted to be on the underlying pathophysiology. During the diagnostic and treatment process, the attention is then clearly on determining the source of the ischaemia and if a brain injury has taken place.

Based on these considerations, the American Heart Association (AHA) [Eas+09] has refined the definition of TIA to: "Transient ischemic attack (TIA): a transient episode of neurological dysfunction caused by focal brain, spinal cord, or retinal ischemia, without acute infarction." TIAs are still defined as brief events, though no time frame is specified, and the important distinction is made with the patient not suffering a permanent cerebral infarction.

Several studies have emphasised and recommended the value of imaging techniques (e.g., diffusion-weighted MRI [Hur+19]) in diagnosing TIAs and assessing the risk for future strokes. Through imaging, damaged or swollen brain tissue can be identified, or the blood flow of arteries supplying the brain can be observed. This can offer valuable insights for diagnosis and guide preventive treatments while demonstrating the value of applying a tissue-based definition in practice [Win+13].

However, the European Stroke Organisation (ESO) [Fon+21] still uses a time-based definition for their guidelines on the management of transient ischaemic attacks. They define a TIA as "an acute loss of focal cerebral or ocular function with symptoms lasting less than 24 hours and which, after adequate investigation, was presumed to be due to embolic or thrombotic vascular disease." They themselves refer to this as a pragmatic definition and state that they chose a time-based one in order to maximise the generalisability of their guidelines. It is recommended that a specialist review be done within these 24 hours and that imaging be used to determine if there is evidence of infarction or other risk factors. Despite the use of a time-based approach, its limitations and the significance of brain imaging are recognised.

An accurate and clinically applicable definition of transient ischaemic attacks is important so as not to have treatment delayed due to a misclassification of a more serious event. TIAs should be treated as an urgent clinical warning for an increased stroke risk and possible future strokes. They provide the opportunity to initiate secondary prophylactic treatments and possibly prevent recurrence or any permanent disabilities. The short- as

well as long-term stroke risk is notably increased following a TIA, with an especially high risk within the first 2 days [Eas+09]. Immediate evaluation and treatment are crucial following a TIA. The aim is to reduce the risk of permanent neurological damage and improve health outcomes for patients.

2.1.2 Risk Stratification

For the assessment of the risk of stroke after a transient ischaemic attack (TIA) clinical tools have been developed. These stratify patients based on the likelihood of them experiencing an imminent stroke [Son+13]. Therefore, they can be used by doctors and care providers to identify high-risk patients, determine the urgency of interventions, and guide their medical decisions accordingly. This ensures that patients with the highest likelihood of having a stroke are able to get the necessary immediate treatment or prevention measures, such as being hospitalised, undergoing imaging, or receiving anticoagulation therapy [Eas+09]. With proper risk stratification the health outcomes for the patients can be improved and healthcare resources can be efficiently allocated.

The two clinical scoring tools covered in this thesis are the ABCD2 and ABCD3-I scores, which are widely used.

ABCD2 Score

In order to provide a more reliable prediction standard, Johnston et al. (2007) [Joh+07] created the ABCD2 score by combining previous scores. It incorporates elements from the California and the ABCD score, which both use clinical factors to assign and then summarise points. The best-performing combinations of their components were termed the ABCD2 score based on the initials of the components (age, blood pressure, clinical features, duration, diabetes). Each of these contributes to the score cumulatively so that the score ranges from 0 to 7. A patient with a score of 6-7 is considered high risk and has a significant probability of having a stroke within 48 hours. The score performed well in tests on its ability to predict patients with high risk (very) early stroke recurrence (2 days, 7 days, and 30 days) after a transient ischaemic attack.

The ABCD2 score has been widely adopted, and studies have shown that it is also capable of identifying patients at high risk of having a late stroke [Yan+10]. However, as this score was intended for the identification of patients in need of urgent care and hospital admission, certain aspects like imaging were not taken into account [Joh+07].

ABCD3-I Score

The aspect of imaging was addressed with the ABCD3-I score by Merwick et al. (2010) [Mer+10], which is an extension of ABCD2 developed to enhance predictive accuracy. In addition to the original five components, the ABCD3-I includes Dual TIA, meaning a second TIA within 7 days of the first event, and two findings from brain and carotid imaging. The score works on a scale from 0-13, with high-risk patients defined as those

with scores from 8-13. The specific components and their assigned points of both scores can be viewed and compared in the Table 2.1.

As a result of the imaging components being included, the ABCD3-I is closely associated with the tissue-based definition of TIAs. The accuracy of risk stratification can be increased by using diffusion-weighted MRIs to identify patients with underlying cerebrovascular pathology that might not be clear during a clinical examination [Mer+10].

Studies have validated the effectiveness of the ABCD3-I score in predicting stroke risk following a TIA [Son+13]. Showing that the addition of imaging findings and recurrent TIA episodes can significantly improve the efficiency of identifying patients with a high risk of strokes.

Compared to the ABCD2 score, the ABCD3-I model provides superior predictive value, particularly in distinguishing individuals who would benefit from immediate hospitalisation, early initiation of secondary prevention strategies, and advanced neurovascular assessments. This is especially the case in hospital-based settings where imaging techniques like MRI are readily available. [Mer+10]

	ABCD2	ABCD3-I
Age $\geq$ 60 years	1	1
Blood pressure $\geq$ 140/90 mm Hg	1	1
Clinical features		
Unilateral weakness	2	2
Speech impairment without weakness	1	1
Duration		
$\geq$ 60 min	2	2
10-59 min	1	1
Diabetes mellitus present	1	1
Dual TIA (TIA prompting medical attention plus at least another TIA in the preceding 7 d)	NA	2
Imaging: ipsilateral $\geq$ 50% stenosis of internal carotid artery	NA	2
Imaging: acute diffusion-weighted imaging hyperintensity	NA	2

Table 2.1: ABCD2 and ABCD3-I Scores

2.2 Gender-Specific Research on Strokes and TIAs

Studies on different patient populations have explored gender-specific differences in incidence rate, clinical features, outcomes, diagnosis, and treatments after strokes and transient ischaemic attacks (TIA).

The temporal trends in TIA incidence were studied in Madsen et al. (2019) [Mad+19], with the focus on sex differences. They examined data from the Greater Cincinnati Northern Kentucky Stroke Study (GCNKSS) between 1993/4 and 2010 and identified 4746 TIA events. While for men the TIA rates decreased significantly over time, they did not for women, remaining rather stable instead. By 2010, for several age groups, women had either similar or higher TIA rates compared to men. Although women had a lower risk of subsequent infarcts and death after a TIA, they highlighted that further research into sex differences in stroke risk would be necessary in order to develop effective stroke preventive methods.

Data from the Spanish (REGITELL) TIA patient registry was used by Purroy et al. (2021) [Pur+21] to examine gender-related differences in clinical as well as neuroimaging characteristics and long-term outcomes after TIA. They studied 723 patients (41.8% women) and found women were older, with more non-definitive TIAs and more events of undetermined causes. However, smoking and other modifiable vascular risk factors were more prevalent among men. They were also more likely to suffer large artery atherosclerosis than women. It was found that the risk for stroke recurrence did not differ between the genders, although men had a higher risk of major vascular events. Notably, only for women a positive diffusion-weighted imaging (DWI) was a predictor for stroke recurrence. The study, though on a rather small population, shows baseline characteristics, symptom presentation, acute ischaemic lesion patterns, and outcome predictors varying by sex.

Acute treatment and early outcomes of acute ischaemic stroke (AIS) were analysed for gender-related differences in Bonkhoff et al. (2021) [Bon+21]. Although the focus was not on TIAs but AISs, the study used the Stroke Registry of Northwestern Germany and worked to identify differences in outcomes and care between female and male patients from 2000 to 2018. It found that after controlling for age, stroke severity, and comorbidities, women had lower rates of in-hospital deaths and better functional outcomes at discharge than men, despite arriving with more severe strokes. In the early years (2000–2009), the treatment of intravenous thrombolysis (IVT) was less common for women, though after adjusting for age and stroke severity, this difference disappeared by 2010–2018. However, even accounting for clinical variables, women remained more likely to undergo intra-arterial therapy (IAT) in both time periods. The study emphasises that reasons and factors leading to this difference in treatment and a better recovery for women despite initial stroke severity need to be explored further. Though they suggest that these gender differences may be explained by stroke location or underlying causes.

In Gocan et al. (2020) [Goc+20] gender disparities in the final diagnosis for patients with a provisional diagnosis of TIA/stroke were explored. The data was obtained in 2015 from Ottawa Hospital's Stroke Prevention Clinic and covered 1770 patients. To evaluate a link between clinical factors and patient's sex, they focused on the possible influence of symptoms and disorder characteristics on diagnosis. They found that symptoms such as sudden onset and event duration were more strongly linked to a definitive TIA/stroke diagnosis in women than in men. However, when women reported or presented with pain

they were less likely to be diagnosed with a TIA or stroke. The study suggests that these discrepancies could point towards biases in diagnosis or variations in how each gender reports symptoms. They do highlight, though, that the study's retrospective nature and exploratory design limit how broadly the findings may be applicable.

Across varying factors, including incidence rates, clinical characteristics, treatment approaches, and outcomes, gender-specific variations in strokes and TIAs have been shown. In the studies on data from various countries and clinical settings, men and women present with different risk profiles, symptom patterns, and reactions to treatment. Factors like age and underlying medical conditions further affect these. The findings emphasise the importance of taking gender into account for clinical decision-making and health outcomes.

2.3 Austrian Stroke Unit Registry

The Austrian Stroke Unit Registry is a national quality assurance initiative managed by Gesundheit Österreich GmbH (GÖG) in cooperation with the Austrian Stroke Society (ÖGSF). It was established in 2003 and is regulated by the Federal Law on GÖG section 15a [Ges24]. Its goal is to enhance the quality of stroke care in Austria by systematically collecting, analysing, and visualising stroke-related data from participating stroke units [Öst].

Since Austria does not operate TIA clinics, patients with strokes or TIAs are typically treated in hospitals. Most are managed in stroke units, which are specialised facilities dedicated to the treatment of acute stroke and transient ischaemic attack (TIA) patients [Kno+16]. Currently there are 38 stroke units across Austria from which data is being collected. This is done in an anonymised fashion, and any scientific analyses have to be approved by an expert committee [Öst]. Information on patient characteristics, treatment procedures, and outcomes is recorded, and the documentation uses standardised definitions and scoring systems to ensure consistency and reliability [Kno+16].

2.4 Risk Stratification Research on the Austrian Stroke Unit Registry

In 2016 a study by Knoflach et al. [Kno+16] examined risk stratification for TIAs and minor strokes on data from the Austrian Stroke Unit Registry. They explored the effectiveness of the ABCD2 and ABCD3-I risk scores in predicting early and 3-month stroke recurrence for patients.

In parts this thesis is based on the approach and results of their study.

2.4.1 TIA Sub-Registry

The Austrian Stroke Unit Registry expanded from December 2010 to January 2014 to include additional variables specifically addressing TIA and minor stroke care. These are not otherwise collected in the registry and therefore the data from this specific time frame forms a sort of TIA sub-registry. With these variables the calculation of various TIA risk scores, including ABCD2 and ABCD3-I, is possible. This was central to the study, as they wanted to analyse these scores predictive value. The population included adult patients with ischaemic stroke (NIH Stroke Scale score of less than 4) or TIA, admitted to the stroke unit within 24 hours of symptom onset with less than 6 hours of in-hospital delay. Patients for whom essential variables were missing were excluded from the study. [Kno+16]

Population 1 of this thesis is based on the data from Knoflach et al. [Kno+16]. Therefore, it uses the data of this TIA sub-registry as well as some data filters from the study. The exact steps of which are detailed in the section on the Data 4.

2.4.2 Study Results

The ABCD2 and ABCD3-I risk scores were evaluated for their predictive ability of early and 3-month stroke recurrence in patients with a transient ischaemic attack (TIA) or minor stroke treated at specialised stroke units. The population covered 5237 patients, 2457 of whom had follow-up data available.

In the study [Kno+16] both risk scores were proven to be effective and useful tools for predicting stroke recurrence, with increasing score points correlating with higher stroke probabilities. However, some components were found to be less relevant in this setting compared to outpatient care. Traditional risk factors such as age, blood pressure, and diabetes had less predictive strength. Instead, imaging results, clinical presentation, and symptom duration stood out as the most significant predictors. The study also emphasises how crucial early and specialised care is for reducing the risk of stroke and the necessity of comprehensive understanding of high-risk patients in order to implement targeted prevention measures.

2.5 Research Gap

In numerous studies gender-specific differences in stroke and transient ischaemic attack (TIA) incidence, symptoms, treatments, and outcomes have been examined. For risk stratification techniques there is a gap for assessment of this kind. The focus has mainly been on validating their predictive performances on general and specific populations, with little attention to how the scores perform across gender groups.

Existing studies such as those by Purroy et al. (2021) [Pur+21] and Gocan et al. (2020) [Goc+20] have highlighted variations in clinical presentation and diagnostic patterns in TIA patients between genders. However, the effectiveness and predictive accuracy of

risk stratification tools, like the ABCD2 and ABCD3-I scores, are not evaluated. Prior research into the Austrian Stroke Unit Registry by Knoflach et al. (2016) [Kno+16] has also validated these risk scores in predicting stroke recurrence on the stroke unit data.

Due to the evidence that gender in part influences TIA presentation and subsequent stroke risk factors, it is important to assess whether existing risk scores are able to provide robust predictions regarding gender. Using data from the Austrian Stroke Unit Registry, this thesis attempts to close this gap through gender-specific analysis of ABCD2 and ABCD3-I scores. It will also explore how gender affects health outcomes and treatment strategies after a TIA in Austrian stroke units. This thesis examines these aspects with the goal of gaining insight into the effect of gender on risk classification, improving targeted stroke prevention strategies, and guiding more equitable healthcare practices.

CHAPTER 3

Methodology

This chapter outlines the methodological approach used in this thesis. The overall structure of the research and analysis process, which is split into three parts, and the evaluation steps for each of the gender-specific differences are presented. The statistical methods used throughout the analyses are explained, with particular attention on ensuring robustness in the results.

3.1 Research Design

The thesis covers three main topics, which are each defined by a distinct research question. To thoroughly explore the gender-specific differences in transient ischaemic attacks of each aspect, the analysis approach is structured into three parts. In separate analyses, health outcomes, risk stratification and treatment methods are examined and investigated for potential differences based on the gender of patients. Due to this, each section utilises at least somewhat different approaches and statistical methods.

3.1.1 Gender-Specific Differences in Health Outcomes

The first part of the analysis focuses on the patient's health and possible additional stroke events after having experienced a TIA. The likelihood of stroke recurrence or complications like epileptic seizures is evaluated and compared between genders. The goal is to determine if men and women have different clinical outcomes after a TIA, while also accounting for confounding factors like age and comorbidities.

For the analysis, logistic regression models will be used with gender and additional covariates to predict the health outcomes of the patients. Odds ratios (ORs) and their confidence intervals (CIs) will be used to assess the strength and direction of correlations between gender and clinical outcomes.

3.1.2 Gender-Specific Differences in the Predictive Accuracy of ABCD2 and ABCD3-I Scores

The second section examines if the ABCD2 and ABCD3-I risk scores, used frequently to predict the probability of stroke recurrence following a TIA, perform equally well for both genders. The aim is to determine whether gender-specific adjustments may be necessary to improve the predictive accuracy of these risk scores.

To evaluate the performance of the scores, the receiver operating characteristic (ROC) curves and the resulting area under the curve (AUC) as well as sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) will be used. With these the risk scores' effectiveness across genders can be compared. Additionally, the effect of other factors like age and comorbidities will also be examined to ensure any differences are truly dependent on gender.

3.1.3 Gender-specific Differences in Diagnostic Methods and Secondary Prophylactic Treatments

In the third analysis possible gender influences on the diagnostic methods used and the secondary prophylactic treatments prescribed after a TIA are investigated. On one hand, the likelihood of patients receiving treatment will be examined for any differences between men and women. On the other hand, the treatment rates for certain conditions like hypertension or atrial fibrillation will be compared.

The first part will be analysed using logistic regression and odds ratios to model the likelihood of receiving specific diagnostic tests or treatments. Again, adjustments will be made for covariates like age and pre-existing conditions. The second part assesses how frequently patients with a specific condition receive treatment, providing insight into how certain illnesses are handled.

3.1.4 Dataset Populations

An important component of the thesis is the use of two dataset populations. While both are from the Austrian Stroke Unit Registry, one forms the TIA sub-registry (2010-2014) 2.4.1 and the other covers more recent but not TIA-centric data. While older, the TIA sub-registry provides more detailed and thorough information on important aspects for this thesis. Specifically, the risk scores - ABCD2 and ABCD3-I - can only be calculated for this population. This means for the risk stratification and analysis of the scores predictive accuracy only the TIA sub-registry data can be used. The health outcomes and treatment methods, however, can be explored across both populations, allowing for extensive analyses and comparisons.

3.2 Statistical Methods

In this section the statistical methods employed to analyse gender-specific differences are outlined. It was explained above which methods are used for each aspect of the three-part analytical approach.

Robustness is an important concept to consider when working with real-world data, which often deviates from theoretical assumptions due to outliers, measurement errors, or model misspecifications. The goal of robust statistics is to provide methods that are resistant to such deviations from ideal conditions. According to 'Robust Statistics' [HR09, pp. 1–6], a robust approach guarantees statistical techniques that produce consistent and reliable results, even in cases when assumptions about the data, such as its distribution, are incorrect. Robust statistical techniques can provide more reliable and broadly applicable conclusions, which in the case of medical data is crucial due to the possible impact of any findings.

In the context of robust statistics, the idea of the 'true model' is also relevant. This refers to the actual underlying process which generates the data, including the genuine relationships between variables and therefore representing the full reality. However, in practice, when dealing with medical data, any interactions are multitudinous and highly complex, with many unknown aspects. Instead of identifying the absolute truth, the goal of statistical modelling is to construct an approximation that best represents the data observed based on finite information [BA02, p. 20]. A model is evaluated on its ability to identify important patterns while remaining robust against false assumptions.

In this thesis, to ensure robustness, multiple logistic regression models with different covariate selections are used. Depending on the section of the analysis, either odds ratios and confidence intervals are then calculated, or ROC curves are generated to evaluate model performance. The covariates used are patients characteristics, meaning variables describing the patient and their condition prior to the TIA event. These can affect the result of the models, as for example a prior heart condition may influence the patient's health outcome as well as the treatment they receive. The specifics of these variables and covariate combinations used for the analyses are explored in Chapter 4.4.3.

Comparing the outcomes of several covariate selections can increase the confidence in the findings. It helps in assessing the stability of effect estimates and identifying covariates which are influential in controlling for bias. Possible issues like confounding or overfitting can be prevented through careful selection of covariates. In medical research, as described in [Bio13], it is particularly important to identify and include relevant risk factors. They increase the probability of a certain health issue but do not guarantee it, which are exactly the type of covariates used in this thesis. Properly adjusting for known risk factors allows for more accurate estimations of model relationships.

3.2.1 Logistic Regression

The formulas and explanations presented in this section are primarily based on the book 'An Introduction to Statistical Learning' [Jam+21], specifically Chapter 4.3.

Logistic regression is a commonly used statistical technique for simulating the probability of a binary or multinomial outcome. The difference to linear regression is that instead of continuous values, the logistic function predicts the likelihood of an observation belonging to a particular category.

Binary Logistic Regression

In the case of the response variable having two possible outcomes, binary logistic regression is used. An example would be predicting whether a patient suffers a certain health condition or not. The response categories can be encoded as 0 and 1, so any predictions fall within this value range. The model uses a set of predictor variables X to predict the probability of an outcome $p(X)$. These values are calculated using the logistic function:

$$p(X) = \frac{e^{\beta_0+\beta_1 X}}{1 + e^{\beta_0+\beta_1 X}} \tag{3.1}$$

The β coefficients are estimated using the maximum likelihood method. The results of this function need to be interpreted, and for this, the odds of an event occurring need to be estimated. These odds are given by:

$$\frac{p(X)}{1 - p(X)} = e^{\beta_0+\beta_1 X} \tag{3.2}$$

This relationship is transformed by taking the logarithm on both sides. The left side of the equation is referred to as the log odds or logit.

$$\log\left(\frac{p(X)}{1 - p(X)}\right) = \beta_0 + \beta_1 X \tag{3.3}$$

Instead of having a direct linear impact on the probability, every change in X results in a multiplicative change of the odds. Because the odds for $X + 1$ are the odds for X multiplied by e^{β_1}.

In this thesis, logistic regression is performed in R using the *glm* function, which allows for flexible model specification and inference [Tea19]:

```
model <- glm(response ~ predictors, data = dataset, family = binomial)
```

Multiple Logistic Regression

When modelling an outcome, multiple predictors, like additional factors or covariates, may need to be considered. Due to the logit transformation, logistic regression can easily be extended to include several independent variables:

$$\log\left(\frac{p(X)}{1 - p(X)}\right) = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p \tag{3.4}$$

where $X_1, X_2, ..., X_p$ are multiple predictors influencing the probability of the outcome.

With this method confounding factors can be taken into account and bias in the estimates may be reduced. This approach will be used in this thesis to include known risk factors in the modelling process and compare the effect they have on the predictions.

Multinomial Logistic Regression

When dealing with a response variable that has more than two categories, multinomial logistic regression needs to be applied. Instead of modelling a single probability like binary logistic regression, multinomial logistic regression estimates separate equations for each category in reference to a baseline category. Out of K classes, this baseline is selected first, for example the Kth class, and then the model predicts the probability of the other k = 1,...,K - 1 categories:

$$Pr(Y = k|X = x) = \frac{e^{\beta_{k0}+\beta_{k1}x_1+\cdots+\beta_{kp}x_p}}{1 + \sum_{l=1}^{K-1} e^{\beta_{l0}+\beta_{l1}x_1+\cdots+\beta_{lp}x_p}} \tag{3.5}$$

It can be demonstrated that the log odds between any two classes for these k = 1,...,K - 1 categories are linear in the features:

$$\log\left(\frac{Pr(Y = k|X = x)}{Pr(Y = K|X = x)}\right) = \beta_{k0} + \beta_{k1}x_1 + \cdots + \beta_{kp}x_p \tag{3.6}$$

The R function *multinom* from the *nnet* package is used for multinomial logistic regression [Rip19]:

```
model <- multinom(response ~ predictors, data = dataset)
```

3.2.2 P-Value and its Limitations

The p-value is a widely used statistical metric to assess how well observed data supports the null hypothesis. Usually, the null hypothesis states that there is no effect, no difference or no relationship between the variables in a population. Assuming the null hypothesis is true, the p-value is the probability of obtaining a result as extreme as or more extreme than the observed one. A small p-value is often seen as evidence against the null hypothesis, since it suggests that such a result would be unlikely if the null hypothesis were right [Jam+21, pp. 67–68].

However, the use of p-values has been heavily criticised due to the potential for misinterpretation. According to Goodman (1999) [Goo99], a false belief exists that the evidential strength of a single result and the probability of mistake over several experiments can both be captured by a single number. This simplifies complicated scientific findings into binary statements of significant or not significant. The p-value is at times also incorrectly interpreted to state the chance that the null hypothesis is false. But as it is based on

the assumption of the null hypothesis being true, it cannot measure the probability of it being false.

Goodman's criticisms are also echoed by Pandis (2013) [Pan13b], who highlights that the problem is not just a matter of misunderstanding but also of the inherent limits of p-values as inference tools. They can cause trivial effects in a large sample to look 'significant' since they don't indicate effect size. Both Goodman and Pandis recommend using confidence intervals (CIs) as indicators for the importance of the observed effect rather than drawing conclusions from p-values.

Furthermore, by offering a range of plausible values for the effect, confidence intervals transform the interpretation from a binary decision to a quantitative evaluation. A confidence interval not only indicates if a result is statistically significant but also its potential precision and clinical relevance. Unlike p-values, which are influenced by sample size, CIs are linked to the actual effect and only grow narrower with additional data without changing the result. Pandis (2013) [Pan13a] suggests the use of confidence intervals and effect estimates to be the better approach.

He had also previously described how effect size metrics, like odds ratios, can improve statistical inference [Pan12]. These measure the extent of difference between two groups, allowing researchers to determine if the difference is clinically significant. A p-value can suggest a statistical difference, but only effect size can demonstrate whether or not it impacts real-world decisions.

3.2.3 Odds Ratio

The odds ratio (OR) is used to assess the relationship between an exposure and an outcome. The exposure refers to a variable or factor – in the case of this thesis, the gender of a patient – that is assumed to have some effect on an outcome. With OR the degree to which an event is more likely to occur for one group of people compared to another can be expressed. It is used in logistic regression, as it can provide an interpretation of the relationship between predictor variables and the outcome of interest. The advantage is that it can be derived naturally from the model coefficients. The odds for logistic regression were already defined by the equation 3.2 above, in the case of a binary variable with the outcomes coded as 0 and 1. Based on this, the odds ratio can then be defined as follows:

$$OR = \frac{p(X = 1)/[1 - p(X = 1)]}{p(X = 0)/[1 - p(X = 0)]} = e^{\beta_1} \tag{3.7}$$

This expression shows that a unit increase in the predictor variables X multiplies the odds by e^{β_1}, which is the coefficient. For multiple logistic regression with several predictors or confounding variables, each coefficient represents the adjusted odds ratio, controlling for the other variables in the model. [KK10, pp. 22–27]

With the odds ratio, both the strength and the direction of the relationship between two variables can be measured. This interpretation depends on the value [Szu10]:

- OR = 1: The exposure has no effect on the odds of the outcome. The odds of the event occurring are the same regardless of the group.
- OR > 1: The exposure is associated with higher odds of the outcome. An OR of 1.5 means the odds of the event are 50% more likely for one (the exposed) group than the other.
- OR < 1: The exposure is associated with lower odds of the outcome. An OR of 0.5 means the odds of the event are half as likely for one (the exposed) group than the other.

Confidence Interval

As described by Szumilas (2010) [Szu10] the precision of the estimated odds ratio can be determined using its confidence interval (CI). A 95% confidence interval shows the range in which the true odds ratio is likely to fall, with 95% certainty. If 1 is included in the confidence interval, then the correlation may not be overly significant, as an odds ratio of 1 suggests no effect. However, if the confidence interval does not include 1, then the width of the interval needs to be considered. A narrow confidence interval suggests more precise estimation of the odds ratio, which may indicate a big sample size or consistent data. Less accuracy is provided by a large confidence interval, which can happen when the sample size is smaller or the data is more variable.

The 95% confidence interval for the odds ratio can be calculated using this formula [KK10, p. 149]:

$$\mathrm{CI} = \exp\left(\hat{\beta} \pm 1.96 \times \mathrm{SE}(\hat{\beta})\right) \tag{3.8}$$

where $\hat{\beta}$ is the estimated regression coefficient, $\mathrm{SE}(\hat{\beta})$ is the standard error of the estimated coefficient and 1.96 is the critical value for a 95% confidence interval.

3.2.4 Receiver Operating Characteristic (ROC) Curve

The receiver operating characteristic (ROC) curve is a graphical tool to evaluate a classifier's performance, in this thesis a risk score that predicts the recurrence of a stroke. It is frequently used in fields like signal detection, machine learning, and medical diagnostics, where the ability to distinguish between two possible outcomes is crucial. Classifiers can determine the classification of an instance at different cut-off values or thresholds, and these different results are shown by the ROC curve.

The ROC curve plots the true positive rate (TPR), often referred to as sensitivity or recall, against the false positive rate (FPR). The FPR is also the same as 1 - specificity, with specificity being the proportion of actual negative instances that are correctly classified as negative. In the graph 3.1 two such ROC curves are displayed. The curve shows the trade-off between the true positive rate (sensitivity) and the false positive rate (1-specificity) over various decision thresholds [KK10, p. 355].

On the ROC graph, a classifier that perfectly distinguishes between positive and negative cases would have a point at (0, 1) since it would have zero false positives and 100% true positives. A classifier that operates at random, on the other hand, will fall on the diagonal line from (0,0) to (1,1). This means the true positive rate and the false positive rate are identical, and the classifier is useless. A well-performing classifier will have higher sensitivity and lower FPR, which will put it closer to the graph's upper left corner. It should properly detect the majority of positives and avoid false positives [Faw06]. For the ROC curves in the graph 3.1, it can be said that the blue curve performs better than the green curve, as it is closer to the top left corner. One important metric derived from the ROC curve is the area under the curve (AUC). It measures the ability of the classifier to discriminate between positive and negative cases and reduces the ROC down to a single comparable value. The AUC can have values from 0 to 1, with 0.5 representing a classifier with random performance and thus no ability to distinguish between classes. This also means that values below 0.5 are worse than random guessing, in which case the classifier's reversed decisions would lead to better results. In general for a classifier an AUC between 0.9 and 1 indicates excellent performance, while a value between 0.8 and 0.9 is considered good. AUC values between 0.7 and 0.8 reflect fair performance, and values between 0.6 and 0.7 are deemed poor [KK10, pp. 356–357]. According to this, the classifier of the blue ROC curve in the graph 3.1 does an excellent job with an AUC of 0.92, but the classifier of the green ROC curve only produces fair classification results with an AUC of 0.73.

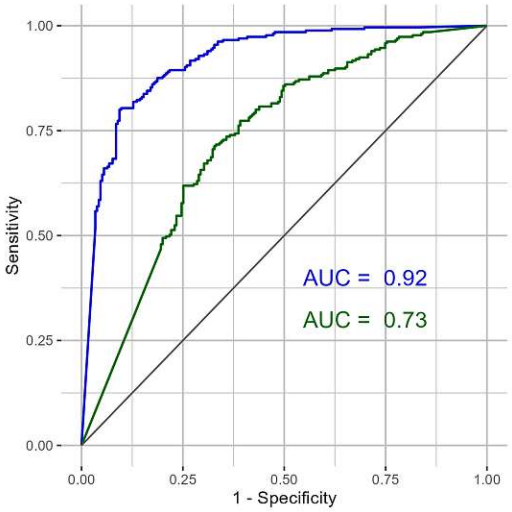


Figure 3.1: Example ROC Curves

Performance Metrics

A common way to evaluate classifiers is to use performance metrics based on the confusion matrix 3.1. With these, it can be quantified, how well a classifier can distinguish between the positive and negative classes. Two of these metrics are sensitivity and specificity, which, as they define the ROC curve, can of course be compared for each threshold value.

	Predicted Positive	**Predicted Negative**
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

Table 3.1: Confusion Matrix

Sensitivity, or the true positive rate (TPR), is the proportion of real positive instances correctly identified by the classifier. In contrast, specificity is the true negative rate (TNR) and, consequently, the proportion of real negative instances that are accurately identified. A high sensitivity reduces the amount of false negatives and shows a good ability to recognise positive cases, while a high specificity reduces the amount of false positives and shows a good ability to recognise negative cases.

$$\text{Sensitivity (TPR)} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{3.9}$$

$$\text{Specificity (TNR)} = \frac{\text{True Negatives}}{\text{True Negatives} + \text{False Positives}} = 1 - FPR \tag{3.10}$$

Additionally, the positive predictive value (PPV), also called precision, and the negative predictive value (NPV) can be used as performance metrics. They each measure how likely an instance – classified as positive or negative, respectively – is to have been identified correctly. A high PPV means that when an instance is predicted to be positive by the classifier, this is more likely to be correct. The NPV works the same but for negative predictions. [Faw06]

$$\text{PPV} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{3.11}$$

$$\text{NPV} = \frac{\text{True Negatives}}{\text{True Negatives} + \text{False Negatives}} \tag{3.12}$$

In this thesis the function *roc* from the *pROC* package [Rob+23] in R is used for ROC analysis. It returns an ROC curve object which can be plotted and also contains the area under the curve (AUC) value. The coordinates of different performance metrics at various thresholds can be retrieved using the *coords* function from the same package.

3.2.5 Data Imputation

In order to deal with missing values in a dataset, data imputation can be used. If missing values are not further addressed, it can in some cases lead to skewed estimates, reduced statistical power and flawed conclusions. Imputation techniques use the observed data to fill in missing entries with plausible values, rather than discarding these, which could lead to substantial data loss [Buu12, pp. 5–6]. The dataset's structure and integrity are preserved for the analysis, though these values are not meant to predict the true missing data but rather reflect a range of potential values. There are various approaches that can be used for data imputation, one of the more effective ones being multiple imputation. It does not only create numerous full datasets by replacing missing values several times but also accounts for uncertainty around the missingness in the statistical analyses afterward [Buu12, pp. 16–18].

In this thesis multiple imputation by chained equations (MICE) is used, with the *mice* package in R. [Tea25] Missing values for each variable are conditionally imputed in MICE,

based on other variables in the dataset [Buu12, p. 109]. To improve model performance and interpretability, the variables were prepared before imputation by ensuring that the data types were appropriate. In this case categorical variables were converted to factors and continuous variables were transformed into categorical formats where necessary. For the imputation model the random forest method was chosen, which is robust to model misspecification and ideal for capturing complex, nonlinear relationships in the data [Buu12, pp. 123–126]. Multiple imputations were created in accordance with standard guidelines to account for the uncertainty present in missing data. This approach uses all of the data that is available while guaranteeing that analyses based on the completed dataset produce reliable statistical results.

```
imputed_data <- mice(data, method = 'rf', m = 15)
```

CHAPTER 4

Data

The data used in this thesis is from the Austrian Stroke-Unit Registry. It is entered in a decentralised fashion by currently 38 stroke units and the patient data is anonymised.

Two populations from the registry are being used, which were collected during different time periods. The first population makes use of the temporary expansion of the registry, where additional TIA and minor stroke specific variables were collected. This same sub-registry was used in the study by Knoflach et al. (2016) [Kno+16] which evaluated the predictive value of the ABCD2 and ABCD3-I scores. Their approach is adapted in this thesis, making use of access to documents specifying their filter steps and formulas used for additional variable calculations. The second population covers more recent data, but it lacks these TIA and minor stroke specific variables. The calculation of the ABCD2 and ABCD3-I scores is not possible with this data, so no analyses of their predictive value can be performed. The research design and the analyses done on each population were discussed in Chapter 3.1 above.

In this chapter the preprocessing steps for Population 1 and Population 2 are explained. Information is provided on all the variables used in the analyses. These are either taken directly from the registry or are additionally created where necessary.

Although the English version of the registry data is being used, some data entries are still in German. These were translated for consistency and understandability. Specifically, "Bereits vorliegend" to "Already on hand", "Vorgesehen" to "Planned", "Ja, bereits bekannt" to "Yes, already known", "Keines" to "None", "Keine Information" to "No information" and "Keine Rehabilitation" to "No Rehabilitation".

4.1 Data Filtering

The data filtering steps for Population 1 were taken from the previous study by Knoflach et al. (2016) [Kno+16]. Their aim was to reduce the data to only relevant TIA and

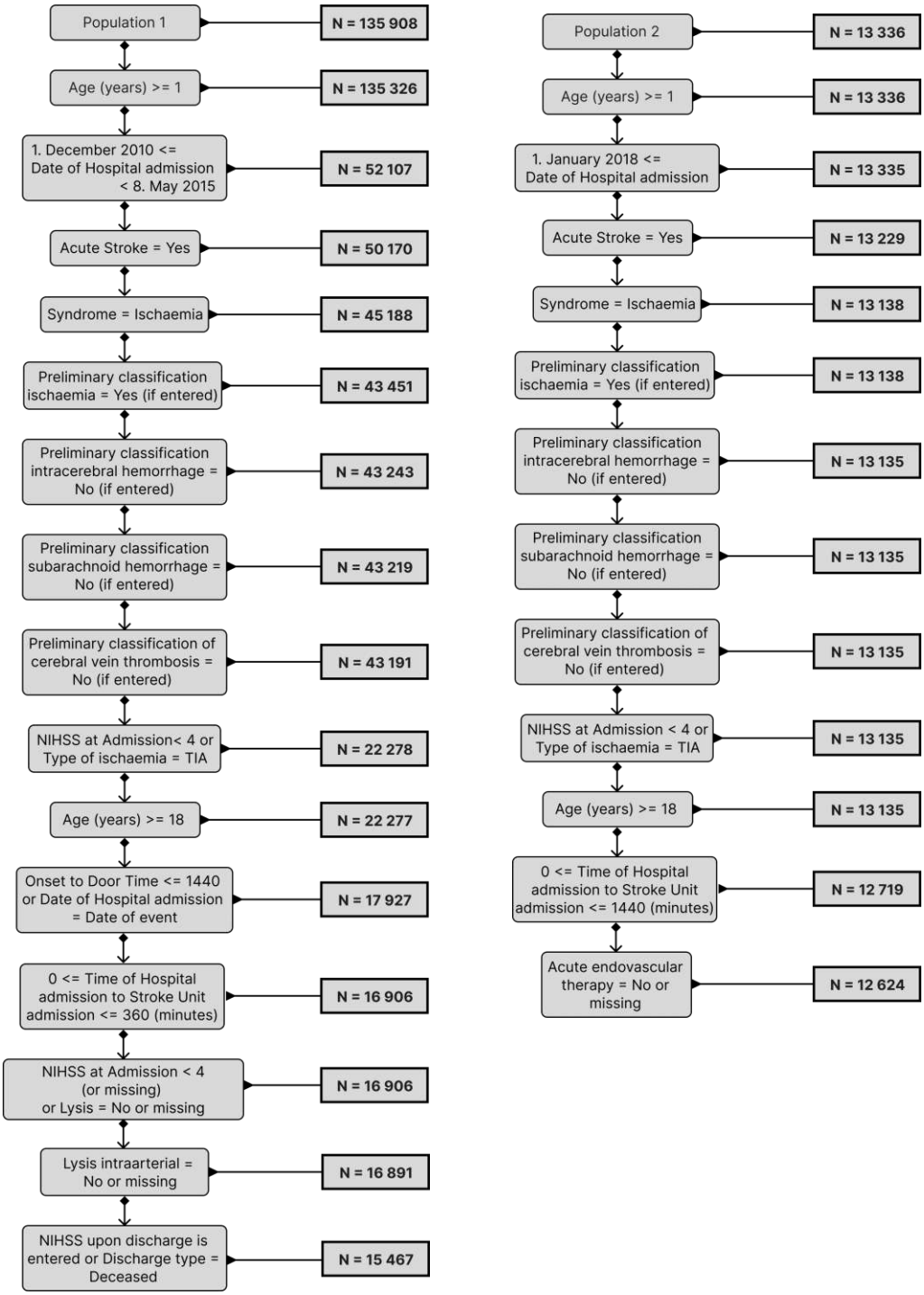


Figure 4.1: Filter Steps Flow Charts for Population 1 and 2

minor strokes, with some additional plausibility checks. They defined the time frame of their data from the first of December 2010 to the eighth of May 2015, this being the time of their download from the registry. In this thesis the same dates were applied as filters for consistency. For Population 2, most of the same filtering steps were used, though the plausibility checks were changed. The data for this population was downloaded on the 27th of October 2024 and had already been pre-filtered according to this thesis' specifications. Therefore, as can be seen in this population's filter flow chart 4.1b, the sample size is only minimally affected by most of the filters, as some had already been applied prior. In contrast, for Population 1, the only pre-filter was for data prior to 2018, which is why the sample size is far larger initially with 135,908 entries. But as can be seen in the flow chart 4.1a, this is quickly reduced by filtering for the required time range.

For both populations the validity of the data is confirmed by checking that patient entries have ages of over one year before filtering for the specific dates. Then through seven filter steps, an acute ischaemic stroke and specifically either a minor one, meaning the NIH Stroke Scale score at admission is below 4, or a TIA are confirmed. It is ensured that the patients are older than 18 in order to limit the thesis to adult patients.

After this, the populations use different plausibility checks. For Population 1 the choice was made not to alter those used in the study [Kno+16], which reduces the sample size by over 6,500 patients. In order to retain more patients of Population 2, it was decided to apply less limiting plausibility checks.

For Population 1, the acuteness of the event is checked using onset and admission times, as well as that it wasn't actually a more severe stroke. This is done by ensuring a high admission NIHSS score wasn't entered and that no lysis therapy was done, which would suggest an at least moderate or severe stroke. Lastly, it is checked that there is an NIHSS score available for the patient upon discharge from the hospital, or if not, that the patient had died. For Population 2, an allowance for a longer time between hospital and stroke unit admission was made to cover patients who suffered a minor stroke or TIA within a day instead of within 6 hours. The last check is that the patient didn't receive acute endovascular therapy, as this would be typically used for severe cases.

This results in sample sizes of 15,467 and 12,624 patients for Population 1 and 2, respectively. For the analyses of the health outcomes of patients and the diagnostic and treatment methods, the full populations will be used. However, for the analysis of the risk stratification, the existence of the ABCD2 and ABCD3-I score is necessary. These need to be calculated, which is only possible for Population 1, and there also only for a subset, as will be shown in the next section.

4.2 ABCD2 & ABCD3-I Scores

The formulas/definitions in this section for the risk scores and their components are those used in the study by Knoflach et al. (2016) [Kno+16]. They are based on the variables available in the Austrian Stroke Unit Registry and, specifically, the additional variables

collected for the TIA sub-registry. These formulas match the general definitions of the scores which were explained in the chapter on Risk Stratification 2.1.2.

In Population 1 an ABCD2 score, which was directly entered into the registry database by care providers, exists. Further, all the necessary variables to re-calculate the ABCD2 and calculate the ABCD3-I score are only available in Population 1. Therefore, any calculations explained here were only done on that dataset.

It is important to note that the accuracy of the provided and the re-calculated ABCD2 score depends on different factors: the existing score relies on the correctness and completeness of its own entries, while the re-calculated score depends on the correctness of the compounded variables. These two scores can be compared in their predictive accuracy. In order to do this they need to be clearly distinguishable. Therefore, going forward the score entered by care providers will be referred to as the *database* score and the other one as the *calculated* score.

The ABCD3-I score has to be calculated regardless, as it doesn't exist within the population. This score is based on the ABCD2 score, so either the *database* or the *calculated* score can be used.

$$\begin{aligned} \text{ABCD} &= \text{A} + \text{B} + \text{C} + \text{D1} \\ \text{ABCD2} &= \text{ABCD} + \text{D2} \\ \text{ABCD3I} &= \text{ABCD2} + \text{D3} + \text{I2} \end{aligned}$$

A ... Age
B ... Blood pressure
C ... Clinical features
D1 ... Duration
D2 ... Diabetes
D3 ... Dual TIA
I2 ... Imaging

The scores can be defined as the elements shown above, with the letters representing the initials. Each component has a point value, which is added up for an ABCD2 score of 0-7 and an ABCD3-I score of 0-13.

For the calculation and comparison of the ABCD2 score, the formulas below were used to define the components. They are based on the variables from the Austrian Stroke Unit Registry and were taken from an internal document of the Knoflach et al. (2016) study [Kno+16].

$$\text{A} = \begin{cases} 1 & \textbf{if } \text{Age} \geq 60 \\ 0 & \textbf{if } \text{Age} < 60 \end{cases}$$

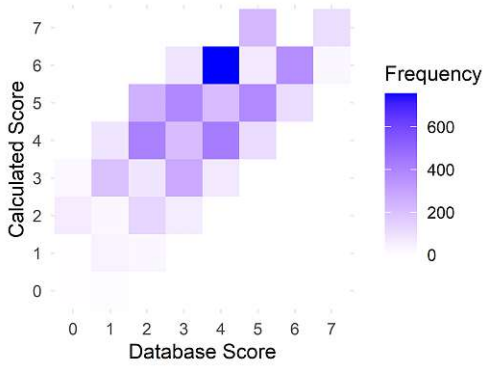
$$\text{B} = \begin{cases} 1 & \textbf{if } \text{Systolic blood pressure (i100005)} \geq 140 \\ & \quad \textbf{or } \text{Diastolic blood pressure (i100006)} \geq 90 \\ 0 & \textbf{if } \text{Systolic blood pressure (i100005)} < 140 \\ & \quad \textbf{and } \text{Diastolic blood pressure (i100006)} < 90 \\ \text{NA} & \textbf{else} \end{cases}$$

$$C = \begin{cases} 2 & \textbf{if one-sided weakness (i100001)} = \text{Yes} \\ 1 & \textbf{if one-sided weakness (i100001)} = \text{No} \\ & \textbf{and} \\ & [\text{Aphasia (i100002)} = \text{Yes} \\ & \quad \textbf{or Dysarthria (i100003)} = \text{Yes}] \\ 0 & \textbf{else} \end{cases}$$

$$D1 = \begin{cases} 0 & \textbf{if } 0 \leq \text{Duration of Symptoms (i100000)} \leq 10 \\ 1 & \textbf{if } 10 < \text{Duration of Symptoms (i100000)} \leq 60 \\ 2 & \textbf{if } 60 < \text{Duration of Symptoms (i100000)} \\ & \quad \textbf{or Duration} > 24 \text{ Hours (i100047)} = \text{Yes} \\ \text{NA} & \textbf{else} \end{cases}$$

$$D2 = \begin{cases} 1 & \textbf{if Diabetes (i9002)} = \text{Yes} \\ 0 & \textbf{if Diabetes (i9002)} = \text{No} \\ \text{NA} & \textbf{else} \end{cases}$$

The *calculated* ABCD2 score can be created using these and then compared to the population's *database* score. The confusion matrix of the two scores' values in the graphic below 4.2 shows that the *calculated* score is generally one to two points higher than the *database* score. The table further confirms this distribution, as well as that 88 fewer patients have data sufficient for calculating an ABCD2 score. This is an important difference, as only 5,331 out of the total 15,467 patients already have a *database* ABCD2 score, so any further reduction in available data should be considered carefully.



Risk Level	*# Database Score*	*# Calculated Score*
0	92	15
1	367	71
2	921	274
3	1040	654
4	1499	1251
5	797	1341
6	482	1303
7	133	334
NA	10136	10224

Figure 4.2: *calculated* vs *database* ABCD2 Score

The observed discrepancies between the *calculated* and *database* ABCD2 scores are significant. Ideally, both should align closely, as they are based on the same fixed criteria.

Yet, the calculated scores tend to be slightly higher. This difference raises questions about potential variations in data entry, interpretation, and processing. Given how important the assessment of a patient's risk for stroke recurrence is, precise calculation of the risk score is crucial.

$$\text{D3} = \begin{cases} 2 & \textbf{if Previous TIA within 7 days (i10018)} = \text{Yes} \\ 0 & \textbf{if Previous TIA within 7 days (i10018)} = \text{No} \\ \text{NA} & \textbf{else} \end{cases}$$

$$\text{I2} = \begin{cases} 4 & \textbf{if (individual corresponding lesion (i100020)} = \text{Yes} \\ & \quad \textbf{or individual not corresponding lesion (i100021)} = \text{Yes} \\ & \quad \textbf{or multiple lesions in a vascular supply area (i100022)} = \text{Yes} \\ & \quad \textbf{or multiple lesions in multiple vascular supply areas (i100023)} = \text{Yes} \\ & \quad \textbf{or lesions of different ages (i100024)} = \text{Yes} \\ & \quad \textbf{or old/subacute lesion on CCT (i100025)} = \text{Yes)} \\ & \textbf{AND ipsilateral carotid stenosis \% (i100045)} \geq 70 \\ 2 & \textbf{if (individual corresponding lesion (i100020)} = \text{Yes} \\ & \quad \textbf{or individual not corresponding lesion (i100021)} = \text{Yes} \\ & \quad \textbf{or multiple lesions in a vascular supply area (i100022)} = \text{Yes} \\ & \quad \textbf{or multiple lesions in multiple vascular supply areas (i100023)} = \text{Yes} \\ & \quad \textbf{or old/subacute lesion on CCT (i100025)} = \text{Yes)} \\ & \textbf{AND ipsilateral carotid stenosis \% (i100045)} < 70 \text{ (or NA)} \\ 2 & \textbf{if (individual corresponding lesion (i100020)} = \text{No} \\ & \quad \textbf{and individual not corresponding lesion (i100021)} = \text{No} \\ & \quad \textbf{and multiple lesions in a vascular supply area (i100022)} = \text{No} \\ & \quad \textbf{and multiple lesions in multiple vascular supply areas (i100023)} = \text{No} \\ & \quad \textbf{and old/subacute lesion on CCT (i100025)} = \text{No)} \\ & \textbf{AND ipsilateral carotid stenosis \% (i100045)} \geq 70 \\ 0 & \textbf{else} \end{cases}$$

The *database* score reflects how information was originally documented in the hospital, while the *calculated* score follows a transparent, reproducible method strictly based on the established definitions of the ABCD2 score. With the exact process used to generate the *database* score being untraceable, it is difficult to determine the source of these differences.

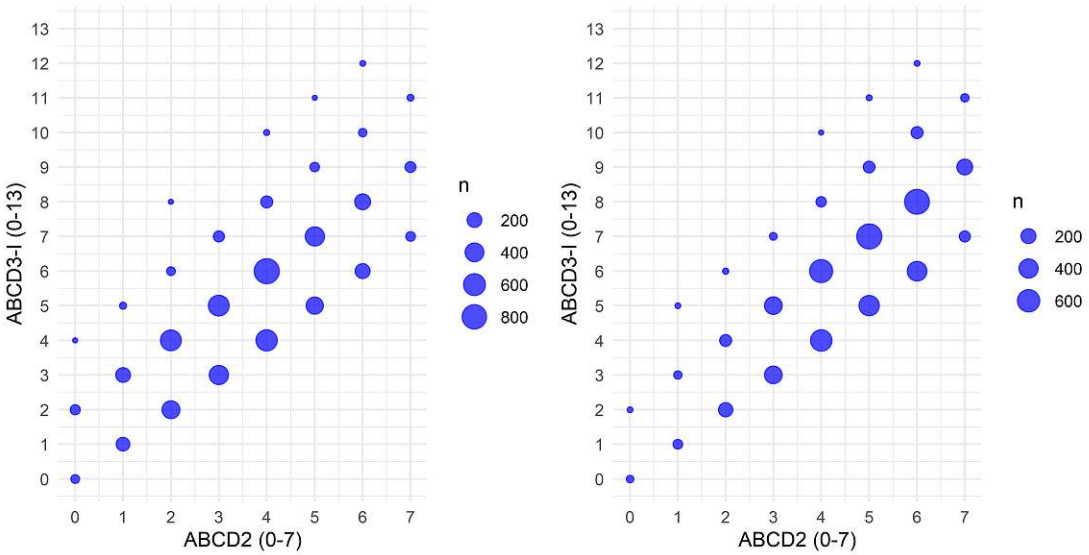
Neither version of the score can be labelled as incorrect, but they do provide slightly different assessments of risk and this is likely to affect their predictive value. Therefore, using both scores in the analysis allows for a direct comparison of the two and a thorough examination of their usefulness in risk stratification. The significance of data quality

in clinical research can be highlighted, and a more comprehensive evaluation of their predictive power can be performed.

With two versions of the ABCD2 score in the population — the *database* score and the *calculated* score — two corresponding versions of the ABCD3-I score can also be created. The ABCD2 score builds the base for the calculation with two additional components - dual TIA (D3) and imaging (I2).

In order to avoid confusion, even though both ABCD3-I scores are calculated, the one based on the ABCD2 score taken directly from the registry will also be referred to as the *database* ABCD3-I score. The other with the *calculated* ABCD2 score as the base will then be called the *calculated* ABCD3-I score. This way the distinction between the two sets of risk scores is clear.

There is a strong correlation between the ABCD2 and ABCD3-I. This relationship can be observed in the plots below. Differences in distribution can also be seen between the *database* 4.3a and the *calculated* scores 4.3b. The trend for slightly higher values in the *calculated* risk scores remains evident. Most *calculated* ABCD2 scores are between 4 and 6, while the majority of the *database* ABCD2 scores are centred around 3 and 4.



(a) Database Scores (b) Calculated Scores

Figure 4.3: Risk Score Scatter-plots

For the analysis of the risk stratification, the population will be limited to the subset of entries for which an ABCD2 and/or ABCD3-I score is available. At most this reduces the sample size to 5,331 patients. For these the *database* ABCD2 score is available, which is the score with the most entries, with only 13 fewer entries for its associated ABCD3-I score. The exact number of values for each score can be seen in this Table 4.1

and it allows for a proper comparison between the *database* and *calculated* scores. It is noteworthy, that while the ABCD3-I score has a range of 0-13, there are no entries with a score of 13 in this population and only two patients with scores of 12. These are in fact the same patients who have the highest scores in both the *database* and the *calculated* ABCD3-I score.

Table 4.1: Population 1. Risk Scores Subset

Variables	**Database** N = 5,331[1]	**Calculated** N = 5,331[1]
ABCD2		
0	92 (1.7%)	15 (0.3%)
1	367 (6.9%)	71 (1.4%)
2	921 (17%)	274 (5.2%)
3	1,040 (20%)	654 (12%)
4	1,499 (28%)	1,251 (24%)
5	797 (15%)	1,341 (26%)
6	482 (9.0%)	1,303 (25%)
7	133 (2.5%)	334 (6.4%)
NA	0 (NA%)	88 (1.7%)
ABCD3-I		
0	33 (0.6%)	13 (0.2%)
1	153 (2.9%)	46 (0.9%)
2	412 (7.7%)	174 (3.3%)
3	632 (12%)	344 (6.6%)
4	1,085 (20%)	638 (12%)
5	858 (16%)	777 (15%)
6	1,082 (20%)	1,076 (21%)
7	556 (10%)	879 (17%)
8	347 (6.5%)	834 (16%)
9	120 (2.3%)	323 (6.2%)
10	28 (0.5%)	101 (1.9%)
11	10 (0.2%)	28 (0.5%)
12	2 (<0.1%)	2 (<0.1%)
NA	13 (0.2%)	96 (1.8%)

[1] n (%)

For the risk stratification analysis 5.2, both versions of the scores will be examined and their predictive value compared. It can be evaluated how the differences between the *database* and *calculated* scores may affect risk assessment. The potential implications for clinical decision-making can then be explored. Additionally, considering both versions

allow for an in-depth analysis of the potential gender differences in risk prediction. The exact number of available values for each score is summarised in Table 4.1, providing a clear overview for comparison.

4.3 Additional Variables

There are additional variables that will be used in the analyses, which can be defined with data from the registry. The first three are also used in the study by Knoflach et al. (2016) [Kno+16] and pertain to patients' health outcomes and stroke recurrence. In both populations the variable for Early Recurrence in Stroke Unit can be defined. However, Early Worsening and Recurrence within 90 days are defined specific to TIA, and the necessary data is only available in the TIA sub-registry, meaning they can only be defined for Population 1.

$$\text{Early Recurrence in Stroke Unit} = \begin{cases} \text{Yes} & \textbf{if Discharge NIHSS (i12020)} > \\ & \text{(Admission NIHSS (i4020)} + 1) \\ & \textbf{or Reinsult (i15001)} = \text{Yes} \\ & \textbf{or Progressive Stroke (i15014)} = \text{Yes} \\ \text{No} & \textbf{else} \end{cases}$$

$$\text{Early Worsening} = \begin{cases} \text{Yes} & \textbf{if } [\text{Discharge NIHSS (i12020)} > (\text{Admission NIHSS (i4020)} + 1)] \\ & \textbf{and not clinically significant (i150001)} \neq \text{Yes} \\ & \textbf{and as part of an infection (i150002)} \neq \text{Yes} \\ & \textbf{OR} \\ & [\ \text{Discharge type (i11003)} = \text{Deceased} \\ & \textbf{and Cause of death (i11004)} = \text{Cerebral edema}\ ] \\ \text{No} & \textbf{else} \end{cases}$$

$$\text{Recurrence within 90 days} = \begin{cases} \text{Yes} & \textbf{if NIHSS at Discharge (i12020)} \\ & > (\text{NIHSS at Admission (i4020)} + 1) \\ & \textbf{OR} \\ & (\text{Recurrent event TIA (i160001)} = \text{Yes} \\ & \textbf{and Date (i160002) - Date of hospital admission (i3006)} \leq 90) \\ & \textbf{OR} \\ & (\text{Recurrent event IS (i160003)} = \text{Yes} \\ & \textbf{and Date (i160004) - Date of hospital admission (i3006)} \leq 90) \\ & \textbf{OR} \\ & (\text{Cause of death (i21030)} = \text{Reinsult or Cerebral edema}) \\ \text{NA} & \textbf{if no Follow-Up (i20200)} \\ \text{No} & \textbf{else} \end{cases}$$

All three variables indicate whether a patient's NIH Stroke Scale score worsened after the hospital admission, as this implies further health issues. Early Recurrence in Stroke Unit and Early Worsening are based on data concerning the patient's hospital stay, specifically information on their condition and health status that was directly observed. For Recurrence within 90 days, additional data collected from the patient or their caregiver during a follow-up is used.

All of the following variables were created specifically for this thesis in consultation with the two neurologists Dr Peter Sommer and Dr Simon Fandler-Höfler.

The Combined Vascular Endpoint is defined as the combination of the three health outcomes described above. This represents patients having experienced either an early worsening of their condition, or stroke recurrence in the hospital or within 90 days after the initial event. In addition, patients who later suffered a cardiac infarction are included in the Cumulative Endpoint.

$$\text{Combined Vascular Endpoint} = \begin{cases} \text{Yes} & \textbf{if Early Worsening (EW)} = \text{Yes} \\ & \quad \textbf{or Early Recurrence in Stroke Unit (ER)} = \text{Yes} \\ & \quad \textbf{or Recurrence within 90 days (R90)} = \text{Yes} \\ \text{No} & \textbf{if Early Worsening (EW)} = \text{No} \\ & \quad \textbf{and Early Recurrence in Stroke Unit (ER)} = \text{No} \\ & \quad \textbf{and Recurrence within 90 days (R90)} = \text{No} \\ \text{NA} & \textbf{else} \end{cases}$$

$$\text{Cumulative Endpoint} = \begin{cases} \text{Yes} & \textbf{if Early Worsening (EW)} = \text{Yes} \\ & \quad \textbf{or Early Recurrence in Stroke Unit (ER)} = \text{Yes} \\ & \quad \textbf{or Recurrence within 90 days (R90)} = \text{Yes} \\ & \quad \textbf{or Cardiac infarction (i15016)} = \text{Yes} \\ \text{No} & \textbf{if Early Worsening (EW)} = \text{No} \\ & \quad \textbf{and Early Recurrence in Stroke Unit (ER)} = \text{No} \\ & \quad \textbf{and Recurrence within 90 days (R90)} = \text{No} \\ & \quad \textbf{and Cardiac infarction (i15016)} = \text{No} \\ \text{NA} & \textbf{else} \end{cases}$$

Some additional treatment variables are also defined for this thesis. Carotid Revascularization specifies whether a patient had a percutaneous transluminal angioplasty (PTA)[Johb] or a carotid endarterectomy (CEA)[Joha]. Both of these are vascular procedures used to treat a narrowing of blood vessels, though CEA is specifically for the carotid artery. In the Austrian Stroke Unit Registry the variables for PTA and CEA have changed, therefore different variables are used for Population 1 than for Population 2, leading to

two separate definitions.

$$\text{Carotid Revascularization (Population 1)} = \begin{cases} \text{Yes} & \textbf{if } [\text{PTA (i17006)} = \text{Yes} \\ & \quad \textbf{or } \text{PTA (i17006)} = \text{Planned}] \\ & \textbf{OR} \\ & [\text{CEA (i17007)} = \text{Yes} \\ & \quad \textbf{or } \text{CEA (i17007)} = \text{Planned}] \\ \text{No} & \textbf{if } \text{PTA (i17006)} = \text{No} \\ & \quad \textbf{and } \text{CEA (i17007)} = \text{No} \\ \text{NA} & \textbf{else} \end{cases}$$

$$\text{Carotid Revascularization (Population 2)} = \begin{cases} \text{Yes} & \textbf{if } [\text{PTA (i10092)} = \text{Yes} \\ & \quad \textbf{or } \text{PTA (i10092)} = \text{Planned}] \\ & \textbf{OR} \\ & [\text{CEA (i10093)} = \text{Yes} \\ & \quad \textbf{or } \text{CEA (i10093)} = \text{Planned}] \\ \text{No} & \textbf{if } \text{PTA (i10092)} = \text{No} \\ & \quad \textbf{and } \text{CEA (i10093)} = \text{No} \\ \text{NA} & \textbf{else} \end{cases}$$

The length of a patients stay can also be considered as a treatment variable. In the registry there are variables for the date of the stroke unit admission and discharge as well as the date of the hospital admission and discharge. To ensure that the duration of the stay relates specifically to the TIA or minor stroke, admission to the stroke unit is used instead of hospital admission. For most patients, these dates are the same. However, for those who were in the hospital when they experienced their event, using the stroke unit admission date provides a more accurate time frame. For the discharge date, the hospital one will be used, as patients may continue to experience complications from the TIA or minor stroke even after leaving the stroke unit, resulting in a prolonged hospital stay. When calculating the number of days between those two dates, some unusual results are returned. For a few patients the number of days in the hospital are either negative or over 365. These values likely result from errors made during data entry, as negative durations are impossible and hospital stays longer than a year are unlikely. To prevent any issues, these entries were set to 'Not Applicable' (NA) instead.

Lastly, the treatment variable Oral Anticoagulants (OAC) at Follow-Up can be defined, which states whether a patient received any OAC medication. For Population 2 this

covers a set of five medications.

$$\text{OAC (Population 2)} = \begin{cases} \text{Yes} & \textbf{if Vitamin K Antagonist (i26020)} = \text{Yes} \\ & \textbf{or Apixaban (i26021)} = \text{Yes} \\ & \textbf{or Dabigatran (i26022)} = \text{Yes} \\ & \textbf{or Edoxaban (i26023)} = \text{Yes} \\ & \textbf{or Rivaroxaban (i26024)} = \text{Yes} \\ \text{No} & \textbf{if Vitamin K Antagonist (i26020)} = \text{No} \\ & \textbf{and Apixaban (i26021)} = \text{No} \\ & \textbf{and Dabigatran (i26022)} = \text{No} \\ & \textbf{and Edoxaban (i26023)} = \text{No} \\ & \textbf{and Rivaroxaban (i26024)} = \text{No} \\ \text{NA} & \textbf{else} \end{cases}$$

However, at the time of the TIA sub-registry only one of these medications was being entered into the database - the Vitamin K Antagonist. So for Population 1 only this variable can be considered, but it only contains 'Yes' or 'NA' as values. To have at least some 'No' entries an additional check is added, to see if a Follow-Up was performed.

$$\text{OAC (Population 1)} = \begin{cases} \text{Yes} & \textbf{if Vitamin K Antagonist (i26020)} = \text{Yes} \\ \text{No} & \textbf{if Vitamin K Antagonist (i26020)} = \text{No} \\ & \textbf{OR} \\ & [\ \text{Vitamin K Antagonist (i26020) missing} \\ & \textbf{and no Follow-Up (i20200)} = \text{Yes}\ ] \\ \text{NA} & \textbf{else} \end{cases}$$

In both populations the mRS at Follow-Up variable was slightly updated to add missing entries. The modified ranking scale is a measure for the degree of disability after a stroke, with the highest value - 6 - meaning that the person died [Sav+21]. Upon follow-up the patient status can also be entered into the Austrian Stroke Unit Registry, where "Deceased" is one of the options. Therefore, any patients with this status should also have an mRS score of 6, but sometimes there was simply no value entered. This is the case for 113 patients in Population 1 and for 25 in Population 2. This was addressed by adapting the mRS at Follow-Up, setting these values to 6.

4.3.1 Recurrence Variables

In order to analyse the predictive accuracy of the ABCD2 and ABCD3-I risk scores, stroke recurrence needs to be defined. There are a number of variables in the registry which work as clinical indicators and additional ones have already been created in this section (e.g., early worsening, early recurrence). However, some variables in the registry

do not have enough data to work as individual indicators, even though they can be used to define recurrence. Almost 20 combinations of variables were tested to find which provided sensible prediction results - in the end, three were selected for thorough analysis in this thesis.

Also, while combined definitions were already created for three of the four individual indicators (cumulative endpoint and combined vascular endpoint), mRS at Follow-Up stands alone. Therefore, a fourth recurrence variable was defined which combines all of the indicators into one. The four definitions are presented below:

$$\text{Recurrence 1} = \begin{cases} \text{Yes} & \textbf{if MRS at Follow-Up} > 4 \text{ (i25001)} = \text{Yes} \\ & \textbf{or Recurrent Stroke (i15001)} = \text{Yes} \\ \text{No} & \textbf{if MRS at Follow-Up} > 4 \text{ (i25001)} = \text{No} \\ & \textbf{and Recurrent Stroke (i15001)} = \text{No} \\ \text{NA} & \textbf{else} \end{cases}$$

$$\text{Recurrence 2} = \begin{cases} \text{Yes} & \textbf{if MRS at Follow-Up} > 4 \text{ (i25001)} = \text{Yes} \\ & \textbf{or Recurrent event (TIA) (i160001)} = \text{Yes} \\ \text{No} & \textbf{if MRS at Follow-Up} > 4 \text{ (i25001)} = \text{No} \\ & \textbf{and Recurrent event (TIA) (i160001)} = \text{No} \\ \text{NA} & \textbf{else} \end{cases}$$

$$\text{Recurrence 3} = \begin{cases} \text{Yes} & \textbf{if MRS at Follow-Up} > 4 \text{ (i25001)} = \text{Yes} \\ & \textbf{or Recurrent Stroke (i15001)} = \text{Yes} \\ & \textbf{or Recurrent event (TIA) (i160001)} = \text{Yes} \\ \text{No} & \textbf{if MRS at Follow-Up} > 4 \text{ (i25001)} = \text{No} \\ & \textbf{and Recurrent Stroke (i15001)} = \text{No} \\ & \textbf{and Recurrent event (TIA) (i160001)} = \text{No} \\ \text{NA} & \textbf{else} \end{cases}$$

$$\text{Recurrence 4} = \begin{cases} \text{Yes} & \textbf{if Early Worsening} = \text{Yes} \\ & \textbf{or Early Recurrence in Stroke Unit} = \text{Yes} \\ & \textbf{or Recurrence within 90 days} = \text{Yes} \\ & \textbf{or MRS at Follow-Up} > 4 \text{ (i25001)} = \text{Yes} \\ \text{No} & \textbf{if Early Worsening} = \text{No} \\ & \textbf{and Early Recurrence in Stroke Unit} = \text{No} \\ & \textbf{and Recurrence within 90 days} = \text{No} \\ & \textbf{and MRS at Follow-Up} > 4 \text{ (i25001)} = \text{No} \\ \text{NA} & \textbf{else} \end{cases}$$

4.4 Data Structure

For transparency and clarity, here is an overview of all the variables used in the thesis - either directly from the Austrian Stroke Unit Registry or defined using them as explained in the section above. The data is split into the two populations, enabling easy comparison and showing which variables are exclusive to each population.

Table 4.2: Populations. All Variables

Variables	**Population 1** N = 15,467[1]	**Population 2** N = 12,624[1]
Gender		
Female	6,884 (45%)	5,940 (47%)
Male	8,583 (55%)	6,684 (53%)
Age in Years	73 (63, 81)	75 (64, 82)
NIHSS at Admission	1 (0, 2)	0 (0, 1)
NA	4 (<0.1%)	0 (0%)
Systolic blood pressure	155 (140, 175)	NA (NA, NA)
NA	10,136 (66%)	12,624 (100%)
Hypertension		
No	2,999 (19%)	2,624 (21%)
Unknown	139 (0.9%)	209 (1.7%)
Yes	12,278 (80%)	9,710 (77%)
NA	51 (0.3%)	81 (0.6%)
Previous stroke		
No	11,643 (76%)	8,992 (72%)
Unknown	377 (2.4%)	581 (4.6%)
Yes	3,396 (22%)	2,970 (24%)
NA	51 (0.3%)	81 (0.6%)
Cardiac infarction		
No	13,775 (89%)	10,823 (86%)
Unknown	399 (2.6%)	646 (5.2%)
Yes	1,242 (8.1%)	1,074 (8.6%)
NA	51 (0.3%)	81 (0.6%)
Hypercholesterolemia		
No	5,817 (38%)	3,098 (25%)
Unknown	628 (4.1%)	520 (4.1%)
Yes	8,971 (58%)	8,925 (71%)
NA	51 (0.3%)	81 (0.6%)
Atrial fibrillation		
de novo (EKG)	0 (0%)	347 (2.8%)
No	11,618 (75%)	9,965 (79%)
Unknown	673 (4.4%)	0 (0%)

Variables	Population 1 N = 15,467[1]	Population 2 N = 12,624[1]
Yes	3,125 (20%)	0 (0%)
Yes, already known	0 (0%)	2,231 (18%)
NA	51 (0.3%)	81 (0.6%)
Smoking		
No	11,659 (76%)	8,627 (69%)
Unknown	1,141 (7.4%)	1,957 (16%)
Yes	2,616 (17%)	1,959 (16%)
NA	51 (0.3%)	81 (0.6%)
Alcohol abuse		
No	13,240 (86%)	9,756 (78%)
Unknown	1,124 (7.3%)	2,121 (17%)
Yes	1,052 (6.8%)	666 (5.3%)
NA	51 (0.3%)	81 (0.6%)
Aetiology		
cardiogen embolic	3,293 (21%)	0 (NA%)
else	316 (2.0%)	0 (NA%)
Macroangiopathy	1,647 (11%)	0 (NA%)
Microangiopathy	4,794 (31%)	0 (NA%)
Unknown	5,417 (35%)	0 (NA%)
NA	0 (0%)	12,624 (100%)
Early Worsening		
No	15,004 (97%)	0 (NA%)
Yes	463 (3.0%)	0 (NA%)
NA	0 (0%)	12,624 (100%)
Early Recurrence in Stroke Unit		
No	14,884 (96%)	12,563 (100%)
Yes	583 (3.8%)	61 (0.5%)
Recurrence within 90 days		
No	14,656 (96%)	0 (NA%)
Yes	607 (4.0%)	0 (NA%)
NA	204 (1.3%)	12,624 (100%)
MRS at Follow-Up		
0	2,926 (50%)	2,698 (65%)
1	1,227 (21%)	626 (15%)
2	565 (9.6%)	257 (6.2%)
3	455 (7.8%)	236 (5.7%)
4	387 (6.6%)	176 (4.2%)
5	60 (1.0%)	53 (1.3%)
6	248 (4.2%)	104 (2.5%)
NA	9,599 (62%)	8,474 (67%)

Variables	Population 1 N = 15,467[1]	Population 2 N = 12,624[1]
Recurrent Stroke (new event, different territory)		
No	15,392 (100%)	12,605 (100%)
Yes	66 (0.4%)	19 (0.2%)
NA	9 (<0.1%)	0 (0%)
Symptomatic intracranial hemorrhage		
No	15,424 (100%)	12,622 (100%)
Yes	34 (0.2%)	2 (<0.1%)
NA	9 (<0.1%)	0 (0%)
ICH or subdural hematoma		
No	3,409 (100%)	0 (NA%)
Yes	7 (0.2%)	0 (NA%)
NA	12,051 (78%)	12,624 (100%)
Epileptic seizures		
No	15,396 (100%)	12,612 (100%)
Yes	62 (0.4%)	12 (<0.1%)
NA	9 (<0.1%)	0 (0%)
Pneumonia		
No	15,328 (99%)	12,571 (100%)
Yes	132 (0.9%)	53 (0.4%)
NA	7 (<0.1%)	0 (0%)
ABCD2 Score		
0	92 (1.7%)	0 (NA%)
1	367 (6.9%)	0 (NA%)
2	921 (17%)	0 (NA%)
3	1,040 (20%)	0 (NA%)
4	1,499 (28%)	0 (NA%)
5	797 (15%)	0 (NA%)
6	482 (9.0%)	0 (NA%)
7	133 (2.5%)	0 (NA%)
NA	10,136 (66%)	12,624 (100%)
ABCD3-I Score		
0	33 (0.6%)	0 (NA%)
1	155 (2.9%)	0 (NA%)
2	413 (7.7%)	0 (NA%)
3	633 (12%)	0 (NA%)
4	1,087 (20%)	0 (NA%)
5	861 (16%)	0 (NA%)
6	1,083 (20%)	0 (NA%)
7	558 (10%)	0 (NA%)
8	348 (6.5%)	0 (NA%)

Variables	Population 1 N = 15,467[1]	Population 2 N = 12,624[1]
9	120 (2.3%)	0 (NA%)
10	28 (0.5%)	0 (NA%)
11	10 (0.2%)	0 (NA%)
12	2 (<0.1%)	0 (NA%)
13	0 (NA%)	0 (NA%)
NA	10,136 (66%)	12,624 (100%)
Recurrence 1		
No	5,540 (94%)	0 (NA%)
Yes	364 (6.2%)	0 (NA%)
NA	9,563 (62%)	12,624 (100%)
Recurrence 2		
No	2,870 (88%)	0 (NA%)
Yes	385 (12%)	0 (NA%)
NA	12,212 (79%)	12,624 (100%)
Recurrence 3		
No	2,859 (87%)	0 (NA%)
Yes	441 (13%)	0 (NA%)
NA	12,167 (79%)	12,624 (100%)
Recurrence 4		
No	14,537 (94%)	0 (NA%)
Yes	930 (6.0%)	0 (NA%)
NA	0 (0%)	12,624 (100%)
Combined vascular endpoint		
No	14,566 (95%)	0 (NA%)
Yes	699 (4.6%)	0 (NA%)
NA	202 (1.3%)	12,624 (100%)
Cumulative endpoint		
No	14,504 (95%)	0 (NA%)
Yes	753 (4.9%)	0 (NA%)
NA	210 (1.4%)	12,624 (100%)
Length of hospital stay	2 (1, 3)	2 (1, 3)
NA	3 (<0.1%)	50 (0.4%)
Cranial computed tomography (CCT)		
Already on hand	515 (3.3%)	0 (0%)
No	2,040 (13%)	0 (0%)
Yes	12,880 (83%)	10,774 (100%)
NA	32 (0.2%)	1,850 (15%)
Magnetic resonance imaging (MRI) - i8004		
Already on hand	122 (0.8%)	0 (0%)
No	12,355 (80%)	0 (0%)

Variables	Population 1 N = 15,467[1]	Population 2 N = 12,624[1]
Yes	2,958 (19%)	2,387 (100%)
NA	32 (0.2%)	10,237 (81%)
Magnetic resonance imaging (MRI) - i27022		
No	3,746 (24%)	0 (NA%)
Planned	1,797 (12%)	0 (NA%)
Yes	9,885 (64%)	0 (NA%)
NA	39 (0.3%)	12,624 (100%)
Transthoracic echocardiography (TTE)		
No	6,270 (41%)	3,500 (28%)
Planned	3,220 (21%)	3,869 (31%)
Yes	5,938 (38%)	5,170 (41%)
NA	39 (0.3%)	85 (0.7%)
Transesophageal echocardiography (TEE)		
No	12,651 (82%)	11,373 (91%)
Planned	1,202 (7.8%)	683 (5.4%)
Yes	1,575 (10%)	483 (3.9%)
NA	39 (0.3%)	85 (0.7%)
Antiplatelet agents		
Dual	0 (0%)	2,525 (21%)
Mono	0 (0%)	6,930 (57%)
No	2,606 (17%)	2,740 (22%)
Yes	12,808 (83%)	0 (NA%)
NA	53 (0.3%)	429 (3.4%)
Heparin (subcutaneous)		
High (>150E)	459 (3.0%)	190 (1.6%)
Low (<75E)	11,256 (73%)	6,678 (55%)
Medium (75-150E)	1,815 (12%)	717 (5.9%)
None	1,884 (12%)	4,588 (38%)
NA	53 (0.3%)	451 (3.6%)
Acetylsalicylic acid		
No	6,472 (42%)	0 (NA%)
Planned	153 (1.0%)	0 (NA%)
Yes	8,698 (57%)	0 (NA%)
NA	144 (0.9%)	12,624 (100%)
Clopidogrel		
No	11,649 (76%)	0 (NA%)
Planned	152 (1.0%)	0 (NA%)
Yes	3,522 (23%)	0 (NA%)
NA	144 (0.9%)	12,624 (100%)
Percut. transluminal angioplasty (PTA)		

Variables	Population 1 N = 15,467[1]	Population 2 N = 12,624[1]
No	15,189 (99%)	12,088 (99%)
Planned	64 (0.4%)	69 (0.6%)
Yes	68 (0.4%)	32 (0.3%)
NA	146 (0.9%)	435 (3.4%)
Carotis endarterectomy (CEA)		
No	14,831 (97%)	11,826 (97%)
Planned	228 (1.5%)	228 (1.9%)
Yes	259 (1.7%)	135 (1.1%)
NA	149 (1.0%)	435 (3.4%)
Carotid Revascularization		
No	14,712 (96%)	11,743 (96%)
Yes	604 (3.9%)	444 (3.6%)
NA	151 (1.0%)	437 (3.5%)
Rehabilitation		
No information	117 (2.1%)	115 (2.8%)
No Rehabilitation	3,924 (70%)	3,494 (86%)
Rehabilitation	1,583 (28%)	434 (11%)
NA	9,843 (64%)	8,581 (68%)
Inpatient rehabilitation		
Yes	1,265 (100%)	303 (100%)
NA	14,202 (92%)	12,321 (98%)
Outpatient rehabilitation		
Yes	84 (100%)	92 (100%)
NA	15,383 (99%)	12,532 (99%)
Platelet inhibitors - i8070		
Dual	0 (NA%)	267 (2.1%)
Mono	0 (NA%)	4,174 (33%)
No	0 (NA%)	8,172 (65%)
NA	15,467 (100%)	11 (<0.1%)
Platelet inhibitors - i26019		
Dual	0 (NA%)	211 (5.2%)
Mono	0 (NA%)	2,774 (69%)
No	0 (NA%)	1,049 (26%)
NA	15,467 (100%)	8,590 (68%)
Oral Anticoagulants (OAC) at Follow-Up		
No	23 (0.4%)	3,184 (80%)
Yes	5,598 (100%)	791 (20%)
NA	9,846 (64%)	8,649 (69%)
Diabetes mellitus		
No	11,649 (76%)	9,116 (73%)

Variables	Population 1 N = 15,467[1]	Population 2 N = 12,624[1]
Unknown	266 (1.7%)	537 (4.3%)
Yes	3,501 (23%)	2,890 (23%)
NA	51 (0.3%)	81 (0.6%)
Regular lipid-lowering drugs		
No	1,782 (32%)	852 (21%)
Unknown	121 (2.2%)	88 (2.2%)
Yes	3,704 (66%)	3,094 (77%)
NA	9,860 (64%)	8,590 (68%)
Regular blood pressure checks		
No	1,318 (24%)	768 (19%)
Unknown	114 (2.0%)	138 (3.4%)
Yes	4,175 (74%)	3,128 (78%)
NA	9,860 (64%)	8,590 (68%)
Regular antihypertensives		
No	1,718 (31%)	1,011 (25%)
Unknown	79 (1.4%)	55 (1.4%)
Yes	3,810 (68%)	2,968 (74%)
NA	9,860 (64%)	8,590 (68%)
Regular antidiabetica		
No	4,031 (72%)	3,233 (80%)
Unknown	84 (1.5%)	64 (1.6%)
Yes	1,492 (27%)	737 (18%)
NA	9,860 (64%)	8,590 (68%)

[1] n (%)

As described in the Research Design 3.1, three separate analyses are performed to address the three research questions. The variables can be grouped according to the analysis they are used in, though some are used in more than one. Not all variables are available in both populations and due to changes in the registry, some variables have different names. For simplicity, the variables, along with their names, will be listed based on the analysis in which they are used.

All variables in the registry (except for Alter) start with an i. The additional variables created for this thesis do not follow this naming convention, making them easy to identify.

4.4.1 Analyses Variables

The first research question is concerned with patients' health outcomes after a transient ischaemic attack. For Population 2, it is not possible to define Early Worsening and

Recurrence within 90 days, and as a result, these variables are missing. The variable covering intracranial hemorrhage (ICH) or subdural hematoma is also unavailable. 4.3

Some health outcome variables were excluded from further analysis due to low number of entries for 'Yes', since establishing a link between occurrence and gender would not be possible. In Population 1 there are only 7 patient's with an ICH or subdural hematoma. While in Population 2 a symptomatic intracranial hemorrhage is recorded for 2 patients. These when even further divided by gender are not statistically meaningful. The results reliability would be greatly reduced and any interpretations would be unstable.

Variable	Population 1	Population 2
Early Worsening	EW	-
Early Recurrence in Stroke Unit	ER	ER
Recurrence within 90 days	R90	-
mRS at Follow-Up	i25001	i25001
Recurrent Stroke (new event, different territory)	i15001	i15001
Symptomatic intracranial hemorrhage	i15002	-
Epileptic seizures	i15004	i15004
Pneumonia	i15011	i15011

Table 4.3: Health Outcomes

The second research question concerns risk stratification. Since the risk scores could only be calculated for Population 1, the analysis is limited to that dataset. In addition to the ABCD2 and ABCD3-I scores, the target variables define stroke recurrence, which presents a clear overlap with health outcomes. 4.4

Variable	Population 1
Early Worsening	EW
Early Recurrence in Stroke Unit	ER
Recurrence within 90 days	R90
mRS at Follow-Up	i25001
Combined vascular endpoint	CVE
Cumulative endpoint	CE
Recurrent 1	R1
Recurrent 2	R2
Recurrent 3	R3
Recurrent 4	R4

Table 4.4: Recurrence Variables

The third research question deals with diagnostic methods and secondary prophylactic treatments, and is split into two separate analyses. First, in the general odds analysis,

many diagnostic methods and treatments are considered separately. The variables for Heparin, PTA and CEA differ between the populations. 4.5

Variable	Population 1	Population 2
Length of hospital stay	LS	LS
Cranial computed tomography (CCT)	i8003	i8003
Magnetic resonance imaging (MRI)	i8004 / i27022	i8004
Transthoracic echocardiography (TTE)	i27024	i27024
Transesophageal echocardiography (TEE)	i27025	i27025
Antiplatelet agents	i10005	i10005
Acetylsalicylic acid	i17001	-
Clopidogrel	i17002	-
Heparin (subcutaneous)	i10004	i10091
Percut. transluminal angioplasty (PTA)	i17006	i10092
Carotid endarterectomy (CEA)	i17007	i10093
Carotid Revascularization	CR	CR
Rehabilitation	i23005	i23005
Inpatient rehabilitation	i23006	i23006
Outpatient rehabilitation	i23007	i23007
Platelet inhibitors	-	i8070 / i26019
Oral Anticoagulants (OAC) at Follow-Up	OAC	OAC

Table 4.5: Diagnostic methods and secondary prophylactic treatments

Variable	Population 1	Population 2
Atrial fibrillation	i9006	i9006
Oral Anticoagulants (OAC) at Follow-Up	OAC	OAC
Hypercholesterolemia	i9005	i9005
Regular lipid-lowering drugs	i26006	i26032
Hypertension	i9001	i9001
Regular blood pressure checks	i26008	i26033
Regular antihypertensives	i26010	i26034
Diabetes mellitus	i9002	i9002
Regular antidiabetica	i26012	i26035

Table 4.6: Illnesses and Treatments

In Population 1, there are two variables indicating whether magnetic resonance imaging (MRI) was performed. While Population 2 has two variables determining whether a patient received platelet inhibitors. Since one variable was recorded during the hospital stay and the other during the follow-up, they cannot be combined. They represent different clinical situations due to the time gap. Population 1 includes two additional

variables - acetylsalicylic acid and clopidogrel - which are specific antiplatelet agents. However, it does not contain a variable for platelet inhibitors.

The second part of this analysis considers treatment rates for certain illnesses. For regular antidiabetica/antihypertensives/blood pressure checks/lipid-lowering drugs, the variables are different between the populations. 4.6

4.4.2 Encoding

For most of the variables, no further preprocessing or encoding is needed. All of the health outcomes and stroke recurrences are binary with either 'Yes' or 'No', except for the mRS score. This variable will either be used as multinomial, with each discrete value on the scale its own category, or by splitting it into two ranges of values. For the analysis of health outcomes, this split will be 0-1 vs. 2-6, representing no or minimal lingering issues versus at least some disability, major disability or death. On the other hand, for stroke recurrence, a different scale separation is necessary. It is only considered a recurrence if severe disability is present or the patient has died. Therefore, the split 0–4 vs. 5–6 will be used.

Missing or NA values will not be considered for any of the health outcomes or stroke recurrence variables, as these mostly represent information that was not collected - for example, if no follow-up was performed for the patient.

In contrast, for the diagnostic methods and secondary prophylactic treatments, further encoding is necessary. For inpatient and outpatient rehabilitation as well as CCT and MRI in Population 2 only 'Yes' or 'NA' values are recorded, meaning that, by default 'NA' must be interpreted as 'No'. This approach is reasonable for most variables in this category, as the absence of a clear entry likely indicates that no treatment was given. This is especially unproblematic since for the majority, 'NA' accounts for less than 1% of all entries and, at most, about 3.5%. The only exceptions are variables recorded through the Follow-Up, as then 'NA' stands for a missing follow-up and should not be changed. For these, about two thirds of entries are missing, which makes sense since a follow-up is only done for about a third of patients. Rehabilitation, platelet inhibitors (i26019), OAC, and all the 'regular' treatment variables are affected by this and are therefore the only ones to retain their missing values.

Some variables also have additional value categories for which encoding needs to be considered. For the 'regular' treatment variables entries of 'Unknown' can be seen as 'NA', they provide no additional information and make up only a very small percentage each. The few 'Unknown' entries for Diabetes mellitus can be assumed to indicate absence of the condition, as a patient not having diabetes is more likely and for the calculation of the treatment rate 'Unknown' is a useless category.

For rehabilitation the entries of 'No information' can be treated as missing, given that with 2.1% and 2.8% this does not represent a meaningful category. Also its other categories 'Rehabilitation' and 'No Rehabilitation' can be changed to 'Yes' and 'No', which ensures consistency with other variables while preserving the original context.

There are a few variables with categories of either 'Already on hand' or 'Planned'. Most of these categories make up only around 1-2% of all entries, meaning they can simply be viewed as additional 'Yes' entries. This is a sensible assumption as either the treatment has already been done or will be done soon. As a result, percutaneous transluminal angioplasty and carotid endarterectomy for both Populations, as well as magnetic resonance imaging (i8004), Acetylsalicylic acid and Clopidogrel for Population 1, can be reduced to binary variables.

However, for transthoracic echocardiography (TTE) and transesophageal echocardiography (TEE) there are many entries of 'Planned', with around 21% and 7.8% in the first population and 31% and 5.4% in the second one. The argument can be made that this category could be significant on its own. Therefore, both a version where this category is retained and one where it is encoded as 'Yes' should be considered. The same is true for CCT with 3.3% 'Already on hand' and MRI (i27022) with 12% of 'Planned' for Population 1. This results in four variables with two encoding versions in Population 1, and two in Population 2.

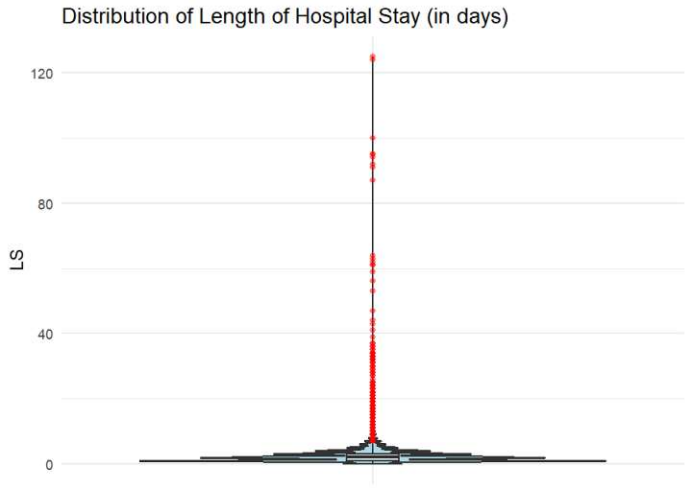


Figure 4.4: Distribution of Length of Hospital Stay

Lastly, for length of hospital stay, it makes more sense to group the number of days into meaningful categories rather than treat it as a continuous integer variable. Table 4.2 shows a median hospital stay of 2 days, with the first and third quartiles at 1 and 3 days, respectively - consistent across both populations. This corresponds to the figure below 4.4, where the distribution is shown with the outliers highlighted. Most values range from 0 to 10 days, with the majority of outliers extending up to 40 days. Based on the distribution, the following classification levels were chosen: 0 days, 1-2 days, 3-7 days, 8-14 days and 14+ days. Stays within one week are divided into three groups, those

between one and two weeks form a separate group, and the final category includes any stays longer than two weeks.

These encoding steps have modified the diagnostic methods and treatment variables. Table 4.7 shows the variable versions that will be used in the analyses.

Table 4.7: Populations. Encoded diagnostic methods and treatments

Variables	**Population 1** N = 15,467[1]	**Population 2** N = 12,624[1]
Length of hospital stay		
0	577 (3.7%)	654 (5.2%)
1-2	8,606 (56%)	8,356 (66%)
3-7	5,619 (36%)	3,338 (27%)
8-14	518 (3.3%)	183 (1.5%)
14+	144 (0.9%)	43 (0.3%)
NA	3 (<0.1%)	50 (0.4%)
Cranial computed tomography (CCT)		
No	2,072 (13%)	1,850 (15%)
Yes	13,395 (87%)	10,774 (85%)
Cranial computed tomography (CCT)		
Already on hand	515 (3.3%)	0 (NA%)
No	2,072 (13%)	0 (NA%)
Yes	12,880 (83%)	0 (NA%)
NA	0 (0%)	12,624 (100%)
Magnetic resonance imaging (MRI) - i8004		
No	12,387 (80%)	10,237 (81%)
Yes	3,080 (20%)	2,387 (19%)
Magnetic resonance imaging (MRI) - i27022		
No	3,785 (24%)	0 (NA%)
Yes	11,682 (76%)	0 (NA%)
NA	0 (0%)	12,624 (100%)
Magnetic resonance imaging (MRI) - i27022		
No	3,785 (24%)	0 (NA%)
Planned	1,797 (12%)	0 (NA%)
Yes	9,885 (64%)	0 (NA%)
NA	0 (0%)	12,624 (100%)
Transthoracic echocardiography (TTE)		
No	6,309 (41%)	3,585 (28%)
Yes	9,158 (59%)	9,039 (72%)
Transthoracic echocardiography (TTE)		
No	6,309 (41%)	3,585 (28%)
Planned	3,220 (21%)	3,869 (31%)

Variables	Population 1 N = 15,467[1]	Population 2 N = 12,624[1]
Yes	5,938 (38%)	5,170 (41%)
Transesophageal echocardiography (TEE)		
No	12,690 (82%)	11,458 (91%)
Yes	2,777 (18%)	1,166 (9.2%)
Transesophageal echocardiography (TEE)		
No	12,690 (82%)	11,458 (91%)
Planned	1,202 (7.8%)	683 (5.4%)
Yes	1,575 (10%)	483 (3.8%)
Antiplatelet agents		
Dual	0 (0%)	2,525 (20%)
Mono	0 (0%)	6,930 (55%)
No	2,659 (17%)	3,168 (25%)
Yes	12,808 (83%)	1 (<0.1%)
Heparin (subcutaneous)		
High (>150E)	459 (3.0%)	190 (1.5%)
Low (<75E)	11,256 (73%)	6,678 (53%)
Medium (75-150E)	1,815 (12%)	717 (5.7%)
None	1,937 (13%)	5,039 (40%)
Acetylsalicylic acid		
No	6,616 (43%)	0 (NA%)
Yes	8,851 (57%)	0 (NA%)
NA	0 (0%)	12,624 (100%)
Clopidogrel		
No	11,793 (76%)	0 (NA%)
Yes	3,674 (24%)	0 (NA%)
NA	0 (0%)	12,624 (100%)
Percut. transluminal angioplasty (PTA)		
No	15,335 (99%)	12,523 (99%)
Yes	132 (0.9%)	101 (0.8%)
Carotis endarterectomy (CEA)		
No	14,980 (97%)	12,261 (97%)
Yes	487 (3.1%)	363 (2.9%)
Carotid Revascularization		
No	14,863 (96%)	12,180 (96%)
Yes	604 (3.9%)	444 (3.5%)
Rehabilitation		
No	3,924 (71%)	3,494 (89%)
Yes	1,583 (29%)	434 (11%)
NA	9,960 (64%)	8,696 (69%)
Inpatient rehabilitation		

Variables	Population 1 N = 15,467[1]	Population 2 N = 12,624[1]
No	14,202 (92%)	12,321 (98%)
Yes	1,265 (8.2%)	303 (2.4%)
Outpatient rehabilitation		
No	15,383 (99%)	12,532 (99%)
Yes	84 (0.5%)	92 (0.7%)
Platelet inhibitors - i8070		
Dual	0 (NA%)	267 (2.1%)
Mono	0 (NA%)	4,174 (33%)
No	0 (NA%)	8,183 (65%)
NA	15,467 (100%)	0 (0%)
Platelet inhibitors - i26019		
Dual	0 (NA%)	211 (5.2%)
Mono	0 (NA%)	2,774 (69%)
No	0 (NA%)	1,049 (26%)
NA	15,467 (100%)	8,590 (68%)
Oral Anticoagulants (OAC) at Follow-Up		
No	23 (0.4%)	3,184 (80%)
Yes	5,598 (100%)	791 (20%)
NA	9,846 (64%)	8,649 (69%)
Diabetes mellitus		
No	11,966 (77%)	9,734 (77%)
Yes	3,501 (23%)	2,890 (23%)
Regular lipid-lowering drugs		
No	1,782 (32%)	852 (22%)
Yes	3,704 (68%)	3,094 (78%)
NA	9,981 (65%)	8,678 (69%)
Regular blood pressure checks		
No	1,318 (24%)	768 (20%)
Yes	4,175 (76%)	3,128 (80%)
NA	9,974 (64%)	8,728 (69%)
Regular antihypertensives		
No	1,718 (31%)	1,011 (25%)
Yes	3,810 (69%)	2,968 (75%)
NA	9,939 (64%)	8,645 (68%)
Regular antidiabetica		
No	4,031 (73%)	3,233 (81%)
Yes	1,492 (27%)	737 (19%)
NA	9,944 (64%)	8,654 (69%)

[1] n (%)

4.4.3 Patient Characteristics & Covariates

Variable	Population 1	Population 2
Age in Years	Alter	Alter
Admission NIHSS	i4020	i4020
Systolic blood pressure	i100005	-
Hypertension	i9001	i9001
Previous stroke	i9003	i9003
Cardiac infarction	i9004	i9004
Hypercholesterolemia	i9005	i9005
Atrial fibrillation	i9006	i9006
Smoking	i9009	i9009
Alcohol abuse	i9010	i9012
Aetiology	aetiology	-

Table 4.8: Covariates: Patient Characteristics

For most of the analyses, the influence of covariates is also considered - specifically, patient characteristics. This includes factors which describe a patient or their health, like pre-existing conditions, which may affect their health outcomes or treatments. Therefore, when examining gender differences, multiple computations are run to assess the effects of other patient characteristics as well. Thus, three approaches are taken: one without covariates, one using only age, and one using the full list of covariates shown in Table 4.8, depending on availability for each population.

In Population 1, this allows for two additional characteristics to be included - systolic blood pressure and aetiology. For alcohol abuse, the variable was changed in the registry; therefore, the newer version is used in Population 2.

Importantly, the variables starting with 'i90' are also referred to as risk factors and describe whether a patient has certain pre-existing conditions or habits that may elevate their risk of a stroke. However, besides 'Yes' and 'No', there are frequent entries of 'Unknown' for these variables. These can be handled in one of two ways. They can be treated as providing no additional context or information, implying the assumption that if a patient's status is unclear, it is more likely that they do not have the risk factor. In this case, any entries marked as 'Unknown' or missing (NA) are set to 'No'. Alternatively, 'Unknown' can be treated as a distinct and valid category, expanded to also include the missing entries. Both of these are reasonable approaches and depend on how much informational value the 'Unknown' category is assumed to provide. For the analyses, both versions of the risk factors will be tested and their influence compared. To enable this, an additional variable is created for each risk factor in which the 'Unknown' category is retained. These variables have the suffix '_unk' appended to their names. The original variables are, in turn, reduced to a binary format with only 'Yes' and 'No' values. However, as shown in Table 4.2, for most risk factors, 'Unknown' and NA entries account for at most five percent of the data. The exceptions are smoking and alcohol

abuse, where in Population 1 more than 7% of entries are missing, and in Population 2 the number rises to over 16% for both variables. These are also risk factors which tend to be more prevalent among men, and are therefore likely to be relevant for the gender-focused analyses in this thesis. A large proportion of entries marked as 'Unknown' or assumed to be 'No' may reduce statistical power. To address this, an additional version of these two risk factors is created, where missing and 'Unknown' values are imputed. Rather than making broad assumptions about the data, reasonable estimates are used for imputation. These are based on other risk factors, as well as the patient's gender, age, and NIHSS score at admission. The resulting variables have the suffix '_imp' appended to their names.

The values for atrial fibrillation in Population 2 present a special case. Instead of simply recording 'Yes', this population distinguishes between 'de novo (EKG)', indicating newly diagnosed atrial fibrillation, and 'Yes, already known'. Since this distinction is not relevant to the analysis and does not exist in Population 1, both categories are recoded as 'Yes' for consistency.

Table 4.9 shows these modified and additional risk factor variables. As a result, there are now three separate versions of each risk factor, which leads to five covariate combinations to be considered in the analyses:

- **None**: No covariates
- **Age**: Only age as covariate
- **All**: All listed covariates - risk factors are binary (Yes/No)
- **All (Imputed)**: All listed covariates - risk factors are binary (Yes/No); Smoking and Alcohol are imputed
- **All (Unknown)**: All listed covariates - risk factors include an 'Unknown' category

Table 4.9: Populations. Risk-Factors

Variables	**Population 1** N = 15,467[1]	**Population 2** N = 12,624[1]
Hypertension		
No	3,189 (21%)	2,914 (23%)
Yes	12,278 (79%)	9,710 (77%)
Hypertension (with Unknown)		
No	2,999 (19%)	2,624 (21%)
Unknown	190 (1.2%)	290 (2.3%)
Yes	12,278 (79%)	9,710 (77%)
Previous stroke		

Variables	Population 1 N = 15,467[1]	Population 2 N = 12,624[1]
No	12,071 (78%)	9,654 (76%)
Yes	3,396 (22%)	2,970 (24%)
Previous stroke (with Unknown)		
No	11,643 (75%)	8,992 (71%)
Unknown	428 (2.8%)	662 (5.2%)
Yes	3,396 (22%)	2,970 (24%)
Cardiac infarction		
No	14,225 (92%)	11,550 (91%)
Yes	1,242 (8.0%)	1,074 (8.5%)
Cardiac infarction (with Unknown)		
No	13,775 (89%)	10,823 (86%)
Unknown	450 (2.9%)	727 (5.8%)
Yes	1,242 (8.0%)	1,074 (8.5%)
Hypercholesterolemia		
No	6,496 (42%)	3,699 (29%)
Yes	8,971 (58%)	8,925 (71%)
Hypercholesterolemia (with Unknown)		
No	5,817 (38%)	3,098 (25%)
Unknown	679 (4.4%)	601 (4.8%)
Yes	8,971 (58%)	8,925 (71%)
Atrial fibrillation		
No	12,342 (80%)	10,046 (80%)
Yes	3,125 (20%)	2,578 (20%)
Atrial fibrillation (with Unknown)		
No	11,618 (75%)	0 (NA%)
Unknown	724 (4.7%)	0 (NA%)
Yes	3,125 (20%)	0 (NA%)
NA	0 (0%)	12,624 (100%)
Smoking		
No	12,851 (83%)	10,665 (84%)
Yes	2,616 (17%)	1,959 (16%)
Smoking (with Unknown)		
No	11,659 (75%)	8,627 (68%)
Unknown	1,192 (7.7%)	2,038 (16%)
Yes	2,616 (17%)	1,959 (16%)
Smoking (Imputed)		
No	12,663 (82%)	10,227 (81%)
Yes	2,804 (18%)	2,397 (19%)
Alcohol abuse		
No	14,415 (93%)	11,958 (95%)

Variables	Population 1 N = 15,467[1]	Population 2 N = 12,624[1]
Yes	1,052 (6.8%)	666 (5.3%)
Alcohol abuse (with Unknown)		
No	13,240 (86%)	9,756 (77%)
Unknown	1,175 (7.6%)	2,202 (17%)
Yes	1,052 (6.8%)	666 (5.3%)
Alcohol abuse (Imputed)		
No	14,318 (93%)	11,722 (93%)
Yes	1,149 (7.4%)	902 (7.1%)

[1] n (%)

CHAPTER 5

Results

This chapter is split into three sections, each covering the results of the analyses addressing one of the research questions. The sections are structured to present the results systematically, beginning with descriptive statistics, followed by detailed analyses of models or performance metrics, and finishing with summaries of the key results and insights.

The results are stratified by gender and, when possible, adjusted for different covariates, such as age and clinical risk factors. This is done to account for potential confounding factors and to interpret the observed differences in their context.

Significant model results and selected visualisations are shown directly in the text to support the interpretation of the findings. While full model outputs, additional tables, and complete sets of figures are available in the Appendix. These additional materials are included to provide full transparency and allow thorough examination of the data and results.

5.1 Gender Differences in Health Outcomes

This section presents the results addressing the first research question: *"Are there gender-specific differences in health outcomes among patients after experiencing a transient ischaemic attack (TIA)?"*

Clinical health outcomes following a TIA are evaluated in this analysis for gender-specific differences in their occurrence. The focus is stroke recurrence, functional recovery as recorded by the modified Rankin Scale (mRS), and complications such as seizures and pneumonia.

The analysis is performed on both the TIA sub-registry (Population 1) and the more recent data from the main stroke registry (Population 2). Due to differences in variable

availability, a more detailed analysis of Population 1 is possible, while the results derived from Population 2 can validate the observations.

The results in this section are derived from logistic regression models. Odds ratios (ORs) with 95% confidence intervals (CIs) are used to compare the likelihood of experiencing specific outcomes between the genders. In all the models the reference group is male, therefore, ORs above 1 indicate higher odds for women, and ORs below 1 indicate lower odds. If the confidence interval crosses 1, the results are considered statistically non-significant. The tables in this section only include significant results, while the tables with all results can be found in the Appendix 7.

Note on mRS Outcome Modeling: Throughout the analysis, the mRS at Follow-Up variable is modelled using multiple logistic regression for each level (mRS = 1 - 6), using mRS = 0 as the reference category. This means the odds are calculated to compare each disability level to a full recovery of the patient. In contrast, the grouped outcome mRS > 1 is modelled as a binary variable, comparing those with some to severe disability (mRS 2–6) to those with better outcomes (mRS 0–1).

5.1.1 Population 1

Descriptive Statistics

An overview of the patient characteristics and risk factors for men and women in Population 1 is provided in the Appendix (Table 1). Since these variables can affect stroke risk and recovery, they are used as covariates for the adjusted models to ensure the independence of the effect gender has on health outcomes.

Some differences can be observed between the genders:

- Women were slightly younger than men (median 73 vs. 75), though the distribution was similar.
- Smoking and alcohol abuse were far more common among men, with 22% vs. 12% and 11% vs. 2.4% after imputation.
- Men were also more likely to have hypercholesterolemia (60% vs. 55%) and a history of cardiac infarction (9.9% vs. 5.7%).
- A previous stroke, hypertension, and macroangiopathy were slightly more common in men.
- Atrial fibrillation, microangiopathy, and cardioembolic aetiology were slightly more common in women.

Most of the variables showed either no notable gender difference or the observed differences were unlikely to have a large impact. However, some of the variables specifically related

to vascular risk profiles and substance-related risk factors did suggest modest gender differences.

The frequencies of the health outcome variables analysed in this section are shown in Table 5.1. It reports the unadjusted outcome distributions for men and women and allows for some observations before evaluating the models:

- Stroke recurrence (both early and within 90 days) as well as early worsening were nearly identical between genders.
- Based on the modified Rankin Scale (mRS), men had higher rates of good functional recovery (mRS 0–1: 75% in men vs. 65% in women).
- Specifically, while men were more likely to have an mRS score of 0, women had higher proportions of 3, 4 and 5 scores.
- Pneumonia was more common in men (1.1% vs. 0.6%).
- Epileptic seizures and intracranial hemorrhage occurred infrequently (less than 0.5% each) in both genders.

Table 5.1: Population 1. Health Outcomes

Gender	**Female** N = 6,884[1]	**Male** N = 8,583[1]
Early Worsening		
No	6,681 (97%)	8,323 (97%)
Yes	203 (2.9%)	260 (3.0%)
Early Recurrence in Stroke Unit		
No	6,632 (96%)	8,252 (96%)
Yes	252 (3.7%)	331 (3.9%)
Recurrence within 90 days		
No	6,514 (96%)	8,142 (96%)
Yes	269 (4.0%)	338 (4.0%)
NA	101 (1.5%)	103 (1.2%)
MRS at Follow-Up (>1)		
0-1	1,676 (65%)	2,477 (75%)
2-6	890 (35%)	825 (25%)
NA	4,318 (63%)	5,281 (62%)
MRS at Follow-Up		
0	1,175 (46%)	1,751 (53%)
1	501 (20%)	726 (22%)
2	253 (9.9%)	312 (9.4%)
3	257 (10%)	198 (6.0%)

Gender	**Female** N = 6,884[1]	**Male** N = 8,583[1]
4	225 (8.8%)	162 (4.9%)
5	41 (1.6%)	19 (0.6%)
6	114 (4.4%)	134 (4.1%)
NA	4,318 (63%)	5,281 (62%)
Recurrent Stroke (new event, different territory)		
No	6,857 (100%)	8,535 (100%)
Yes	25 (0.4%)	41 (0.5%)
NA	2 (<0.1%)	7 (<0.1%)
Symptomatic intracranial hemorrhage		
No	6,864 (100%)	8,560 (100%)
Yes	18 (0.3%)	16 (0.2%)
NA	2 (<0.1%)	7 (<0.1%)
Epileptic seizures		
No	6,848 (100%)	8,548 (100%)
Yes	34 (0.5%)	28 (0.3%)
NA	2 (<0.1%)	7 (<0.1%)
Pneumonia		
No	6,842 (99%)	8,486 (99%)
Yes	40 (0.6%)	92 (1.1%)
NA	2 (<0.1%)	5 (<0.1%)

[1] n (%)

Unadjusted Outcome Rates by Gender

In Table 5.2 the results from the unadjusted logistic regression models are presented.

Variable	OR [95% CI]	P-Value	Interpretation
Pneumonia	0.54 [0.37-0.78]	0.001	Women had lower odds.
mRS at Follow-Up (>1)	1.59 [1.42-1.79]	< 0.001	Women had higher odds.
mRS at Follow-Up = 2	1.21 [1.01-1.45]	0.041	Women had higher odds.
mRS at Follow-Up = 3	1.93 [1.58-2.36]	< 0.001	Women had higher odds.
mRS at Follow-Up = 4	2.07 [1.67-2.57]	< 0.001	Women had higher odds.
mRS at Follow-Up = 5	3.21 [1.86-5.57]	< 0.001	Women had higher odds.

Table 5.2: Population 1 - Unadjusted Outcome Odds Ratios by Gender (Significant Results Only)

Women had significantly lower odds of pneumonia (OR = 0.54), suggesting they were less likely to experience this complication after a TIA. This is the only complication for

which a significant gender difference was found, and it matches the observed difference in the frequency table.

The results show that functional recovery was consistently worse for women in this unadjusted model. The odds of some to severe disability (mRS > 1) were 59% higher for women compared to men. When observing the mRS separately, the only scores for which women didn't have higher odds were the lowest and the highest one, which equals the death of a patient. Slight disabilities (mRS = 2) were only slightly more common for women, with and OR of 1.21. While based on the confidence intervals, women were at least around 1.6 times more likely to have moderate (mRS = 3) or moderately severe disabilities (mRS = 4) after a TIA. For severe disabilities (mRS = 5) the odds were more than three times higher for women, though the CI ranges from 1.86 to 5.57 more likely. As the mRS score 5 does not occur as often as the rest, this wide CI is to be expected.

These models do not control for covariates, and this means some of the observed differences may actually be dependent on underlying clinical or demographic factors. The direction and magnitude of the effects will be further assessed in the adjusted models that follow.

Adjusted Analysis: Controlling for Age

Table 5.3 shows the odds ratios for women compared to men after adjusting only for age. This helps identify if the previously observed gender differences were influenced by the age distributions of the genders.

Variable	OR [95% CI]	P-Value	Interpretation
Pneumonia	0.41 [0.28-0.6]	< 0.001	Women had lower odds.
mRS at Follow-Up (>1)	1.21 [1.07-1.37]	0.002	Women had higher odds.
mRS at Follow-Up = 3	1.42 [1.15-1.75]	0.001	Women had higher odds.
mRS at Follow-Up = 4	1.36 [1.08-1.71]	0.008	Women had higher odds.
mRS at Follow-Up = 5	2.2 [1.26-3.87]	0.006	Women had higher odds.

Table 5.3: Population 1 - Outcome odds ratios by gender controlled for Age (Significant Results Only)

Women's lower odds for pneumonia remained significant even after adjusting for age, though the effect was slightly reduced from an OR of 0.54 to one of 0.41. This indicates that the difference cannot just be explained by age.

The functional recovery differences between the genders persisted, though with lower effect sizes than in the unadjusted model. The grouped outcome mRS > 1 remained significant, with 21% higher odds for women than men. The gender difference for mRS = 2 was no longer significant in this model, but the higher mRS scores still showed clearly higher odds for women. Moderate disability (mRS = 3) was 1.4 times, moderately severe disability (mRS = 4) 1.36 times, and severe disability over two times more likely in women. For mRS = 6 (death), the odds remained unaffected by gender.

Overall, these results suggest that age alone did not account for the observed gender differences in functional outcomes after a TIA, especially in the moderate to severe range.

Adjusted Analysis: Controlling for Patient Characteristics and Risk Factors

In the following models, additional covariates were taken into account to isolate the effect of gender as best as possible. These included relevant clinical and stroke risk factor variables such as smoking, atrial fibrillation, alcohol abuse and previous stroke.

Variable	OR [95% CI]	P-Value	Interpretation
Pneumonia	0.45 [0.2-0.96]	0.046	Women had lower odds.
mRS at Follow-Up (>1)	1.31 [1.06-1.61]	0.012	Women had higher odds.
mRS at Follow-Up = 3	1.7 [1.21-2.38]	0.002	Women had higher odds.
mRS at Follow-Up = 5	3.11 [1.23-7.86]	0.017	Women had higher odds.

Table 5.4: Population 1 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors (Significant Results Only)

As shown in Table 5.4, the lower odds of pneumonia for women remained (OR = 0.45), although the 95% CI widened to [0.20–0.96]. The difference was still statistically significant, even after adjusting for the risk profiles, but the wider interval implied more uncertainty around the true effect size.

The grouped mRS outcome (mRS > 1) also remained statistically significant, with women showing higher odds of worse functional recovery. Although the lower bound of the CI was fairly close to 1 (CI: [1.06–1.61]), the effect remained consistent with direction in previous models.

This pattern continued for the individual mRS levels, with lower ORs and wider CIs. Moderate (mRS = 3) and severe disabilities (mRS = 5) still had clearly higher odds of affecting women than men. However, mRS = 4 no longer reached statistical significance with the covariates included. Its CI crossed 1 and it had a p-value of 0.107 (clearly above 0.05).

Variable	OR [95% CI]	P-Value	Interpretation
Pneumonia	0.45 [0.2-0.97]	0.048	Women had lower odds.
mRS at Follow-Up (>1)	1.32 [1.07-1.62]	0.010	Women had higher odds.
mRS at Follow-Up = 3	1.71 [1.21-2.4]	0.002	Women had higher odds.
mRS at Follow-Up = 5	3.21 [1.24-8.3]	0.016	Women had higher odds.

Table 5.5: Population 1 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors with Unknown included (Significant Results Only)

Including the unknown values in the covariates (Table 5.5) slightly shifted the estimates but did not change the overall findings. The odds ratios of pneumonia, mRS > 1, and

mRS = 3 remained the same, and so they kept their significance. For mRS = 5, the OR showed a slightly lower effect with 3.06 instead of 3.21, but women were still far more likely to suffer severe disabilities after a TIA than men.

Variable	OR [95% CI]	P-Value	Interpretation
Pneumonia	0.45 [0.2-0.96]	0.045	Women had lower odds.
mRS at Follow-Up (>1)	1.32 [1.07-1.63]	0.009	Women had higher odds.
mRS at Follow-Up = 3	1.71 [1.22-2.39]	0.002	Women had higher odds.
mRS at Follow-Up = 5	3.06 [1.21-7.76]	0.018	Women had higher odds.

Table 5.6: Population 1 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors with Smoking and Alcohol Abuse imputed (Significant Results Only)

Finally, in the imputed model (Table 5.6), where missing values for smoking and alcohol abuse were handled through imputation, the pattern of results remained stable.

Women were still less likely to develop pneumonia as a complication (OR = 0.45), but they were also less likely to have a full or almost full functional recovery (OR = 1.32). mRS scores of 3 and 5 were significantly more common for women, both being at least 1.2 times more likely based on the CIs.

Summary

Across all models two gender-specific patterns consistently emerged in Population 1:

- **Pneumonia was significantly less likely for women** than for men. Regardless of the model, this difference remained stable and statistically significant, although the effect size slightly decreased and the confidence interval widened.
- **Women were more likely to experience worse functional outcomes** after a TIA, especially at moderate and severe disability levels (mRS = 3 and 5). The grouped outcome (mRS > 1) was significant across all models and showed a clear trend toward poorer recovery for women.

Some individual mRS levels (mRS = 2 and 4) did not stay significant in the (fully) adjusted models. This indicates that the observed differences for these scores could, in part, be explained by age or clinical characteristics. However, as the higher odds for moderate and severe disability remained after the adjustments, gender clearly had an effect on the functional outcome of patients and their mRS scores after a TIA.

In contrast, outcomes such as stroke recurrence, early worsening, seizures, or death (mRS = 6) consistently showed no significant gender differences in any model. This suggests that these clinical outcomes were not affected by gender in this population.

5.1.2 Population 2

Descriptive Statistics

Table 2 in the Appendix shows the patient characteristics and risk factors for both men and women in Population 2. There are fewer variables available than in Population 1, but all the important risk factors are covered. These are again used as covariates in the adjusted models to isolate the effect gender has on health outcomes.

The genders differed in a few ways:

- Women were notably older than men (median 78 vs. 73), with a clearly different age distribution.
- Smoking and alcohol abuse were far more common among men, with 23% vs. 14% and 10% vs. 3.7% after imputation.
- Men were more likely to have suffered a cardiac infarction (11% vs. 5.8%).
- A previous stroke and hypercholesterolemia were slightly more common in men.
- Hypertension was slightly more common in women.

Atrial fibrillation showed no gender difference at all, while some of the mentioned differences were negligibly small and therefore unlikely to have any effect on the models. It is also interesting that the age difference in this population was not only more pronounced but also the other way around compared to Population 1, where men were older.

In Table 5.7, the health outcome distributions between the genders are shown. While fewer variables are available than in Population 1, for those covered in both, similar patterns can be observed:

- Early as well as different-territory stroke recurrence happened with very similar rates between genders.
- Men had higher rates of functional recovery (mRS 0–1: 84% in men vs. 75% in women).
- While mRS scores of 5 and 6 were equally common between both genders, the rates for scores between 2 and 4 were clearly higher for women. Women also had a lower rate of full recovery (mRS = 0).
- Epileptic seizures were very rare, though slightly more common in women.
- The occurrence rates of pneumonia were almost the same, but only a total of 19 women vs. 34 men suffered this complication.

Table 5.7: Population 2. Health Outcomes

Gender	**Female** N = 5,940[1]	**Male** N = 6,684[1]
Early Recurrence in Stroke Unit		
No	5,909 (99%)	6,654 (100%)
Yes	31 (0.5%)	30 (0.4%)
MRS at Follow-Up (>1)		
0-1	1,410 (75%)	1,914 (84%)
2-6	462 (25%)	364 (16%)
NA	4,068 (68%)	4,406 (66%)
MRS at Follow-Up		
0	1,135 (61%)	1,563 (69%)
1	275 (15%)	351 (15%)
2	138 (7.4%)	119 (5.2%)
3	133 (7.1%)	103 (4.5%)
4	114 (6.1%)	62 (2.7%)
5	31 (1.7%)	22 (1.0%)
6	46 (2.5%)	58 (2.5%)
NA	4,068 (68%)	4,406 (66%)
Recurrent Stroke (new event, different territory)		
No	5,929 (100%)	6,676 (100%)
Yes	11 (0.2%)	8 (0.1%)
Epileptic seizures		
No	5,932 (100%)	6,680 (100%)
Yes	8 (0.1%)	4 (<0.1%)
Pneumonia		
No	5,921 (100%)	6,650 (99%)
Yes	19 (0.3%)	34 (0.5%)

[1] n (%)

Unadjusted Outcome Rates by Gender

Table 5.8 presents the unadjusted logistic regression results for Population 2.

It can be seen that women were significantly more likely to experience worse functional outcomes after a TIA. The odds of suffering from any level of disability (mRS > 1) were 72% higher for women than for men. The mRS scores 2 through 5 were all significantly more likely to occur in women, meaning that of the poorer functional outcomes, only death (mRS = 6) was equally likely between the genders. The clearest difference was seen with mRS = 4, which was 2.5 times more likely for women than for men.

Variable	OR [95% CI]	P-Value	Interpretation
mRS at Follow-Up (>1)	1.72 [1.48-2.01]	< 0.001	Women had higher odds.
mRS at Follow-Up = 2	1.6 [1.24-2.06]	< 0.001	Women had higher odds.
mRS at Follow-Up = 3	1.78 [1.36-2.33]	< 0.001	Women had higher odds.
mRS at Follow-Up = 4	2.53 [1.84-3.48]	< 0.001	Women had higher odds.
mRS at Follow-Up = 5	1.94 [1.11-3.36]	0.019	Women had higher odds.

Table 5.8: Population 2 - Unadjusted outcome odds ratios by gender (Significant Results Only)

No statistically significant differences were observed in stroke recurrence or any of the complications. However, while the result was not statistically significant, pneumonia followed the same trend toward lower odds for women (OR = 0.63, CI: [0.35–1.09]) as found in Population 1.

Adjusted Analysis: Controlling for Age

In Table 5.9, age was added as a covariate to evaluate whether the observed differences were related to the older average age of women.

Variable	OR [95% CI]	P-Value	Interpretation
Pneumonia	0.52 [0.29-0.9]	0.023	Women had lower odds.
mRS at Follow-Up (>1)	1.39 [1.18-1.64]	< 0.001	Women had higher odds.
mRS at Follow-Up = 2	1.4 [1.08-1.81]	0.012	Women had higher odds.
mRS at Follow-Up = 3	1.38 [1.05-1.83]	0.022	Women had higher odds.
mRS at Follow-Up = 4	1.9 [1.37-2.65]	< 0.001	Women had higher odds.

Table 5.9: Population 2 - Outcome odds ratios by gender controlled for Age (Significant Results Only)

The effect on functional recovery was slightly reduced in strength after adjusting for age, but the pattern remained. Women still clearly tended to have worse functional outcomes, with mRS scores over 1 being 39% more likely for them. The gender differences for severe disabilities (mRS = 5) were not statistically significant in this model. However, the mRS scores of 2, 3 and 4 all remained significantly more likely for women.

Interestingly, pneumonia became statistically significant when accounting for age. Although its CI was very close to 1, women had clearly lower odds of this complication (OR = 0.52). This effect was not only similar in direction but also in size to the findings from Population 1.

Adjusted Analysis: Controlling for Patient Characteristics and Risk Factors

The next three models added further adjustments for additional patient characteristics and risk factors.

Variable	OR [95% CI]	P-Value	Interpretation
mRS at Follow-Up (>1)	1.56 [1.31-1.86]	< 0.001	Women had higher odds.
mRS at Follow-Up = 2	1.53 [1.17-2.01]	0.002	Women had higher odds.
mRS at Follow-Up = 3	1.69 [1.26-2.27]	< 0.001	Women had higher odds.
mRS at Follow-Up = 4	2.13 [1.52-3]	< 0.001	Women had higher odds.

Table 5.10: Population 2 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors (Significant Results Only)

In the fully adjusted model (Table 5.10), the pattern of worse functional recovery for women remained clearly present. The grouped outcome mRS > 1 and the separate scores 2, 3 and 4 were still significantly more likely in women. In fact, in comparison to the age-adjusted model, the effect size for all four of these had increased, though the CI ranges also widened slightly.

In this model, the difference in pneumonia was no longer statistically significant (CI: [0.33–1.08]), but the OR (0.61) still suggested a trend toward lower incidence for women.

Variable	OR [95% CI]	P-Value	Interpretation
mRS at Follow-Up (>1)	1.57 [1.32-1.87]	< 0.001	Women had higher odds.
mRS at Follow-Up = 2	1.53 [1.17-2]	0.002	Women had higher odds.
mRS at Follow-Up = 3	1.71 [1.27-2.29]	< 0.001	Women had higher odds.
mRS at Follow-Up = 4	2.18 [1.55-3.07]	< 0.001	Women had higher odds.

Table 5.11: Population 2 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors with Unknown included (Significant Results Only)

Viewing the unknown values as a separate group in the risk factors produced minimal change, which can be seen in Table 5.11. The OR values and CIs for mRS > 1 and individual scores 2 to 4 remained significant and stable. The biggest change was for mRS = 4, from an OR of 2.13 to one of 2.18, but this hardly increased the effect.

Variable	OR [95% CI]	P-Value	Interpretation
mRS at Follow-Up (>1)	1.54 [1.3-1.84]	< 0.001	Women had higher odds.
mRS at Follow-Up = 2	1.51 [1.15-1.97]	0.003	Women had higher odds.
mRS at Follow-Up = 3	1.68 [1.25-2.24]	< 0.001	Women had higher odds.
mRS at Follow-Up = 4	2.11 [1.5-2.97]	< 0.001	Women had higher odds.

Table 5.12: Population 2 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors with Smoking and Alcohol Abuse imputed (Significant Results Only)

In the final model (Table 5.12), with imputation of missing smoking and alcohol abuse data, the results remained robust. Women continued to have significantly worse functional

outcomes, especially at mRS = 2 through 4. mRS > 1 had an OR of 1.54, showing a persistent effect even after controlling for all available confounding factors.

The trend for pneumonia was still present in the last two models (OR = 0.61), but the effect was not significant. This was likely due to the lower sample size for this health outcome.

Summary

The results from Population 2 strongly mirrored the patterns observed in Population 1, even with fewer available variables and different patient characteristics.

- **Women consistently showed worse functional recovery** following a TIA. This effect was strongest in the moderate disability range (mRS = 2–4) and persisted across all adjusted models. The grouped mRS > 1 outcome was also significantly higher for women in every model.
- **Pneumonia showed a possible gender difference**, with lower odds for women, but this only reached statistical significance in the model adjusted for age. In the fully adjusted models, the effect remained suggestive but fell short of significance.

In the unadjusted model the mRS score 5 also showed a significant difference, but this did not remain throughout the other models. For stroke recurrence, early worsening, and seizure risk, no gender differences were observed.

Overall, these findings supported and reinforced the conclusions from Population 1, suggesting a consistent gender disparity in functional recovery outcomes following a TIA that persisted even in this more recent population.

5.2 Predictive Accuracy of the ABCD2 and ABCD3-I Scores

This section presents the results concerning the second research question: *"Are there gender-specific differences in the predictive accuracy of the ABCD2 and ABCD3-I scores for recurrence of ischaemic events after a transient ischaemic attack (TIA)?"*

The goal is to assess whether the predictive performance of the ABCD2 and ABCD3-I risk scores varied by gender. This can provide information on whether these clinical tools are equally accurate in determining recurrence risk for both men and women.

This analysis is done on a filtered subset of Population 1 (the TIA sub-registry), which includes only patients for whom at least an ABCD2 score is available. The performance of the scores was evaluated using receiver operating characteristic (ROC) curves, the area under the curve (AUC) metric and related performance metrics. Two different versions of the risk scores were examined and compared for their accuracy and possible gender

differences. First, the *database* ABCD2 score directly entered by care providers and the ABCD3-I score based on it. Second, the *calculated* ABCD2 and ABCD3-I scores, which are re-calculated with data from the registry. While the AUC values of all outcomes are discussed in this section, only relevant selected ROC curves and performance metrics plots are included. Additional tables and figures are available in Appendix 7.

In addition to examining the risk scores in isolation, each model was also evaluated under adjustment for additional covariates — the patient characteristics and clinical risk factors. This helps assess not only the discrimination ability of the scores but also how their performance changed when patient data was added.

It is important to note that while the inclusion of covariates typically will increase predictive accuracy, this does not reflect the quality of the risk score itself. However, it does show how well the score integrates into broader clinical profiles. Outcomes that remain within poor discrimination range (AUC < 0.7) even after full covariate adjustment, can be assumed to have limited practical value. If the recurrence risk cannot be meaningfully predicted even when including detailed patient information, the outcome is unlikely to be a reliable target for risk stratification.

Note on mRS Outcome Modeling: In this analysis mRS at Follow-Up is defined as signifying stroke recurrence when patients have a score of 5 or 6. This means that patients are either suffering from severe disabilities greatly limiting their motor functions or have died after their hospital stay.

Note on Performance Metric Plots: The performance metrics were calculated across the full range of possible decision thresholds - that is, different cut-off values of the ABCD2 and ABCD3-I scores used to classify individuals as high or low risk. The thresholds were placed between actual score values (e.g. 2.5, 3.5, 4.5,...), so that at each threshold, patients scoring above the value are classified as "high risk". In the performance plots these thresholds are shown on the x-axis.

5.2.1 Descriptive Overview

Since this analysis is based on a filtered subset of Population 1, the subset is slightly smaller and demographically different from the full cohort. Table 3 in the Appendix shows the patient characteristics and clinical risk factors for each gender.

Several differences can be observed between male and female patients:

- Women were older than men in this subset (median 75 vs. 70). This is very different from the full Population 1, where women were slightly younger.
- Smoking and alcohol abuse were again much more common among men. After imputation, 25% vs. 14% were smokers and 14% vs. 3.5% were classified with alcohol abuse.

- Men had higher rates of hypercholesterolemia (63% vs. 59%) and cardiac infarction (9.6% vs. 5.4%).
- Atrial fibrillation (21% vs. 17%) and cardioembolic aetiology (21% vs. 18%) were somewhat more common in women.
- A previous stroke was reported slightly more often in men (22%) than in women (20%).
- Macroangiopathy was more frequently seen in men (12% vs. 6.9%), while microangiopathy rates were nearly identical between the genders.

Most other variables showed either no notable differences, or the observed differences were small enough to likely have limited impact. However, the gender gap in age and risk factors - particularly smoking and alcohol - may influence the stroke recurrence outcomes.

The distribution of the ABCD2 and ABCD3-I scores for both male and female patients is shown in Table 5.13. The two versions of the scores - *calculated* and *database* - are split, so they can be compared with each other. For the most part the scores were similarly distributed across gender, but some differences stand out:

- With the **calculated ABCD2 score**, women were slightly more likely to be assigned a score of 4 (25% vs 23%), while men more frequently received scores of 7 (6.7% vs. 6.0%).
- The **database ABCD2 scores** showed a similar pattern. A larger proportion of women had scores of 4 (30% vs. 27%), whereas men had slightly elevated proportions at scores 3, 5, and 6.
- For the **calculated ABCD3-I scores**, men were more often classified into the higher risk brackets. They were the only group to reach the maximum recorded score of 12, and they had slightly higher proportions at scores 10 and 11. Women, by contrast, were more often assigned mid-level scores between 4 and 6.
- A comparable pattern is seen in the **database ABCD3-I scores**, with men again showing a trend toward values in the upper score range.

Table 5.13: Population 1. Risk Stratification Scores

Gender	**Female** N = 2,369[1]	**Male** N = 2,962[1]
ABCD2 - Calculated		
0	6 (0.3%)	9 (0.3%)
1	26 (1.1%)	45 (1.5%)

Gender	**Female** N = 2,369[1]	**Male** N = 2,962[1]
2	118 (5.1%)	156 (5.4%)
3	269 (12%)	385 (13%)
4	574 (25%)	677 (23%)
5	611 (26%)	730 (25%)
6	589 (25%)	714 (25%)
7	139 (6.0%)	195 (6.7%)
NA	37 (1.6%)	51 (1.7%)
ABCD2 - Database		
0	34 (1.4%)	58 (2.0%)
1	150 (6.3%)	217 (7.3%)
2	412 (17%)	509 (17%)
3	432 (18%)	608 (21%)
4	712 (30%)	787 (27%)
5	341 (14%)	456 (15%)
6	229 (9.7%)	253 (8.5%)
7	59 (2.5%)	74 (2.5%)
ABCD3-I - Calculated		
0	5 (0.2%)	8 (0.3%)
1	15 (0.6%)	31 (1.1%)
2	81 (3.5%)	93 (3.2%)
3	158 (6.8%)	186 (6.4%)
4	290 (12%)	348 (12%)
5	347 (15%)	430 (15%)
6	511 (22%)	565 (19%)
7	378 (16%)	501 (17%)
8	359 (15%)	475 (16%)
9	141 (6.1%)	182 (6.3%)
10	38 (1.6%)	63 (2.2%)
11	5 (0.2%)	23 (0.8%)
12	0 (0%)	2 (<0.1%)
NA	41 (1.7%)	55 (1.9%)
ABCD3-I - Database		
0	11 (0.5%)	22 (0.7%)
1	71 (3.0%)	82 (2.8%)
2	191 (8.1%)	221 (7.5%)
3	267 (11%)	365 (12%)
4	512 (22%)	573 (19%)
5	353 (15%)	505 (17%)
6	515 (22%)	567 (19%)
7	238 (10%)	318 (11%)

Gender	**Female** N = 2,369[1]	**Male** N = 2,962[1]
8	143 (6.0%)	204 (6.9%)
9	49 (2.1%)	71 (2.4%)
10	13 (0.5%)	15 (0.5%)
11	1 (<0.1%)	9 (0.3%)
12	0 (0%)	2 (<0.1%)
NA	5 (0.2%)	8 (0.3%)

[1] n (%)

Since the data is pre-filtered to the scores availability, the proportion of missing values was very low.

The distribution of the outcome variables related to stroke recurrence also needs to be examined. In Table 5.14 the clinically defined indicators (such as Early worsening or Recurrence within 90 days) and the four combined recurrence definitions are shown.

Overall, recurrence was relatively uncommon across all definitions. For the indicators Early worsening, Early recurrence in the Stroke Unit and Recurrence within 90 days fewer than 5% of patients experienced the event, regardless of gender. The outcome based on mRS at Follow-Up greater than 4 also had low event rates (5.0% in women and 4.1% in men), but with considerable missingness. This is due to Follow-Up data being only available for roughly one third of patients.

Recurrence 1 through Recurrence 4 also showed low event frequencies, typically between 5% and 9%, depending on how recurrence was defined. The highest rates were observed in Recurrence 3, with 9.1% of women and 8.4% of men experiencing recurrent stroke events. Even though Recurrence 4 includes the broadest set of indicators, it had lower absolute rates (6.2% in women and 6.5% in men).

Table 5.14: Population 1. Recurrence Variables

Gender	**Female** N = 2,369[1]	**Male** N = 2,962[1]
Early Worsening		
No	2,322 (98%)	2,884 (97%)
Yes	47 (2.0%)	78 (2.6%)
Early Recurrence in Stroke Unit		
No	2,294 (97%)	2,849 (96%)
Yes	75 (3.2%)	113 (3.8%)
Recurrence within 90 days		
No	2,182 (96%)	2,736 (95%)

Gender	**Female** N = 2,369[1]	**Male** N = 2,962[1]
Yes	102 (4.5%)	137 (4.8%)
NA	85 (3.6%)	89 (3.0%)
MRS at Follow-Up > 4		
No	974 (95%)	1,249 (96%)
Yes	51 (5.0%)	53 (4.1%)
NA	1,344 (57%)	1,660 (56%)
Combined vascular endpoint		
No	2,168 (95%)	2,719 (95%)
Yes	116 (5.1%)	155 (5.4%)
NA	85 (3.6%)	88 (3.0%)
Cumulative endpoint		
No	2,157 (94%)	2,714 (94%)
Yes	127 (5.6%)	160 (5.6%)
NA	85 (3.6%)	88 (3.0%)
Recurrence 1		
No	970 (94%)	1,244 (95%)
Yes	58 (5.6%)	67 (5.1%)
NA	1,341 (57%)	1,651 (56%)
Recurrence 2		
No	901 (92%)	1,165 (93%)
Yes	83 (8.4%)	82 (6.6%)
NA	1,385 (58%)	1,715 (58%)
Recurrence 3		
No	897 (91%)	1,162 (92%)
Yes	90 (9.1%)	96 (7.6%)
NA	1,382 (58%)	1,704 (58%)
Recurrence 4		
No	2,221 (94%)	2,769 (93%)
Yes	148 (6.2%)	193 (6.5%)

[1] n (%)

Across all variables, the gender differences in event rates were small and inconsistent in direction. In some cases, women had slightly higher recurrence rates, in others, men did. However, given the relatively low number of observed recurrence events, as well as the substantial missingness for certain outcomes like the mRS at Follow-Up, even small fluctuations can noticeably shift percentages. Meaning observed gender differences may be more reflective of data sparsity rather than actual underlying differences.

Finally, the distribution of the recurrence events across the ABCD2 and ABCD3-I scores

showed some differences between the *calculated* and *database* versions (Appendix 7). In the *calculated* scores, recurrence was clustered more at higher score values, particularly around the mid-to-high ranges. In contrast, for the *database* scores, events were spread more broadly across the middle scores, with slightly fewer events at the upper ends. The distribution of *calculated* scores obviously matches more closely with what would be expected of the risk scores. These differences likely reflect a better consistency in the re-calculated scores, as well as differences in how care providers assigned scores in clinical practice compared to the strict algorithmic definitions used for the *calculated* versions.

5.2.2 Evaluation of 'Database' Risk Scores

The AUC values of the *database* ABCD2 and ABCD3-I risk scores are visualised in Figure 5.1, while tables with the exact values can be found in Appendix 7. These values were evaluated across all the recurrence outcomes and levels of covariate adjustment, with separate results for women, men and the full cohort.

Without the inclusion of additional patient data, both scores demonstrated limited ability to discriminate between recurrence or no recurrence. The ABCD2 scores fell within the 0.5-0.6 value range across all outcomes and genders, with only one exception. This corresponds to failed discrimination and prediction results close to randomness. Based on these results, the score alone was, in most cases, insufficient for accurate recurrence prediction. The one exception was the mRS at Follow-Up > 4 outcome for women, where the AUC was 0.606 and so barely passed as a poor performance.

While most values still remained below 0.6, the ABCD3-I score had slightly better results. Though still within the range of failed discrimination, more values tended towards the upper limit. Additionally, three outcomes presented with results for poor discrimination in at least one gender - Early worsening, mRS and Recurrence 1. The better performance of the ABCD3-I score was, of course, expected, as it builds on the ABCD2 score and adds additional information. On average the ABCD3-I score outperformed the ABCD2 by +0.049, and this difference was larger for men (+0.055) than for women (+0.041).

Adjusting for only age generally resulted in gains in predictive accuracy, though the degree of it varied strongly. In the ABCD2 score, for example, Early recurrence in the Stroke Unit showed no difference overall, while the value increased for women by +0.003 and for men by +0.01. In contrast, the AUC value of Recurrence 2 increased by +0.073 and +0.168 for women and men, respectively. Interestingly, this adjustment benefitted especially the predictive accuracy for men, with on average +0.089 higher values compared to women with only +0.037. While less pronounced, this pattern was also true for the ABCD3-I score and values increased on average by +0.07 and +0.028. Due to the adjustment for age, several outcome variables now had poor (0.6-0.7) or even fair (0.7-0.8) performances, as can be seen in the plot 5.1. For both risk scores, Recurrence 1, Recurrence 2 and mRS crossed the 0.7 threshold, with the last one almost reaching 0.8 for men.

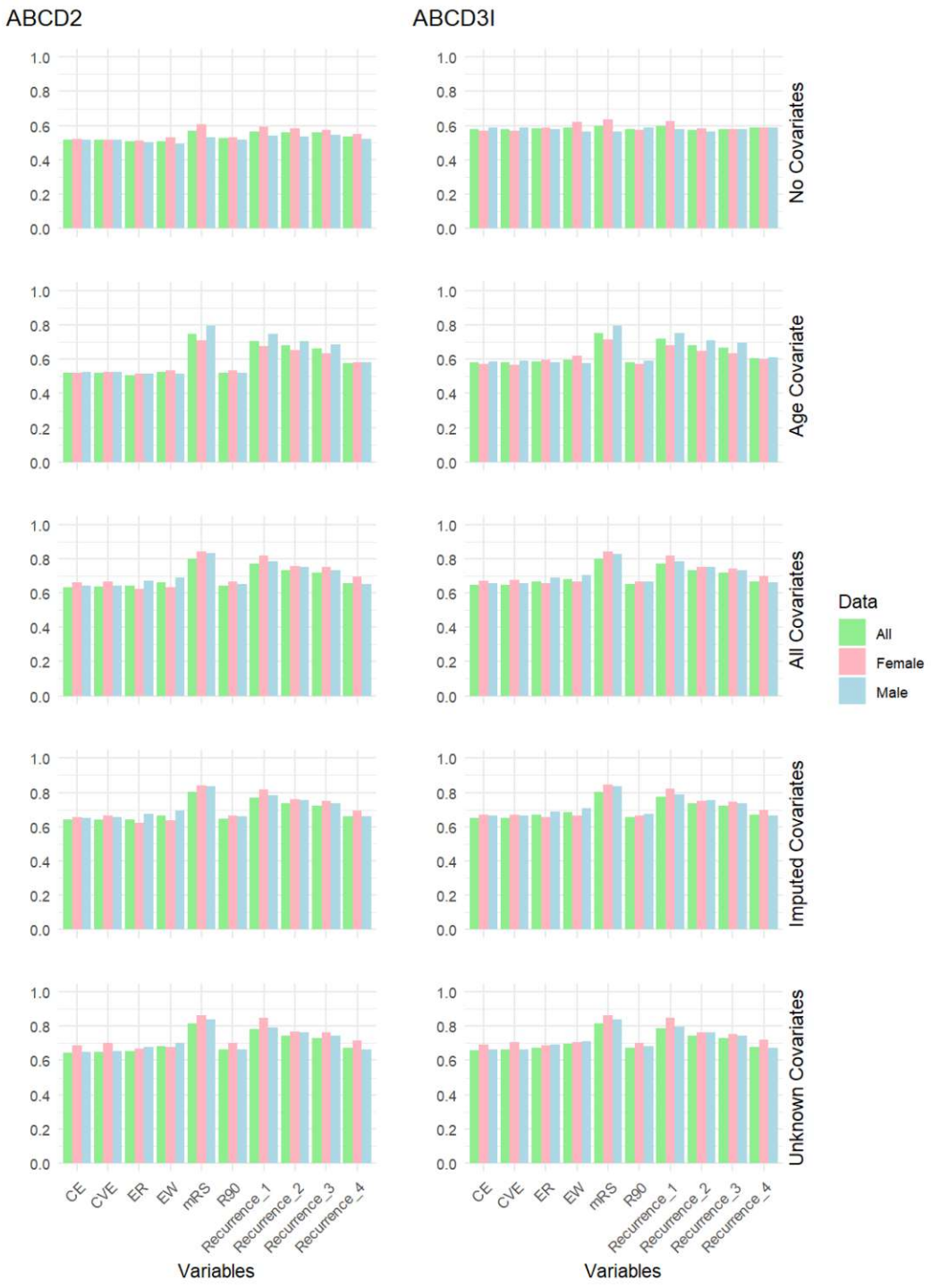


Figure 5.1: AUC Values - ABCD2

When adjusting for all available covariates, meaning clinical risk factors and patient characteristics, the predictive accuracy of both scores notably improved again. As expected, adding more patient information allowed the models to better capture underlying risk patterns, though the extent of improvement still varied by outcome, score, and gender.

For the ABCD2 score, this pushed several outcomes finally out of the failed discrimination range (0.5-0.6) and mRS at Follow-Up > 4 and Recurrence 1 across the 0.8 threshold to a good predictive performance. The highest AUC of 0.842 was observed for women in the mRS outcome. While only adjusting for age had biased the predictive accuracy towards a better performance for men in some of the variables, the full adjustment balanced this out again. Still, men gained more predictive accuracy (+0.182) than women (+0.159) between no adjustment and adjusting for all covariates.

What can also be observed is that the ABCD3-I score no longer clearly outperformed the ABCD2 score. On average only +0.009 lies between the two AUC values, which barely makes a difference. Due to this, any observations about the ABCD2 score are also true for the ABCD3-I score. Most outcomes are now firmly in the poor or fair performance range, with mRS and Recurrence 1 even surpassing this. The mRS AUC reached 0.843 for women and 0.828 for men, representing the highest predictive accuracy so far.

Despite these improvements, some outcomes remained difficult to predict even with all covariates included. For instance, Early recurrence in the Stroke Unit, Recurrence within 90 days, the Combined vascular endpoint, the Cumulative endpoint and Recurrence 4 all still failed to exceed the 0.7 threshold, only achieving poor discrimination across both scores. This pattern is largely due to the limited predictability of just two components: Early recurrence in the Stroke Unit and Recurrence within 90 days, since these are part of the other three composite outcomes.

Imputing values for the risk factors, smoking and alcohol abuse, hardly affected the predictive accuracy of the scores at all, with only an average increase of +0.002. Interestingly, imputation did not always result in improvement, instead marginally reducing some AUC values. For example, in the ABCD3-I score for women in outcomes like Recurrence within 90 days and the Combined vascular endpoint, the AUC dropped by -0.001. While clearly not a difference of relevant size, it is interesting that the additional values reduced the predictive accuracy.

Including *Unknown* as a category for the risk factors did increase the overall discrimination ability compared to the fully adjusted results. The effect was similar for both risk scores, therefore, still keeping them at about the same AUC values for the outcomes. On average for the ABCD2 score the increase was +0.028 for women and +0.01 for men, while for the ABCD3-I score it was +0.025 and +0.09, respectively. The classification of the predictive strength did not change for most outcomes. However, a few did manage to finally cross the 0.7 threshold, though each only for one gender - women. For example, Recurrence within 90 days reached fair discrimination (AUC 0.701) with the ABCD2 score, as did the Combined vascular endpoint and Recurrence 4.

Gender Differences in Predictive Performance

Across most outcomes and adjustment levels, gender differences in the predictive performances of the ABCD2 and ABCD3-I scores were small. There was no consistent advantage for either women or men, and the relative differences tended to shift slightly depending on the adjustment.

In the unadjusted models, women generally had slightly higher AUC values. However, these differences were in most cases marginal. After adjusting for age, this trend temporarily reversed, and men saw greater accuracy increases. While age is incorporated in both risk scores through a binary cut-off (age > 60), the additional granularity from modelling age as a continuous covariate clearly added value — especially for men, who were on average younger.

Once all covariates were included, the gender gap narrowed and women tended to have slightly higher AUC values again. In some cases the difference was large enough that one gender had a better-rated predictive performance than the other. Specifically, women had good, while men only had fair discrimination for the Recurrence 1 outcome in the ABCD2 score. The same was true in the ABCD3-I score, though additionally Early worsening had a fair performance for men versus a poor one for women.

Imputing the missing data for smoking and alcohol abuse had virtually no impact on gender disparities, though in some cases it even slightly reduced performances. The additional *Unknown* category led to minor predictive accuracy improvements. This caused some gender differences, as only women crossed the 0.7 threshold for a few outcomes.

While gender differences in predictive accuracy were generally small and inconsistent, they did emerge in specific outcomes and under certain model adjustments. To better understand these and to assess whether they may be significant, the next section takes a closer look at a few outcomes where differences between genders or between models were most pronounced.

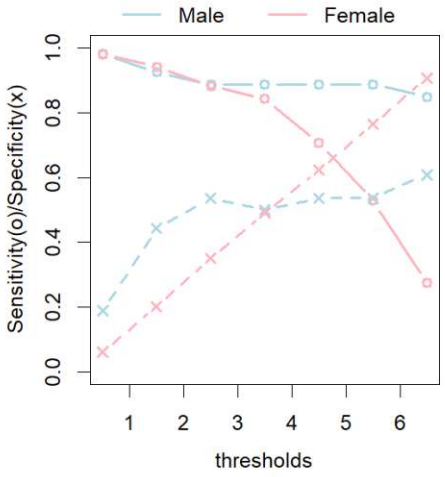
Gender-Specific Score Performance in Selected Outcomes

For the analysis ROC curves and performance metric plots were used. Only some of the recurrence outcomes are examined in this section and for those only a handful of all possible graphs were included. In the Appendix all 7 ROC curves and all performance plots 7 can be found.

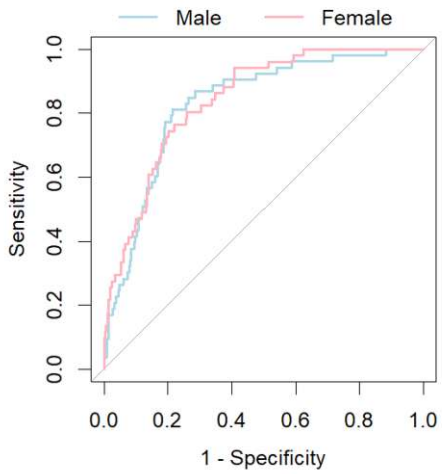
The modified Rankin Scale (mRS) at Follow-Up > 4 was one of the strongest-performing outcomes in the *database* score analysis, particularly under full covariate adjustment. Figure 5.2 shows three plots which further illustrate the behaviour already suggested by the AUC values.

The sensitivity and specificity curves in Figure 5.2a reveal how adjusting for age affected the ABCD2 score across thresholds. As expected for both genders, sensitivity began high and then decreased while specificity increased. However, for men sensitivity was far

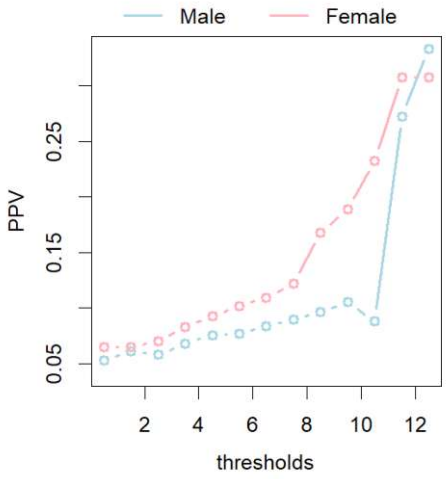
more stable across thresholds compared to women, whose sensitivity declined sharply, particularly after threshold 4.5. The better balance between sensitivity and specificity for men directly mirrors the far higher AUC that was observed for men (0.797 vs. 0.712).



(a) ABCD2 - Age Covariate: Sensitivity/Specificity



(b) ABCD2 - All Covariates: ROC Curves



(c) ABCD3-I - All Covariates (with Unknown): PPV

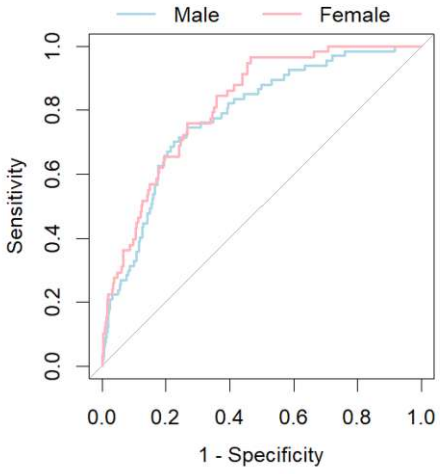
Figure 5.2: 'Database' Scores - mRS at Follow-Up > 4: Performance Metrics

On the one hand, this difference speaks to the age adjustment more strongly affecting men, since their curves did not really show the expected trade-off between sensitivity and specificity. On the other hand, for women this suggests a tighter clustering of high mRS values in women with the ABCD2 score of 4, who were then incorrectly classified

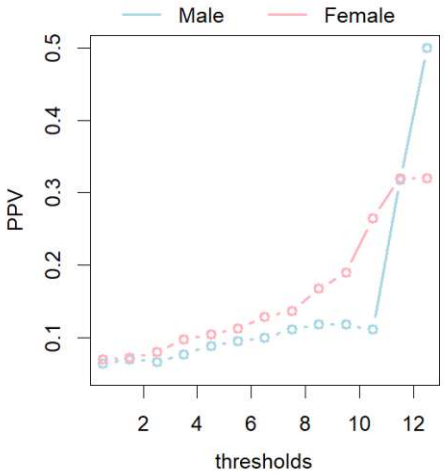
as "low" risk. The score distribution for mRS at Follow-Up > 4 (Appendix 7) confirms this, though men also show a similar pattern. However, for them, the age adjustment was able to suppress the effect.

This better predictive ability was not consistent, and after the full adjustment, different behaviour was observed. The ROC curves in Figure 5.2b demonstrate the strong performance of the ABCD2 score in the fully adjusted case. Both the male and female curves approach the top left corner of the plot, confirming good discrimination. But the curve for women lies above that of men in most of the plot, especially in the high-sensitivity region. This aligns with the slightly higher AUC seen in women (0.842) compared to men (0.831). Despite the earlier imbalance in sensitivity and specificity in the age-adjusted case, under full adjustment the risk patterns were captured slightly more effectively in women.

Lastly, Figure 5.2c highlights the positive predictive value (PPV) across thresholds for the ABCD3-I score when *Unknown* was included as its own risk factor category. The curves show a mostly steady rise in PPV for both genders, but around threshold 7.5 this quickly becomes far steeper for women. Although there were clusters of poor outcomes at ABCD3-I scores 6 and 7 (Appendix 7), the PPV only increased afterwards. This suggests that in the adjusted model, many women with scores in this range were not classified as high-risk due to moderating covariate effects. Based on the abrupt PPV increase, the predicted risk aligned with the recurrence outcome only at higher score values.



(a) ABCD2 - All Covariates: ROC Curves



(b) ABCD3-I - All Covariates (with Unknown): PPV

Figure 5.3: 'Database' Scores - Recurrence 1: Performance Metrics

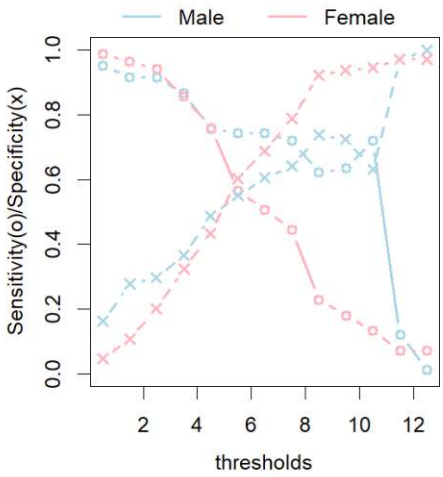
Throughout the AUC analysis, Recurrence 1 was very similar in performance patterns to the mRS. It had an especially strong discriminative ability in the fully adjusted case, which can be seen in Figure 5.3a. The ROC curves for the ABCD2 score clearly

show that both men and women benefitted from the full covariate inclusion. Still the performance for women was better, with their curve lying consistently above the one for men. This matches the behaviour observed for mRS and of course corresponds to the AUC values for Recurrence 1. However, this caused different ratings for the discrimination performances. With an AUC of 0.819 women had a good, while men with 0.784 only had a fair discrimination. Even though the ROC curves seem very similar, this threshold distinction leads to a stronger interpretation of the observed difference.

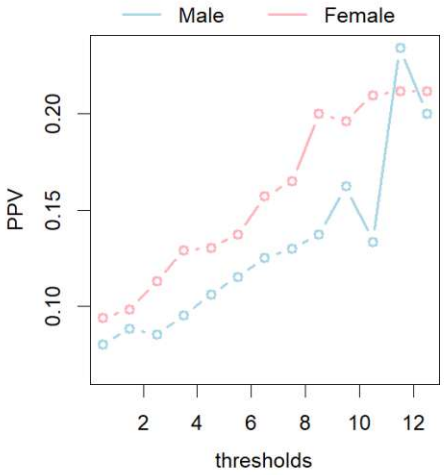
Further similarities between Recurrence 1 and mRS at Follow-Up can be seen in the positive predictive value plot in Figure 5.3b. The PPV across the ABCD3-I thresholds for the fully adjusted model (with *Unknown* included) was, for the most part, slightly higher for women. The difference increased around score 9, where the line rose for women but flattened for men until score 11. This suggests that at lower thresholds, the ABCD3-I score was already able to more reliably identify true positives in women. The AUC values confirm this as again women reached a good predictive performance (0.847) compared to the fair one observed in men (0.795). Despite the genders having similarly distributed outcomes (Appendix 7), the adjustment allowed for better identification of relevant predictors of recurrence in women.

Compared to mRS at Follow-Up > 4 and Recurrence 1, the predictive accuracy for Recurrence 2 was generally lower. However, after adjustment this outcome showed substantial improvements. The sensitivity and specificity plot in Figure 5.4a of the ABCD3-I score after age adjustment illustrates the first clear shift in this outcome. While both genders follow the expected inverse relationship between sensitivity and specificity, the curve for men shows a notable separation from the earlier failed performance observed in the unadjusted AUC (0.567), with sensitivity values now consistently very high across thresholds. While for women the trade-off was very typical and happened at threshold 5.5, for men it happened at 8.5 and then also reversed again for a single threshold at 10.5. This further reinforces the strong effect age had on the predictive ability, even pushing the AUC for men across two thresholds to a fair discrimination (0.078).

The second shift for Recurrence 2 was for the fully adjusted case. The differences to the unadjusted AUC values were immense, with both genders now having similar and clearly fair performances (0.750 vs. 0.752). Figure 5.4b shows the PPV for the ABCD3-I score increased steadily for both genders, suggesting an improved prediction accuracy at higher scores. However, even with almost identical AUC values, the PPV lines show distinct differences, especially in the steepness of the increase. Up until threshold 8.5 the PPVs developed very similarly, with women having slightly higher values, but then the PPV for men begins to rise sharply. Still, the total difference in PPV values is rather low, which explains why men only have an AUC that is higher by +0.002.



(a) ABCD3-I - Age Covariate: Sensitivity/Specificity



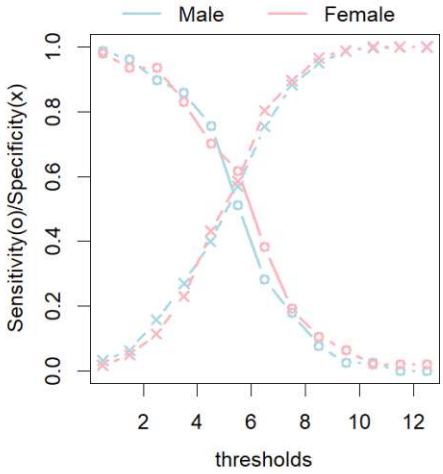
(b) ABCD3-I - All Covariates: PPV

Figure 5.4: 'Database' Scores - Recurrence 2: Performance Metrics

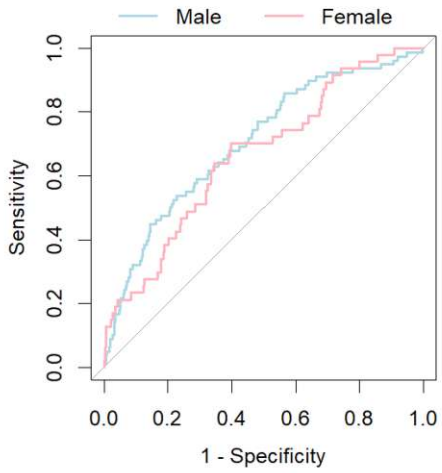
Early worsening differs from the outcomes examined so far. Not only was it not affected by the age adjustment, but it is also one of the few outcomes where, with all covariates included, men had a higher predictive accuracy. In the figures in 5.5 these behaviours are analysed for the ABCD3-I score.

The sensitivity and specificity plot in Figure 5.5a shows that the expected trade-off pattern for both genders is almost identical after age adjustment. The lines cross over at the 5.5 threshold, which is also where this trade-off happens when no adjustment is applied. However, in contrast to the previously observed behaviour for this adjustment, there was no improvement to the discrimination ability. The AUC values stayed almost the same and women were slightly favoured (0.621 vs. 0.577). This suggests that age, as a sole covariate, did not meaningfully enhance the score's ability to identify Early worsening and rather preserved the initial performance pattern.

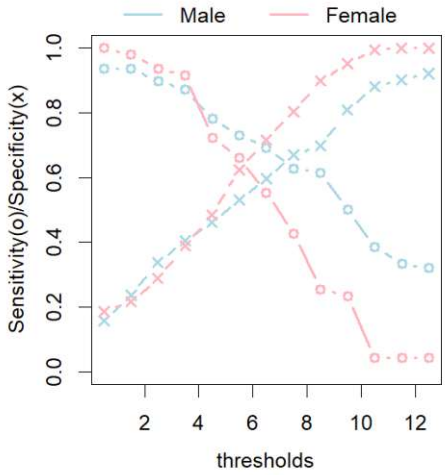
The inclusion of all covariates did, however, have an impact, which can be seen in Figure 5.5b. Both curves show at least decent discrimination, but with the reversals that the curve for men is now consistently above that of women. For men sensitivity increased more steeply, while for women the curve almost has a step pattern where it flattened at times. Rather than just closing the gap as observed in other outcomes, here the adjustment pushed the performance for men, leading to a fair AUC (0.705) compared to a still poor AUC for women (0.665).



(a) ABCD3-I - Age Covariate: Sensitivity/Specificity



(b) ABCD3-I - All Covariates: ROC Curves



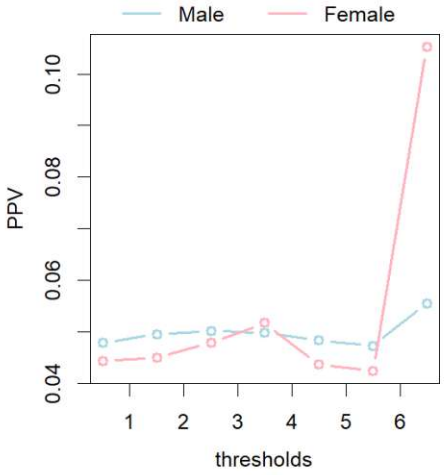
(c) ABCD3-I - All Covariates (with Unknown): Sensitivity/Specificity

Figure 5.5: 'Database' Scores - Early Worsening: Performance Metrics

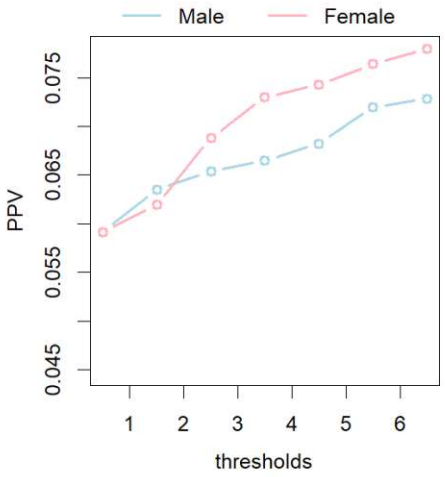
The difference caused by this adjustment and the additional inclusion of the *Unknown* category is visible in Figure 5.5c. While sensitivity and specificity still follow the inverse pattern compared to the plot with only age adjustment (5.5a), a clear performance shift can be seen. For men the sensitivity declines more gradually, while the specificity increases more steadily, resulting in a more balanced trade-off. A slight improvement is also noticeable for women, though the change is subtler. These differences also led to both genders crossing the fair discrimination threshold, with AUC values of 0.712 for men and 0.705 for women. This narrowed the performance gap but still kept the trend

of male patients benefitting in this outcome.

Despite having a weak performance overall, Recurrence within 90 days is of interest. Only when fully adjusted with the *Unknown* category included was a fair discrimination reached, and then only for women. For ABCD2 in the unadjusted case, both genders failed at determining recurrence, with AUC values of 0.533 for women and 0.519 for men. But in the final version the AUC for women reached 0.701, while men remained lower at 0.665. This jump in performance can also be seen in the PPV plots.



(a) ABCD2 - No Covariate: PPV



(b) ABCD2 - All Covariates (with Unknown): PPV

Figure 5.6: 'Database' Scores - Recurrence within 90 Days: Performance Metrics

Without adjustment (Figure 5.6a), the PPV remained mostly flat across thresholds for both genders, staying below 0.06. At the final threshold the value for men rose only slightly, while the one for women rose above 0.1. However, this is caused only by a few observations (4 for men and 6 for women - Appendix 7). In contrast, after full adjustment and with *Unknown* included (Figure 5.6b), the PPV rises consistently across the thresholds. While the absolute PPV remains low, which is to be expected based on the AUCs, for women a meaningful difference can be observed. The divergence of the curves reflects the better alignment between prediction and outcome in women, supported by the AUC difference.

5.2.3 Evaluation of 'Calculated' Risk Scores

The AUC values for the *calculated* ABCD2 and ABCD3-I scores are shown in Figure 5.7, and tables with the results can be found in Appendix 7. The evaluation was performed across all recurrence outcomes and levels of covariate adjustment, for both the entire population and for the genders separately.

Compared to the *database* scores, the *calculated* versions demonstrated consistently higher predictive performances. On average the unadjusted *calculated* ABCD2 scores were +0.076 and the ABCD3-I scores +0.049 better. This improvement was observed across all outcomes and was slightly more pronounced in men (+0.081 for ABCD2 and +0.055 for ABCD3-I) than in women. As before, on average the ABCD3-I outperformed ABCD2, with the gap particularly evident among men (+0.029) compared to women (+0.013). Most AUC values fell into the poor discrimination range (0.6–0.7), unlike the *database* results, which mostly classify as failed discrimination (0.5-0.6).

An important difference can be observed in which outcomes performed best. While in the *database* scores mRS at Follow-Up > 4 and Recurrence 1 stood out as the strongest early performers, in the *calculated* version, these outcomes, along with Recurrence 2 and Recurrence 3, were the weakest in the unadjusted models — all with AUCs below 0.6. Since Recurrence 1-3 all include the mRS score they are heavily dependent on it, and their weak performance can likely be attributed to it as well. In contrast, other outcomes such as Early worsening and Early recurrence in the Stroke Unit showed far better predictive strength. Notably, Early worsening had the highest unadjusted AUC value among all outcomes, reaching 0.702 for women using the ABCD3-I score — the only unadjusted model to cross into the fair discrimination range.

Adjusting for age boosted predictive performance across all outcomes, especially for men. Average AUC increases were +0.048 for ABCD2 and +0.044 for ABCD3-I, with gains for men nearly twice those seen in women (+0.068 vs. +0.026 for ABCD2; +0.061 vs. +0.023 for ABCD3-I). For the ABCD2 score, most outcomes remained in the poor discrimination range, but mRS at Follow-Up > 4 and Recurrence 1 both crossed into fair discrimination — despite starting as some of the weakest performing outcomes. The ABCD2 score AUC jumped from 0.581 to 0.796 in the mRS outcome for men, which is an immense shift of +0.215, exceeding any other. Smaller but similar increases happened for Recurrence 1-3, which are all dependent on the mRS score. The ABCD3-I score showed similar patterns, with mRS and Recurrence 1–3 again standing out due to their considerable gains with age adjustment. These results suggest that while the mRS outcome on its own lacks discriminative power in the unadjusted *calculated* scores, it becomes highly informative once age is considered — particularly for male patients.

After adjusting for all available covariates, predictive accuracy improved further, with more outcomes reaching fair or even good levels of discrimination. The ABCD2 score achieved an average increase of +0.101 over the unadjusted version, while ABCD3-I rose by +0.087. Women saw slightly larger average improvements in both scores (+0.110 for ABCD2 and +0.103 for ABCD3-I), but overall the gender balance was relatively even, with neither gender consistently outperforming the other. Just like before, the differences between the two scores were now minimal. The ABCD3-I score only kept a slight overall advantage (+0.009), with a slightly wider gap for men (+0.012) than for women (+0.005).

Compared to the *database* scores, the *calculated* versions continued to perform better, but the margin narrowed. For ABCD2, the difference to the *database* version was now

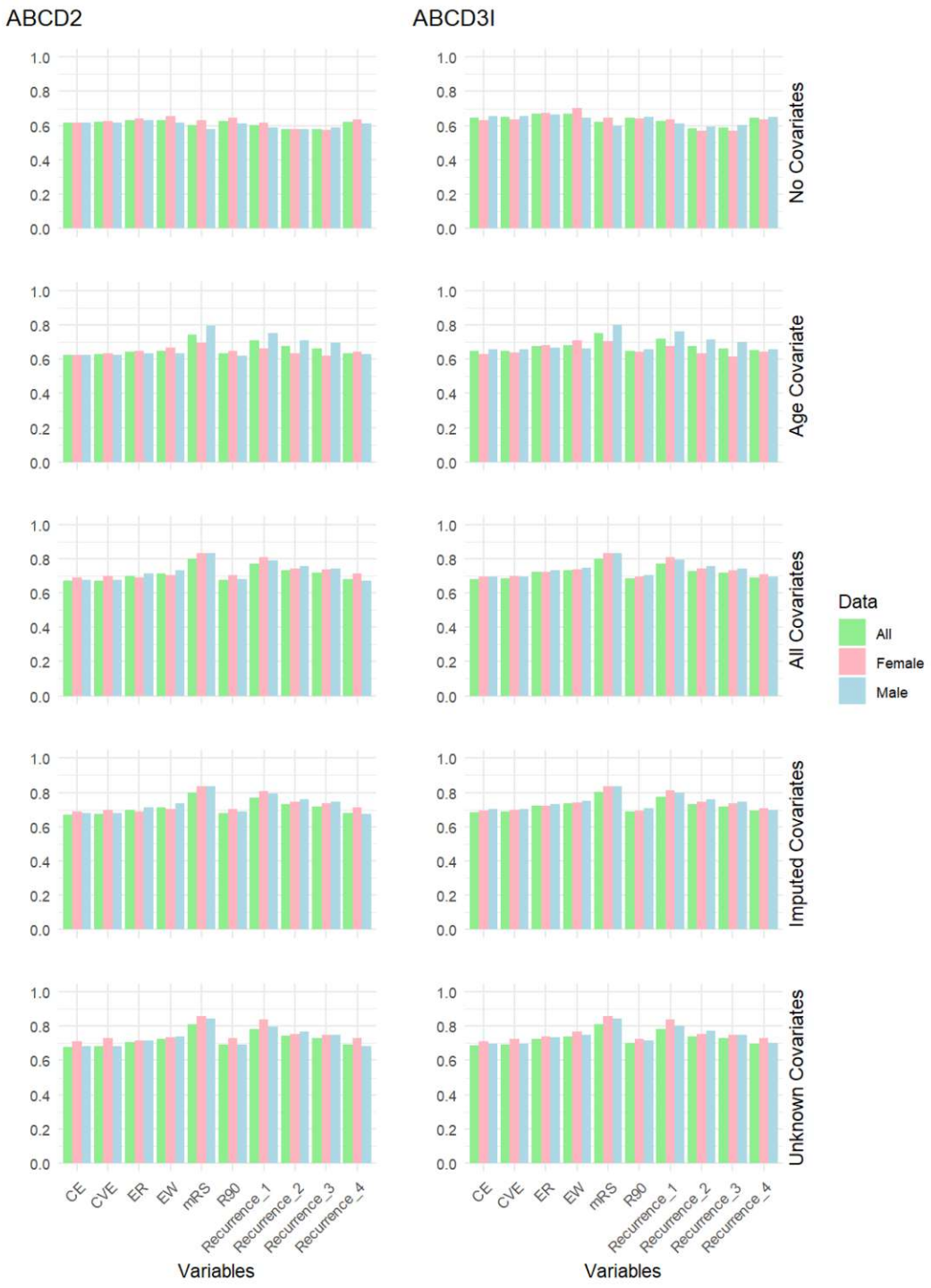


Figure 5.7: AUC Values

+0.023; for ABCD3-I, it was +0.022. These smaller differences suggest that much of the early advantage seen with the *calculated* scores can be explained by better baseline discrimination — once detailed covariate information is included, both scoring approaches reach a similar ceiling. However, again, mRS at Follow-Up > 4 stands out, even though it had low performance without adjustment, now it reached well into the good performance range (AUC > 0.83) for both genders in both scores.

Imputing missing values for the smoking and alcohol abuse risk factors did not notably affect predictive accuracy. The average differences in AUC values were negligible (+0.001 for ABCD2 and +0.002 for ABCD3-I), and no changes in discrimination classification were observed. In fact, some outcomes saw a slightly decreased performance, echoing findings from the *database* scores. These results further support that imputation added little meaningful information.

Treating *Unknown* as its own category for the risk factors led to clearly better predictive accuracies. On average, AUC values increased by +0.011 for ABCD2 and +0.010 for ABCD3-I, with larger gains seen for women (+0.023 and +0.022, respectively). Similarly to the *database* scores, this occasionally pushed women's scores over the 0.7 threshold. Overall the *calculated* ABCD2 and ABCD3-I remained better by +0.020 each, which is a far smaller difference than for the unadjusted and age-adjusted versions, but very close to the previous fully adjusted one.

In both scores, additional outcomes reached a fair discrimination after this final adjustment. While mRS clearly had the highest AUC, other outcomes were noteworthy due to a sustained and stable performance. Early worsening and Early recurrence in the Stroke Unit both steadily improved from a poor to fair discrimination as covariates were introduced.

Gender Differences in Predictive Performance

As with the *database* scores, gender differences in the *calculated* ABCD2 and ABCD3-I scores were generally small and inconsistent.

In the unadjusted models, women had, in general, slightly higher predictive values, but the differences were also not universal. This advantage largely disappeared after adjustment for age, where men benefitted more strongly - likely due to the age distribution differences. However, in mRS and Recurrence 1–3, the effect for male patients was disproportionate, pushing their AUCs from failed into fair or even good ranges.

After full adjustment, the gender gap narrowed considerably, and many outcomes reached similar levels of discrimination for both men and women. In some cases women had slightly higher AUC values, but in others men did, though this direction was mostly consistent for an outcome across the two scores. While most differences remained too small to be meaningful, a few did result in crossing different rating thresholds (e.g., good versus fair).

The final model again produced slight gains in predictive accuracy, pushing a few outcomes over threshold values. This happened more often for women than men, which created a few additional rating gaps. However, these differences were minimal in size and mostly limited to the outcomes that were already near the threshold to begin with.

Overall, the gender-specific differences that were present remained small and were highly dependent on context. In most cases, men and women performed similarly, and shifts in advantage were often tied to specific covariate adjustments or the characteristics of an outcome. The next section examines some specific outcomes to determine where differences might be more meaningful.

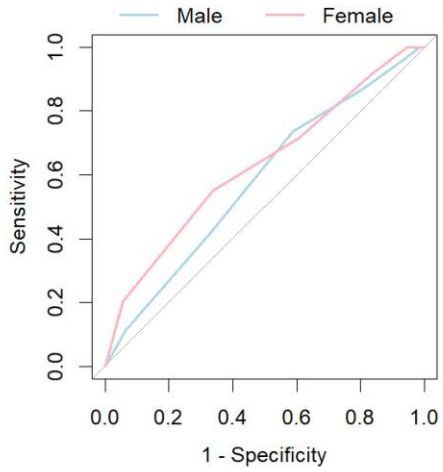
Gender-Specific Score Performance in Selected Outcomes

In the *calculated* scores without adjustment, the modified Rankin Scale (mRS) at Follow-Up > 4 stood out as one of the outcomes with the weakest predictive ability. Even though AUCs were comparable to the *database* version, other outcomes, like Early worsening, already outperformed mRS. However, after adjustments, it became the highest-performing outcome, which can be seen in the plots 5.8.

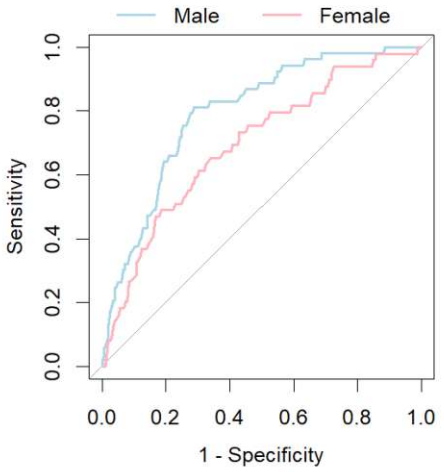
The ROC curves for the ABCD2 score (Figures 5.8a and 5.8b) show the big difference between no covariate adjustment and age adjustment. Initially, both genders were close to the diagonal line, which indicates almost random prediction, with AUC values of 0.63 for women and 0.581 for men. After adjusting for age alone, both curves improved dramatically, especially for men, with a significant shift towards the top-left corner (AUC values: 0.796 vs. 0.696). While the behaviour was consistent with the observations from the *database* scores, the disproportionateness of the increase speaks to how weak the predictive ability for only mRS itself is.

For the ABCD3-I score similar improvements were observed, and the resulting negative predictive values (NPV) can be seen in Figure 5.8c. The NPV was overall very high for both genders, but for men the values stayed consistently above 0.98, while for women they dropped to 0.95. The gap between the genders is not consistent and widens for higher thresholds. This again shows that men benefitted more from the inclusion of age than women, though the difference is not as pronounced for the NPV as for other metrics.

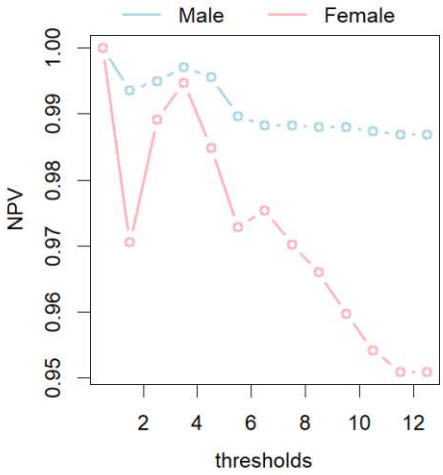
Lastly, the highest performance over all outcomes was observed for the mRS at Follow-Up with the fully adjusted (including *Unknown*) ABCD3-I score. The PPV (Figure 5.8d) visualises the continued linear improvement across thresholds. Women in general had higher PPV values, and while for men the line flattened at higher thresholds, women had an even steeper increase at these. Despite men having better results after age adjustment, the full adjustment leads to a better predictive performance for women, which is consistent with the 'database' results.



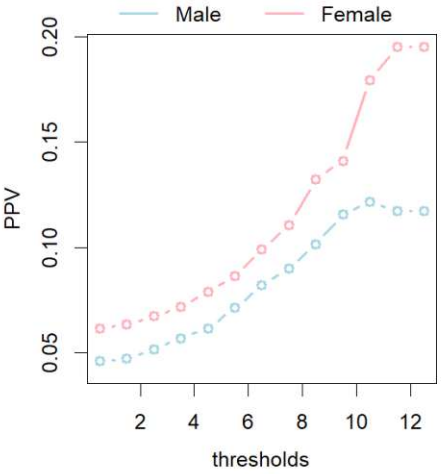
(a) ABCD2 Score - No Covariates: ROC Curves



(b) ABCD2 Score - Age Covariate: ROC Curves



(c) ABCD3-I Score - Age Covariate: NPV



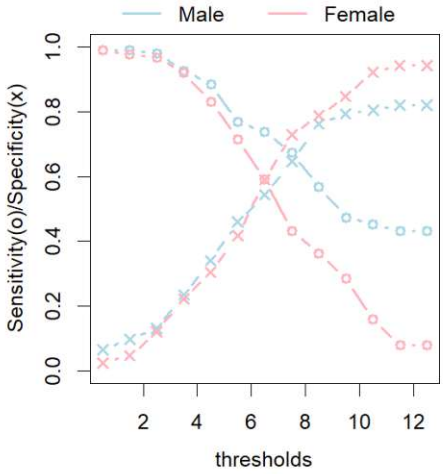
(d) ABCD3-I Score - All Covariates (with Unknown): PPV

Figure 5.8: 'Calculated' Scores - mRS at Follow-Up > 4: Performance Metrics

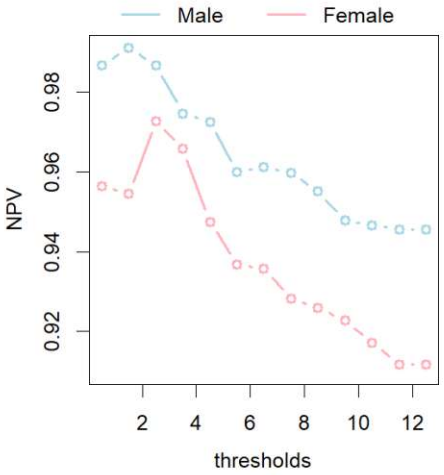
The behaviour observed for mRS at Follow-Up > 4 is also present for Recurrence 1-3, though slightly weaker and with lower predictive performances. Recurrence 3 was chosen to illustrate the strong gender-specific improvement after age adjustment for the ABCD3-I score. Without adjustment the outcome had weak predictive accuracy, with AUC values of 0.571 for women and 0.604 for men. The age adjustment pushed the AUC to 0.703 for men and therefore to fair discrimination, while women only had poor discrimination with 0.618.

Figure 5.9a shows that while both genders followed the expected trade-off pattern, a better balance between sensitivity and specificity was achieved for men. After the 5.5

threshold the sensitivity declined more sharply for women, while for men the line was flatter and the values remained consistently higher. This reflects the stronger predictive ability for men after age adjustment.



(a) ABCD3-I Score - Age Covariate: Sensitivity/Specificity



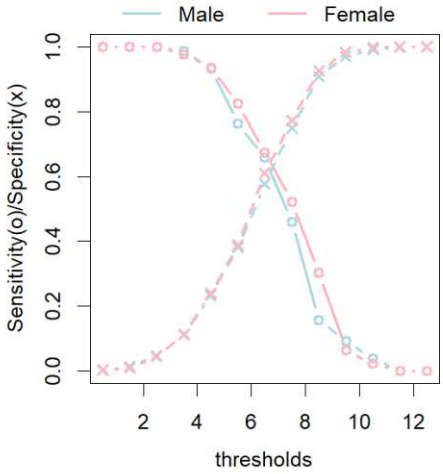
(b) ABCD3-I Score - Age Covariate: NPV

Figure 5.9: 'Calculated' Scores - Recurrence 3: Performance Metrics

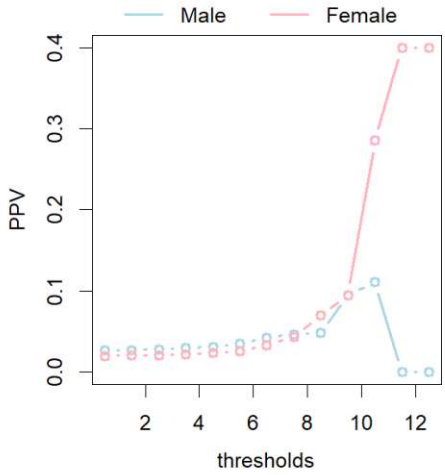
Additionally, the NPV curves in Figure 5.9b show that for men, higher negative predictive values were achieved across the thresholds. Although the absolute differences were small, the trend was stable and mirrored the AUC advantage. This suggests that after accounting for age, ruling out recurrence worked slightly better for men.

Early worsening started out with the best unadjusted discrimination - the ABCD3-I score achieved an AUC of 0.702 for women and 0.648 for men. The performance metrics seen in Figure 5.10a behaved very similarly for both genders, though the sensitivity had slightly higher values for women across the middle thresholds. The trade-off happened at around 6.5, which matches the score distribution (Appendix 7), where most affected patients were clustered at mid-to-high ABCD3-I scores.

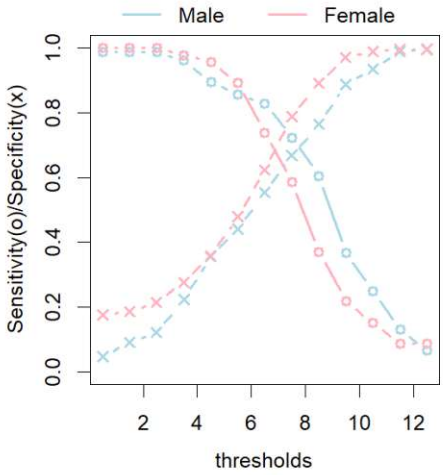
While the previously observed outcomes required age adjustment to improve, the predictive performance of Early worsening was mostly unaffected by it. The positive predictive value for the age-adjusted ABCD3-I score (Figure 5.10b) steadily increased for both genders up until threshold 9.5, afterwards women had a clear advantage. For men the PPV reached just above 0.1 before dropping down again, but for women it went up to 0.4 and remained there. Unlike for mRS and Recurrence 3, no sudden advantage for men was observed after adjusting for age. Instead, the PPV plot shows better predictive accuracy for women, which is confirmed by the AUC values of 0.708 for women and 0.663 for men. However, these are only slightly higher than the unadjusted values.



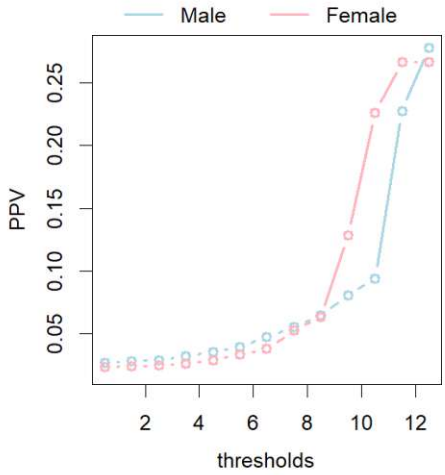
(a) ABCD3-I Score - No Covariates: Sensitivity/Specificity



(b) ABCD3-I Score - Age Covariate: PPV



(c) ABCD3-I Score - All Covariates (with Unknown): Sensitivity/Specificity



(d) ABCD3-I Score - All Covariates (with Unknown): PPV

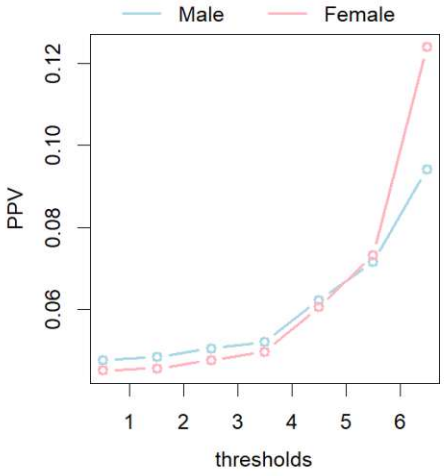
Figure 5.10: 'Calculated' Scores - Early Worsening: Performance Metrics

When adjusting for all covariates, this advantage is actually switched around, with men having a better predictive performance, but upon including the *Unknown* category, this changes back again. In Figure 5.10c the sensitivity and specificity curves of this final adjustment show a more stable trade-off compared to the unadjusted version. For both metrics the lines are less steep, though the difference between genders is more pronounced. Meanwhile, the PPV curves (Figure 5.10d) for both genders look now very similar, though the top value reached is lower than before, at around 0.28.

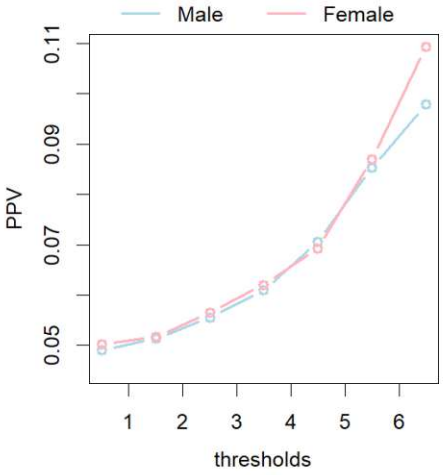
The behaviour of Early worsening was very similar to the *database* score version, but

with an earlier and stronger predictive ability, especially for women.

While for the *calculated* scores, Recurrence within 90 days had a slightly better predictive ability, its behaviour was still comparable to the *database* counterpart. For the ABCD2 score the AUC values steadily increased, women consistently had higher values, and only for them was a fair discrimination achieved through full adjustment. Specifically, the AUC values increased from 0.645 for women and 0.613 for men with the unadjusted ABCD2 score to 0.731 and 0.694 with the inclusion of *Unknown*.



(a) ABCD2 Score - No Covariates: PPV



(b) ABCD2 Score - All Covariates (with Unknown): PPV

Figure 5.11: 'Calculated' Scores - Recurrence within 90 days: Performance Metrics

This change can be observed through a comparison of the positive predictive value plots. Without any adjustment (Figure 5.11a), the PPV started relatively flat across thresholds for both genders before suddenly increasing at threshold 4.5. This matches the score distribution (Appendix 7), since most patients with recurrence had scores of over 4. For the last threshold the lines separate, and women reach a higher PPV than men. This can also still be observed in the fully adjusted model (Figure 5.11b). However, overall the PPV was almost the same for both genders and increased far steadier across all thresholds. The curves were smoother and the underlying distribution was not as evident due to the adjustment.

5.2.4 Summary

Overall, the predictive performance of both the ABCD2 and ABCD3-I scores was limited when used on their own, with most outcomes falling into the failed (AUC 0.5–0.6) or poor discrimination (AUC 0.6–0.7) ranges. Predictive accuracy improved notably after adjusting for age and further with full covariate adjustment, including clinical risk factors.

While including imputed data for smoking and alcohol abuse either hardly increased or even reduced discrimination ability, treating *Unknown* as its own category for the risk factors did positively affect it. However, still only a few outcomes achieved fair (AUC 0.7–0.8) or good (AUC 0.8 - 0.9) discrimination.

Gender-specific differences in score performance were generally small and inconsistent across outcomes and adjustment levels. Without adjustment, women often had slightly higher predictive accuracy, while, after adjusting for age, men tended to benefit more strongly. This was particularly true for mRS at Follow-Up > 4 and three of its composite variables (Recurrence 1-3). However, with full adjustment these differences largely evened out, although for some outcomes women reached higher discrimination ratings (e.g., Recurrence within 90 days and Recurrence 4). Some differences in positive and negative predictive values were also noted, but these were usually modest.

The *calculated* versions of the ABCD2 and ABCD3-I scores consistently outperformed the *database* versions across all adjustment levels and outcomes. This difference was especially important and evident in the unadjusted case, where the *calculated* scores showed substantial baseline discrimination improvements. This is likely related to the underlying distribution of the recurrence events at the different score values. Since for the *database* scores these are far more evenly spread out than would be expected from a risk score, the recurrence prediction at higher score thresholds is simply more likely to fail. The *calculated* scores, which are more internally consistent and algorithmically defined, seem to offer a cleaner starting point for prediction. However, after full adjustment for covariates, the performance gap between the two versions narrowed considerably, showing that rich clinical data eventually levels the playing field between score versions.

The predictive accuracy for the base unadjusted scores was not very high, neither were the gender-specific differences that were observed. On the one hand, the results showed that the inclusion of age as a simple binary cut-off (>60 years) in the ABCD2 and ABCD3-I scores fails to capture a lot of predictive value of age seen when it is modelled continuously. On the other hand, the full covariate adjustment shows the theoretical ceiling of achievable predictive performance, which, on average at most is actually a 0.8 AUC value. The actual discrimination possible when relying solely on the unadjusted scores remained substantially lower, but it provides additional context and proper values to compare these performances.

5.3 Gender Differences in Diagnostics and Treatments

This section presents the results for the third research question: *"Are there gender-specific differences in diagnostic methods and secondary prophylactic treatments after a transient ischaemic attack (TIA)?"*

The aim is to determine whether male and female patients receive different diagnostic procedures or differ in the initiation of secondary prophylactic treatments following a TIA. In addition to overall treatment comparisons, treatment likelihoods for atrial fibrillation,

hypertension, hypercholesterolemia, and diabetes mellitus are also evaluated. This allows for the identification of potential disparities in care based on a patient's gender.

As in the first analysis, data from both Population 1 (TIA sub-registry) and Population 2 (main stroke registry) are used. Both populations contain relevant diagnostic and treatment data, though the available variables differ slightly. Since the two populations represent different time periods, it can be determined whether gender-related differences are consistent across contexts.

The results are mainly derived from logistic regression models. For diagnostic and therapeutic interventions, odds ratios (ORs) with 95% confidence intervals (CIs) are compared by gender. Since men are the reference group, ORs above 1 indicate higher odds for women, and ORs below 1 indicate lower odds. Statistically non-significant differences are defined by confidence intervals crossing 1 and are not interpreted further. As before, only significant results are included in the main tables, with all results available in Appendix 7.

In addition to model-based analysis, observed treatment rates by gender are compared, both overall and stratified by age group. These comparisons are presented through summary tables and mosaic plots. The underlying data for these plots can be found in the appendix.

5.3.1 Population 1

Descriptive Statistics

This analysis again uses the full Population 1 dataset, as in the first research question. The same patient characteristics and risk factors describe the data and are included as covariates in the adjusted models. As previously discussed (Section 5.1.1), some notable gender differences can be observed: men were more likely to smoke, abuse alcohol, have suffered a cardiac infarction, or have hypercholesterolemia. Women were slightly more likely to have atrial fibrillation, microangiopathy, or a cardio embolic aetiology.

A full overview of these variables is provided in Table 1 in the Appendix.

Table 5.15 reports the unadjusted frequencies of diagnostic procedures and secondary treatments for men and women. It allows for some observations before evaluating the model-based results:

- Length of hospital stay, antiplatelet agents, acetylsalicylic acid, and heparin (subcutaneous) had nearly identical rates across genders.
- Cranial computed tomography (CCT) was performed in the majority of patients, with a slightly higher rate in women (88%) than in men (86%).
- Magnetic resonance imaging (MRI) showed a small gender difference, with men being more likely to have an MRI in both available variables. For the *i8004* variable,

MRI was performed in 21% of men vs. 18% of women. In the other (*i27022*) variable, 66% of men had an MRI performed or planned compared to 61% of women.

- Echocardiographic procedures were less frequently performed on women. Transthoracic echocardiography (TTE) was done or planned in 60% of men and 58% of women. Transesophageal echocardiography (TEE) showed a slightly more notable gap: 19.4% of men vs. 15.6% of women had the procedure performed or planned.
- Clopidogrel was prescribed more frequently to men (25%) than to women (22%). While overall antiplatelet use was equal, a small gender difference was present for this specific antiplatelet selection.
- Carotid interventions were more common among men. Percutaneous transluminal angioplasty (PTA) was performed in 1.0% of men vs. 0.6% of women; carotid endarterectomy (CEA) in 4.1% vs. 1.9%; and the combined carotid revascularization variable in 5.1% of men vs. 2.4% of women.
- Rehabilitation data had high missingness (around 64–65%), limiting interpretation. Among those with data, men were slightly more likely to receive rehabilitation (30% vs. 28%), including inpatient rehabilitation (8.7% vs. 7.5%). Outpatient rehabilitation was rare in both genders (<1%).
- Oral anticoagulants (OAC) at Follow-Up were prescribed equally and to almost all patients with available data. However, Follow-Up data were only present in about 36% of patients.

Overall, while many diagnostic and therapeutic interventions were similar between men and women, subtle gender differences appeared in several areas — particularly in the use of MRI, echocardiography, clopidogrel, and carotid interventions.

Table 5.15: Population 1. Diagnostic methods and secondary prophylactic treatments

Gender	**Female** N = 6,884[1]	**Male** N = 8,583[1]
Length of hospital stay		
0	268 (3.9%)	309 (3.6%)
1-2	3,858 (56%)	4,748 (55%)
14+	54 (0.8%)	90 (1.0%)
3-7	2,495 (36%)	3,124 (36%)
8-14	207 (3.0%)	311 (3.6%)
NA	2 (<0.1%)	1 (<0.1%)
Cranial computed tomography (CCT)		
No	851 (12%)	1,221 (14%)
Yes	6,033 (88%)	7,362 (86%)

Gender	Female N = 6,884[1]	Male N = 8,583[1]
Cranial computed tomography (CCT)		
Already on hand	225 (3.3%)	290 (3.4%)
No	851 (12%)	1,221 (14%)
Yes	5,808 (84%)	7,072 (82%)
Magnetic resonance imaging (MRI) - i8004		
No	5,627 (82%)	6,760 (79%)
Yes	1,257 (18%)	1,823 (21%)
Magnetic resonance imaging (MRI) - i27022		
No	1,896 (28%)	1,889 (22%)
Yes	4,988 (72%)	6,694 (78%)
Magnetic resonance imaging (MRI) - i27022		
No	1,896 (28%)	1,889 (22%)
Planned	785 (11%)	1,012 (12%)
Yes	4,203 (61%)	5,682 (66%)
Transthoracic echocardiography (TTE)		
No	2,868 (42%)	3,441 (40%)
Yes	4,016 (58%)	5,142 (60%)
Transthoracic echocardiography (TTE)		
No	2,868 (42%)	3,441 (40%)
Planned	1,440 (21%)	1,780 (21%)
Yes	2,576 (37%)	3,362 (39%)
Transesophageal echocardiography (TEE)		
No	5,811 (84%)	6,879 (80%)
Yes	1,073 (16%)	1,704 (20%)
Transesophageal echocardiography (TEE)		
No	5,811 (84%)	6,879 (80%)
Planned	482 (7.0%)	720 (8.4%)
Yes	591 (8.6%)	984 (11%)
Antiplatelet agents		
No	1,195 (17%)	1,464 (17%)
Yes	5,689 (83%)	7,119 (83%)
Acetylsalicylic acid		
No	2,950 (43%)	3,666 (43%)
Yes	3,934 (57%)	4,917 (57%)
Clopidogrel		
No	5,361 (78%)	6,432 (75%)
Yes	1,523 (22%)	2,151 (25%)
Heparin subcutaneous		
High (>150E)	194 (2.8%)	265 (3.1%)
Low (<75E)	5,032 (73%)	6,224 (73%)

Gender	Female N = 6,884[1]	Male N = 8,583[1]
Medium (75-150E)	817 (12%)	998 (12%)
None	841 (12%)	1,096 (13%)
Percut. transluminal angioplasty (PTA)		
No	6,840 (99%)	8,495 (99%)
Yes	44 (0.6%)	88 (1.0%)
Carotid endarterectomy (CEA)		
No	6,753 (98%)	8,227 (96%)
Yes	131 (1.9%)	356 (4.1%)
Carotid Revascularization		
No	6,716 (98%)	8,147 (95%)
Yes	168 (2.4%)	436 (5.1%)
Rehabilitation		
No	1,744 (72%)	2,180 (70%)
Yes	667 (28%)	916 (30%)
NA	4,473 (65%)	5,487 (64%)
Inpatient rehabilitation		
No	6,367 (92%)	7,835 (91%)
Yes	517 (7.5%)	748 (8.7%)
Outpatient rehabilitation		
No	6,846 (99%)	8,537 (99%)
Yes	38 (0.6%)	46 (0.5%)
Oral Anticoagulants (OAC) at Follow-Up		
No	10 (0.4%)	13 (0.4%)
Yes	2,447 (100%)	3,151 (100%)
NA	4,427 (64%)	5,419 (63%)

[1] n (%)

Unadjusted Likelihood of Diagnostics and Treatments by Gender

Table 5.16 shows the results of the unadjusted logistic regression models comparing the likelihood of receiving specific diagnostics or treatments between men and women.

In the unadjusted model, several gender differences were observed in diagnostic and treatment procedures following TIA. Women were slightly less likely to have longer hospital stays in the 8–14 day range, with an odds ratio of 0.77 and a confidence interval narrowly below 1 (CI: [0.60–0.98]), indicating a modest but statistically significant difference. It should be mentioned that for the longest stay category (14+ days), women also had lower odds, though the CI was wide and just touched 1 (CI: [0.48–1.01]), making this difference neither robust nor significant.

Cranial computed tomography (CCT) was more frequently performed on women, with an 18% increase in odds compared to men. In contrast, magnetic resonance imaging (MRI) was consistently less common in women. Depending on the variable, they had between 17% and 26% lower odds of receiving an MRI. Planned procedures also followed this pattern, with women having 23% lower odds of an MRI being scheduled.

Variable	OR [95% CI]	P-Value	Interpretation
Length of hospital stay = 8-14	0.77 [0.6-0.98]	0.031	Women had lower odds.
CCT	1.18 [1.07-1.29]	< 0.001	Women had higher odds.
CCT = Yes	1.18 [1.07-1.29]	< 0.001	Women had higher odds.
MRI - i8004	0.83 [0.76-0.9]	< 0.001	Women had lower odds.
MRI - i27022	0.74 [0.69-0.8]	< 0.001	Women had lower odds.
MRI - i27022 = Planned	0.77 [0.69-0.86]	< 0.001	Women had lower odds.
MRI - i27022 = Yes	0.74 [0.68-0.8]	< 0.001	Women had lower odds.
TTE = Yes	0.92 [0.86-0.99]	0.021	Women had lower odds.
TEE	0.75 [0.69-0.81]	< 0.001	Women had lower odds.
TEE = Planned	0.79 [0.7-0.89]	< 0.001	Women had lower odds.
TEE = Yes	0.71 [0.64-0.79]	< 0.001	Women had lower odds.
Clopidogrel	0.85 [0.79-0.92]	< 0.001	Women had lower odds.
PTA	0.62 [0.43-0.89]	0.010	Women had lower odds.
CEA	0.45 [0.36-0.55]	< 0.001	Women had lower odds.
Carotid Revascularization	0.47 [0.39-0.56]	< 0.001	Women had lower odds.
Inpatient rehabilitation	0.85 [0.76-0.96]	0.007	Women had lower odds.

Table 5.16: Population 1 - Unadjusted treatment odds ratios by gender (Significant Results Only)

Echocardiographic procedures also showed gender differences. For transthoracic echocardiography (TTE), women had slightly lower odds of receiving the procedure (OR = 0.92). The confidence interval in this case was right at 1 (CI: [0.86-0.99]), so the effect was marginal and right at the threshold of statistical non-significance. The difference was far more pronounced with transesophageal echocardiography (TEE). Women had substantially lower odds of undergoing TEE, both for planned (OR = 0.79) and completed procedures (OR = 0.71).

Clopidogrel was prescribed less often to women, who had 15% lower odds of receiving it. No significant gender differences were found in the use of general antiplatelet therapy or acetylsalicylic acid.

Carotid interventions were clearly more common in men. Women were 38% less likely to undergo percutaneous transluminal angioplasty (PTA), and 55% less likely to receive a carotid endarterectomy (CEA). For the combined carotid revascularization variable, this meant women had 53% lower odds of receiving either of the two treatments.

Inpatient rehabilitation was slightly less common among women (OR = 0.85), though

overall rehabilitation and outpatient rehabilitation did not show a significant gender difference.

Adjusted Analysis: Controlling for Age

In Table 5.17, the results of the logistic regression models adjusted for age are presented. The goal is to assess whether the observed gender differences in diagnostic and treatment procedures persist after accounting for this important demographic factor.

Variable	OR [95% CI]	P-Value	Interpretation
Length of hospital stay = 8-14	0.76 [0.59-0.96]	0.024	Women had lower odds.
Length of hospital stay = 14+	0.65 [0.44-0.95]	0.025	Women had lower odds.
TEE	0.89 [0.82-0.98]	0.013	Women had lower odds.
TEE = Yes	0.85 [0.76-0.95]	0.004	Women had lower odds.
Acetylsalicylic acid	1.17 [1.09-1.25]	< 0.001	Women had higher odds.
Clopidogrel	0.78 [0.73-0.85]	< 0.001	Women had lower odds.
PTA	0.62 [0.42-0.89]	0.010	Women had lower odds.
CEA	0.41 [0.34-0.51]	< 0.001	Women had lower odds.
Carotid Revascularization	0.44 [0.36-0.53]	< 0.001	Women had lower odds.
Inpatient rehabilitation	0.88 [0.78-0.99]	0.033	Women had lower odds.

Table 5.17: Population 1 - Treatment odds ratios by gender controlled for Age (Significant Results Only)

The gender gap for the length of hospital stays became slightly more pronounced. Women remained 24% less likely to stay between 8 and 14 days, but were now also 35% less likely to stay for 14 days or longer. While both odds ratios were about the same as before, the latter was now statistically significant, with its CI no longer crossing 1.

After adjusting for age, the significance of the gender differences in cranial computed tomography (CCT) and magnetic resonance imaging (MRI) disappeared. The odds ratios decreased slightly in magnitude, and the confidence intervals now included 1 (e.g., CCT – CI: [0.96–1.16]; MRI i8004 – CI: [0.85–1.00]).

Similarly, transthoracic echocardiography (TTE) showed no significant gender difference in any category after adjustment. However, women were still 11% less likely to receive transesophageal echocardiography (TEE) overall and 15% less likely to have the procedure performed. Both of these differences remained statistically significant, though the effect sizes were reduced compared to the unadjusted model.

Antiplatelet use overall remained non-significant, but acetylsalicylic acid (ASA) use became significantly more likely in women, with a 17% increase in odds compared to men. In contrast, the lower likelihood of clopidogrel use in women persisted and became even more pronounced after adjusting for age, with a 22% decrease in odds instead of only 15%.

For carotid interventions, the gender differences remained with only minimal changes in effect sizes. Women continued to have lower odds of undergoing PTA, CEA, and carotid revascularization in general.

Similarly, the odds for rehabilitation remained mostly the same. Inpatient rehabilitation was still less likely for women, though now only by 12% instead of 15%.

Adjusted Analysis: Controlling for Patient Characteristics and Risk Factors

In the next models, additional clinical and stroke risk factor covariates were included to better isolate the independent effect of gender. These included variables such as smoking, alcohol abuse, atrial fibrillation, prior stroke, and other vascular risk factors.

Variable	OR [95% CI]	P-Value	Interpretation
Clopidogrel	0.8 [0.7-0.93]	0.003	Women had lower odds.
CEA	0.56 [0.38-0.81]	0.002	Women had lower odds.
Carotid Revascularization	0.68 [0.48-0.94]	0.020	Women had lower odds.

Table 5.18: Population 1 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors (Significant Results Only)

As shown in Table 5.18, most previously observed gender differences lost statistical significance once these covariates were introduced. Further, differences which had already lost significance after adjusting for age also remained non-significant.

Longer hospital stays showed no sign of a gender difference anymore. The CIs for 8–14 days (CI: [0.57–1.37]) and over 14 days (CI: [0.43–1.61]) spanned far across the significance threshold. Transesophageal echocardiography (TEE) also had no longer a significant difference, with a reduced effect size and confidence intervals now including 1.

Acetylsalicylic acid use, which had shown a higher likelihood for women in the age-adjusted model, also lost significance in this fully adjusted model. However, women's lower odds of receiving clopidogrel remained consistent. They were still around 20% less likely to receive it compared to men.

Gender differences partially remained for carotid interventions. Women still had significantly lower odds of receiving a carotid endarterectomy (CEA) (OR: 0.56), but the difference for PTA was no longer statistically significant. However, the combined variable for carotid revascularization still showed women having 32% lower odds of undergoing either intervention.

The previously observed gender difference in inpatient rehabilitation was no longer statistically significant after adjusting for risk profiles.

Variable	OR [95% CI]	P-Value	Interpretation
Clopidogrel	0.8 [0.7-0.92]	0.002	Women had lower odds.
CEA	0.54 [0.37-0.78]	0.001	Women had lower odds.
Carotid Revascularization	0.66 [0.47-0.92]	0.016	Women had lower odds.

Table 5.19: Population 1 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors with Unknown included (Significant Results Only)

When including unknown values as a separate category (Table 5.19), the results remained stable. No new variables became significant and the overall pattern remained unchanged. The odds ratio for clopidogrel remained at 0.80, and both carotid interventions continued to show significantly lower odds for women.

Variable	**OR [95% CI]**	**P-Value**	**Interpretation**
Clopidogrel	0.81 [0.7-0.93]	0.003	Women had lower odds.
CEA	0.56 [0.38-0.81]	0.003	Women had lower odds.
Carotid Revascularization	0.68 [0.48-0.94]	0.021	Women had lower odds.

Table 5.20: Population 1 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors with Smoking and Alcohol Abuse imputed (Significant Results Only)

In the final model (Table 5.20), missing values for smoking and alcohol abuse were handled through imputation.

Women were less likely to receive clopidogrel by 19%, carotid endarterectomy by 44%, and carotid revascularization in general by 32%. The results closely matched the previous two models, with hardly any changes.

Treatment Likelihood for Specific Conditions

Treatments	**Gender**	**No**	**Yes**	**Treatment Rate**
Oral Anticoagulants (OAC)	Male	0	578	100%
for Atrial fibrillation	Female	2	494	99.6%
Regular lipid-lowering drugs	Male	443	1487	77.05%
for Hypercholesterolemia	Female	343	1043	75.25%
Regular blood pressure checks	Male	524	1949	78.81%
for Hypertension	Female	384	1568	80.33%
Regular antihypertensives	Male	708	1752	71.22%
for Hypertension	Female	622	1324	68.04%
Regular antidiabetica	Male	290	423	59.33%
for Diabetes mellitus	Female	208	301	59.14%

Table 5.21: Population 1 - Treatment Rates per gender

This section evaluates whether treatment likelihood differed by gender for patients with specific secondary conditions following a TIA. Four common comorbidities were considered: atrial fibrillation, hypercholesterolemia, hypertension (with two treatments: blood pressure checks and antihypertensive use), and diabetes mellitus. Determining treatment rates by gender helps identify inequities in care provision, specifically in stations where clinical need is the same.

As shown in Table 5.21, treatment rates were generally high across conditions for both men and women. Oral anticoagulants (OAC) for patients with atrial fibrillation stood out with nearly universal treatment in both genders (100% in men, 99.6% in women), leaving little room for disparity. For the other conditions, subtle differences could be identified. Men were slightly more likely to receive lipid-lowering medication. In order to help treat hypertension, blood pressure checks were slightly more common among women, while the opposite was true for antihypertensive prescriptions. The use of antidiabetic medication, however, was virtually identical between genders.

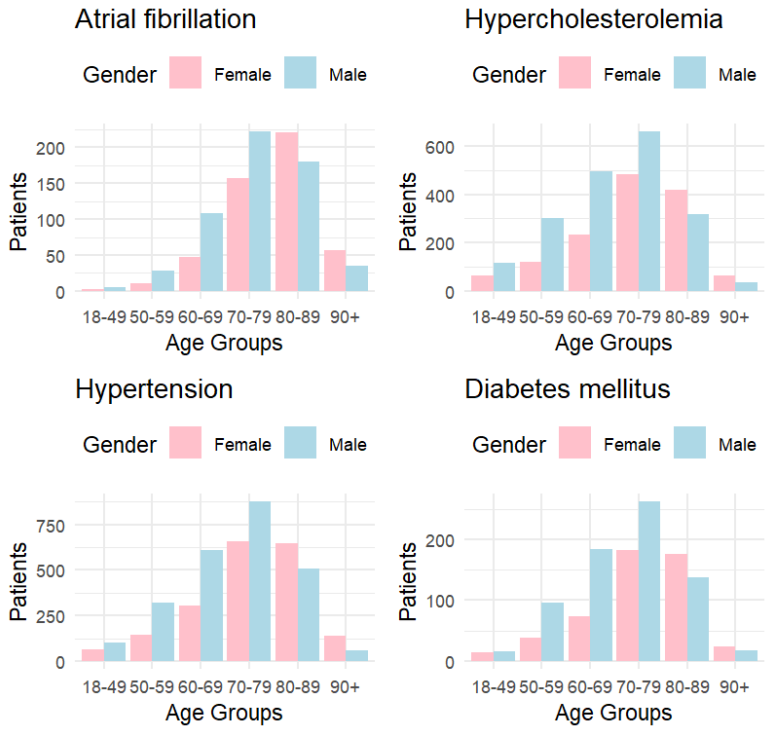


Figure 5.12: Distribution of conditions by age group and gender - Population 1

Since chronic conditions tend to become more common with age, this must be considered when assessing gender-based treatment rate differences. Figure 5.12 visualises how each condition is distributed by age group and gender. As expected, older patients were more affected by all four conditions, but the distributions did somewhat differ.

Atrial fibrillation was most common in the oldest age groups. Among women, the highest concentration was seen in the 80–89 range, whereas for men, the peak occurred slightly earlier, in the 70–79 group. Hypercholesterolemia was consistently more common among men across almost all age groups and peaked in both genders between 70 and 79. The distribution of hypertension was fairly balanced between men and women, though men were more affected in most age groups. Meanwhile, diabetes mellitus had a slightly younger profile, with noticeably more men being affected, especially among those aged 50 to 69.

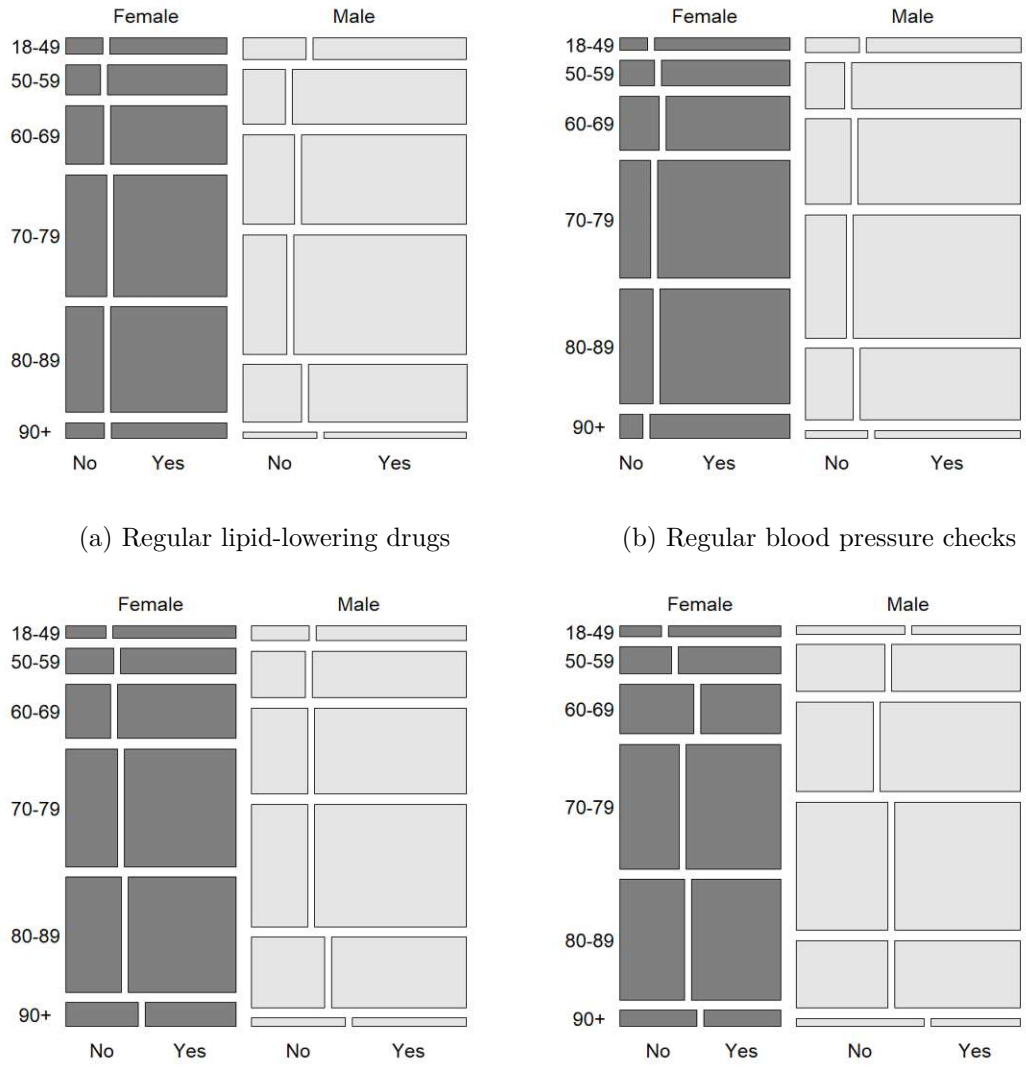
From age 80 forward a clear shift can be seen for all four conditions, as women become more frequently affected. This likely reflects the older age distribution among women in the dataset.

To determine whether gender differences in treatment rates persist across age groups, Figure 5.13 presents mosaic plots of the treatments. Since oral anticoagulants (OAC) were prescribed to nearly all patients with atrial fibrillation, this condition is excluded from the plots. The breakdown of absolute numbers of patients and percentages by age group and gender are provided in Appendix 7.

For lipid-lowering treatment in patients with hypercholesterolemia, the mosaic plot 5.13a shows a fairly balanced pattern across most age groups, with no strong or consistent gender disparity. However, some apparent differences stand out in a few age groups. In the 70–79 age group, men (79.70%) were nearly 6% more likely to receive treatment than women (73.97%). A similar difference was present in the 90+ group, where 75.00% of women were treated versus only 65.71% of men, reversing the pattern. While these differences are not extreme, they are very noticeable, as the other age groups are far closer aligned in treatment rates.

The mosaic plot 5.13b for regular blood pressure monitoring among patients with hypertension shows mostly comparable treatment rates between men and women, though again some age-specific disparities are apparent. Among patients aged 90 and older, women had a substantially higher treatment rate (86.03%) than men (70.00%), a gap of 16 percentage points. Since there are only 60 male patients but 136 female patients in this age group, this difference is very noticeable. A smaller disparity also occurred in the youngest group (18–49), where 83.33% of women received regular blood pressure checks compared to 74.26% of men. For all other ages, treatment rates were closely aligned between men and women, with no more than a 2–3 percentage point difference.

The other treatment for patients with hypertension is antihypertensive medication, as shown in the mosaic plot 5.13c. A modest difference is seen in the age group 50-59 with men (74.39%) being a bit more likely to receive treatment than women (70.92%). In the 70-79 age group this gap is more noticeable, due to a difference of nearly 5 percentage points in favour of men. In all other age groups, treatment rates were fairly similar, with differences of only 1–2 percentage points. In the oldest group (90+), rates were nearly identical between men and women, both just under 56%.



(a) Regular lipid-lowering drugs

(b) Regular blood pressure checks

(c) Regular antihypertensives

(d) Regular antidiabetica

Figure 5.13: Population 1 - Treatment Mosaic Plots

The final mosaic plot 5.13d shows the use of antidiabetic medication in patients with diabetes mellitus. Treatment rates varied more noticeably between age groups than between genders. In the 50–59 group, 66.67% of women received treatment compared to 59.38% of men — a difference of over 7 percentage points. This gap widened in the 60–69 group to over 12 percentage points, with 64.67% of men and just 52.05% of women treated. In the 70–79 and 80–89 groups, treatment rates were nearly identical between genders (around 58%). The youngest (18–49) and oldest (90+) age groups also had large differences, but the total number of patients was small in both cases. As such,

while younger women and older men had slightly higher rates in their age groups, these differences need to be considered with caution.

Summary

From the analyses of gender differences in diagnostics and secondary treatments, two key factors remained across all models for Population 1:

- **Women had lower odds of receiving clopidogrel**. This effect was stable across all models, with odds ratios remaining around 0.8. This indicates that women were about 20% less likely to receive this antiplatelet therapy than men.
- **Carotid interventions, particularly CEA, are less commonly done for women.** Women remained 44–47% less likely to undergo carotid endarterectomy, even after full adjustment for clinical risk profiles.

Most of the initial differences disappeared after adjustment. Notably, the longer hospital stays observed in men and the lower likelihood of women receiving MRI or TEE were all explained by age or other clinical characteristics.

The additional analysis of treatment rates for secondary conditions also revealed that while overall treatment levels were generally high and mostly similar between genders, smaller disparities remained in certain age groups. These age-dependent patterns suggest that subtle differences in treatment provision may still exist — particularly for antihypertensive and antidiabetic medication in middle-aged patients and for lipid-lowering therapy or blood pressure monitoring in the oldest age groups.

5.3.2 Population 2

Descriptive Statistics

This analysis is based on the complete Population 2 dataset. The patient characteristics and risk factors are again included as covariates in the adjusted models. These are discussed in detail in Section 5.1.2, with a full overview provided in Table 2 in the Appendix.

Women were notably older than men and were more frequently diagnosed with hypertension. However, the risk factors smoking, alcohol abuse, cardiac infarction, hypercholesterolemia and a previous stroke all were more common among men.

Table 5.22 shows the unadjusted frequencies of diagnostic methods and secondary prophylactic treatments in Population 2, allowing for some observations before the model-based analysis:

- For length of hospital stay, cranial computed tomography (CCT), magnetic resonance imaging (MRI), heparin dosage, and oral anticoagulants (OAC) at Follow-Up, there were no notable gender differences in the distribution of values.

- Echocardiographic procedures showed some gender variation. Transthoracic echocardiography (TTE) was performed or planned on 73% of men compared to 70% of women. While transesophageal echocardiography (TEE) far less common in general, men were still more likely to receive or have the procedure planned (11% vs. 7.3%).
- Antiplatelet agent use was high in both groups. Dual therapy was more common among men (22%) than women (18%), while mono therapy was slightly more common in women (57% vs. 53%).
- Carotid interventions were performed infrequently in both genders but showed a clear gender pattern. Carotid endarterectomy (CEA) was done in 3.7% of men versus 2.0% of women, and overall carotid revascularization in 4.4% of men versus 2.5% of women.
- Rehabilitation data had high levels of missingness (around 68–70%), limiting interpretation. Among patients with available data, men were slightly more likely to receive rehabilitation in general (12% vs. 9.9%) and inpatient rehabilitation (2.8% vs. 2.0%). Outpatient rehabilitation remained rare in both groups, at under 1%.
- Platelet inhibitor prescriptions (which were recorded in two separate variables *i8070* and *i26019*) showed small differences. Dual therapy was slightly more common in men in both variables, and women were marginally more likely to receive no platelet inhibitors or only monotherapy.

Overall, this population showed broadly similar diagnostic and treatment patterns between men and women, with subtle gender differences in echocardiographic procedures, dual antiplatelet therapy, carotid interventions and platelet inhibitors.

Table 5.22: Population 2. Diagnostic methods and secondary prophylactic treatments

Gender	**Female** N = 5,940[1]	**Male** N = 6,684[1]
Length of hospital stay		
0	306 (5.2%)	348 (5.2%)
1-2	3,948 (67%)	4,408 (66%)
14+	23 (0.4%)	20 (0.3%)
3-7	1,549 (26%)	1,789 (27%)
8-14	90 (1.5%)	93 (1.4%)
NA	24 (0.4%)	26 (0.4%)
Cranial computed tomography (CCT)		
No	868 (15%)	982 (15%)
Yes	5,072 (85%)	5,702 (85%)
Magnetic resonance imaging (MRI)		

Gender	Female N = 5,940[1]	Male N = 6,684[1]
No	4,817 (81%)	5,420 (81%)
Yes	1,123 (19%)	1,264 (19%)
Transthoracic echocardiography (TTE)		
No	1,759 (30%)	1,826 (27%)
Yes	4,181 (70%)	4,858 (73%)
Transthoracic echocardiography (TTE)		
No	1,759 (30%)	1,826 (27%)
Planned	1,848 (31%)	2,021 (30%)
Yes	2,333 (39%)	2,837 (42%)
Transesophageal echocardiography (TEE)		
No	5,508 (93%)	5,950 (89%)
Yes	432 (7.3%)	734 (11%)
Transesophageal echocardiography (TEE)		
No	5,508 (93%)	5,950 (89%)
Planned	260 (4.4%)	423 (6.3%)
Yes	172 (2.9%)	311 (4.7%)
Antiplatelet agents		
Dual	1,071 (18%)	1,454 (22%)
Mono	3,366 (57%)	3,564 (53%)
No	1,503 (25%)	1,665 (25%)
NA	0 (0%)	1 (<0.1%)
Heparin		
High (>150E)	73 (1.2%)	117 (1.8%)
Low (<75E)	3,124 (53%)	3,554 (53%)
Medium (75-150E)	350 (5.9%)	367 (5.5%)
None	2,393 (40%)	2,646 (40%)
PTA/Stent of the ACI		
No	5,902 (99%)	6,621 (99%)
Yes	38 (0.6%)	63 (0.9%)
Carotid endarterectomy (CEA)		
No	5,823 (98%)	6,438 (96%)
Yes	117 (2.0%)	246 (3.7%)
Carotid Revascularization		
No	5,793 (98%)	6,387 (96%)
Yes	147 (2.5%)	297 (4.4%)
Rehabilitation		
No	1,597 (90%)	1,897 (88%)
Yes	175 (9.9%)	259 (12%)
NA	4,168 (70%)	4,528 (68%)
Inpatient rehabilitation		

Gender	**Female** N = 5,940[1]	**Male** N = 6,684[1]
No	5,822 (98%)	6,499 (97%)
Yes	118 (2.0%)	185 (2.8%)
Outpatient rehabilitation		
No	5,909 (99%)	6,623 (99%)
Yes	31 (0.5%)	61 (0.9%)
Platelet inhibitors - i8070		
Dual	101 (1.7%)	166 (2.5%)
Mono	1,835 (31%)	2,339 (35%)
No	4,004 (67%)	4,179 (63%)
Platelet inhibitors - i26019		
Dual	76 (4.2%)	135 (6.1%)
Mono	1,270 (70%)	1,504 (68%)
No	474 (26%)	575 (26%)
NA	4,120 (69%)	4,470 (67%)
Oral Anticoagulants (OAC) at Follow-Up		
No	1,441 (81%)	1,743 (80%)
Yes	348 (19%)	443 (20%)
NA	4,151 (70%)	4,498 (67%)

[1] n (%)

Unadjusted Likelihood of Diagnostics and Treatments by Gender

Table 5.23 shows the results of the unadjusted logistic regression models comparing men's and women's likelihood of receiving diagnostic procedures and secondary prophylactic treatments after a TIA.

Women had lower odds of undergoing echocardiographic procedures. For transthoracic echocardiography (TTE), they were between 11% and 15% less likely to receive the procedure. The difference was even more pronounced for transesophageal echocardiography (TEE), where women had 36% lower odds of receiving the procedure overall. Specifically, they were also 34% less likely to have TEE planned and 40% less likely to have it actually performed than men.

Overall antiplatelet agent use did not differ significantly, but women had 18% lower odds of receiving dual antiplatelet therapy. Similarly, for treatment with heparin, only high-dose heparin showed a difference, with women less likely to receive high doses (OR = 0.69).

Carotid interventions were more commonly performed on men. Women had 47% lower odds of receiving a carotid endarterectomy (CEA), and 45% lower odds of undergoing

either carotid intervention, as measured by the combined carotid revascularization variable. PTA itself narrowly missed conventional significance thresholds, with an odds ratio of 0.68, a 95% CI of [0.45–1.01] and a p-value of 0.058.

Variable	OR [95% CI]	P-Value	Interpretation
TTE	0.89 [0.83-0.96]	0.004	Women had lower odds.
TTE = Yes	0.85 [0.78-0.93]	< 0.001	Women had lower odds.
TEE	0.64 [0.56-0.72]	< 0.001	Women had lower odds.
TEE = Planned	0.66 [0.57-0.78]	< 0.001	Women had lower odds.
TEE = Yes	0.6 [0.49-0.72]	< 0.001	Women had lower odds.
Antiplatelet agents = Dual	0.82 [0.73-0.91]	< 0.001	Women had lower odds.
Heparin = High	0.69 [0.51-0.93]	0.014	Women had lower odds.
CEA	0.53 [0.42-0.66]	< 0.001	Women had lower odds.
Carotid Revascularization	0.55 [0.44-0.67]	< 0.001	Women had lower odds.
Rehabilitation	0.8 [0.65-0.98]	0.034	Women had lower odds.
Inpatient rehabilitation	0.71 [0.56-0.9]	0.004	Women had lower odds.
Outpatient rehabilitation	0.57 [0.36-0.87]	0.011	Women had lower odds.
Platelet inhibitors - i8070 = Mono	0.82 [0.76-0.88]	< 0.001	Women had lower odds.
Platelet inhibitors - i8070 = Dual	0.64 [0.49-0.82]	< 0.001	Women had lower odds.
Platelet inhibitors - i26019 = Dual	0.68 [0.5-0.93]	0.015	Women had lower odds.

Table 5.23: Population 2 - Unadjusted treatment odds ratios by gender (Significant Results Only)

Gender differences were also observed in rehabilitation. Women had 20% lower odds of receiving any rehabilitation, 29% lower odds for inpatient rehabilitation, and 43% lower odds for outpatient rehabilitation.

For platelet inhibitors, both available variables reflected similar trends. Women had significantly lower odds of receiving either monotherapy (18% less likely) or dual therapy (36% less likely) in the *i8070* variable. In the *i26019* version women were 32% less likely to receive dual therapy.

Adjusted Analysis: Controlling for Age

Table 5.24 presents the logistic regression results adjusted for age, allowing for evaluation of whether gender differences in treatment and diagnostics remain after accounting for this important demographic factor.

While the overall gender gap in transthoracic echocardiography (TTE) was no longer statistically significant after adjusting for age, the procedure being performed — and not just planned — was still less likely for women (OR = 0.90). For transesophageal echocardiography (TEE), gender differences persisted with the effect size slightly reduced. Women had 23% lower odds overall, were 20% less likely to have the procedure planned, and 27% less likely to have it performed.

Variable	OR [95% CI]	P-Value	Interpretation
TTE = Yes	0.9 [0.83-0.98]	0.018	Women had lower odds.
TEE	0.77 [0.67-0.87]	< 0.001	Women had lower odds.
TEE = Planned	0.8 [0.68-0.94]	0.006	Women had lower odds.
TEE = Yes	0.73 [0.6-0.88]	0.001	Women had lower odds.
Antiplatelet agents = Mono	1.24 [1.14-1.35]	< 0.001	Women had higher odds.
Heparin = High	0.7 [0.52-0.94]	0.020	Women had lower odds.
CEA	0.51 [0.41-0.64]	< 0.001	Women had lower odds.
Carotid Revascularization	0.53 [0.43-0.65]	< 0.001	Women had lower odds.
Inpatient rehabilitation	0.74 [0.58-0.93]	0.011	Women had lower odds.
Outpatient rehabilitation	0.57 [0.37-0.88]	0.013	Women had lower odds.
Platelet inhibitors - i8070 = Mono	0.74 [0.68-0.79]	< 0.001	Women had lower odds.
Platelet inhibitors - i8070 = Dual	0.63 [0.49-0.81]	< 0.001	Women had lower odds.
OAC at Follow-Up	0.8 [0.68-0.94]	0.008	Women had lower odds.

Table 5.24: Population 2 - Treatment odds ratios by gender controlled for Age (Significant Results Only)

With the age adjustment, the gender difference in antiplatelet treatment notably changed. Women were now 24% more likely to receive mono therapy, while the previous difference in dual therapy was no longer statistically significant.

For heparin treatments, the effect stayed consistent with the unadjusted model. Women had 30% lower odds of receiving high doses of heparin.

Gender differences in carotid interventions also remained. The odds of women receiving CEA (OR = 0.51) or any carotid revascularization (OR = 0.53) were still significantly lower than for men. PTA once again narrowly missed the conventional significance threshold (OR = 0.67, CI: [0.44–1.00]).

For rehabilitation the effect sizes shifted slightly compared to the unadjusted model. The odds for women receiving inpatient rehabilitation decreased from 29% to 26%, while the 43% lower odds for outpatient rehabilitation remained unchanged. However, the difference for general rehabilitation was no longer statistically significant.

Platelet inhibitor prescriptions continued to show significant gender differences. In the *i8070* variable, the gap for mono therapy became more pronounced, with women now 26% less likely to receive it instead of 18%. Women also remained 37% less likely to receive dual therapy. However, in the *i26019* variable, no significant differences were present any more.

An additional gender difference was also observed, with women less likely to receive oral anticoagulants (OAC) at Follow-Up than men.

Adjusted Analysis: Controlling for Patient Characteristics and Risk Factors

In the following models, additional patient characteristics and risk factors were included as covariates to better isolate the independent effect of gender.

Variable	OR [95% CI]	P-Value	Interpretation
TTE = Yes	0.9 [0.82-0.98]	0.015	Women had lower odds.
TEE	0.74 [0.65-0.85]	< 0.001	Women had lower odds.
TEE = Planned	0.77 [0.65-0.91]	0.002	Women had lower odds.
TEE = Yes	0.71 [0.58-0.86]	< 0.001	Women had lower odds.
Antiplatelet agents = Dual	0.83 [0.73-0.94]	0.004	Women had lower odds.
Heparin = High	0.71 [0.52-0.96]	0.029	Women had lower odds.
CEA	0.53 [0.42-0.66]	< 0.001	Women had lower odds.
Carotid Revascularization	0.56 [0.45-0.69]	< 0.001	Women had lower odds.
Inpatient rehabilitation	0.78 [0.61-0.98]	0.038	Women had lower odds.
Outpatient rehabilitation	0.57 [0.36-0.87]	0.011	Women had lower odds.
Platelet inhibitors - i8070 = Mono	0.76 [0.69-0.82]	< 0.001	Women had lower odds.
Platelet inhibitors - i8070 = Dual	0.72 [0.55-0.94]	0.014	Women had lower odds.

Table 5.25: Population 2 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors (Significant Results Only)

As shown in Table 5.25, most significant gender differences observed in earlier models remained stable even after controlling for clinical risk factors.

This includes the consistently lower odds for women receiving transesophageal echocardiography (TEE) — both planned and performed — with effect sizes similar to those in the age-adjusted model. For transthoracic echocardiography (TTE), the gender difference for the performed procedure also persisted.

The pattern of lower use of dual antiplatelet therapy among women returned in this model after having lost significance in the age-only adjusted model. The effect size remained in the same range as previously seen, with women having 17% lower odds than men. Mono antiplatelet therapy, however, no longer differed significantly between genders.

Heparin treatment also continued to show lower odds for women at high dosages (OR = 0.71), while differences for medium and low doses remained non-significant.

As in earlier models, women continued to have significantly lower odds of receiving a carotid endarterectomy (CEA) or any carotid revascularization. PTA remained non-significant, consistent with earlier models.

The gender differences for patients receiving outpatient and inpatient rehabilitation have also remained consistent, with women having lower odds for both.

Prescription patterns of platelet inhibitors remained largely consistent with earlier models. For the *i8070* variable, women had previously been 26% less likely to receive mono therapy,

which remained nearly unchanged in this model at 24% lower odds. Though for dual therapy the gender gap narrowed slightly from 37% lower odds to only 28%.

The previously observed difference in OAC prescription at Follow-Up was no longer significant in this model.

Variable	OR [95% CI]	P-Value	Interpretation
TTE = Yes	0.89 [0.82-0.98]	0.012	Women had lower odds.
TEE	0.75 [0.65-0.85]	< 0.001	Women had lower odds.
TEE = Planned	0.77 [0.65-0.91]	0.002	Women had lower odds.
TEE = Yes	0.72 [0.59-0.88]	0.001	Women had lower odds.
Antiplatelet agents = Dual	0.83 [0.73-0.94]	0.004	Women had lower odds.
Heparin = High	0.7 [0.52-0.95]	0.024	Women had lower odds.
CEA	0.53 [0.42-0.67]	< 0.001	Women had lower odds.
Carotid Revascularization	0.56 [0.46-0.69]	< 0.001	Women had lower odds.
Inpatient rehabilitation	0.77 [0.61-0.98]	0.037	Women had lower odds.
Outpatient rehabilitation	0.56 [0.36-0.86]	0.010	Women had lower odds.
Platelet inhibitors - i8070 = Mono	0.76 [0.7-0.83]	< 0.001	Women had lower odds.
Platelet inhibitors - i8070 = Dual	0.72 [0.56-0.94]	0.017	Women had lower odds.

Table 5.26: Population 2 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors with Unknown included (Significant Results Only)

Including unknown values as a separate category in the covariates (Table 5.26) did not change the pattern of results. All previously significant differences persisted with very similar odds ratios and confidence intervals.

Women continued to have lower odds of receiving echocardiographic procedures (TTE and TEE), dual antiplatelet therapy, and high-dose heparin. Carotid revascularization rates also remained significantly lower among women, as did their likelihood of receiving inpatient and outpatient rehabilitation. Platelet inhibitor differences were again confirmed in the *i8070* variable, while no new differences emerged.

The final model (Table 5.27), which included imputed values for smoking and alcohol abuse, again confirmed the earlier findings. The gender differences remained consistent overall, with slight changes in effect sizes.

The outpatient rehabilitation difference persisted, with women having 44% lower odds of receiving this form of care. However, inpatient rehabilitation was no longer statistically significant, with the CI now just overlapping 1 (CI: [0.62–1.00], p = 0.055).

All other previously significant differences — including lower odds for echocardiographic procedures, high-dose heparin, carotid interventions, and platelet inhibitor prescriptions in the *i8070* variable — remained stable in direction and significance.

Variable	OR [95% CI]	P-Value	Interpretation
TTE = Yes	0.91 [0.83-0.99]	0.032	Women had lower odds.
TEE	0.75 [0.66-0.86]	< 0.001	Women had lower odds.
TEE = Planned	0.77 [0.65-0.9]	0.002	Women had lower odds.
TEE = Yes	0.73 [0.6-0.89]	0.002	Women had lower odds.
Antiplatelet agents = Dual	0.84 [0.74-0.95]	0.007	Women had lower odds.
Heparin = High	0.7 [0.51-0.95]	0.021	Women had lower odds.
CEA	0.53 [0.42-0.66]	< 0.001	Women had lower odds.
Carotid Revascularization	0.57 [0.46-0.7]	< 0.001	Women had lower odds.
Outpatient rehabilitation	0.56 [0.35-0.86]	0.009	Women had lower odds.
Platelet inhibitors - i8070 = Mono	0.76 [0.7-0.83]	< 0.001	Women had lower odds.
Platelet inhibitors - i8070 = Dual	0.71 [0.55-0.93]	0.012	Women had lower odds.

Table 5.27: Population 2 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors with Smoking and Alcohol Abuse imputed (Significant Results Only)

Treatment Likelihood for Specific Conditions

This section investigates whether gender differences existed in the likelihood of receiving treatment for common comorbidities following a TIA. As in the previous analysis for Population 1, the focus is on four clinical conditions frequently managed as part of secondary prevention: atrial fibrillation, hypercholesterolemia, hypertension (with blood pressure monitoring and antihypertensives as treatments), and diabetes mellitus.

Treatments	**Gender**	**No**	**Yes**	**Treatment Rate**
Oral Anticoagulants (OAC)	Male	89	340	79.25%
for Atrial fibrillation	Female	86	270	75.84%
Regular lipid-lowering drugs	Male	200	1379	87.33%
for Hypercholesterolemia	Female	227	1043	82.13%
Regular blood pressure checks	Male	200	1418	87.64%
for Hypertension	Female	169	1219	87.82%
Regular antihypertensives	Male	264	1377	83.91%
for Hypertension	Female	287	1108	79.43%
Regular antidiabetica	Male	149	380	71.83%
for Diabetes mellitus	Female	112	250	69.06%

Table 5.28: Population 2 - Treatment Rates per gender

Table 5.28 summarises the overall treatment rates by gender. In contrast to Population 1, the differences here are more pronounced in some areas. Notably, treatment with oral anticoagulants (OAC) can be meaningfully analysed in this population, as not all patients with atrial fibrillation received the medication. Women were slightly less likely than men to be prescribed OACs. The same was true for hypercholesterolemia, where 5%

fewer women received lipid-lowering drugs. For hypertension, blood pressure checks were nearly equally common across genders, but a prescription of antihypertensives was again received more often by men. Antidiabetic medication had only a small gender difference, but again in favour of men.

The prevalence of chronic conditions is strongly influenced by age and must be considered when examining gender-based treatment patterns. Figure 5.14 presents the distribution of patients by age group and gender for each of the four conditions.

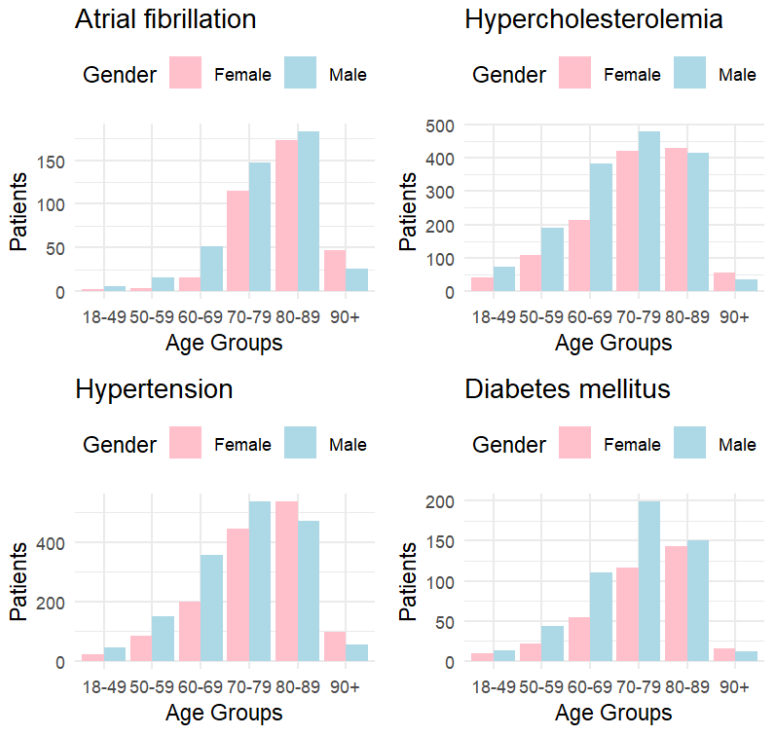


Figure 5.14: Distribution of conditions by age group and gender - Population 2

While all conditions became more common with increasing age of the patients, the exact patterns varied slightly across conditions and gender. Atrial fibrillation was most common in the 80–89 group for both men and women, though men had higher counts overall. Up until peaking in the 70–79 range, hypercholesterolemia was more common for men than women. Though for patients ages 80-89, women were only slightly more affected. Hypertension followed a similar but slightly more balanced distribution. Men aged 70-79 and women aged 80-89 are the largest groups and make up about the same amount of patients. Diabetes mellitus affected men consistently more than women, especially in the 70–79 group, where the difference was most pronounced.

The treatment rates across age groups are shown in Figure 5.15, which displays mosaic plots for each of the five treatments. Unlike in Population 1, oral anticoagulants (OAC)

are included here, as the treatment was not universally prescribed and thus allows for meaningful comparisons. Patient counts and exact percentages by gender and age group can be found in Appendix 7.

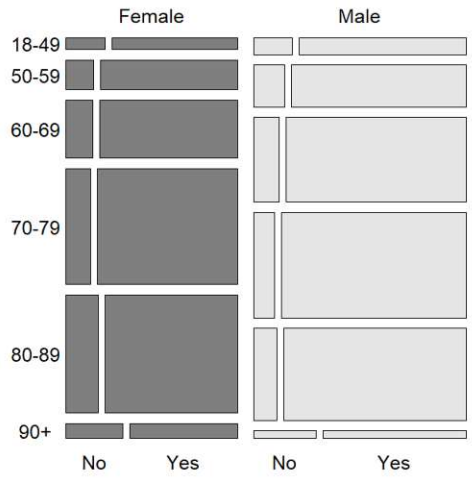
The first mosaic plot 5.15a shows lipid-lowering drug treatment among patients with hypercholesterolemia. While overall treatment rates were high for both genders, men consistently had higher rates in every age group. In the 80–89 age group, 88.46% of men and only 80.74% of women received treatment — a gap of nearly 8%. Across other age groups the differences were not as extreme but still notable, with all groups except 50–59 having a gender gap of more than 3.5%.

For regular blood pressure checks among patients with hypertension, the mosaic plot 5.15b shows high treatment rates across all age groups and both genders. In most age groups, men and women received monitoring at nearly equal rates, with differences typically below 3 percentage points and no consistent gender trend. Men were slightly more likely to receive treatment in some groups, women in others, and for the age group 80-89, the treatment rate was almost identical. The only substantial difference was for the youngest patients (18–49), where 86.96% of women received regular checks compared to just 67.39% of men — a gap of nearly 20%. This age group is relatively small, but the magnitude of the difference still stands out.

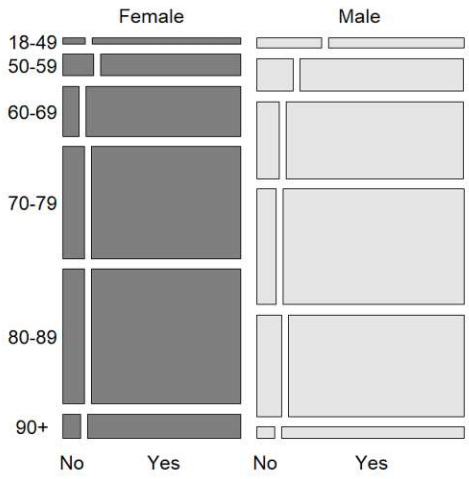
The mosaic plot 5.15c displays treatment rates for antihypertensive medication among patients with hypertension. Men were more likely to receive treatment in almost every age group, though the size of the gap varied. In the age ranges 60-69 and 70-79, treatment rates were nearly identical, with differences of less than 2 percentage points between men and women. The only group where women were slightly more likely to receive antihypertensives was 50–59, though the difference was minimal (84.71% vs. 82.17%).

For the youngest patients, the gap was far wider, with more than a 9% difference. For the 90+ age group, the disparity went up to 12%, with just 56.52% of women treated, compared to 67.27% of men. Similarly, in the 80–89 group, 77.55% of women received antihypertensives, while 84.53% of men did — a gap of 7 percentage points. This shows that gender differences in treatment were most pronounced among the oldest and youngest patients.

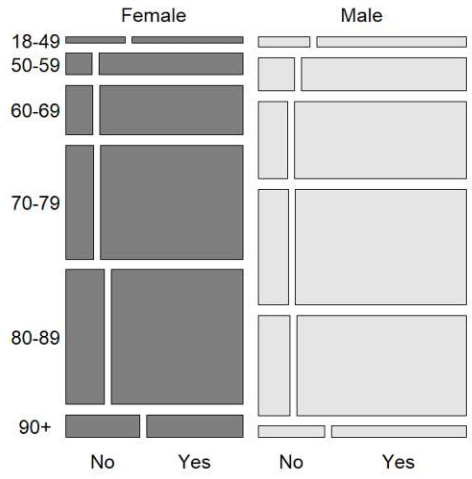
Antidiabetic medication use among patients with diabetes mellitus (Figure 5.15d) shows a more irregular pattern compared to the previous treatments. Although treatment coverage was solid for most age groups, there was no consistent gender difference - neither in direction or size. In the 70–79 age group, men had a noticeably higher treatment rate (81.91%) than women (65.81%), resulting in a gap of over 16%. However, in both the 80–89 and 90+ age groups, this pattern reversed, meaning women were more likely to receive treatment. In the 90+ group, the gap was even larger at almost 23%, with 81.25% of women treated versus just 58.33% of men. However, this group had relatively few patients overall, making these treatment rates unstable. For patients under 70, treatment rates were more balanced, with no consistent gender trend and differences staying below 7%.



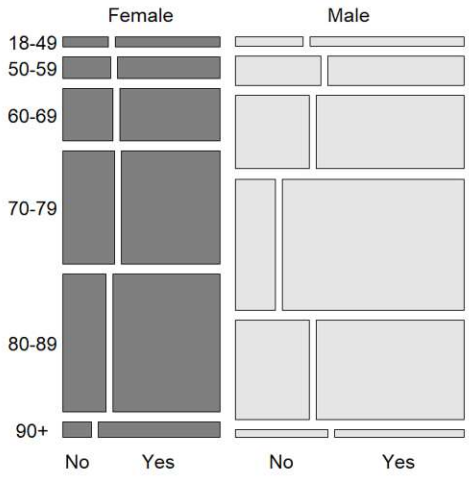
(a) Regular lipid-lowering drugs



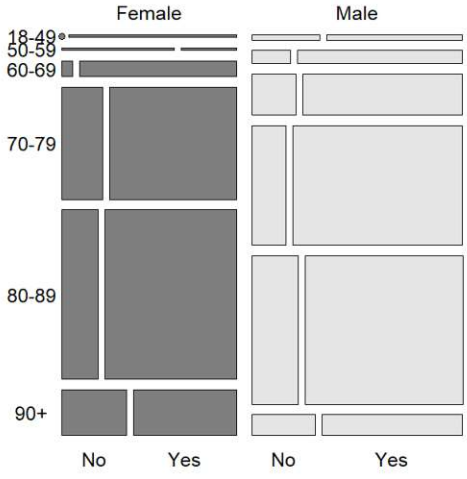
(b) Regular blood pressure checks



(c) Regular antihypertensives



(d) Regular antidiabetica



(e) Oral Anticoagulants (OAC)

Figure 5.15: Population 2 - Treatment Mosaic Plots

The final mosaic plot 5.15e displays oral anticoagulant (OAC) treatment among patients with atrial fibrillation. The youngest two age groups (18–49 and 50–59) contain only 5 women combined, making it difficult to draw any meaningful conclusions about gender differences there. The 60-69 age group shows a clear disparity, with 93.75% of women receiving OACs compared to only 78.43% of men - a gap of over 15%. However, this trend reversed in the next age group, with women 8% less likely to receive treatment than men. In the 80–89 group, this gap disappears almost completely and turns into men having just about 1% worse odds than women. Among the oldest patients, women were again less likely to receive OACs, with a treatment rate of 61.70% compared to 69.23% in men — a difference of 6.5 percentage points.

Summary

The analysis of gender differences in diagnostics and secondary treatments in Population 2 revealed a broader and more persistent pattern than in Population 1. Several gender differences remained significant even after adjusting for age and clinical risk factors.

- **Women were consistently less likely to receive echocardiographic procedures, especially TEE.** While upon adjustment the gap for transthoracic echocardiography (TTE) clearly narrowed, transesophageal echocardiography — both planned and performed — remained far more common for men.
- **Lower odds of carotid interventions among women were observed.** As in Population 1, women were significantly less likely to receive carotid endarterectomy (CEA), and this difference remained robust across all model adjustments.
- **Gender differences in antiplatelet treatment were somewhat variable.** Women were less likely to receive dual antiplatelet therapy across most models, except when only adjusting for age. Instead, mono antiplatelet therapy was then significantly more common for men.
- **Women were less likely to receive high-dose heparin treatments.** However, the odds were about equal for both genders receiving either low or medium doses instead of no heparin treatment at all.
- **Women were less likely to receive rehabilitation.** Across the models differences for inpatient and outpatient rehabilitation could be observed, with women always having lower odds. However, the significance for inpatient rehabilitation was lost in the final imputed model.
- **Women had lower odds of receiving platelet inhibitors.** This was most consistently observed in the *i8070* variable, where women were less likely to receive either mono or dual therapy.

The analysis of treatment rates for specific conditions highlighted further differences. Compared to Population 1, more pronounced gender disparities were observed. Women were consistently less likely to receive lipid-lowering drugs across nearly all age groups. For antihypertensive and antidiabetic medication, differences were present, but they were not consistent in favouring one gender. Still, the oldest and youngest age groups were clearly most affected. In contrast, regular blood pressure checks showed only minor variations, with the one notable being that among the youngest patients, men were far less likely to be monitored. Oral anticoagulant (OAC) use had also no consistent pattern, with women having higher treatment rates in some age groups and lower in others.

Overall, while Population 1 showed only a few consistent gender gaps, the findings in Population 2 suggest a broader pattern of gender-based differences in secondary care.

CHAPTER 6

Discussion

In this chapter, the results for the three research questions are discussed in the context of existing clinical knowledge and relevant literature. Each research question is covered separately, with a summary and interpretation of the findings. The consistency of results across the two populations is examined, and relevant limitations, such as small sample sizes, are considered.

6.1 Research Question 1: Gender Differences in Health Outcomes

In the analysis of health outcomes after a transient ischaemic attack (TIA), two main gender-specific differences were observed. In both populations women had worse functional recoveries, and especially in Population 1, men had higher odds of suffering from pneumonia.

Across both populations, women were consistently more likely to have mRS at Follow-Up scores larger than 1, therefore suffering from at least some form of disability. In Population 2 (recent main stroke registry) this effect was more pronounced, with women having 1.54 times higher odds than men, compared to Population 1 (TIA sub-registry), where the odds ratio was 1.32. (ORs under full adjustment with imputation)

The gender difference in functional recovery was observed throughout all models and covariate adjustments. Although adjusting for age slightly decreased the effect size compared to the unadjusted model, it remained statistically significant. This suggests that while age affected a patient's mRS score it did not fully explain the disparity between the genders. The effect persisted after further adjustment for patient characteristics and risk factors such as pre-existing conditions, smoking and alcohol abuse. The severity of symptoms at admission was also accounted for using the NIH Stroke Scale, still women remained at higher risk of poorer functional recovery following a TIA.

Which specific mRS at Follow-Up scores were more likely for women differed slightly between the two populations. In Population 1, higher odds were found for moderate and severe disabilities (mRS = 3 and mRS = 5), while in Population 2, women were more likely to experience slight to moderately severe disabilities (mRS = 2, 3, and 4). This is consistent with the stronger overall effect for mRS at Follow-Up > 1 observed in Population 2.

In the literature, results for gender differences in functional recovery after stroke events vary. For acute ischaemic strokes, Bonkhoff et al. (2021) [Bon+21] found that while women presented with worse functional status at admission, they had either equal or better recovery outcomes than men. Synhaeve et al. (2016) [Syn+16] found that for patients between 18 and 50 years old, women had worse very long-term outcomes after an ischaemic stroke. While for general acute strokes, Kim et al. (2010) [Kim+10] reported that both at three months and one year after the event, women had worse functional outcomes. All these studies dealt with acute strokes rather than transient ones, making their results not directly transferable since patients presented with milder symptoms at admission, with no gender difference.

Though interestingly Kim et al. (2010) [Kim+10] also found no gender difference in mortality, which aligns with results of Purroy et al. (2010) [Pur+21] in a TIA population and matches this thesis' findings that mRS scores of 6 (death) were equally common between genders.

These studies highlight that results on gender differences in recovery depend on the population and type of cerebrovascular event studied. Still, the worse functional outcomes observed for women after TIAs are consistent with previous findings in more severe stroke populations.

For the occurrence of pneumonia after a TIA, a potential gender difference was observed. Men were consistently more likely to suffer pneumonia across all models in Population 1. When adjusting for age, a similar trend was present in Population 2, but it did not remain statistically significant after accounting for the patient characteristics and risk factors.

Across both populations, pneumonia was a rare complication, with only 0.9% of patients affected in the first and 0.4% in the second. The low incidence rate affects the statistical significance of any findings and limits the ability to reliably identify gender differences, especially in Population 2.

Previous research has indicated that men are often more vulnerable than women to respiratory infections and inflammatory lung diseases, possibly as a result of differences in the immune response [Cha+17; Yan+01]. Therefore, the observed trend of more men catching pneumonia may be explainable, even though the gender difference can not be consistently confirmed in the data.

6.2 Research Question 2: Predictive Accuracy of the ABCD2 and ABCD3-I Scores

For patients after a transient ischaemic attack (TIA), no substantial gender-specific differences were observed in the predictive accuracy of the ABCD2 and ABCD3-I scores. Across both risk scores and all analysed recurrence outcomes, the discrimination ability was broadly similar between men and women. Although small differences occurred for some outcomes and under certain covariate adjustments, they were generally minor, and there was no consistent trend.

Both the *database* scores, of which the ABCD2 was directly entered by care providers during clinical assessment, and the *calculated* scores, constructed algorithmically from registry data using the score definitions, were evaluated. For the *calculated* scores, the recurrence events were shifted towards higher score values, as would be expected for a risk assessment tool. In contrast, recurrence events in the *database* scores were more evenly spread across score values. This suggests a certain degree of variability in clinical score assignment, though the reasons for this discrepancy are unclear. The better alignment of the *calculated* scores with the expected risk stratification supports the validity of their higher predictive accuracy across all recurrence outcomes and adjustment levels. However, with the inclusion of patient characteristics and risk factors, the performance gap narrowed considerably. This is to be expected, as the detailed information of a patient's risk profile is likely to capture most relevant aspects for the recurrence prediction.

The predictive accuracies observed in the fully adjusted case can also provide additional context for interpretation. The AUC values did not increase indiscriminately but rather hit a ceiling of predictive performance. For the more accurate *calculated* scores, this was on average 0.72 across both ABCD2 and ABCD3-I, respectively. Some outcomes did achieve better results, but these were outliers and not representative of the majority of the recurrence event data. The average base, unadjusted AUC values were 0.61 for ABCD2 and 0.64 for ABCD3-I, which, based on the standard definition, puts them in the poor discrimination range. But in the context of the best results, which could be reached using the complete risk profile, these values are fairly decent, considering the limited number of clinical variables used for the scores.

These results also align well with the findings of Knoflach et al. (2016) [Kno+16], who evaluated the predictive accuracy of the risk scores in the same Austrian stroke unit population. They reported for the ABCD3-I score AUC values of 0.664 (95% CI: 0.618–0.709) for early stroke recurrence and 0.646 (95% CI: 0.592–0.700) for 3-month stroke recurrence. These values are directly comparable to those obtained for the *calculated* scores in this thesis. This supports the robustness of the findings since the same population and score definitions were used. It also suggests that the limited predictive ability of the scores is a feature of the clinical context rather than the study design.

The predictive performance was in most cases the same or very similar for both genders.

In the base analyses, while differences were observed, they were inconsistent in size and direction, with sometimes women and sometimes men having higher AUC values. While women had slightly better predictive performances on average, the difference was only +0.0158 for the ABCD2 score and +0.0013 for the ABCD3-I score. These are not substantial and are unlikely to affect the risk prediction in practice.

Larger AUC gaps were observed for Early worsening and mRS at Follow-Up > 4 in both scores, with differences around +0.048. However, these outcomes had the smallest sample sizes among all recurrence definitions, with only 47 women versus 78 men affected for Early worsening, and 51 women versus 53 men for mRS. Due to this and the fact that these differences were not present across other recurrence outcomes, the robustness and generalisability of these findings are limited.

To investigate whether patient characteristics and risk factors might differently influence predictive performance across genders, the scores were also adjusted. Including age benefitted male patients more, particularly in outcomes dependent on the modified Rankin Scale, such as Recurrence 1–3. This effect likely depends on the underlying distribution difference of age in this data, since men tended to be younger. Although age is already included as a binary threshold (>60 years) in both scores, this did not capture the difference between the genders. This suggests that introducing additional age thresholds could potentially enhance risk stratification for the ABCD2 and ABCD3-I scores.

After full adjustment for clinical risk factors and patient characteristics, AUC values were fairly similar between the genders. In a few cases, women achieved slightly higher discrimination classifications, such as reaching good discrimination for Recurrence 1 and Recurrence within 90 days, while men remained at fair performance. However, these differences were dependent on the adjustment and not consistent across outcomes.

In the evaluation of the ABCD2 and ABCD3-I scores, their predictive performance was found to be mostly comparable between men and women after a TIA. While small gender differences were observed for individual outcomes, these were either minor, inconsistent or not robust. However, given the relatively small sample size for recurrence events, further research with larger cohorts would be valuable to confirm the observed gender comparability.

6.3 Research Question 3: Gender Differences in Diagnostics and Treatment

Across both populations, multiple gender-specific differences in the diagnostic procedures and secondary prophylactic treatments provided after a transient ischaemic attack (TIA) were observed. Some of these were consistent across populations, but most were limited to the more recent registry data.

In Population 1 the gender differences in diagnostics and secondary treatments were relatively limited after adjusting for age and clinical risk factors. Still, two consistent

patterns remained, with women being less likely to receive clopidogrel, an antiplatelet medicine, and less likely to undergo carotid revascularisation procedures.

In contrast, Population 2 showed broader and more persistent gender disparities. Even after full adjustment, women were less likely to undergo echocardiographic procedures (particularly transesophageal echocardiography), less likely to receive dual antiplatelet therapy, high-dose heparin, carotid interventions, and rehabilitation services, even after full adjustment. This suggests that between these two populations some gender-based differences in care may have become more pronounced or more noticeable over time.

In the following, each of the affected procedures and treatments is discussed in more detail, alongside relevant supporting findings from the existing literature.

Women in both populations were almost half as likely as men to undergo carotid revascularisation procedures, particularly carotid endarterectomy (CEA). These differences were consistent and remained even after adjusting for patient characteristics and risk factors. Ramkumar et al. (2022) [Ram+22] reported similar gender discrepancies, finding that women were less likely to undergo CEA or carotid artery stenting. However, following these procedures, a higher five-year risk of stroke for women was also identified. This is supported by Kremer et al. (2023) [Kre+23], who found that women treated with CEA had a significantly higher rate of stroke or death at four months. It is possible that concerns about increased clinical risk may have contributed to the lower treatment rates among women, but an underlying bias can't be ruled out.

In Population 1, women were less commonly prescribed clopidogrel, while in Population 2, they were less likely to receive dual antiplatelet therapy after a TIA. A general differences in antiplatelet use is supported by Arrich et al. (2008) [Arr+08], who found within 48 hours of being admitted for a stroke women had a much lower likelihood of receiving the drugs. While specifically for dual antiplatelet medication after a TIA or mild stroke Solomonow et al. (2023) [Sol+23] identified higher prescription rates for men. This aligns with the findings from Population 2, supporting the observation of a gender bias for intensive antiplatelet therapy.

Echocardiographic procedures were also not equally used in Population 2, specifically transesophageal echocardiography (TEE) and transthoracic echocardiography (TEE). For women the odds of undergoing TEE were consistently and substantially lower, while for TTE the difference was less pronounced after covariate inclusion, but it was still notable. In the literature, the treatment rates for echocardiographic investigations after an ischaemic stroke have been examined. Gall et al. (2010) [Gal+10] reported that women had lower rates, although adjusting for age and stroke severity diminished the effect. A smaller but consistent gender difference was identified by Leslie-Mazwi et al. (2007) [Les+07] for the use of either TTE or TTE. The lower treatment rates for women found in this thesis, which for TEE persisted after age adjustment, are therefore in line with previous observations.

For women, rehabilitation was less common in Population 2, with outpatient rehabilitation remaining consistently far less likely. In contrast, for inpatient rehabilitation the difference

disappeared with the inclusion of the imputed smoking and alcohol abuse data. In Ott et al. (2022) [Ott+22], the effect of gender on rehabilitation for different conditions was discussed. They highlighted that for women, worse access, use and outcomes were observed and that these were influenced by several social and systemic factors. It is possible that these factors also affected the rehabilitation rates after a TIA.

High doses of heparin as part of secondary prevention were prescribed less to women in Population 2. Roosendaal et al. (2022) [Roo+22] showed that after standard heparin dosage during non-cardiac arterial procedures, women had higher anticoagulation levels and experienced more bleeding complications. Due to these different responses to heparin in women, a more cautious approach in dosing may have been chosen, leading to the observed gender disparity.

Additionally, treatment rates for a few secondary prevention conditions were analysed for possible gender differences while taking into account patients' ages. This included the prescription of drugs and monitoring for atrial fibrillation, hypertension, hypercholesterolemia and diabetes mellitus.

In Population 1 treatment rates were generally high for both genders, with no major disparities. Any observed differences were small and varied by age group, with no gender being favoured consistently. However, slightly higher treatment rates for women were usually observed over the age of 80.

The more recent data from Population 2 showed more pronounced gender differences. Women were less likely to receive lipid-lowering therapy and antihypertensives in several age groups. Although differences in antidiabetic medication use varied and lacked a clear gender trend, the youngest and oldest patients were most affected. Similar shifting gender differences were observed for oral anticoagulants (OAC), with younger women more likely to receive OACs but also older men having higher prescription rates.

Across most groups, the observed differences were relatively small, and some were based on only a limited number of patients. Therefore, any interpretation needs to be done with caution, and factors such as clinical considerations and patient preferences need to be taken into account.

CHAPTER 7

Conclusion

This thesis investigated gender differences in health outcomes, risk stratification performance, diagnostic methods and secondary prophylactic treatments following transient ischaemic attacks (TIA) using clinical registry data from Austrian stroke units.

Women were found more likely to experience worse functional recovery three months after a TIA, even when accounting for confounding factors, like age and stroke risk profiles. However, no gender difference was observed for mortality, which was consistent with previous findings in acute stroke populations.

The predictive accuracy of the ABCD2 and ABCD3-I risk scores was broadly comparable between men and women. While for certain outcomes small differences in the discrimination ability were observed, these were found to be inconsistent, minor, and not robust, and therefore, unlikely to provide meaningful insights. The comparison between the risk scores entered by care providers and those calculated from the registry data showed that the calculated versions achieved better discrimination ability across recurrence outcomes. This suggests that the formulas used for the score calculations provided a systematic improvement to the predictive accuracy of the risk stratification tools.

In the earlier TIA sub-registry, gender-specific differences in diagnostic procedures and secondary prophylactic treatments were relatively limited. Carotid interventions were less commonly used on women, and they were also less likely to be prescribed clopidogrel as part of an antiplatelet therapy. In comparison, the more recent registry data showed additional and more persistent gender disparities. Women were less likely to undergo carotid interventions and echocardiographic procedures. They were also less likely to receive dual antiplatelet therapy, high-dose heparin and rehabilitation services.

Clinical factors reported in the literature may explain and support some of these differences. For instance, different reactions to heparin and increased risks and recurrence rates in women following carotid procedures have been reported. However, the consistency of the observed differences across several treatments could indicate a more systemic cause.

Unconscious gender biases or simply disparities in access to care may contribute to unequal treatment of patients after a TIA.

Due to the use of two distinct patient populations – one with detailed TIA-related variables and another covering more recent care practices – a detailed analysis of gender differences over time was possible. Since the data is from a real-world registry, it reflects routine clinical practices, making the findings relevant and comparable to the Austrian healthcare context.

However, the data only includes information from the routine patient care and does not cover the clinical decision-making processes or patient details, such as the socio-economic background. The observational nature of registries also means that no causal links can be established in the analyses. Additionally, small sample sizes for some treatments and recurrence outcomes, especially after stratification by gender, limited the detection and evaluation of subtle differences. Missing data, particularly for follow-up data, also affected some analyses. These limitations raise the need for cautious interpretation while also highlighting areas where future research with more extensive data would be useful.

Appendix

Patient Characteristics and Risk Factors

Population 1 — Used for RQ1 and RQ3

Table 1: Population 1. Patient Characteristics and Risk Factors

Gender	**Female** N = 6,884[1]	**Male** N = 8,583[1]
Age in Years	73 (63, 81)	75 (64, 82)
NIHSS at admission	1 (0, 2)	1 (0, 2)
NA	2 (<0.1%)	2 (<0.1%)
Systolic blood pressure	158 (140, 180)	154 (140, 171)
NA	4,515 (66%)	5,621 (65%)
Aetiology		
cardiogen embolic	1,497 (22%)	1,796 (21%)
else	116 (1.7%)	200 (2.3%)
Macroangiopathy	543 (7.9%)	1,104 (13%)
Microangiopathy	2,271 (33%)	2,523 (29%)
Unknown	2,457 (36%)	2,960 (34%)
Hypertension		
No	1,365 (20%)	1,824 (21%)
Yes	5,519 (80%)	6,759 (79%)
Hypertension (with Unknown		
No	1,299 (19%)	1,700 (20%)
Unknown	66 (1.0%)	124 (1.4%)
Yes	5,519 (80%)	6,759 (79%)
Previous stroke		
No	5,454 (79%)	6,617 (77%)
Yes	1,430 (21%)	1,966 (23%)
Previous stroke (with Unknown)		
No	5,250 (76%)	6,393 (74%)
Unknown	204 (3.0%)	224 (2.6%)
Yes	1,430 (21%)	1,966 (23%)
Cardiac infarction		
No	6,492 (94%)	7,733 (90%)
Yes	392 (5.7%)	850 (9.9%)

Gender	Female N = 6,884[1]	Male N = 8,583[1]
Cardiac infarction (with Unknown)		
No	6,295 (91%)	7,480 (87%)
Unknown	197 (2.9%)	253 (2.9%)
Yes	392 (5.7%)	850 (9.9%)
Hypercholesterolemia		
No	3,073 (45%)	3,423 (40%)
Yes	3,811 (55%)	5,160 (60%)
Hypercholesterolemia (with Unknown)		
No	2,739 (40%)	3,078 (36%)
Unknown	334 (4.9%)	345 (4.0%)
Yes	3,811 (55%)	5,160 (60%)
Atrial fibrillation		
No	5,404 (79%)	6,938 (81%)
Yes	1,480 (21%)	1,645 (19%)
Atrial fibrillation (with Unknown)		
No	5,078 (74%)	6,540 (76%)
Unknown	326 (4.7%)	398 (4.6%)
Yes	1,480 (21%)	1,645 (19%)
Smoking		
No	6,077 (88%)	6,774 (79%)
Yes	807 (12%)	1,809 (21%)
Smoking (with Unknown)		
No	5,561 (81%)	6,098 (71%)
Unknown	516 (7.5%)	676 (7.9%)
Yes	807 (12%)	1,809 (21%)
Smoking (Imputed)		
No	6,033 (88%)	6,656 (78%)
Yes	851 (12%)	1,927 (22%)
Alcohol abuse		
No	6,728 (98%)	7,687 (90%)
Yes	156 (2.3%)	896 (10%)
Alcohol abuse (with Unknown)		
No	6,282 (91%)	6,958 (81%)
Unknown	446 (6.5%)	729 (8.5%)
Yes	156 (2.3%)	896 (10%)
Alcohol abuse (Imputed)		
No	6,717 (98%)	7,598 (89%)
Yes	167 (2.4%)	985 (11%)

[1] Median (Q1, Q3); n (%)

Population 2 — Used for RQ1 and RQ3

Table 2: Population 2. Patient Characteristics and Risk Factors

Gender	**Female** N = 5,940[1]	**Male** N = 6,684[1]
Age in Years	78 (67, 84)	73 (62, 81)
NIHSS at admission	0 (0, 1)	0 (0, 1)
Hypertension		
No	1,301 (22%)	1,613 (24%)
Yes	4,639 (78%)	5,071 (76%)
Hypertension (with Unknown)		
No	1,184 (20%)	1,440 (22%)
Unknown	117 (2.0%)	173 (2.6%)
Yes	4,639 (78%)	5,071 (76%)
Previous stroke		
No	4,642 (78%)	5,012 (75%)
Yes	1,298 (22%)	1,672 (25%)
Previous stroke (with Unknown)		
No	4,327 (73%)	4,665 (70%)
Unknown	315 (5.3%)	347 (5.2%)
Yes	1,298 (22%)	1,672 (25%)
Cardiac infarction		
No	5,595 (94%)	5,955 (89%)
Yes	345 (5.8%)	729 (11%)
Cardiac infarction (with Unknown)		
No	5,250 (88%)	5,573 (83%)
Unknown	345 (5.8%)	382 (5.7%)
Yes	345 (5.8%)	729 (11%)
Hypercholesterolemia		
No	1,816 (31%)	1,883 (28%)
Yes	4,124 (69%)	4,801 (72%)
Hypercholesterolemia (with Unknown)		
No	1,530 (26%)	1,568 (23%)
Unknown	286 (4.8%)	315 (4.7%)
Yes	4,124 (69%)	4,801 (72%)
Atrial fibrillation		
No	4,732 (80%)	5,314 (80%)
Yes	1,208 (20%)	1,370 (20%)
Smoking		
No	5,276 (89%)	5,389 (81%)
Yes	664 (11%)	1,295 (19%)
Smoking (with Unknown)		
No	4,334 (73%)	4,293 (64%)
Unknown	942 (16%)	1,096 (16%)
Yes	664 (11%)	1,295 (19%)
Smoking (Imputed)		
No	5,127 (86%)	5,150 (77%)

Gender	Female N = 5,940[1]	Male N = 6,684[1]
Yes	813 (14%)	1,534 (23%)
Alcohol abuse		
No	5,791 (97%)	6,167 (92%)
Yes	149 (2.5%)	517 (7.7%)
Alcohol abuse (with Unknown)		
No	4,805 (81%)	4,951 (74%)
Unknown	986 (17%)	1,216 (18%)
Yes	149 (2.5%)	517 (7.7%)
Alcohol abuse (Imputed)		
No	5,719 (96%)	5,992 (90%)
Yes	221 (3.7%)	692 (10%)

[1] Median (Q1, Q3); n (%)

Filtered Population 1 — Used for RQ2

Table 3: Population 1 Risk-Scores Subset. Patient Characteristics and Risk Factors

Gender	Female N = 2,369[1]	Male N = 2,962[1]
Age in Years	75 (65, 83)	70 (59, 77)
NIHSS at admission	1 (0, 2)	1 (0, 2)
Systolic blood pressure	158 (140, 180)	154 (140, 171)
Aetiology		
cardiogen embolic	487 (21%)	523 (18%)
else	36 (1.5%)	70 (2.4%)
Macroangiopathy	163 (6.9%)	348 (12%)
Microangiopathy	769 (32%)	917 (31%)
Unknown	914 (39%)	1,104 (37%)
Hypertension		
No	498 (21%)	652 (22%)
Yes	1,871 (79%)	2,310 (78%)
Hypertension (with Unknown)		
No	482 (20%)	609 (21%)
Unknown	16 (0.7%)	43 (1.5%)
Yes	1,871 (79%)	2,310 (78%)
Previous stroke		
No	1,896 (80%)	2,322 (78%)
Yes	473 (20%)	640 (22%)
Previous stroke (with Unknown)		
No	1,829 (77%)	2,254 (76%)
Unknown	67 (2.8%)	68 (2.3%)
Yes	473 (20%)	640 (22%)

Gender	Female N = 2,369[1]	Male N = 2,962[1]
Cardiac infarction		
No	2,240 (95%)	2,678 (90%)
Yes	129 (5.4%)	284 (9.6%)
Cardiac infarction (with Unknown)		
No	2,187 (92%)	2,601 (88%)
Unknown	53 (2.2%)	77 (2.6%)
Yes	129 (5.4%)	284 (9.6%)
Hypercholesterolemia		
No	962 (41%)	1,105 (37%)
Yes	1,407 (59%)	1,857 (63%)
Hypercholesterolemia (with Unknown)		
No	882 (37%)	1,020 (34%)
Unknown	80 (3.4%)	85 (2.9%)
Yes	1,407 (59%)	1,857 (63%)
Atrial fibrillation		
No	1,860 (79%)	2,448 (83%)
Yes	509 (21%)	514 (17%)
Atrial fibrillation (with Unknown)		
No	1,778 (75%)	2,344 (79%)
Unknown	82 (3.5%)	104 (3.5%)
Yes	509 (21%)	514 (17%)
Smoking		
No	2,046 (86%)	2,247 (76%)
Yes	323 (14%)	715 (24%)
Smoking (with Unknown)		
No	1,897 (80%)	2,069 (70%)
Unknown	149 (6.3%)	178 (6.0%)
Yes	323 (14%)	715 (24%)
Smoking (Imputed)		
No	2,032 (86%)	2,213 (75%)
Yes	337 (14%)	749 (25%)
Alcohol abuse		
No	2,286 (96%)	2,576 (87%)
Yes	83 (3.5%)	386 (13%)
Alcohol abuse (with Unknown)		
No	2,141 (90%)	2,361 (80%)
Unknown	145 (6.1%)	215 (7.3%)
Yes	83 (3.5%)	386 (13%)
Alcohol abuse (Imputed)		
No	2,285 (96%)	2,549 (86%)
Yes	84 (3.5%)	413 (14%)

[1] Median (Q1, Q3); n (%)

Gender Differences in Health Outcomes

Population 1 - Health Outcomes Odds Ratios

Variable	OR [95% CI]	P-Value	Interpretation
Early Recurrence in Stroke Unit	0.95 [0.8-1.12]	0.525	-
Early Worsening	0.97 [0.81-1.17]	0.771	-
Recurrence within 90 days	0.99 [0.84-1.17]	0.950	-
Recurrent Stroke	0.76 [0.46-1.24]	0.278	-
Symptomatic intracranial hemorrhage	1.4 [0.71-2.78]	0.325	-
Epileptic seizures	1.52 [0.92-2.52]	0.104	-
Pneumonia	0.54 [0.37-0.78]	0.001	Women had lower odds.
mRS at Follow-Up (>1)	1.59 [1.42-1.79]	< 0.001	Women had higher odds.
mRS at Follow-Up = 1	1.03 [0.9-1.18]	0.686	-
mRS at Follow-Up = 2	1.21 [1.01-1.45]	0.041	Women had higher odds.
mRS at Follow-Up = 3	1.93 [1.58-2.36]	< 0.001	Women had higher odds.
mRS at Follow-Up = 4	2.07 [1.67-2.57]	< 0.001	Women had higher odds.
mRS at Follow-Up = 5	3.21 [1.86-5.57]	< 0.001	Women had higher odds.
mRS at Follow-Up = 6	1.27 [0.98-1.65]	0.074	-

Table 4: Population 1 - Unadjusted outcome odds ratios by gender

Variable	OR [95% CI]	P-Value	Interpretation
Early Recurrence in Stroke Unit	0.91 [0.77-1.08]	0.287	-
Early Worsening	0.94 [0.77-1.13]	0.493	-
Recurrence within 90 days	0.96 [0.81-1.13]	0.602	-
Recurrent Stroke	0.79 [0.47-1.29]	0.347	-
Symptomatic intracranial hemorrhage	1.31 [0.66-2.62]	0.447	-
Epileptic seizures	1.44 [0.86-2.4]	0.164	-
Pneumonia	0.41 [0.28-0.6]	< 0.001	Women had lower odds.
mRS at Follow-Up (>1)	1.21 [1.07-1.37]	0.002	Women had higher odds.
mRS at Follow-Up = 1	0.97 [0.84-1.11]	0.629	-
mRS at Follow-Up = 2	1.04 [0.86-1.25]	0.689	-
mRS at Follow-Up = 3	1.42 [1.15-1.75]	0.001	Women had higher odds.
mRS at Follow-Up = 4	1.36 [1.08-1.71]	0.008	Women had higher odds.
mRS at Follow-Up = 5	2.2 [1.26-3.87]	0.006	Women had higher odds.
mRS at Follow-Up = 6	0.84 [0.64-1.11]	0.217	-

Table 5: Population 1 - Outcome odds ratios by gender controlled for Age

Variable	OR [95% CI]	P-Value	Interpretation
Early Recurrence in Stroke Unit	0.85 [0.62-1.16]	0.314	-
Early Worsening	0.81 [0.55-1.19]	0.288	-
Recurrence within 90 days	0.98 [0.74-1.29]	0.874	-
Recurrent Stroke	0.94 [0.41-2.09]	0.872	-
Symptomatic intracranial hemorrhage	1.09 [0.36-3.28]	0.882	-
Epileptic seizures	1.84 [0.74-4.86]	0.200	-
Pneumonia	0.45 [0.2-0.96]	0.046	Women had lower odds.
mRS at Follow-Up (>1)	1.31 [1.06-1.61]	0.012	Women had higher odds.
mRS at Follow-Up = 1	0.95 [0.76-1.19]	0.631	-
mRS at Follow-Up = 2	1.05 [0.76-1.44]	0.775	-
mRS at Follow-Up = 3	1.7 [1.21-2.38]	0.002	Women had higher odds.
mRS at Follow-Up = 4	1.39 [0.93-2.08]	0.107	-
mRS at Follow-Up = 5	3.11 [1.23-7.86]	0.017	Women had higher odds.
mRS at Follow-Up = 6	0.7 [0.42-1.18]	0.186	-

Table 6: Population 1 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors

Variable	OR [95% CI]	P-Value	Interpretation
Early Recurrence in Stroke Unit	0.86 [0.63-1.17]	0.343	-
Early Worsening	0.82 [0.56-1.2]	0.316	-
Recurrence within 90 days	0.99 [0.75-1.3]	0.923	-
Recurrent Stroke	0.98 [0.42-2.21]	0.954	-
Symptomatic intracranial hemorrhage	1.06 [0.35-3.19]	0.921	-
Epileptic seizures	1.83 [0.73-4.86]	0.205	-
Pneumonia	0.45 [0.2-0.97]	0.048	Women had lower odds.
mRS at Follow-Up (>1)	1.32 [1.07-1.62]	0.010	Women had higher odds.
mRS at Follow-Up = 1	0.93 [0.75-1.17]	0.557	-
mRS at Follow-Up = 2	1.05 [0.76-1.45]	0.767	-
mRS at Follow-Up = 3	1.71 [1.21-2.4]	0.002	Women had higher odds.
mRS at Follow-Up = 4	1.37 [0.91-2.05]	0.131	-
mRS at Follow-Up = 5	3.21 [1.24-8.3]	0.016	Women had higher odds.
mRS at Follow-Up = 6	0.75 [0.44-1.27]	0.280	-

Table 7: Population 1 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors with Unknown included

Variable	OR [95% CI]	P-Value	Interpretation
Early Recurrence in Stroke Unit	0.85 [0.62-1.16]	0.313	-
Early Worsening	0.82 [0.56-1.2]	0.318	-
Recurrence within 90 days	0.99 [0.75-1.31]	0.961	-
Recurrent Stroke	0.93 [0.4-2.07]	0.855	-
Symptomatic intracranial hemorrhage	1.07 [0.36-3.24]	0.899	-
Epileptic seizures	1.82 [0.73-4.83]	0.207	-
Pneumonia	0.45 [0.2-0.96]	0.045	Women had lower odds.
mRS at Follow-Up (>1)	1.32 [1.07-1.63]	0.009	Women had higher odds.
mRS at Follow-Up = 1	0.94 [0.76-1.18]	0.620	-
mRS at Follow-Up = 2	1.05 [0.76-1.45]	0.759	-
mRS at Follow-Up = 3	1.71 [1.22-2.39]	0.002	Women had higher odds.
mRS at Follow-Up = 4	1.43 [0.96-2.14]	0.083	-
mRS at Follow-Up = 5	3.06 [1.21-7.76]	0.018	Women had higher odds.
mRS at Follow-Up = 6	0.72 [0.43-1.22]	0.225	-

Table 8: Population 1 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors with Smoking and Alcohol Abuse imputed

Population 2 - Health Outcomes Odds Ratios

Variable	OR [95% CI]	P-Value	Interpretation
Early Recurrence in Stroke Unit	1.16 [0.7-1.93]	0.555	-
Recurrent Stroke	1.55 [0.63-4]	0.347	-
Epileptic seizures	2.25 [0.71-8.44]	0.185	-
Pneumonia	0.63 [0.35-1.09]	0.105	-
mRS at Follow-Up (>1)	1.72 [1.48-2.01]	< 0.001	Women had higher odds.
mRS at Follow-Up = 1	1.08 [0.9-1.28]	0.397	-
mRS at Follow-Up = 2	1.6 [1.24-2.06]	< 0.001	Women had higher odds.
mRS at Follow-Up = 3	1.78 [1.36-2.33]	< 0.001	Women had higher odds.
mRS at Follow-Up = 4	2.53 [1.84-3.48]	< 0.001	Women had higher odds.
mRS at Follow-Up = 5	1.94 [1.11-3.36]	0.019	Women had higher odds.
mRS at Follow-Up = 6	1.09 [0.74-1.62]	0.662	-

Table 9: Population 2 - Unadjusted outcome odds ratios by gender

Variable	OR [95% CI]	P-Value	Interpretation
Early Recurrence in Stroke Unit	0.99 [0.59-1.66]	0.972	-
Recurrent Stroke	1.49 [0.6-3.9]	0.393	-
Epileptic seizures	2.08 [0.64-7.87]	0.239	-
Pneumonia	0.52 [0.29-0.9]	0.023	Women had lower odds.
mRS at Follow-Up (>1)	1.39 [1.18-1.64]	< 0.001	Women had higher odds.
mRS at Follow-Up = 1	1 [0.84-1.19]	0.975	-
mRS at Follow-Up = 2	1.4 [1.08-1.81]	0.012	Women had higher odds.
mRS at Follow-Up = 3	1.38 [1.05-1.83]	0.022	Women had higher odds.
mRS at Follow-Up = 4	1.9 [1.37-2.65]	< 0.001	Women had higher odds.
mRS at Follow-Up = 5	1.42 [0.81-2.49]	0.223	-
mRS at Follow-Up = 6	0.78 [0.52-1.18]	0.236	-

Table 10: Population 2 - Outcome odds ratios by gender controlled for Age

Variable	OR [95% CI]	P-Value	Interpretation
Early Recurrence in Stroke Unit	1.11 [0.66-1.88]	0.685	-
Recurrent Stroke	1.63 [0.64-4.34]	0.311	-
Epileptic seizures	2.38 [0.72-9.21]	0.171	-
Pneumonia	0.61 [0.33-1.08]	0.093	-
mRS at Follow-Up (>1)	1.56 [1.31-1.86]	< 0.001	Women had higher odds.
mRS at Follow-Up = 1	1.07 [0.89-1.28]	0.469	-
mRS at Follow-Up = 2	1.53 [1.17-2.01]	0.002	Women had higher odds.
mRS at Follow-Up = 3	1.69 [1.26-2.27]	< 0.001	Women had higher odds.
mRS at Follow-Up = 4	2.13 [1.52-3]	< 0.001	Women had higher odds.
mRS at Follow-Up = 5	1.68 [0.94-2.99]	0.081	-
mRS at Follow-Up = 6	0.87 [0.57-1.33]	0.514	-

Table 11: Population 2 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors

Variable	OR [95% CI]	P-Value	Interpretation
Early Recurrence in Stroke Unit	1.11 [0.66-1.88]	0.687	-
Recurrent Stroke	1.62 [0.63-4.36]	0.319	-
Epileptic seizures	2.39 [0.72-9.23]	0.169	-
Pneumonia	0.61 [0.33-1.08]	0.094	-
mRS at Follow-Up (>1)	1.57 [1.32-1.87]	< 0.001	Women had higher odds.
mRS at Follow-Up = 1	1.07 [0.89-1.28]	0.474	-
mRS at Follow-Up = 2	1.53 [1.17-2]	0.002	Women had higher odds.
mRS at Follow-Up = 3	1.71 [1.27-2.29]	< 0.001	Women had higher odds.
mRS at Follow-Up = 4	2.18 [1.55-3.07]	< 0.001	Women had higher odds.
mRS at Follow-Up = 5	1.65 [0.92-2.95]	0.093	-
mRS at Follow-Up = 6	0.88 [0.58-1.34]	0.555	-

Table 12: Population 2 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors with Unknown included

Variable	OR [95% CI]	P-Value	Interpretation
Early Recurrence in Stroke Unit	1.11 [0.66-1.87]	0.699	-
Recurrent Stroke	1.58 [0.62-4.22]	0.339	-
Epileptic seizures	2.31 [0.7-8.95]	0.184	-
Pneumonia	0.61 [0.33-1.08]	0.097	-
mRS at Follow-Up (>1)	1.54 [1.3-1.84]	< 0.001	Women had higher odds.
mRS at Follow-Up = 1	1.06 [0.88-1.27]	0.559	-
mRS at Follow-Up = 2	1.51 [1.15-1.97]	0.003	Women had higher odds.
mRS at Follow-Up = 3	1.68 [1.25-2.24]	< 0.001	Women had higher odds.
mRS at Follow-Up = 4	2.11 [1.5-2.97]	< 0.001	Women had higher odds.
mRS at Follow-Up = 5	1.64 [0.92-2.92]	0.096	-
mRS at Follow-Up = 6	0.85 [0.56-1.3]	0.454	-

Table 13: Population 2 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors with Smoking and Alcohol Abuse imputed

Predictive Accuracy of the ABCD2 and ABCD3-I Scores

Recurrence Outcome Distribution

Table 14: 'Database' Scores Distribution for Early Worsening

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	57	33	1	1
1	213	148	4	2
2	494	405	15	7
3	591	425	17	7
4	763	695	24	17
5	447	334	9	7
6	248	227	5	2
7	71	55	3	4
ABCD3-I				
0	22	11	0	0
1	81	70	1	1
2	216	188	5	3
3	358	266	7	1
4	563	503	10	9
5	490	348	15	5
6	543	503	24	12
7	313	229	5	9
8	198	140	6	3
9	68	46	3	3
10	15	13	0	0
11	7	0	2	1
12	2	0	0	0

Table 15: 'Database' Scores Distribution for Early Recurrence in Stroke Unit

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	57	32	1	2
1	213	145	4	5
2	486	401	23	11
3	583	421	25	11
4	755	685	32	27
5	439	331	17	10
6	246	225	7	4
7	70	54	4	5
ABCD3-I				
0	22	11	0	0
1	81	68	1	3
2	215	187	6	4
3	355	265	10	2
4	556	497	17	15
5	485	344	20	9
6	537	493	30	22
7	305	226	13	12
8	194	140	10	3
9	68	45	3	4
10	15	13	0	0
11	6	0	3	1
12	2	0	0	0

Table 16: 'Database' Scores Distribution for Recurrence within 90 days

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	55	29	2	2
1	202	137	6	5
2	474	384	23	13
3	554	403	30	15
4	723	638	39	40
5	426	320	22	15
6	234	220	11	6
7	68	51	4	6
ABCD3-I				
0	22	11	0	0
1	81	62	1	3
2	209	171	8	8
3	335	255	12	6

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
4	537	478	17	19
5	462	329	26	10
6	511	465	37	28
7	297	218	16	18
8	188	132	12	5
9	64	44	4	3
10	14	12	1	1
11	6	0	3	1
12	2	0	0	0

Table 17: 'Database' Scores Distribution for Combined vascular endpoint

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	55	29	2	2
1	201	136	7	6
2	470	380	27	17
3	551	400	33	18
4	719	634	43	44
5	422	320	27	15
6	233	219	12	7
7	68	50	4	7
ABCD3-I				
0	22	11	0	0
1	81	62	1	3
2	209	171	8	8
3	334	254	13	7
4	533	474	21	23
5	458	326	30	13
6	507	460	41	33
7	294	218	20	18
8	187	132	13	5
9	64	43	4	4
10	14	12	1	1
11	6	0	3	1
12	2	0	0	0

Table 18: 'Database' Scores Distribution for Cumulative Endpoint

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	55	29	2	2
1	201	135	7	7
2	469	377	28	20
3	549	400	35	18
4	718	631	44	47
5	422	319	27	16
6	232	216	13	10
7	68	50	4	7
ABCD3-I				
0	22	11	0	0
1	81	62	1	3
2	209	169	8	10
3	334	253	13	8
4	532	473	22	24
5	456	326	32	13
6	506	457	42	36
7	294	217	20	19
8	186	129	14	8
9	64	43	4	4
10	14	12	1	1
11	6	0	3	1
12	2	0	0	0

Table 19: 'Database' Scores Distribution for Recurrence 1

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	24	16	0	1
1	104	57	4	1
2	212	167	10	7
3	248	171	14	7
4	344	313	19	22
5	168	132	13	9
6	112	93	5	5
7	32	21	2	6
ABCD3-I				
0	8	4	0	0
1	35	31	1	0
2	89	69	1	4
3	157	93	7	5

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
4	240	233	10	8
5	201	146	13	3
6	248	224	18	19
7	124	94	8	9
8	98	53	5	4
9	32	17	3	5
10	7	6	0	0
11	3	0	1	1
12	1	0	0	0

Table 20: 'Database' Scores Distribution for Recurrence 2

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	23	17	0	0
1	99	51	6	2
2	205	155	9	10
3	228	156	19	13
4	313	292	26	31
5	164	126	13	12
6	103	87	9	8
7	30	17	0	7
ABCD3-I				
0	8	4	0	0
1	33	30	2	0
2	86	67	2	4
3	149	81	10	10
4	232	214	10	16
5	188	136	16	6
6	223	211	24	23
7	117	88	7	13
8	87	50	7	6
9	32	15	3	4
10	6	5	1	0
11	2	0	0	1
12	1	0	0	0

Table 21: 'Database' Scores Distribution for Recurrence 3

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	23	16	0	1
1	99	51	6	2
2	204	155	12	10
3	228	155	22	15
4	312	290	29	34
5	163	126	15	12
6	103	87	10	8
7	30	17	2	8
ABCD3-I				
0	8	4	0	0
1	33	30	2	0
2	86	66	2	5
3	149	81	11	10
4	231	213	13	17
5	188	135	18	8
6	222	210	27	25
7	116	88	9	13
8	87	50	8	6
9	32	15	4	5
10	6	5	1	0
11	2	0	1	1
12	1	0	0	0

Table 22: 'Database' Scores Distribution for Recurrence 4

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	56	32	2	2
1	207	143	10	7
2	478	393	31	19
3	567	410	41	22
4	732	657	55	55
5	422	320	34	21
6	237	217	16	12
7	70	49	4	10
ABCD3-I				
0	22	11	0	0
1	80	68	2	3
2	213	182	8	9
3	347	255	18	12

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
4	548	486	25	26
5	468	339	37	14
6	514	471	53	44
7	294	215	24	23
8	187	135	17	8
9	66	42	5	7
10	14	12	1	1
11	6	0	3	1
12	2	0	0	0

Table 23: 'Calculated' Scores Distribution for Early Worsening

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	9	6	0	0
1	45	26	0	0
2	156	117	0	1
3	376	267	9	2
4	667	568	10	6
5	709	597	21	14
6	685	575	29	14
7	188	130	7	9
ABCD3-I				
0	8	5	0	0
1	31	15	0	0
2	93	81	0	0
3	185	157	1	1
4	344	288	4	2
5	417	342	13	5
6	557	504	8	7
7	486	371	15	7
8	452	349	23	10
9	177	130	5	11
10	59	36	4	2
11	20	4	3	1
12	2	0	0	0

Table 24: 'Calculated' Scores Distribution for Early Recurrence in Stroke Unit

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	9	6	0	0
1	45	26	0	0
2	156	116	0	2
3	375	265	10	4
4	661	563	16	11
5	698	591	32	20
6	675	566	39	23
7	182	126	13	13
ABCD3-I				
0	8	5	0	0
1	31	15	0	0
2	93	81	0	0
3	184	157	2	1
4	344	285	4	5
5	415	339	15	8
6	550	497	15	14
7	477	367	24	11
8	446	344	29	15
9	171	125	11	16
10	56	36	7	2
11	20	4	3	1
12	2	0	0	0

Table 25: 'Calculated' Scores Distribution for Recurrence within 90 days

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	8	5	1	0
1	45	26	0	0
2	150	111	2	1
3	355	245	15	8
4	640	549	18	14
5	673	564	36	27
6	644	526	45	34
7	173	120	18	17
ABCD3-I				
0	8	5	0	0
1	31	15	0	0
2	90	77	2	0
3	178	143	3	4

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
4	330	274	6	8
5	394	319	19	14
6	531	484	17	16
7	460	353	28	15
8	425	317	33	21
9	163	120	15	18
10	54	31	8	4
11	18	4	4	1
12	2	0	0	0

Table 26: 'Calculated' Scores Distribution for Combined vascular endpoint

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	8	5	1	0
1	45	26	0	0
2	150	110	2	2
3	353	242	17	11
4	638	541	20	22
5	667	560	43	31
6	637	520	52	40
7	170	118	21	19
ABCD3-I				
0	8	5	0	0
1	31	15	0	0
2	90	77	2	0
3	178	142	3	5
4	330	272	6	10
5	392	315	21	18
6	528	477	20	23
7	453	352	36	16
8	420	311	38	27
9	161	117	17	21
10	53	31	9	4
11	18	4	4	1
12	2	0	0	0

Table 27: 'Calculated' Scores Distribution for Cumulative Endpoint

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	8	5	1	0
1	45	26	0	0
2	150	110	2	2
3	353	242	17	11
4	638	541	20	22
5	667	560	43	31
6	637	520	52	40
7	170	118	21	19
ABCD3-I				
0	8	5	0	0
1	31	15	0	0
2	90	77	2	0
3	178	142	3	5
4	330	272	6	10
5	392	315	21	18
6	528	477	20	23
7	453	352	36	16
8	420	311	38	27
9	161	117	17	21
10	53	31	9	4
11	18	4	4	1
12	2	0	0	0

Table 28: 'Calculated' Scores Distribution for Recurrence 1

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	4	4	0	0
1	15	8	0	0
2	58	42	2	0
3	158	99	7	6
4	272	226	8	11
5	319	256	21	9
6	322	272	19	19
7	76	52	9	11
ABCD3-I				
0	3	3	0	0
1	10	4	0	0
2	31	32	1	0
3	73	59	1	3

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
4	130	109	4	3
5	184	137	9	5
6	237	215	6	13
7	216	164	18	6
8	225	170	18	13
9	76	47	7	11
10	30	15	0	1
11	9	4	2	1
12	0	0	0	0

Table 29: 'Calculated' Scores Distribution for Recurrence 2

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	3	4	0	0
1	15	7	1	0
2	56	40	3	1
3	150	86	7	11
4	256	211	11	16
5	297	245	25	13
6	298	253	27	26
7	73	46	8	14
ABCD3-I				
0	2	3	0	0
1	10	4	1	0
2	30	31	1	0
3	71	53	1	6
4	128	99	6	8
5	172	124	13	10
6	223	203	9	16
7	201	156	18	10
8	204	157	24	18
9	72	43	6	11
10	26	15	1	1
11	9	4	2	1
12	0	0	0	0

Table 30: 'Calculated' Scores Distribution for Recurrence 3

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	3	4	0	0
1	15	7	1	0
2	56	40	3	1
3	150	84	9	13
4	256	211	12	17
5	296	244	29	14
6	297	252	30	28
7	73	46	11	15
ABCD3-I				
0	2	3	0	0
1	10	4	1	0
2	30	31	1	0
3	71	53	2	6
4	128	99	6	8
5	172	122	14	12
6	223	202	10	18
7	200	155	22	11
8	203	157	27	19
9	72	43	8	12
10	26	15	1	1
11	9	4	3	1
12	0	0	0	0

Table 31: 'Calculated' Scores Distribution for Recurrence 4

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
ABCD2				
0	8	6	1	0
1	45	26	0	0
2	153	116	3	2
3	367	256	18	13
4	650	552	27	22
5	674	577	56	34
6	652	540	62	49
7	172	115	23	24
ABCD3-I				
0	8	5	0	0
1	31	15	0	0
2	91	81	2	0
3	183	151	3	7

Score	Variable: No Male	Variable: No Female	Variable: Yes Male	Variable: Yes Female
4	338	280	10	10
5	406	329	24	18
6	541	484	24	27
7	455	358	46	20
8	426	327	49	32
9	162	116	20	25
10	55	34	8	4
11	19	4	4	1
12	2	0	0	0

AUC Tables

Table 32: 'Database' ABCD2 AUC Values

Gender	No Covariates	Age Covariate	All Covariates	All Covariates (with Imputed)	All Covariates (with Unknown)
Early Worsening					
Female	0.533	0.536	0.633	0.634	0.680
Male	0.494	0.515	0.690	0.694	0.703
All	0.508	0.527	0.663	0.666	0.681
Early Recurrence in Stroke Unit					
Female	0.511	0.514	0.623	0.622	0.669
Male	0.506	0.516	0.673	0.673	0.679
All	0.507	0.507	0.643	0.643	0.654
Recurrence within 90 days					
Female	0.533	0.535	0.665	0.664	0.701
Male	0.519	0.522	0.651	0.660	0.665
All	0.525	0.522	0.642	0.646	0.663
MRS at Follow-Up > 4					
Female	0.606	0.712	0.842	0.842	0.863
Male	0.531	0.797	0.831	0.836	0.840
All	0.568	0.746	0.800	0.802	0.813
Combined vascular endpoint					
Female	0.520	0.524	0.666	0.665	0.701
Male	0.519	0.524	0.645	0.654	0.652
All	0.519	0.523	0.636	0.639	0.651
Cumulative Endpoint					
Female	0.521	0.523	0.659	0.658	0.688
Male	0.518	0.524	0.643	0.653	0.650
All	0.519	0.523	0.635	0.640	0.647
Recurrence 1					
Female	0.592	0.677	0.819	0.819	0.846
Male	0.540	0.750	0.784	0.786	0.793
All	0.564	0.708	0.769	0.770	0.782

Gender	No Covariates	Age Covariate	All Covariates	All Covariates (with Imputed)	All Covariates (with Unknown)
Recurrence 2					
Female	0.582	0.655	0.758	0.758	0.768
Male	0.538	0.706	0.752	0.756	0.764
All	0.561	0.681	0.734	0.735	0.745
Recurrence 3					
Female	0.574	0.636	0.751	0.751	0.762
Male	0.544	0.689	0.731	0.734	0.743
All	0.559	0.663	0.720	0.721	0.732
Recurrence 4					
Female	0.551	0.584	0.696	0.695	0.718
Male	0.524	0.580	0.651	0.659	0.664
All	0.535	0.579	0.659	0.661	0.672

Table 33: 'Database' ABCD3-I AUC Values

Gender	No Covariates	Age Covariate	All Covariates	All Covariates (with Imputed)	All Covariates (with Unknown)
Early Worsening					
Female	0.620	0.621	0.665	0.665	0.705
Male	0.566	0.577	0.705	0.707	0.712
All	0.586	0.598	0.682	0.684	0.696
Early Recurrence in Stroke Unit					
Female	0.590	0.597	0.655	0.655	0.688
Male	0.579	0.581	0.689	0.690	0.693
All	0.584	0.588	0.668	0.668	0.674
Recurrence within 90 days					
Female	0.572	0.572	0.667	0.665	0.702
Male	0.587	0.593	0.668	0.677	0.680
All	0.580	0.583	0.651	0.654	0.671
MRS at Follow-Up > 4					
Female	0.636	0.717	0.843	0.843	0.864
Male	0.565	0.795	0.828	0.834	0.838
All	0.599	0.752	0.801	0.803	0.814
Combined vascular endpoint					
Female	0.569	0.568	0.674	0.672	0.706
Male	0.588	0.590	0.658	0.666	0.665
All	0.580	0.582	0.649	0.652	0.661
Cumulative Endpoint					
Female	0.571	0.573	0.670	0.668	0.693
Male	0.588	0.589	0.657	0.665	0.663
All	0.581	0.583	0.649	0.652	0.657
Recurrence 1					

Gender	No Covariates	Age Covariate	All Covariates	All Covariates (with Imputed)	All Covariates (with Unknown)
Female	0.625	0.683	0.820	0.821	0.847
Male	0.579	0.753	0.784	0.787	0.795
All	0.600	0.718	0.772	0.773	0.785
Recurrence 2					
Female	0.583	0.651	0.750	0.750	0.761
Male	0.567	0.708	0.752	0.755	0.764
All	0.574	0.682	0.733	0.735	0.744
Recurrence 3					
Female	0.579	0.632	0.744	0.744	0.756
Male	0.577	0.695	0.734	0.737	0.747
All	0.578	0.668	0.720	0.720	0.732
Recurrence 4					
Female	0.587	0.600	0.699	0.698	0.720
Male	0.590	0.610	0.661	0.667	0.673
All	0.589	0.605	0.668	0.669	0.680

Table 34: 'Calculated' ABCD2 AUC Values

Gender	No Covariates	Age Covariate	All Covariates	All Covariates (with Imputed)	All Covariates (with Unknown)
Early Worsening					
Female	0.656	0.667	0.704	0.705	0.734
Male	0.615	0.635	0.735	0.736	0.740
All	0.630	0.651	0.713	0.714	0.724
Early Recurrence in Stroke Unit					
Female	0.640	0.651	0.691	0.691	0.717
Male	0.629	0.636	0.714	0.714	0.717
All	0.633	0.643	0.699	0.699	0.705
Recurrence within 90 days					
Female	0.645	0.650	0.703	0.702	0.731
Male	0.613	0.623	0.682	0.687	0.694
All	0.627	0.634	0.677	0.679	0.693
MRS at Follow-Up > 4					
Female	0.630	0.696	0.834	0.834	0.856
Male	0.581	0.796	0.831	0.835	0.842
All	0.604	0.745	0.798	0.800	0.811
Combined vascular endpoint					
Female	0.629	0.633	0.701	0.699	0.729
Male	0.619	0.624	0.675	0.679	0.681
All	0.623	0.629	0.673	0.675	0.684
Cumulative Endpoint					
Female	0.617	0.623	0.692	0.690	0.712

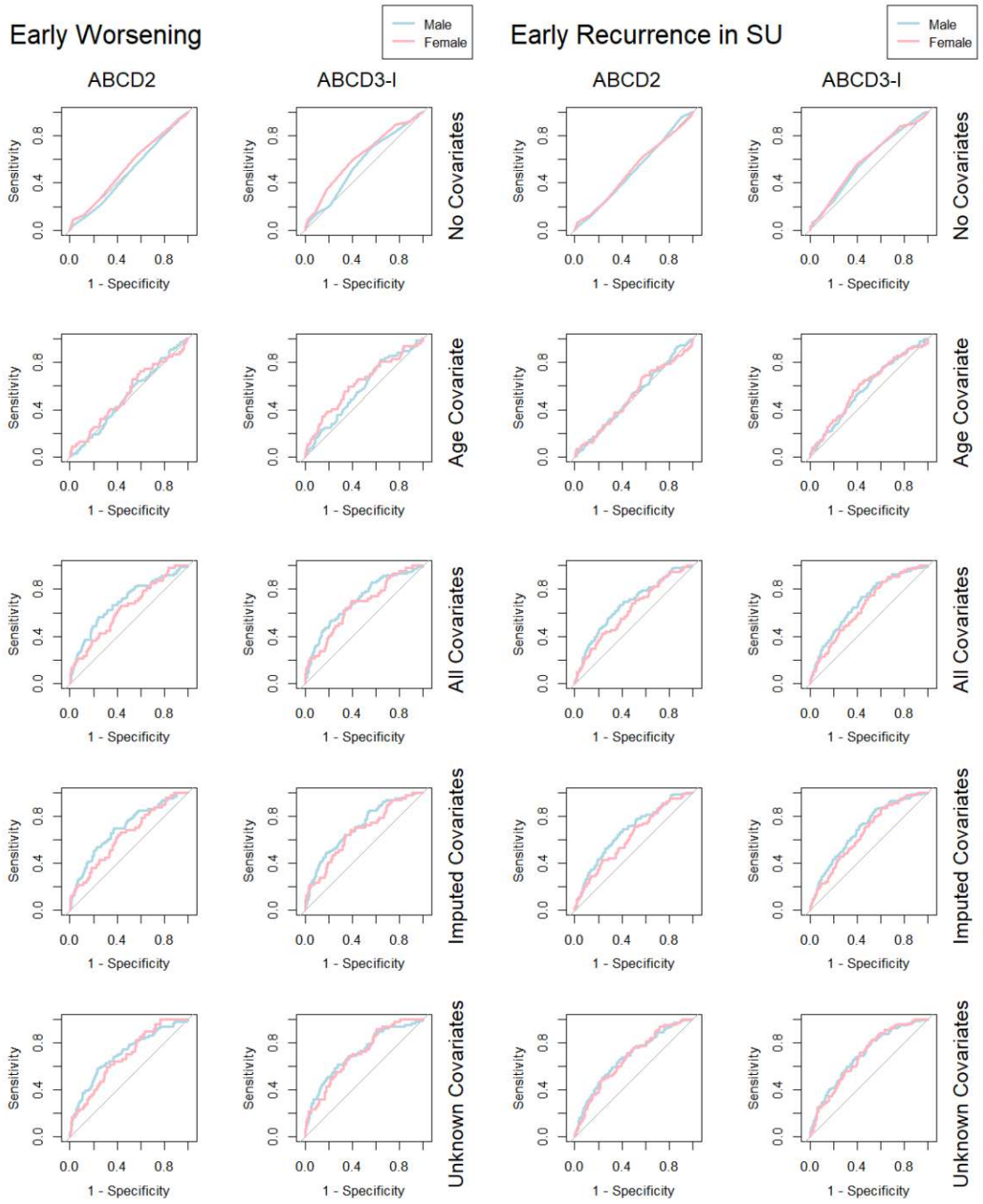
Gender	No Covariates	Age Covariate	All Covariates	All Covariates (with Imputed)	All Covariates (with Unknown)
Male	0.619	0.623	0.675	0.680	0.681
All	0.618	0.624	0.670	0.671	0.679
Recurrence 1					
Female	0.615	0.665	0.809	0.809	0.839
Male	0.589	0.754	0.790	0.793	0.796
All	0.601	0.710	0.769	0.770	0.781
Recurrence 2					
Female	0.580	0.634	0.744	0.744	0.754
Male	0.581	0.709	0.756	0.760	0.769
All	0.581	0.677	0.732	0.733	0.742
Recurrence 3					
Female	0.575	0.618	0.736	0.736	0.748
Male	0.587	0.695	0.741	0.744	0.748
All	0.581	0.661	0.719	0.719	0.729
Recurrence 4					
Female	0.636	0.643	0.712	0.711	0.731
Male	0.614	0.628	0.672	0.676	0.681
All	0.624	0.632	0.679	0.680	0.691

Table 35: 'Calculated' ABCD3-I AUC Values

Gender	No Covariates	Age Covariate	All Covariates	All Covariates (with Imputed)	All Covariates (with Unknown)
Early Worsening					
Female	0.702	0.708	0.738	0.739	0.767
Male	0.648	0.663	0.748	0.752	0.750
All	0.669	0.684	0.732	0.735	0.741
Early Recurrence in Stroke Unit					
Female	0.673	0.683	0.721	0.722	0.740
Male	0.664	0.668	0.733	0.734	0.734
All	0.669	0.677	0.721	0.721	0.724
Recurrence within 90 days					
Female	0.641	0.642	0.696	0.695	0.725
Male	0.649	0.656	0.704	0.710	0.714
All	0.646	0.650	0.686	0.688	0.700
MRS at Follow-Up > 4					
Female	0.648	0.703	0.834	0.834	0.857
Male	0.600	0.799	0.831	0.836	0.843
All	0.623	0.752	0.799	0.801	0.813
Combined vascular endpoint					
Female	0.636	0.638	0.701	0.700	0.727
Male	0.657	0.659	0.695	0.702	0.699

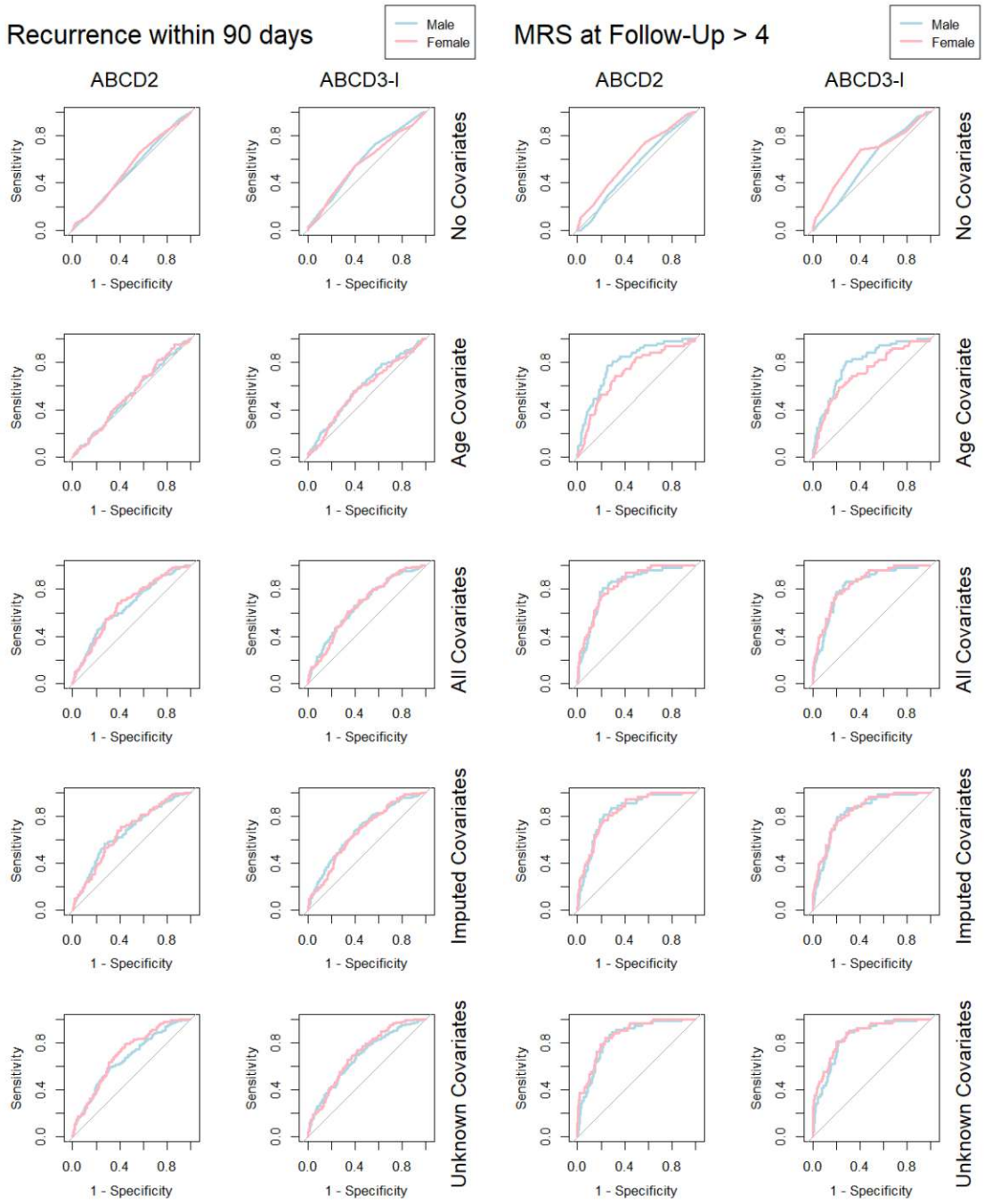
Gender	No Covariates	Age Covariate	All Covariates	All Covariates (with Imputed)	All Covariates (with Unknown)
All	0.648	0.651	0.685	0.687	0.693
Cumulative Endpoint					
Female	0.630	0.631	0.693	0.692	0.711
Male	0.656	0.658	0.695	0.702	0.699
All	0.644	0.647	0.682	0.684	0.689
Recurrence 1					
Female	0.638	0.677	0.810	0.810	0.839
Male	0.613	0.761	0.793	0.796	0.800
All	0.625	0.720	0.772	0.773	0.783
Recurrence 2					
Female	0.572	0.633	0.744	0.744	0.756
Male	0.593	0.715	0.757	0.761	0.771
All	0.582	0.679	0.730	0.731	0.742
Recurrence 3					
Female	0.571	0.618	0.735	0.735	0.748
Male	0.604	0.703	0.744	0.747	0.751
All	0.588	0.665	0.718	0.718	0.728
Recurrence 4					
Female	0.638	0.646	0.708	0.708	0.728
Male	0.652	0.660	0.693	0.698	0.701
All	0.646	0.653	0.690	0.691	0.699

ROC Curves



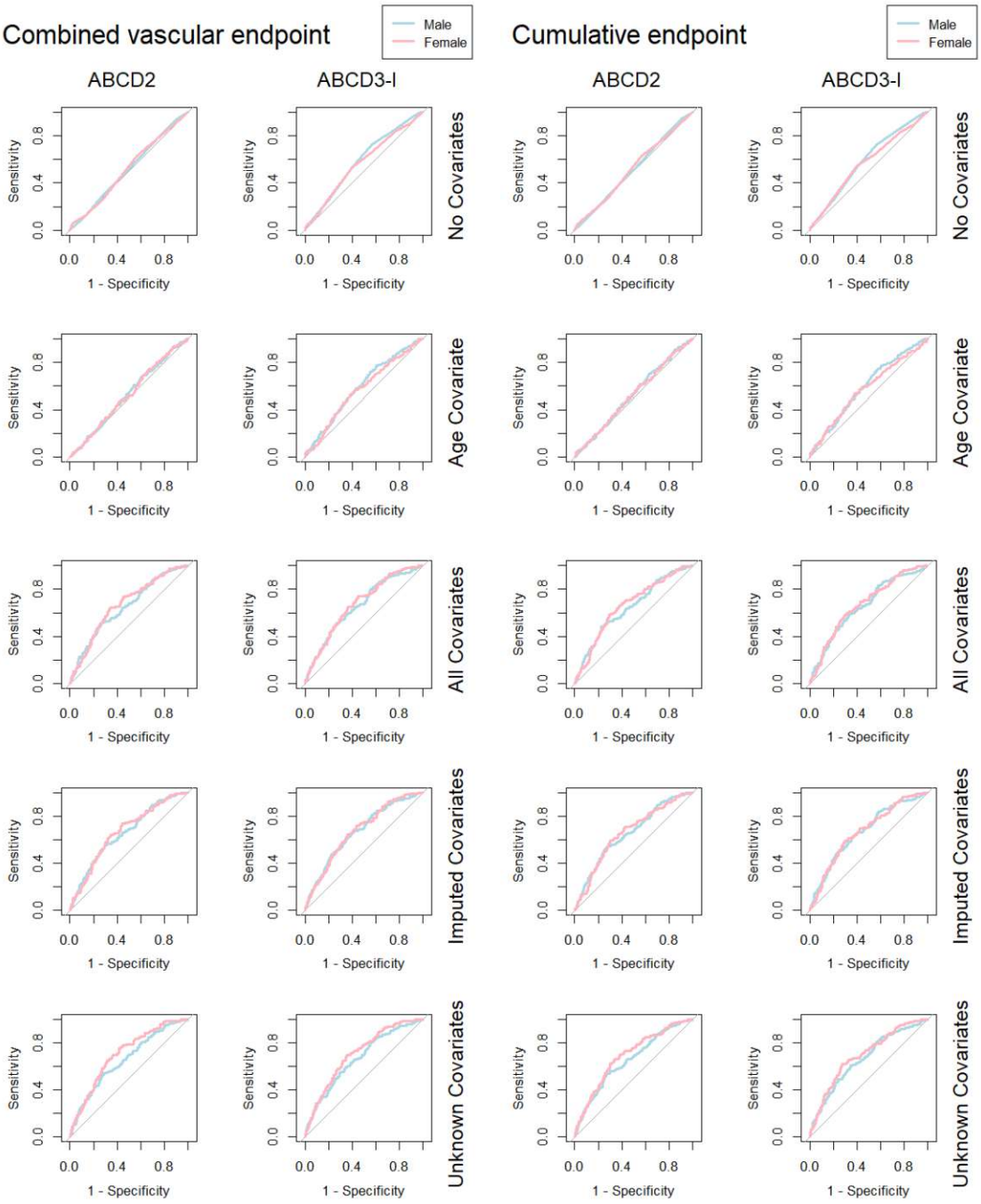
(a) Early Worsening (b) Early Recurrence in Stroke Unit

Figure 1: 'Database' Risk Scores ROC Curves



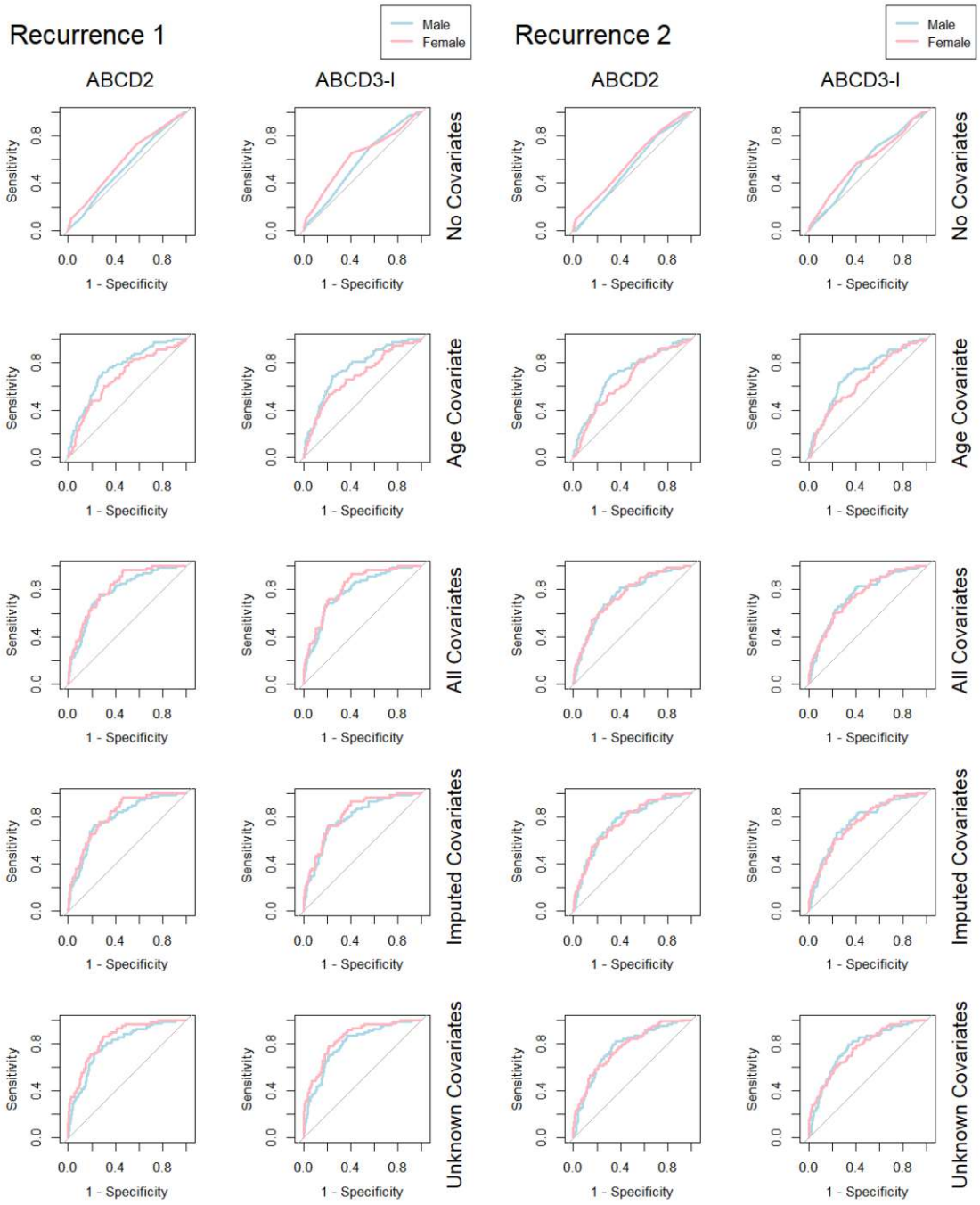
(c) Recurrence within 90 days

(d) MRS at Follow-Up > 4



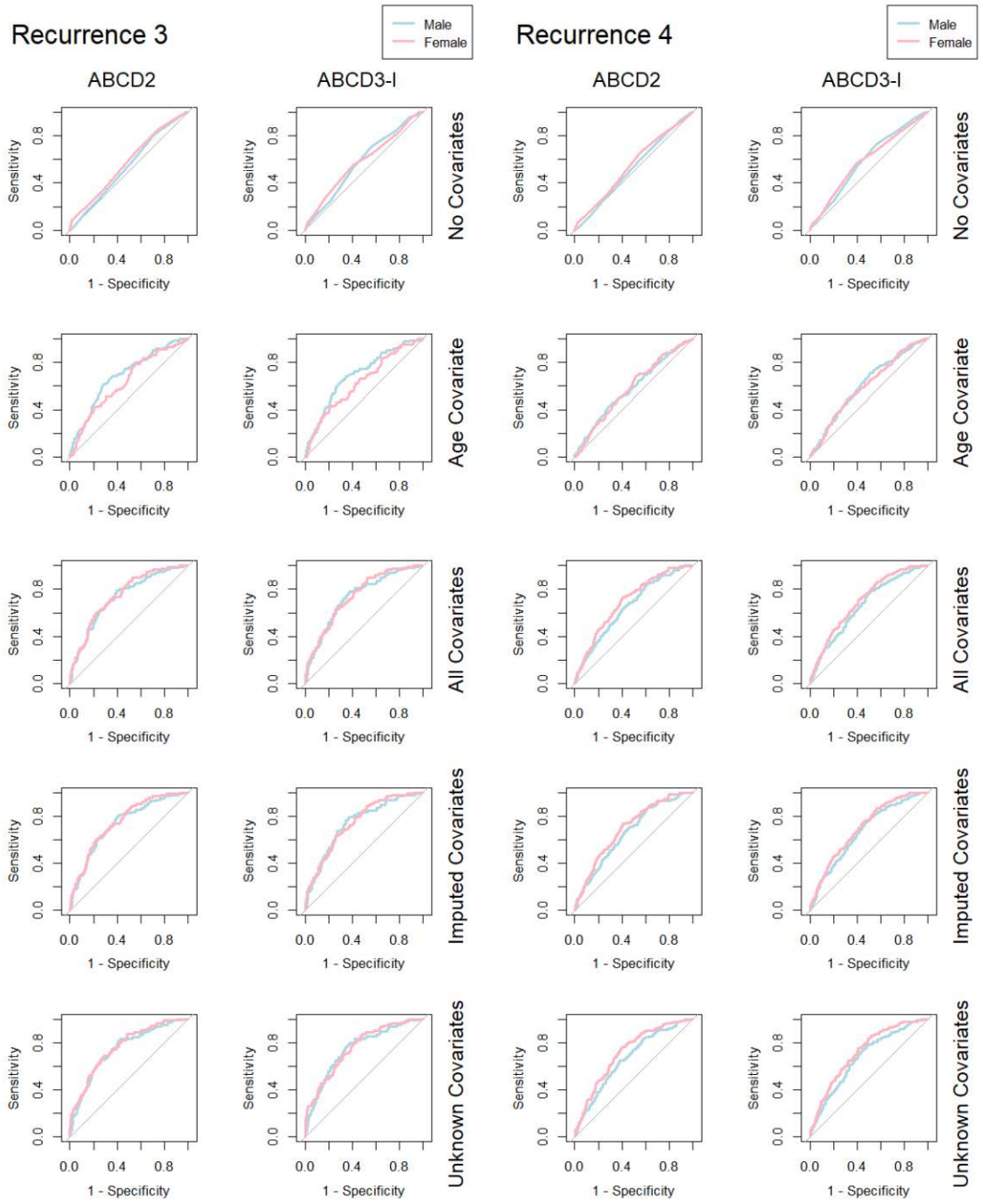
(e) Combined vascular endpoint

(f) Cumulative endpoint



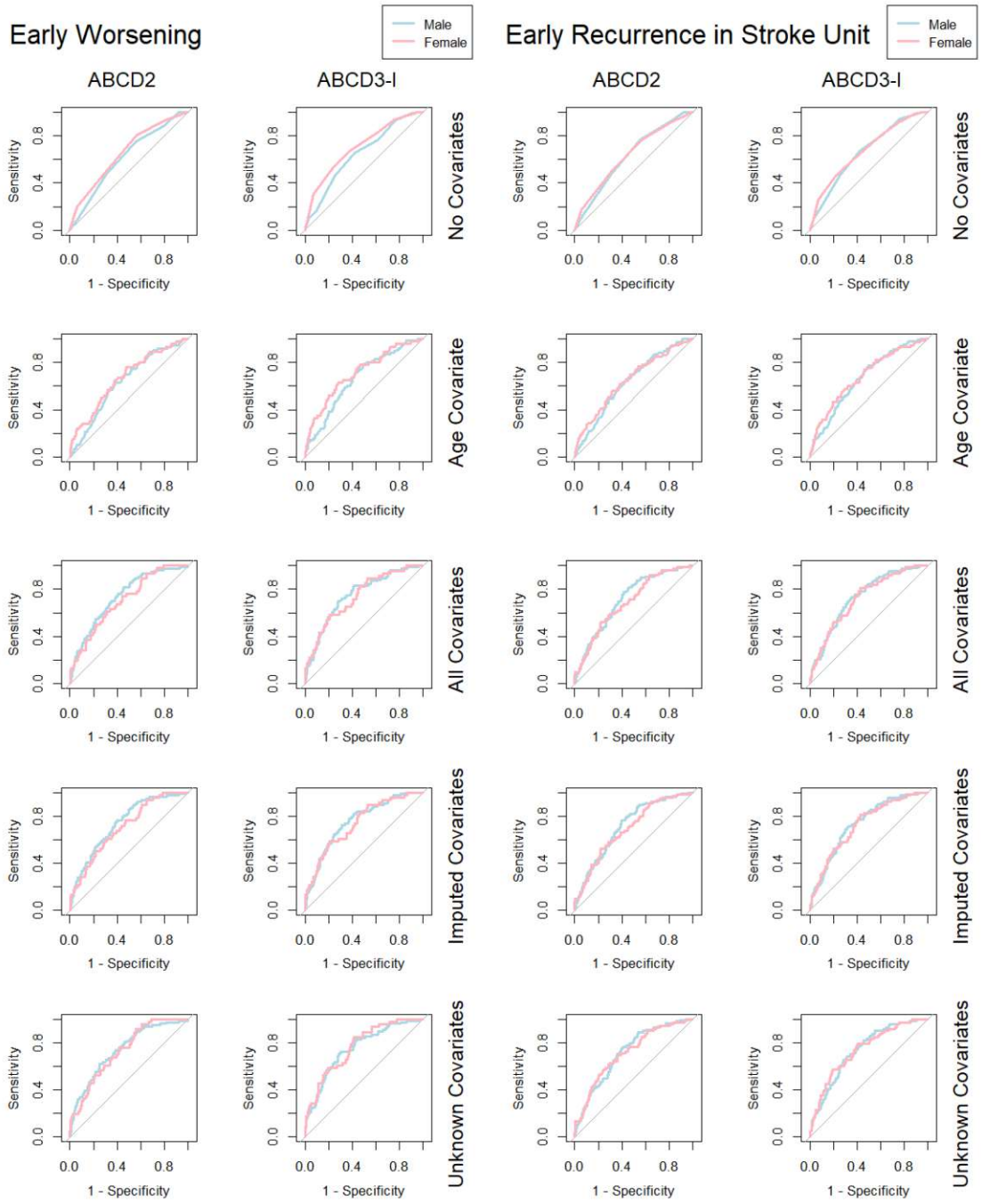
(g) Recurrence 1

(h) Recurrence 2



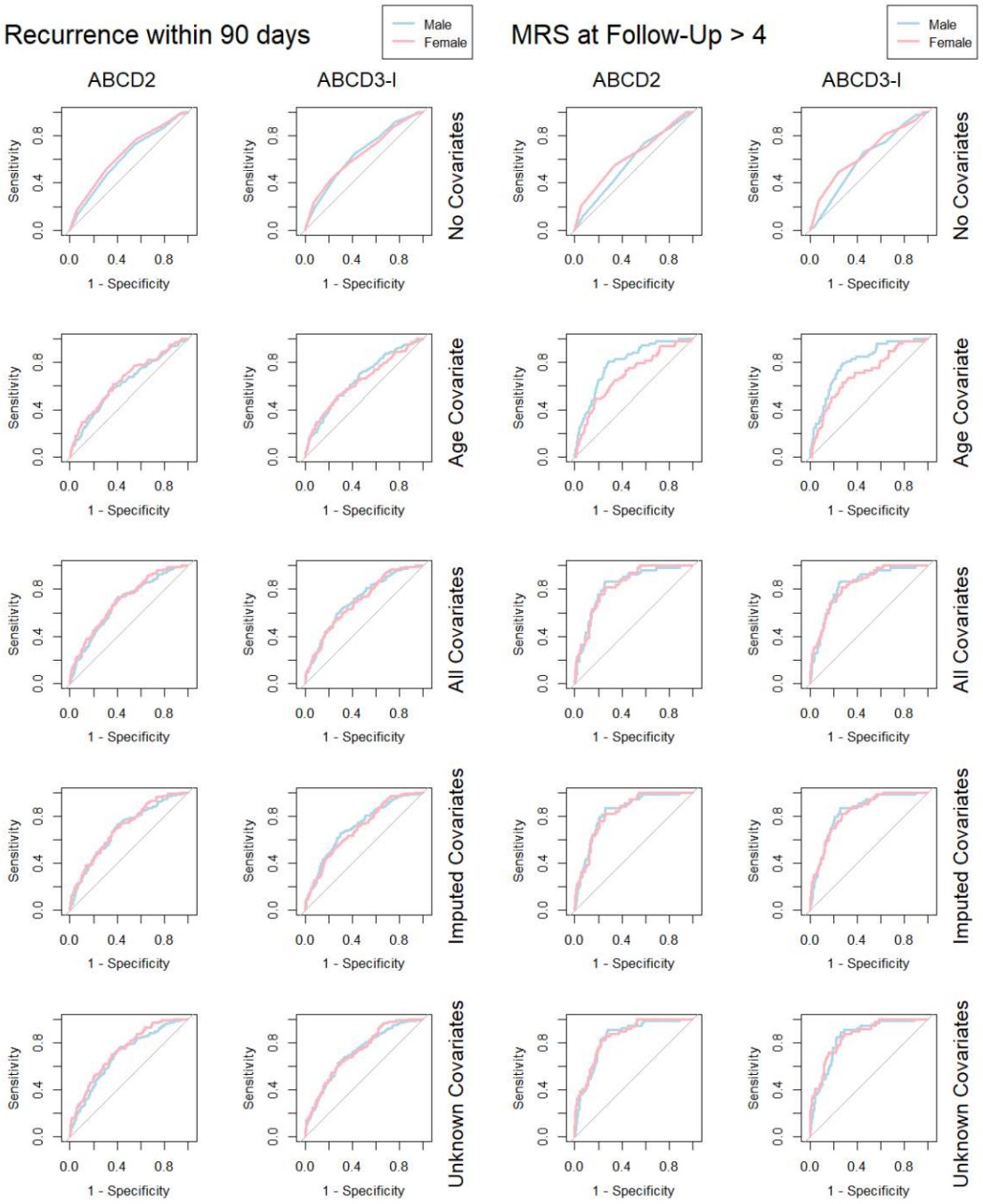
(i) Recurrence 3

(j) Recurrence 4



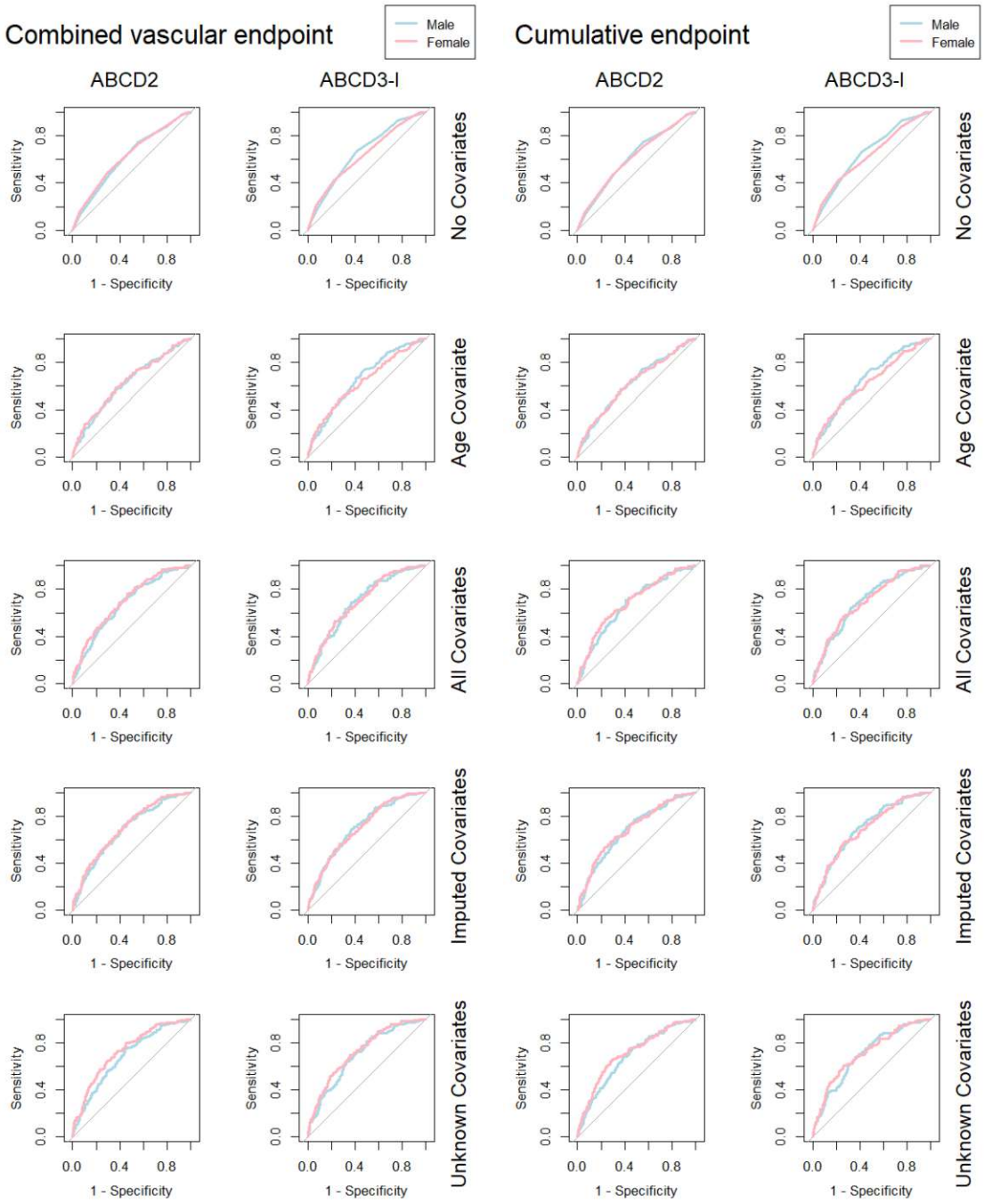
(a) Early Worsening (b) Early Recurrence in Stroke Unit

Figure 2: 'Calculated' Risk Scores ROC Curves



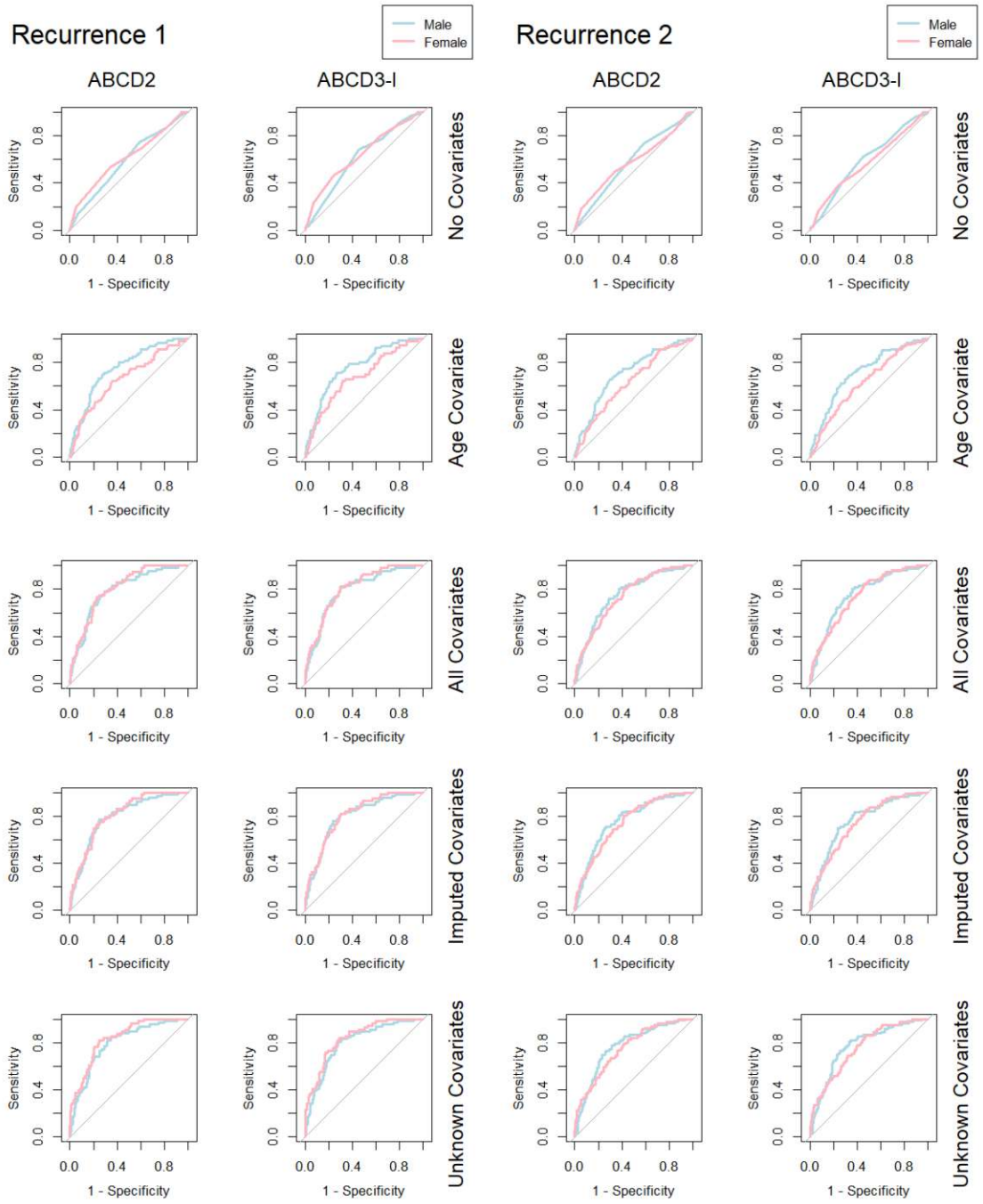
(c) Recurrence within 90 days

(d) MRS at Follow-Up > 4



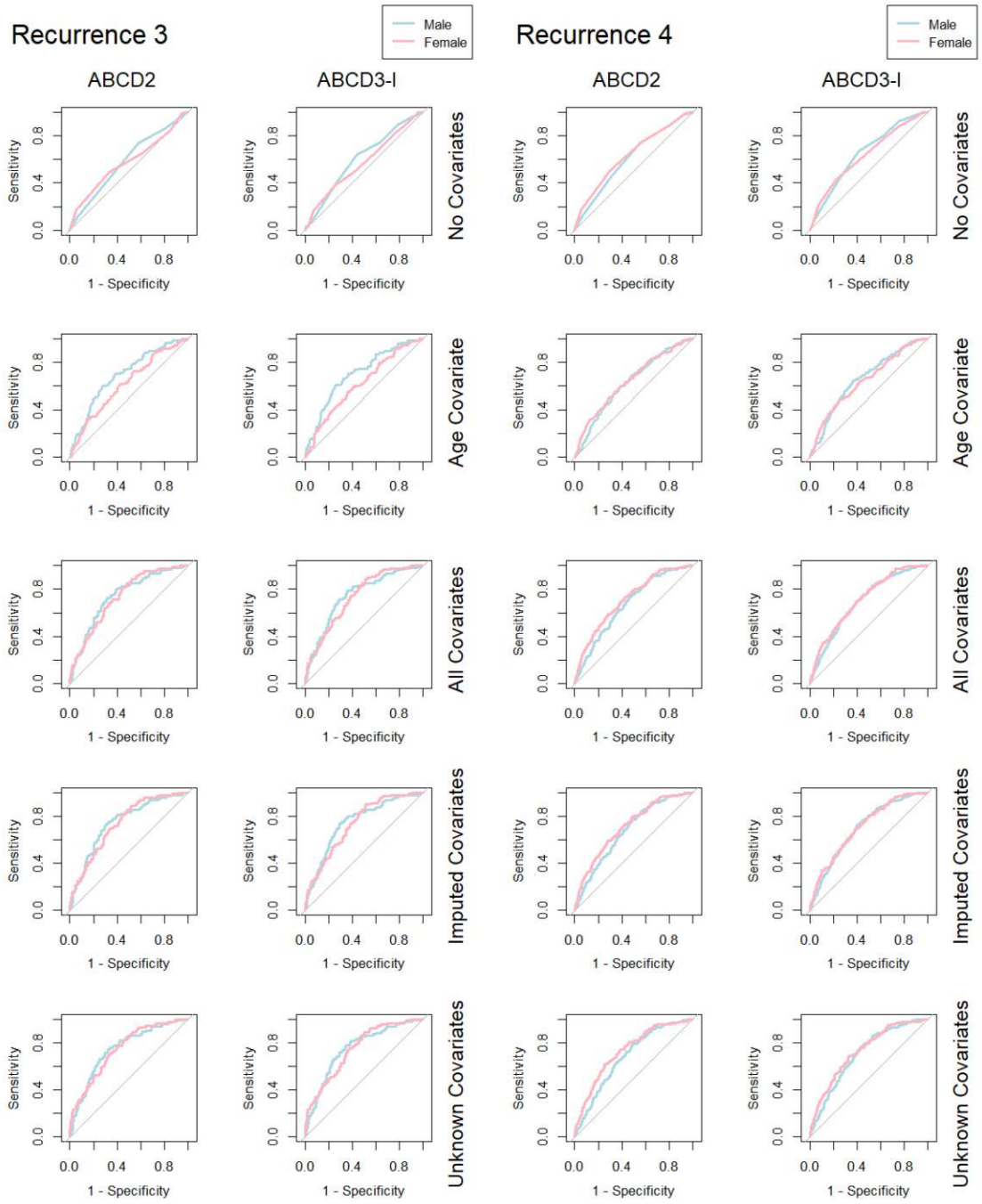
(e) Combined vascular endpoint

(f) Cumulative endpoint



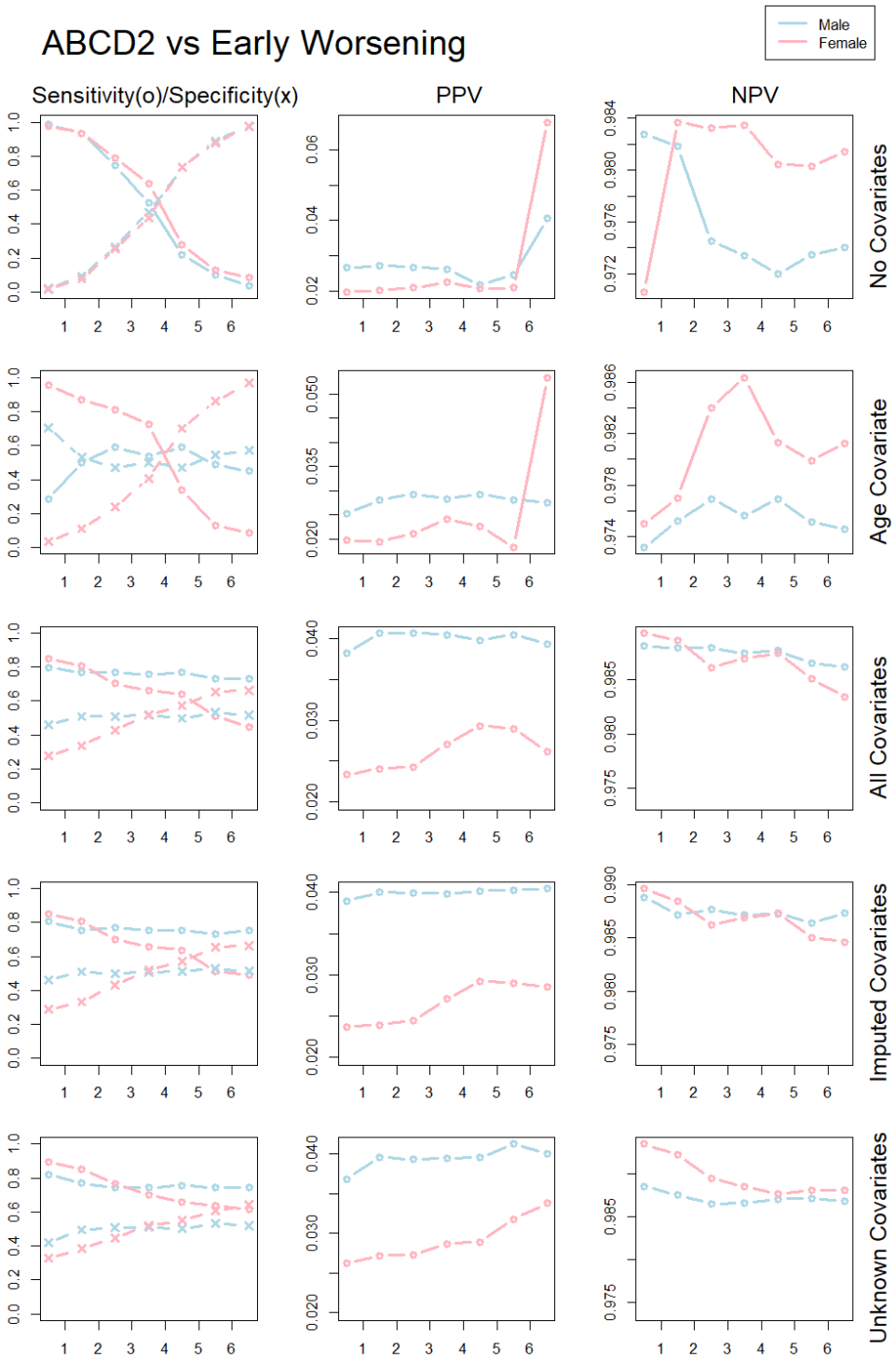
(g) Recurrence 1

(h) Recurrence 2



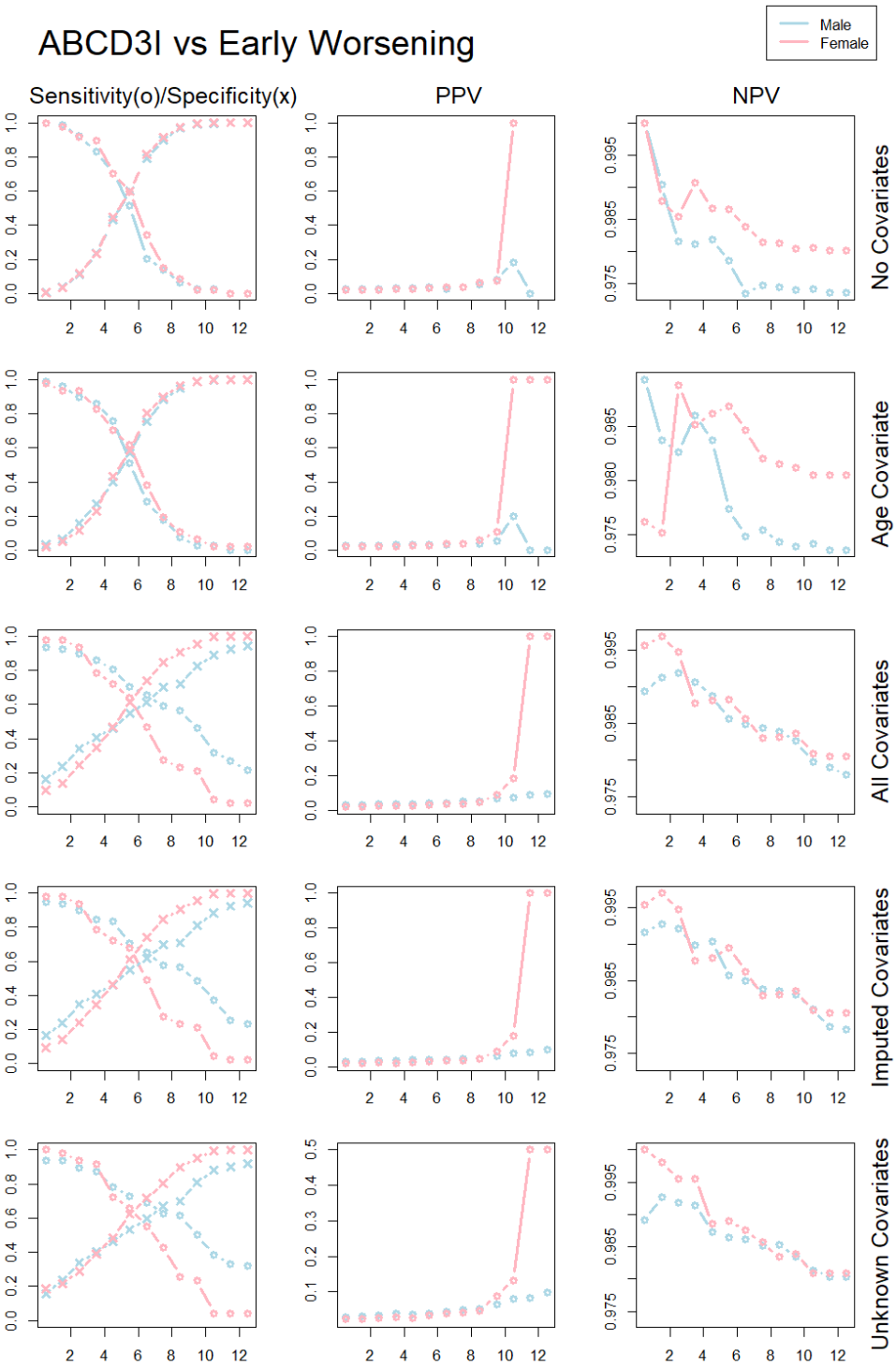
(i) Recurrence 3

(j) Recurrence 4

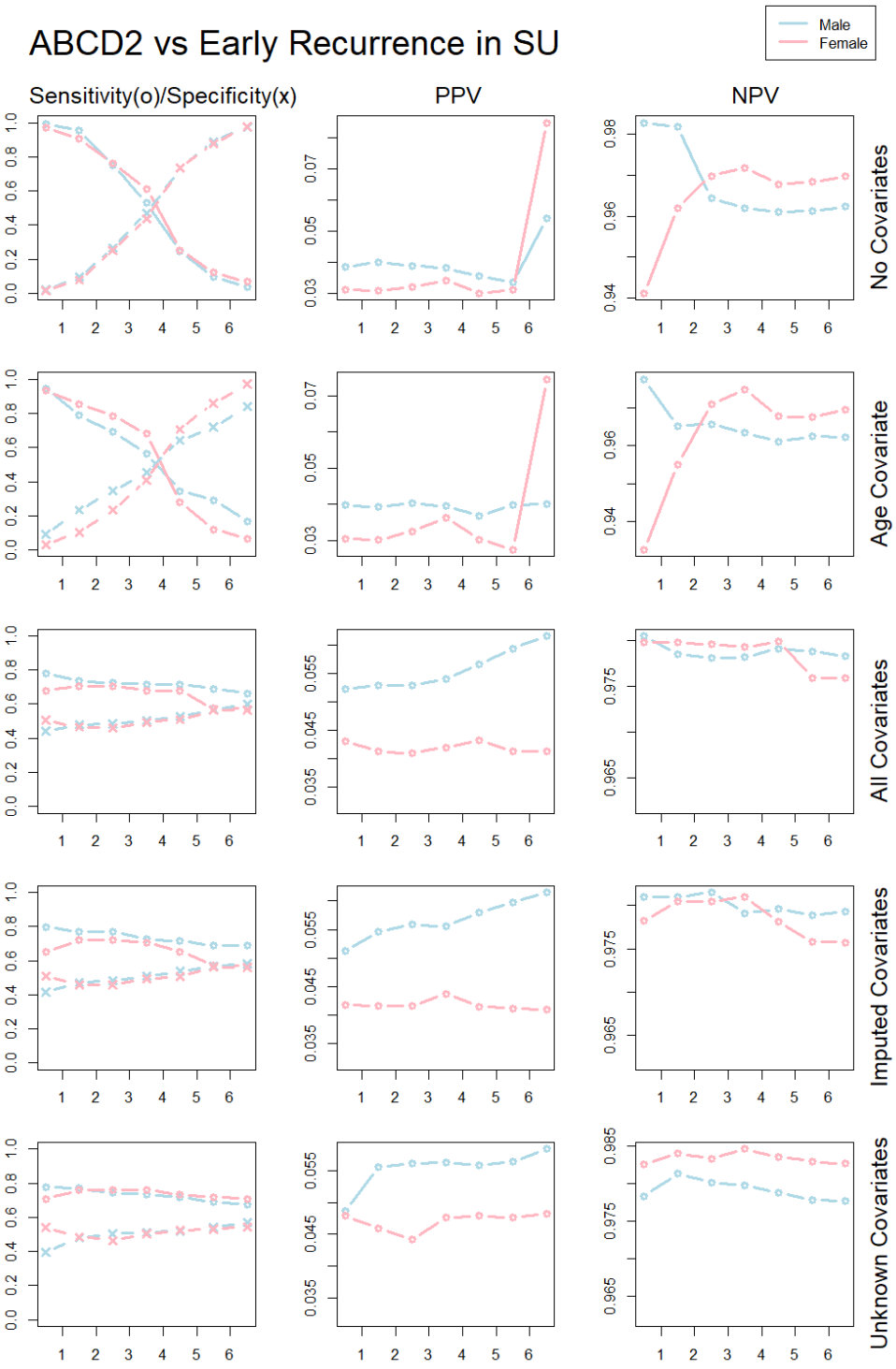


(a) ABCD2 - Early Worsening

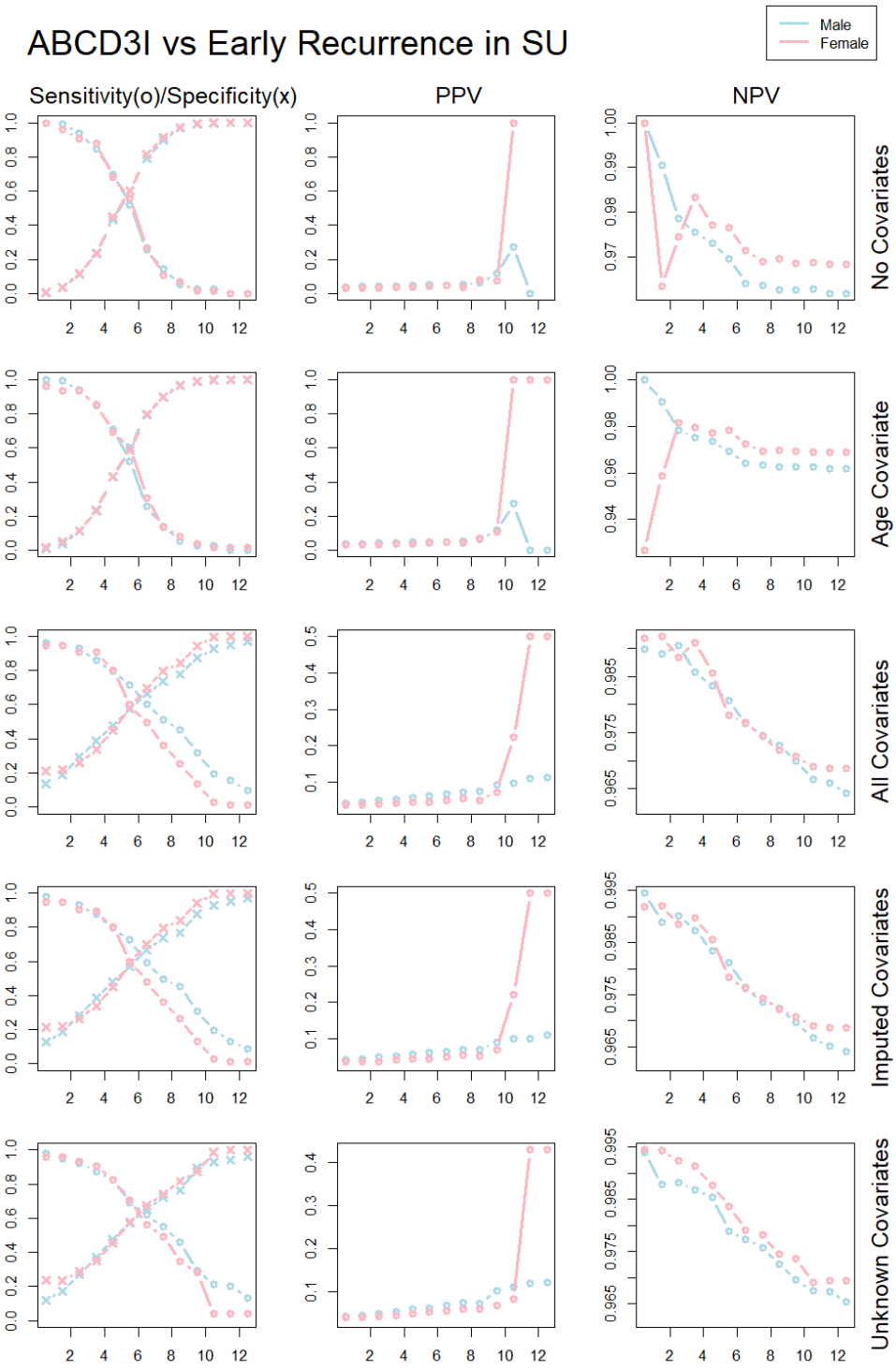
Figure 3: 'Database' Risk Scores Performance Metrics



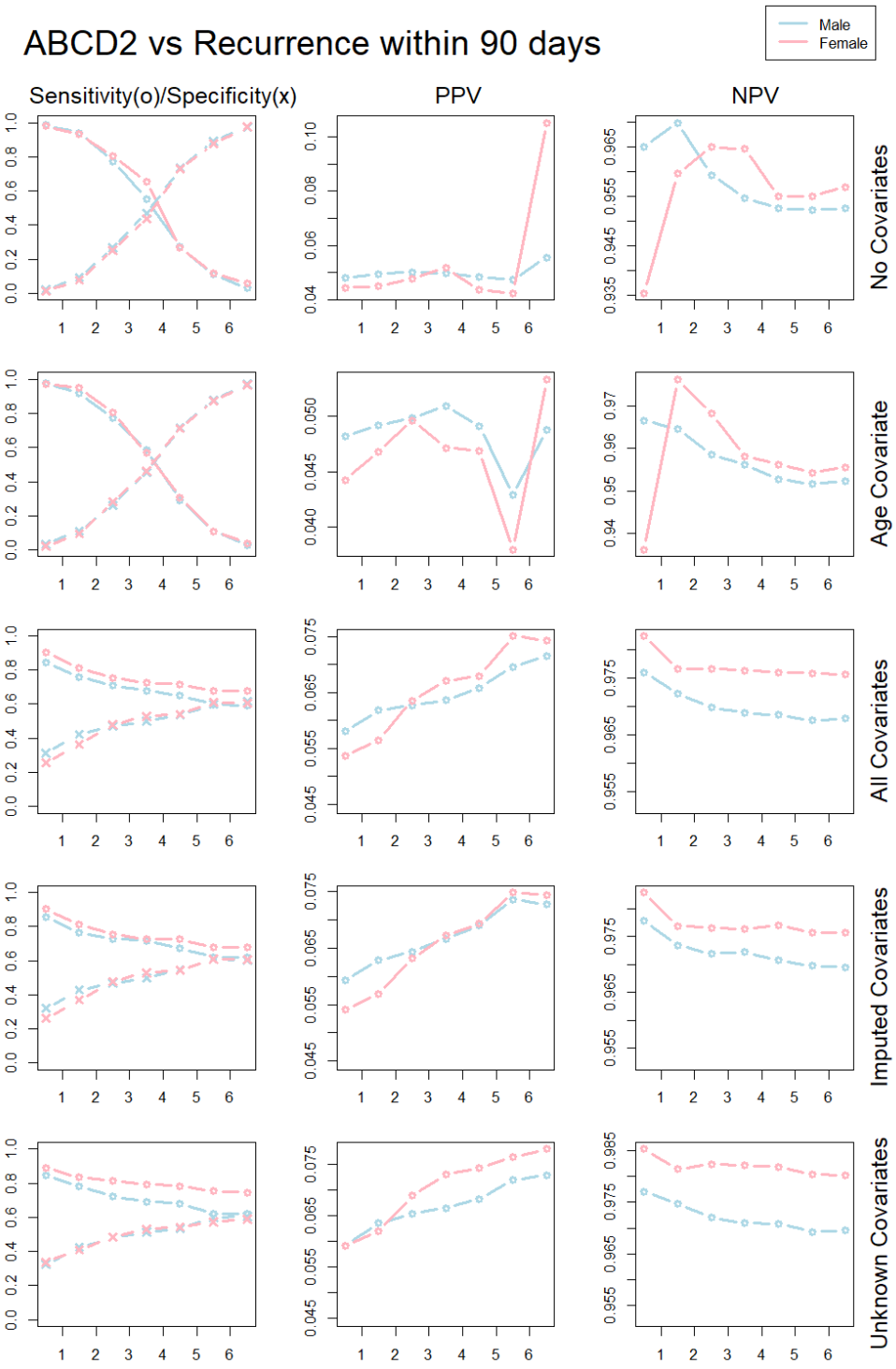
(b) ABCD3-I - Early Worsening



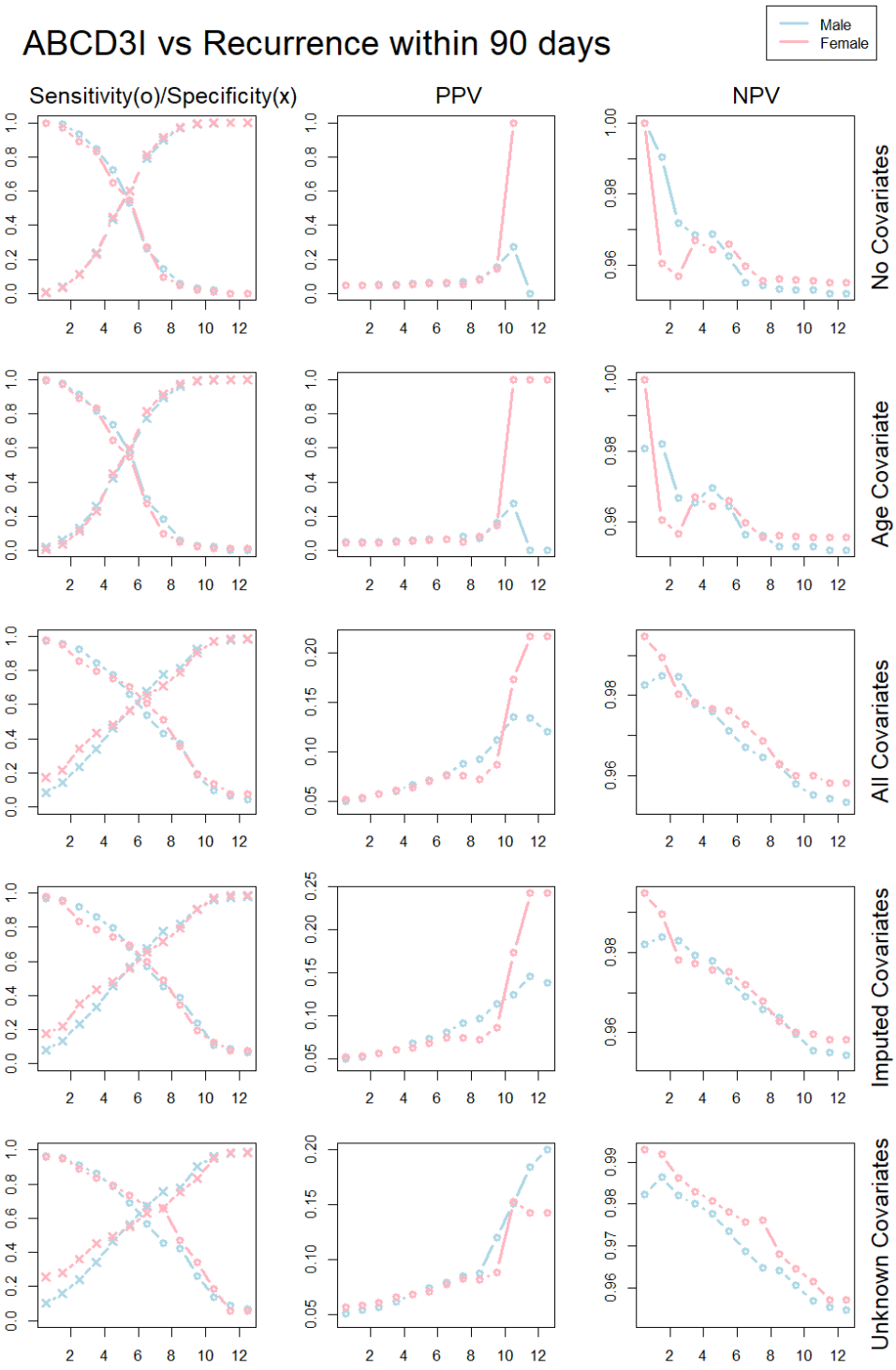
(c) ABCD2 - Early Recurrence in Stroke Unit



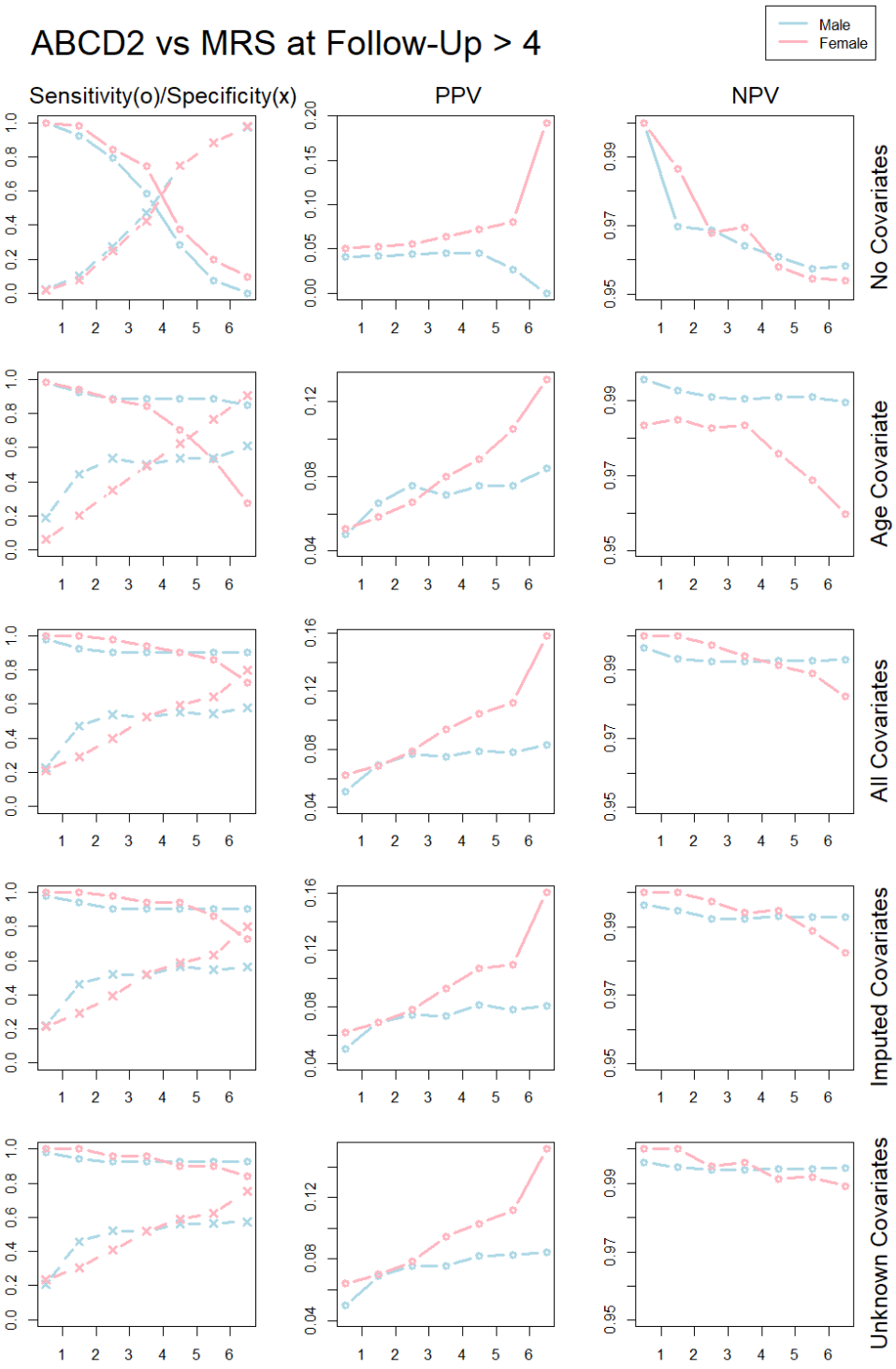
(d) ABCD3-I - Early Recurrence in Stroke Unit



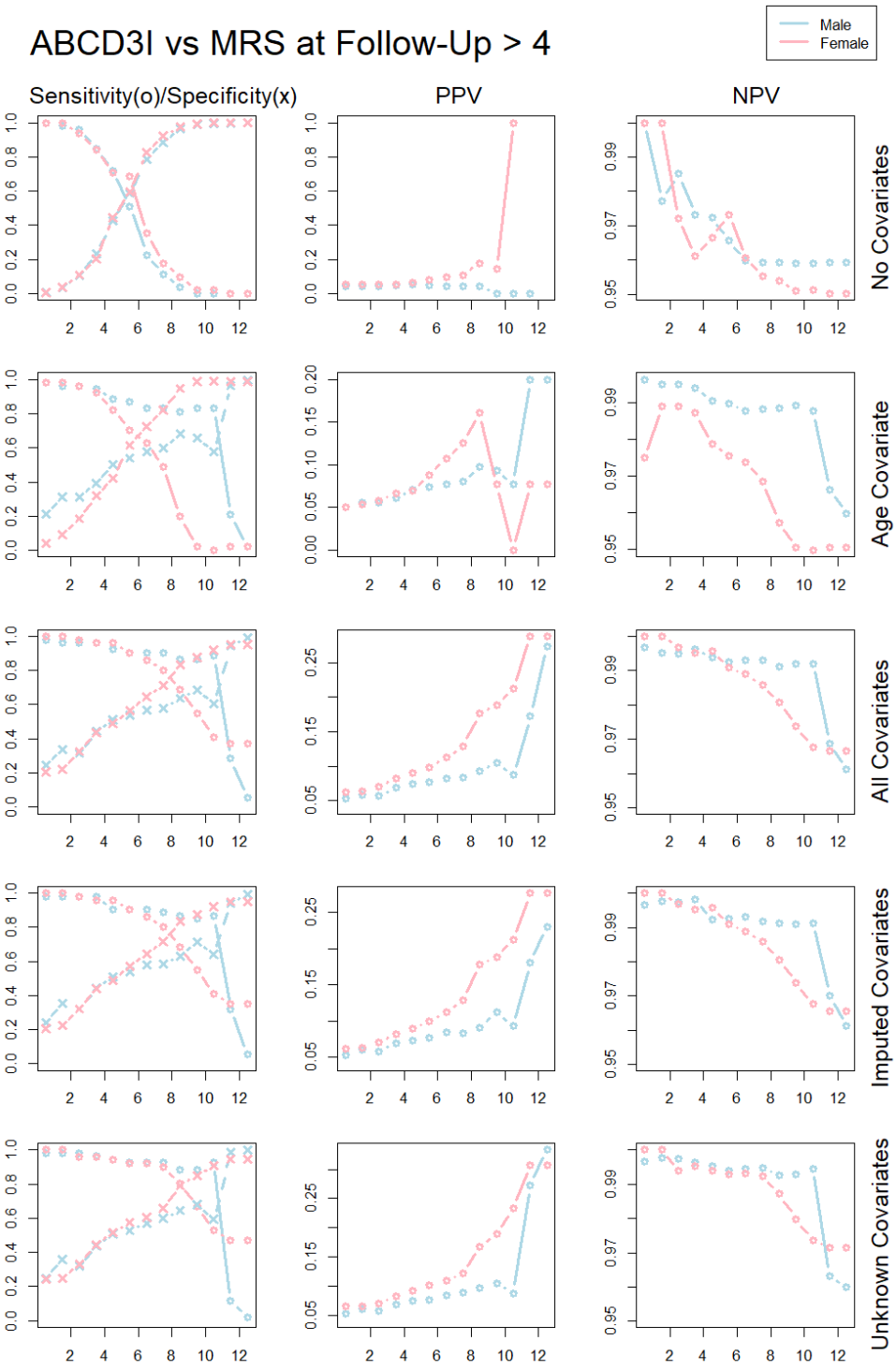
(e) ABCD2 - Recurrence within 90 days



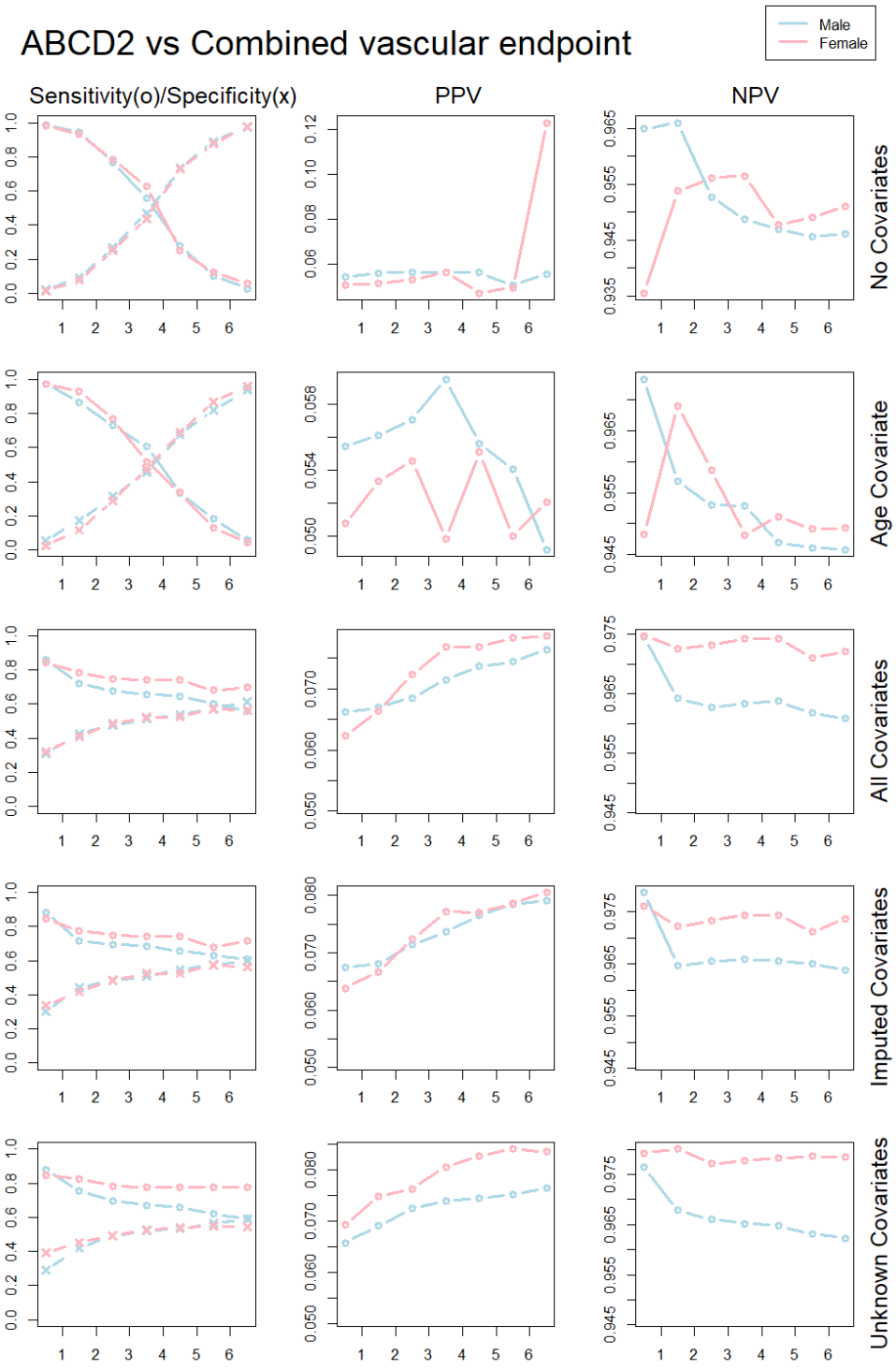
(f) ABCD3-I - Recurrence within 90 days



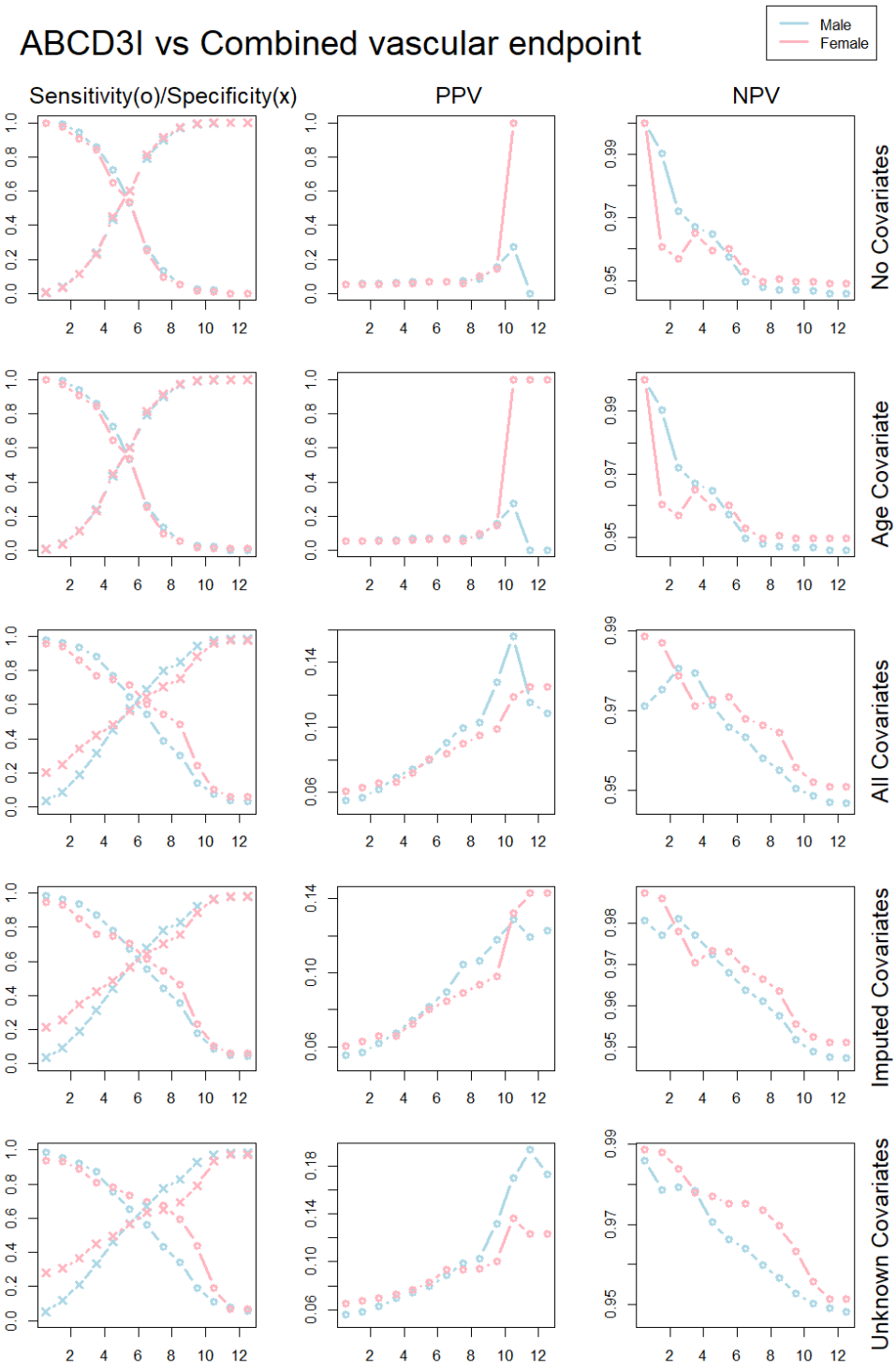
(g) ABCD2 - MRS at Follow-Up > 4



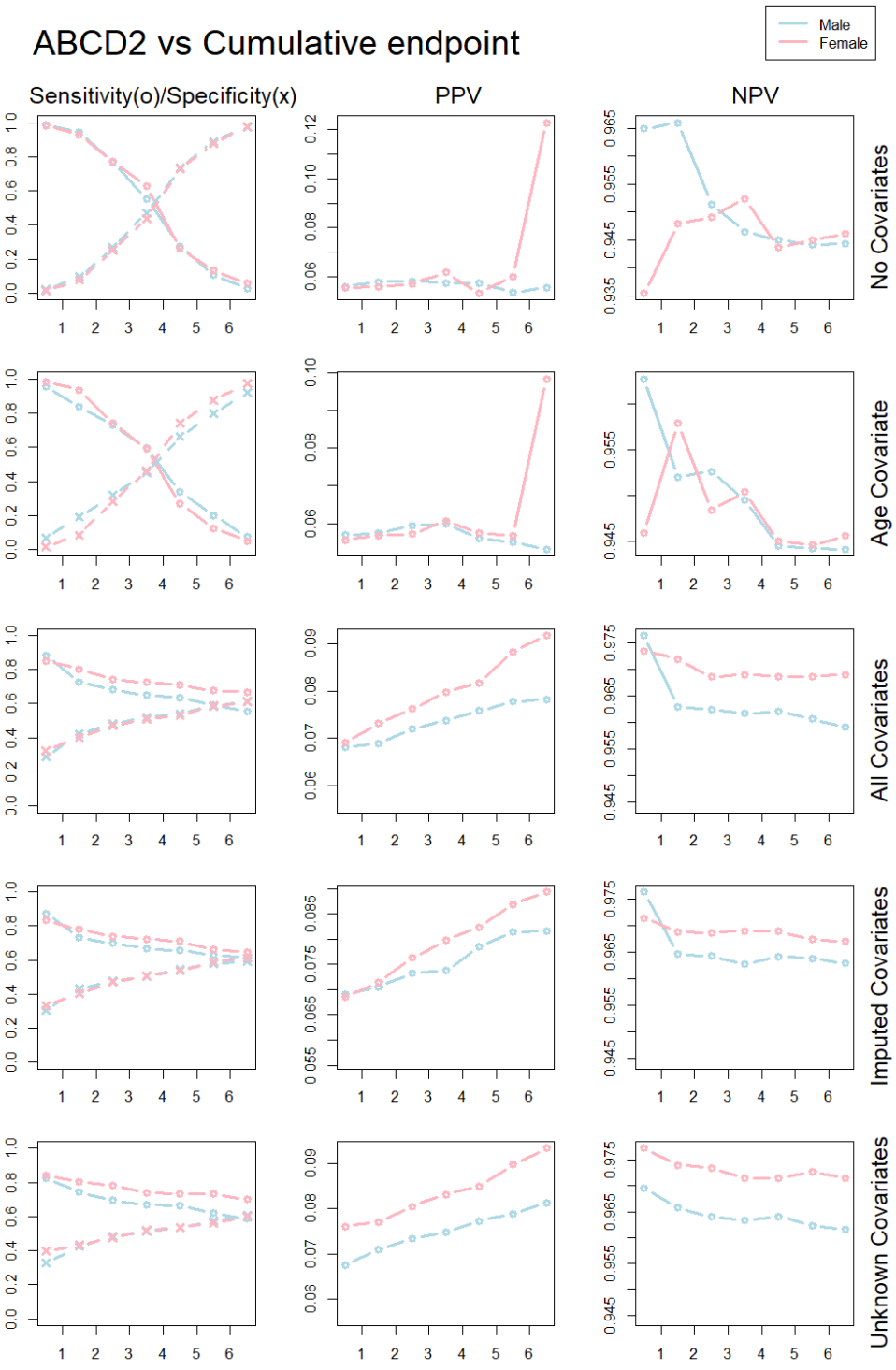
(h) ABCD3-I - MRS at Follow-Up > 4



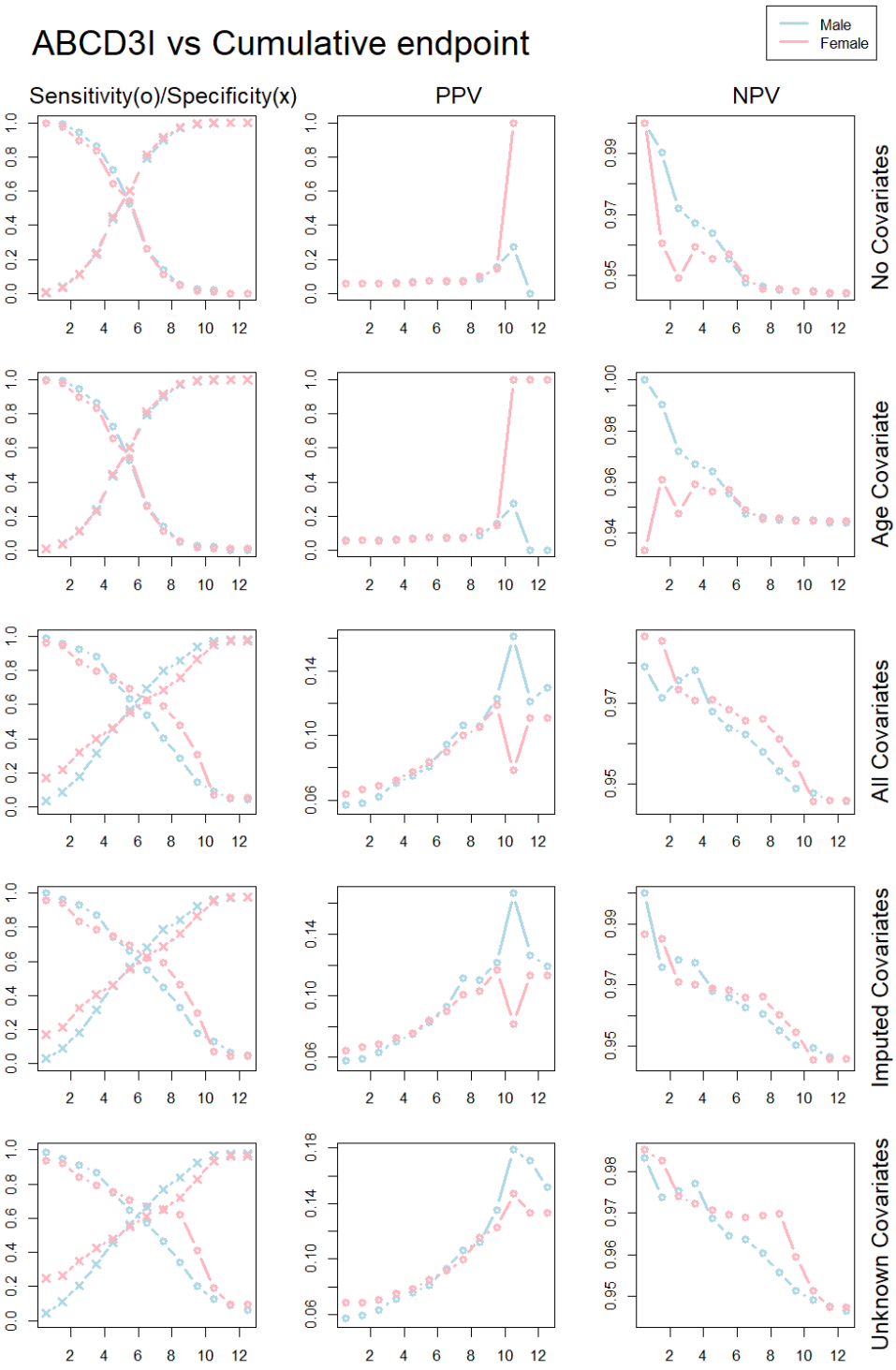
(i) ABCD2 - Combined vascular endpoint



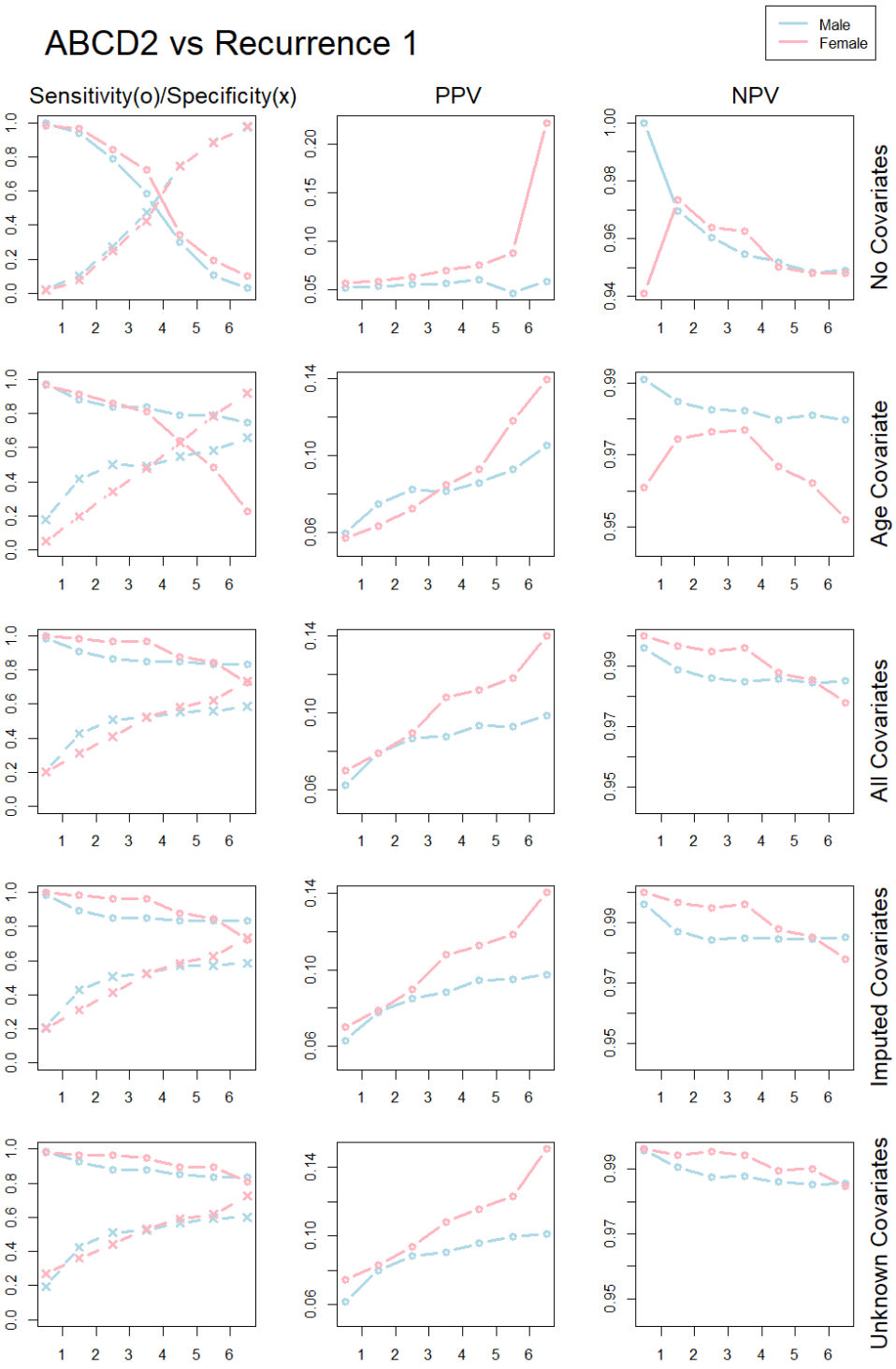
(j) ABCD3-I - Combined vascular endpoint



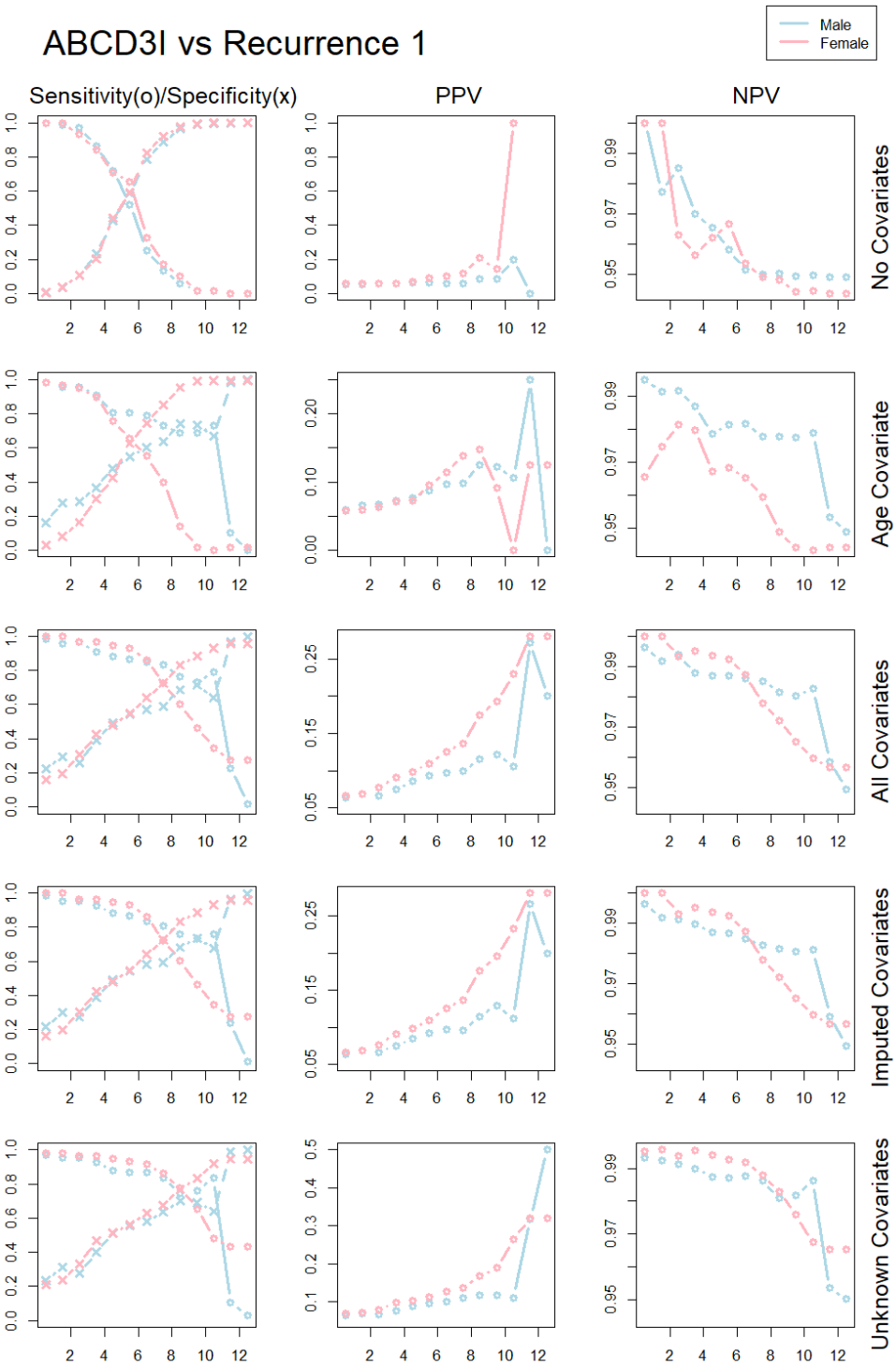
(k) ABCD2 - Cumulative endpoint



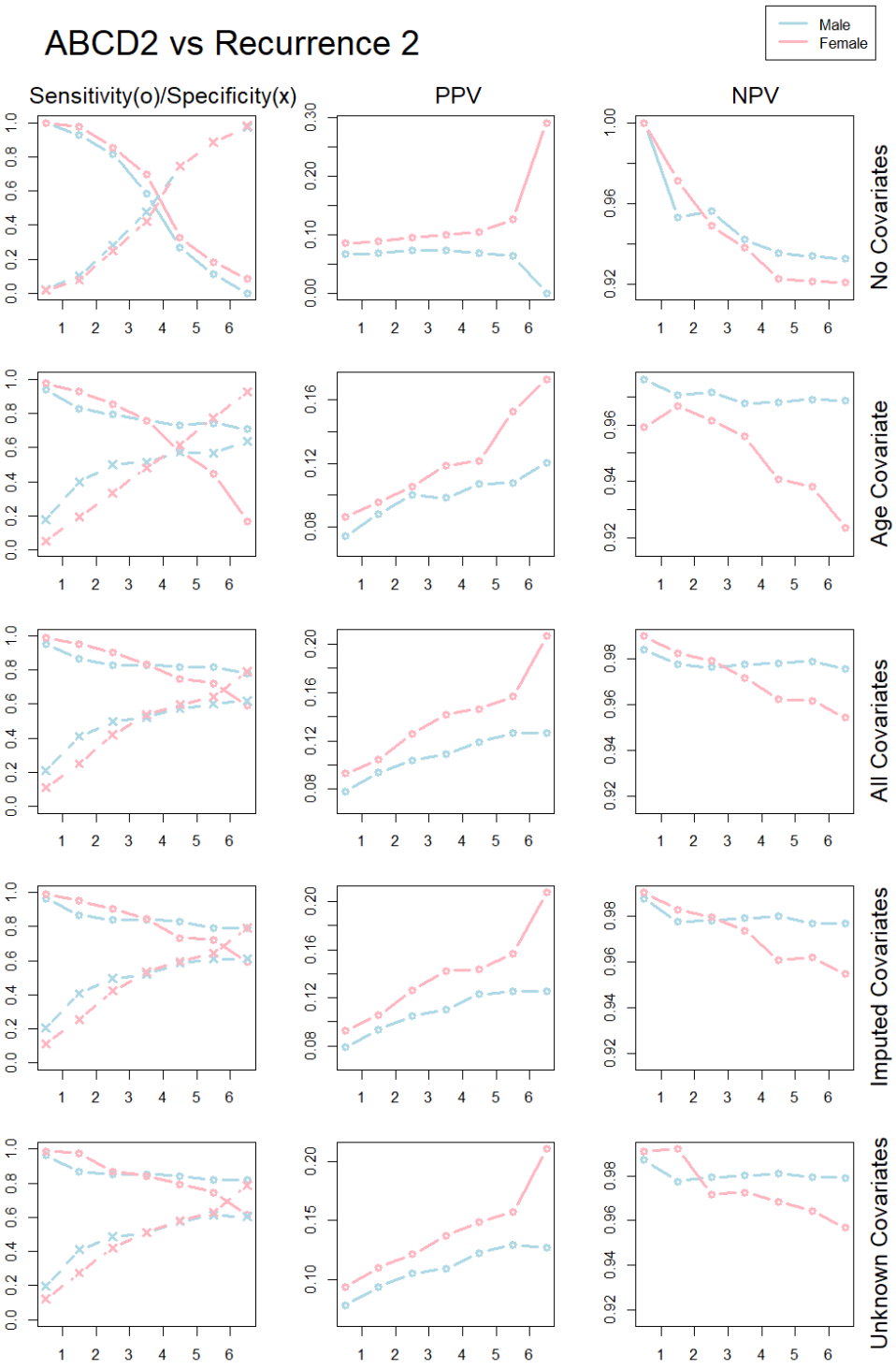
(l) ABCD3-I - Cumulative endpoint



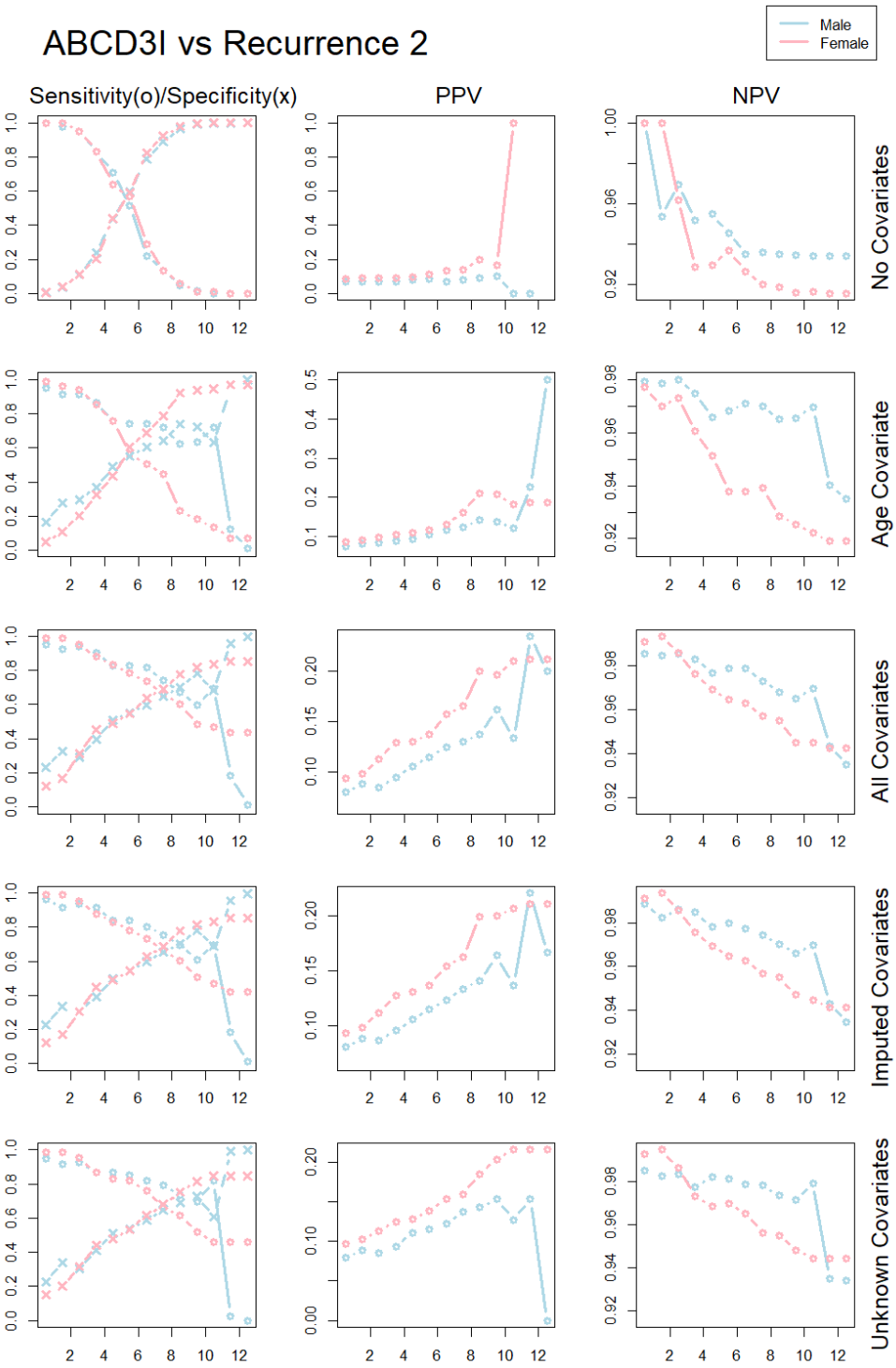
(m) ABCD2 - Recurrence 1



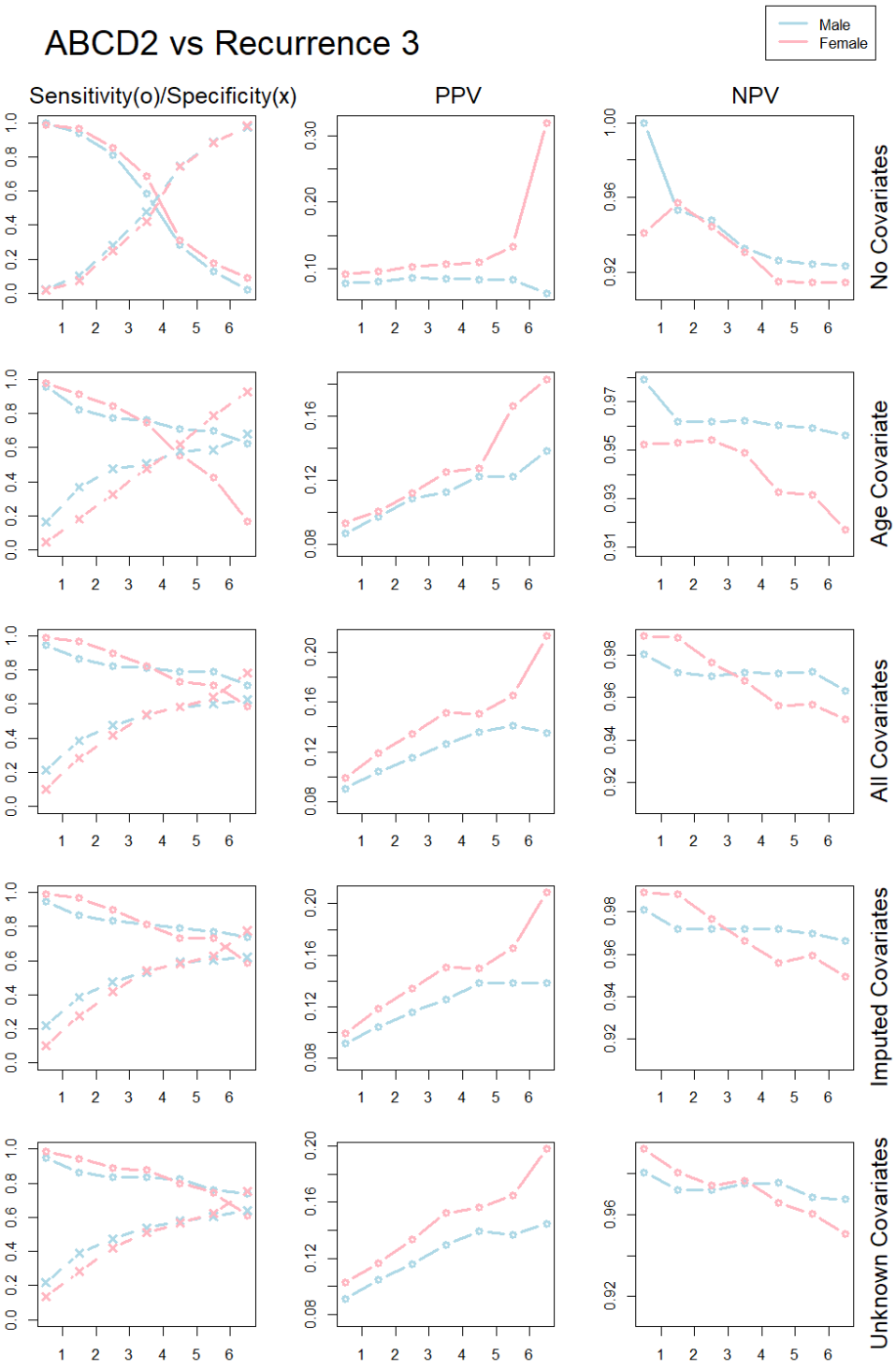
(n) ABCD3-I - Recurrence 1



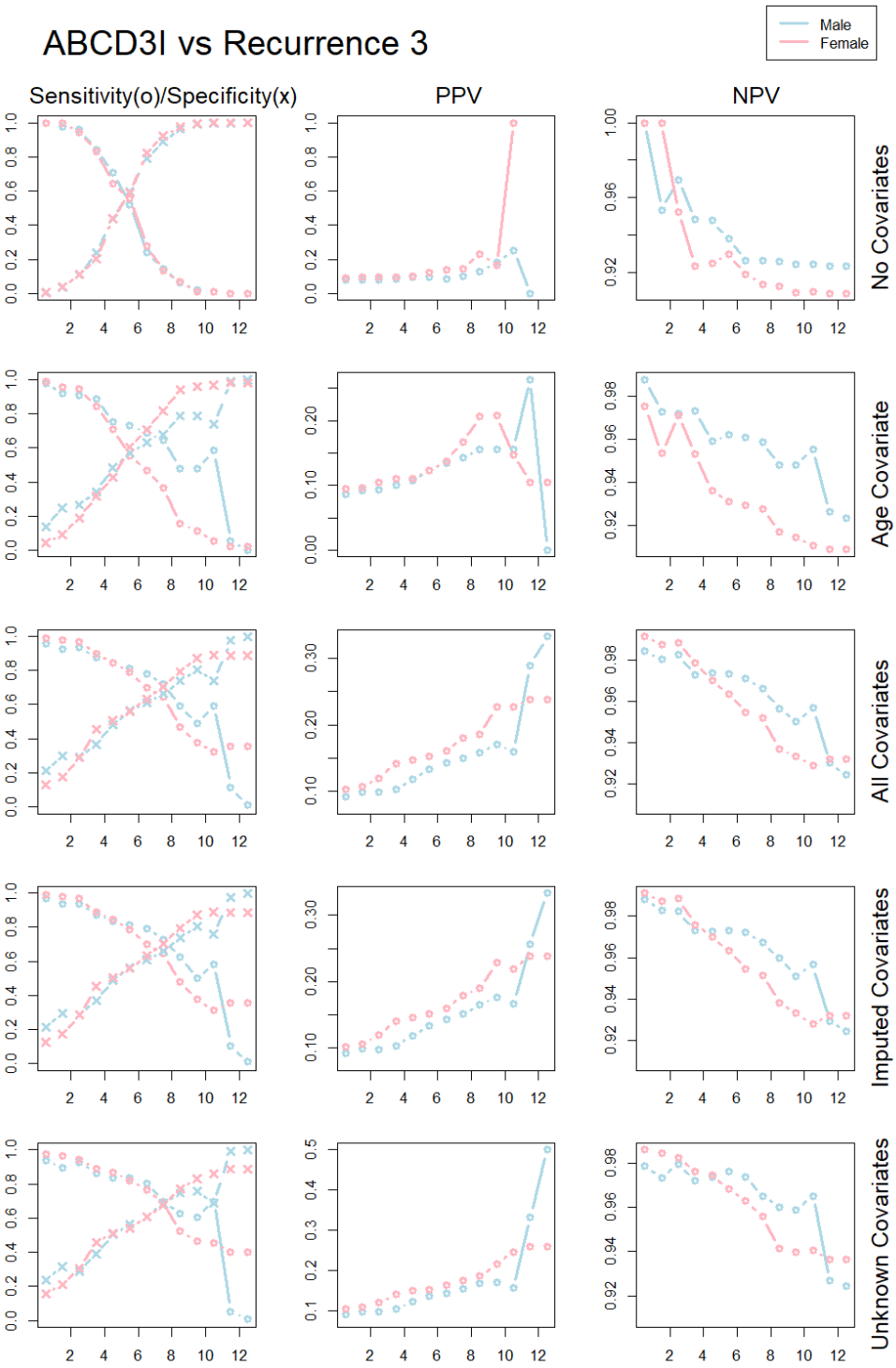
(o) ABCD2 - Recurrence 2



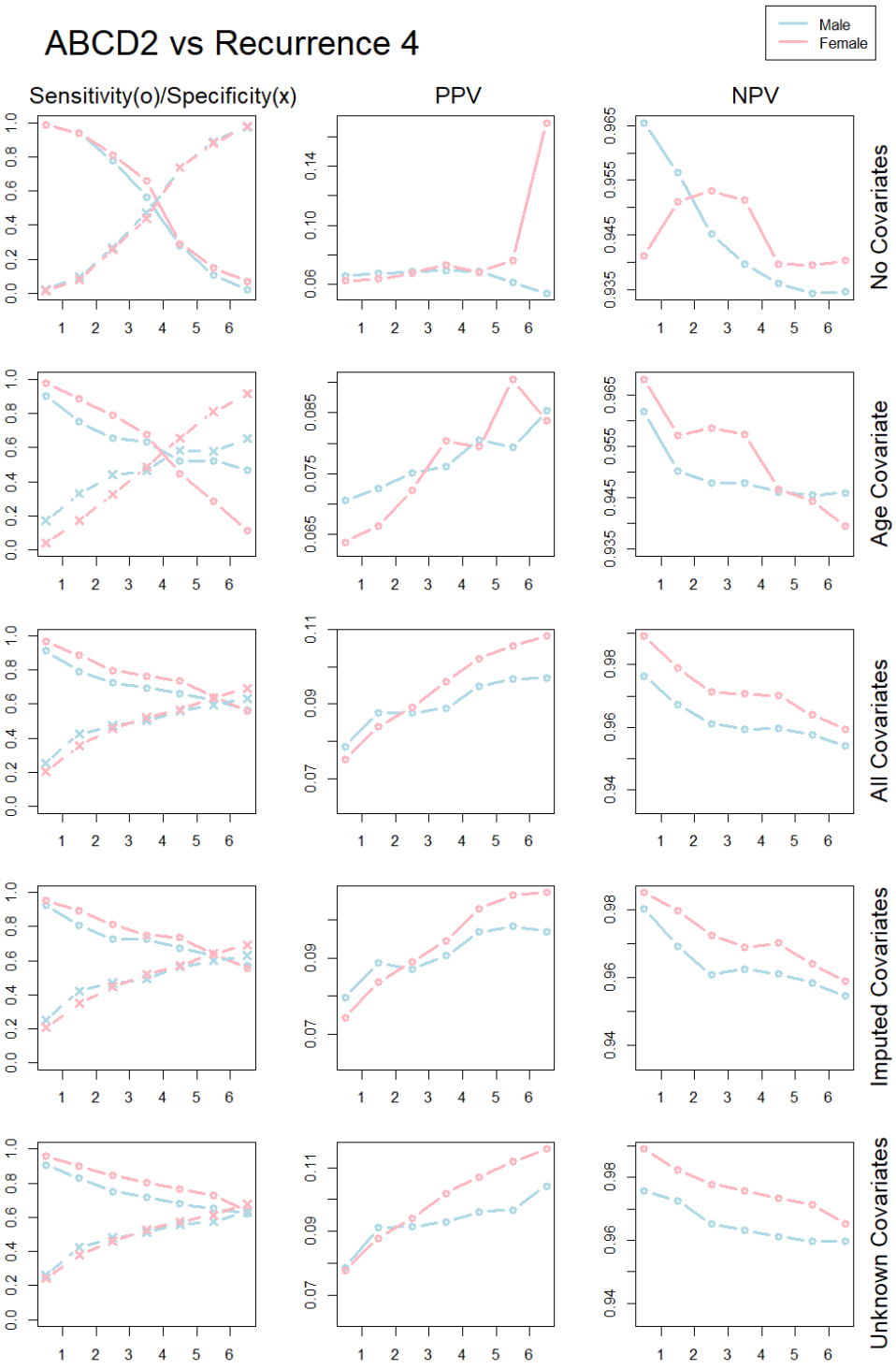
(p) ABCD3-I - Recurrence 2



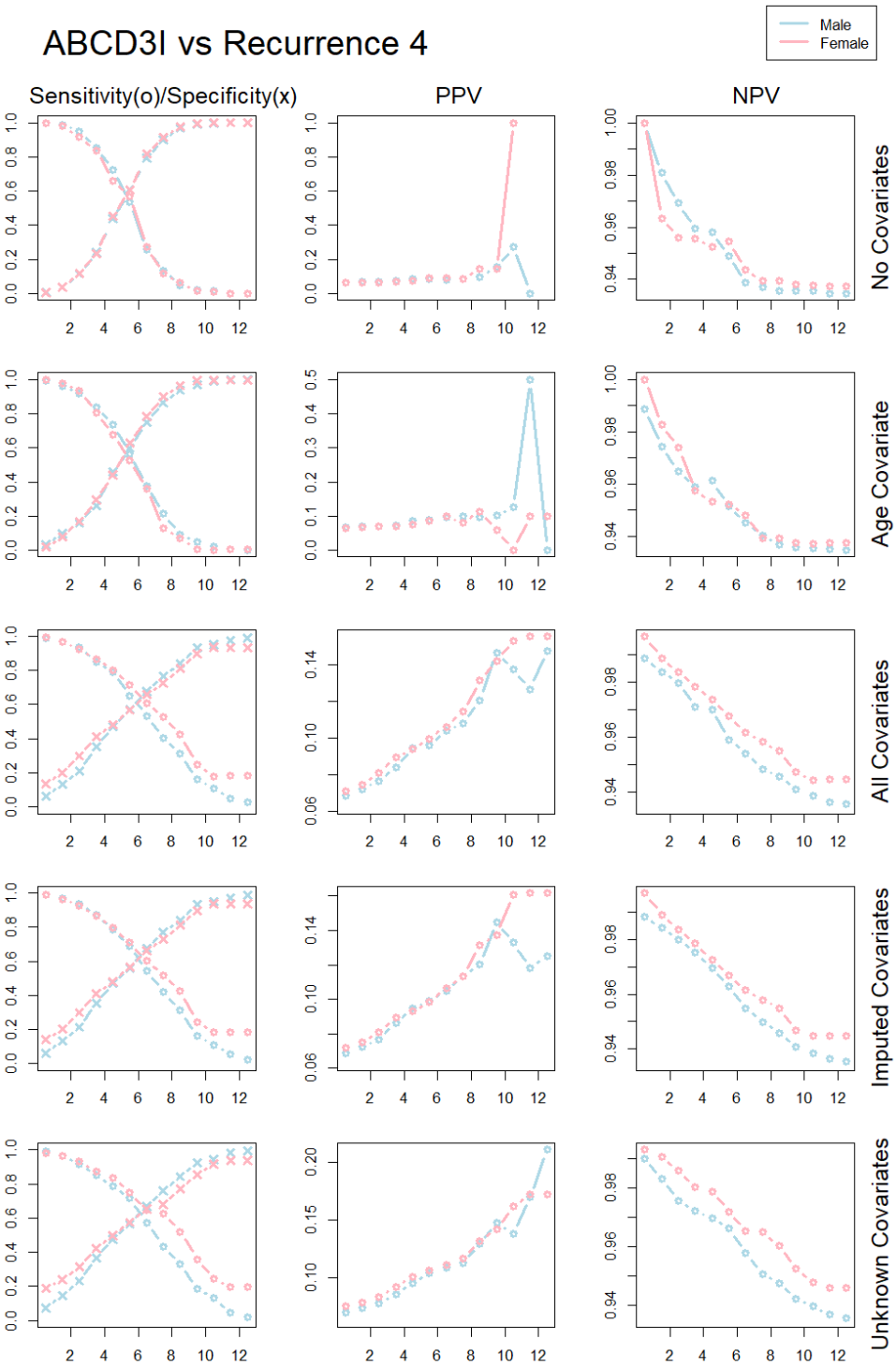
(q) ABCD2 - Recurrence 3



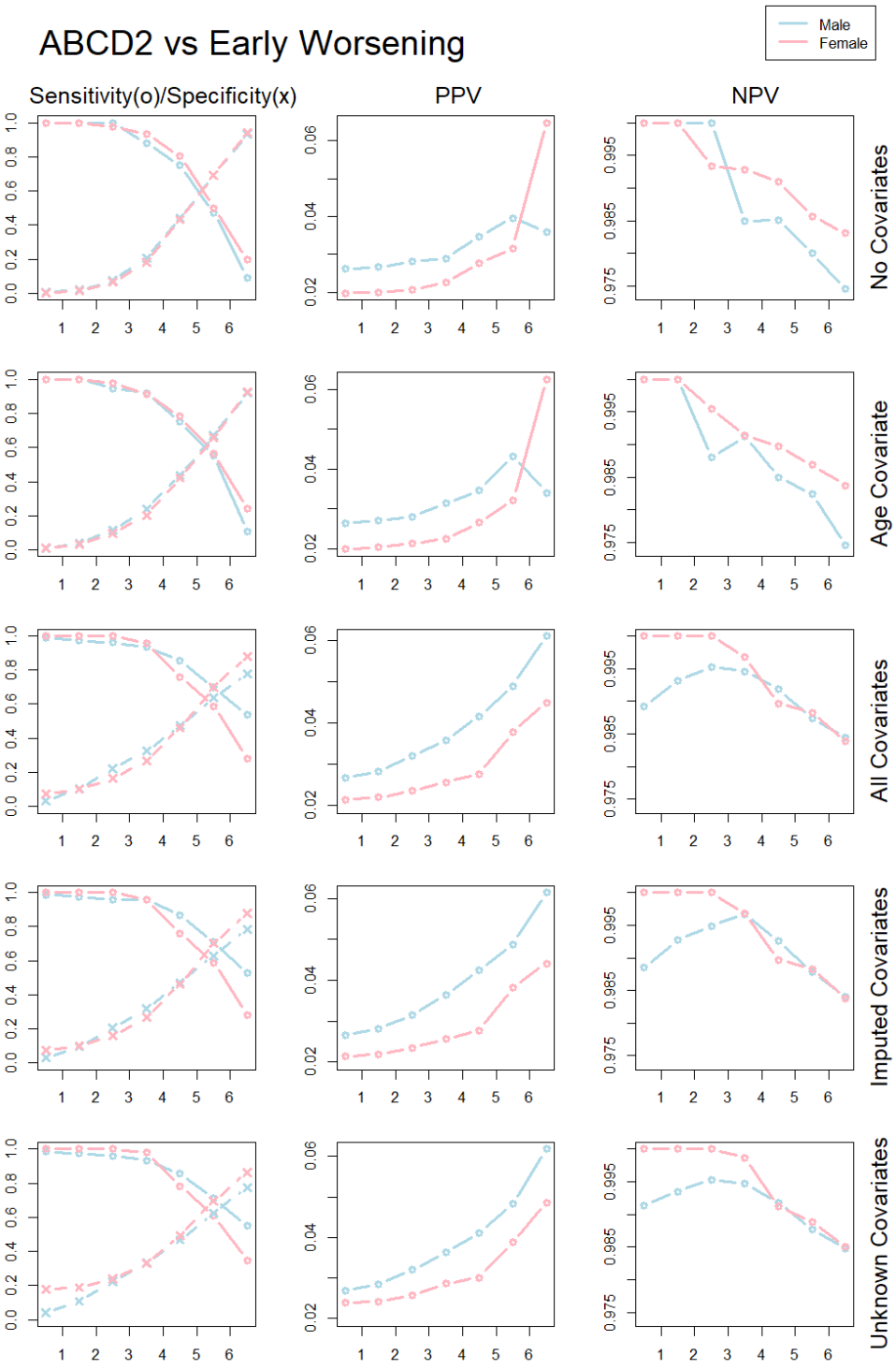
(r) ABCD3-I - Recurrence 3



(s) ABCD2 - Recurrence 4

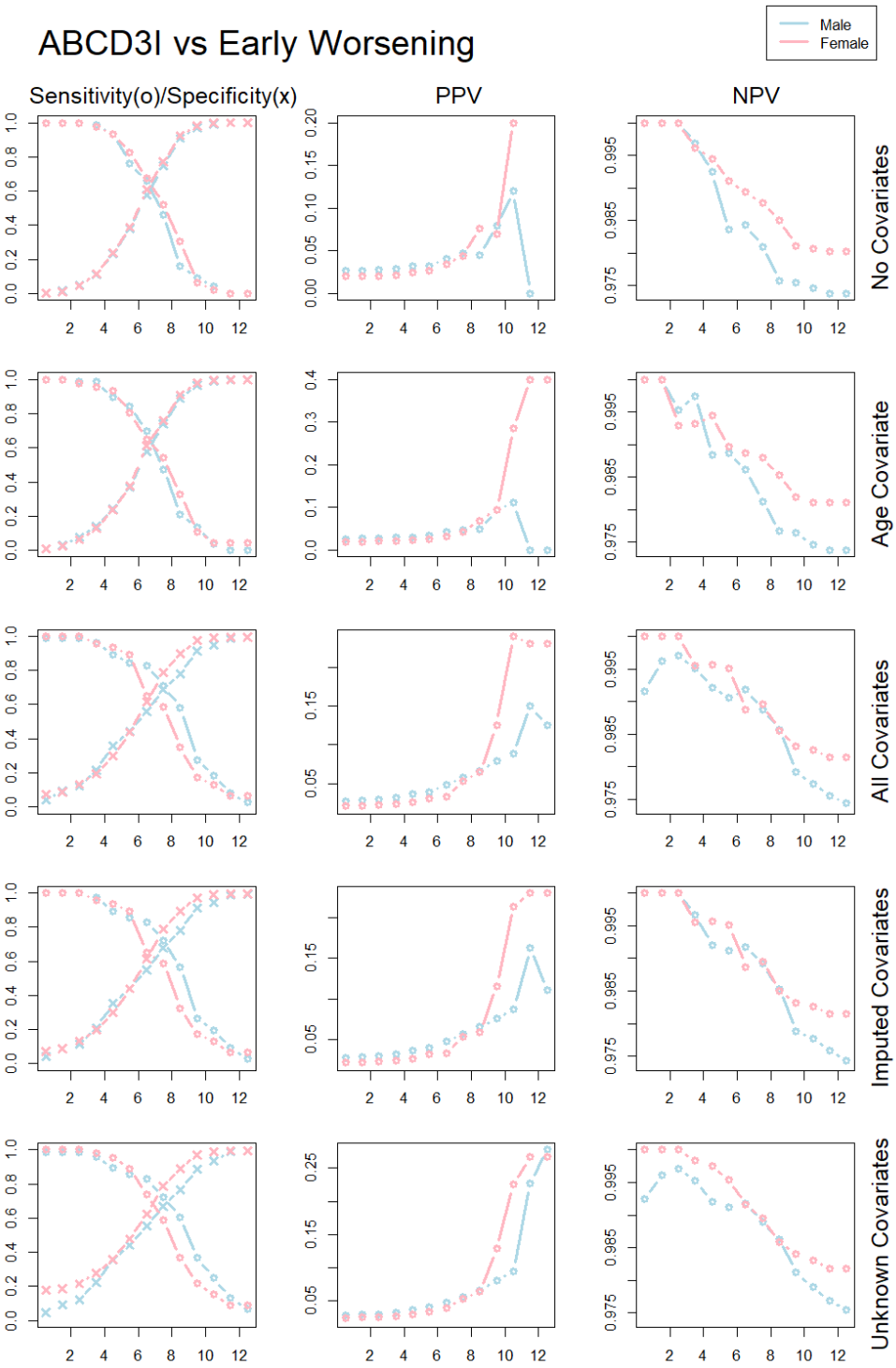


(t) ABCD3-I - Recurrence 4

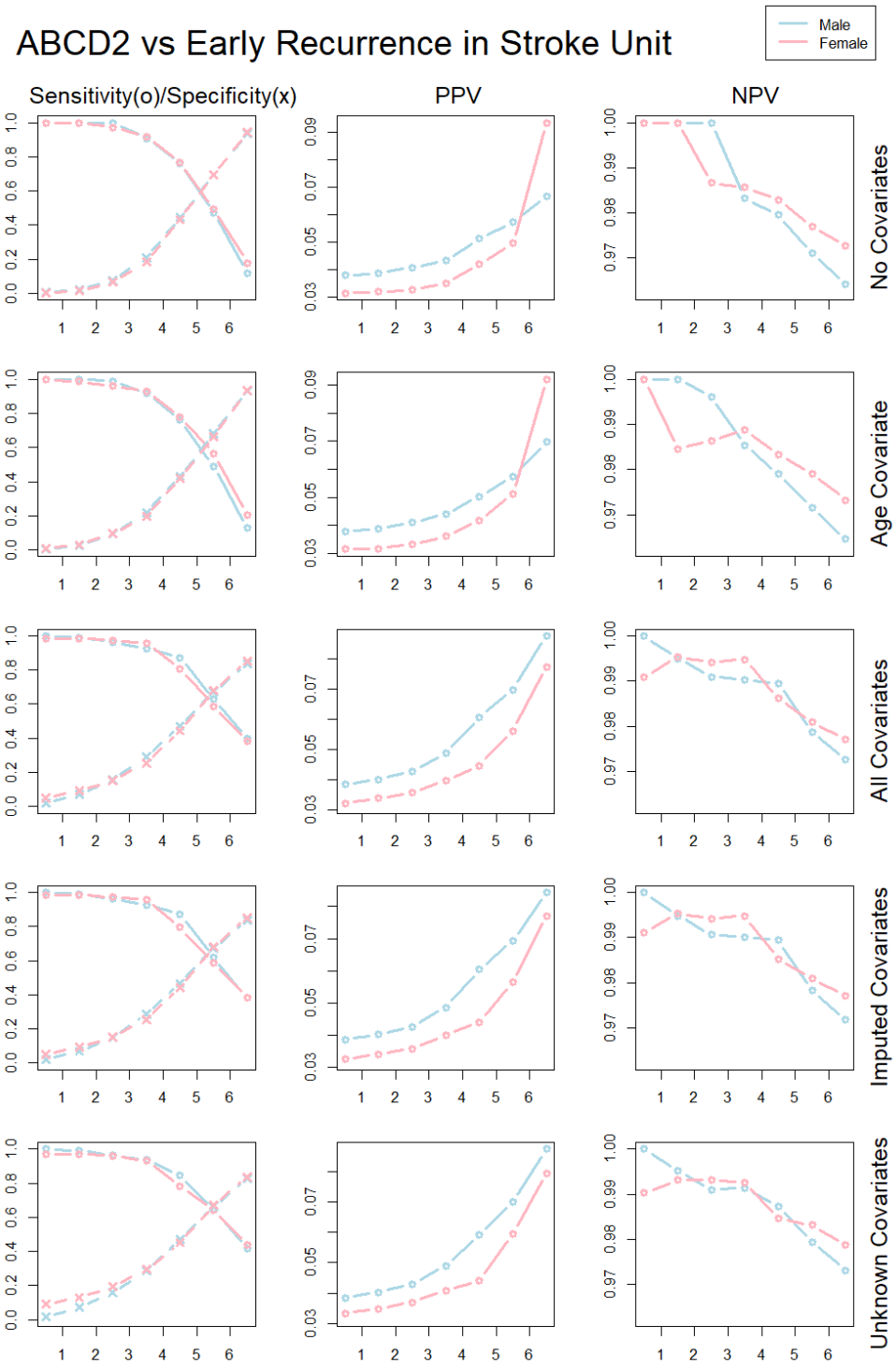


(a) ABCD2 - Early Worsening

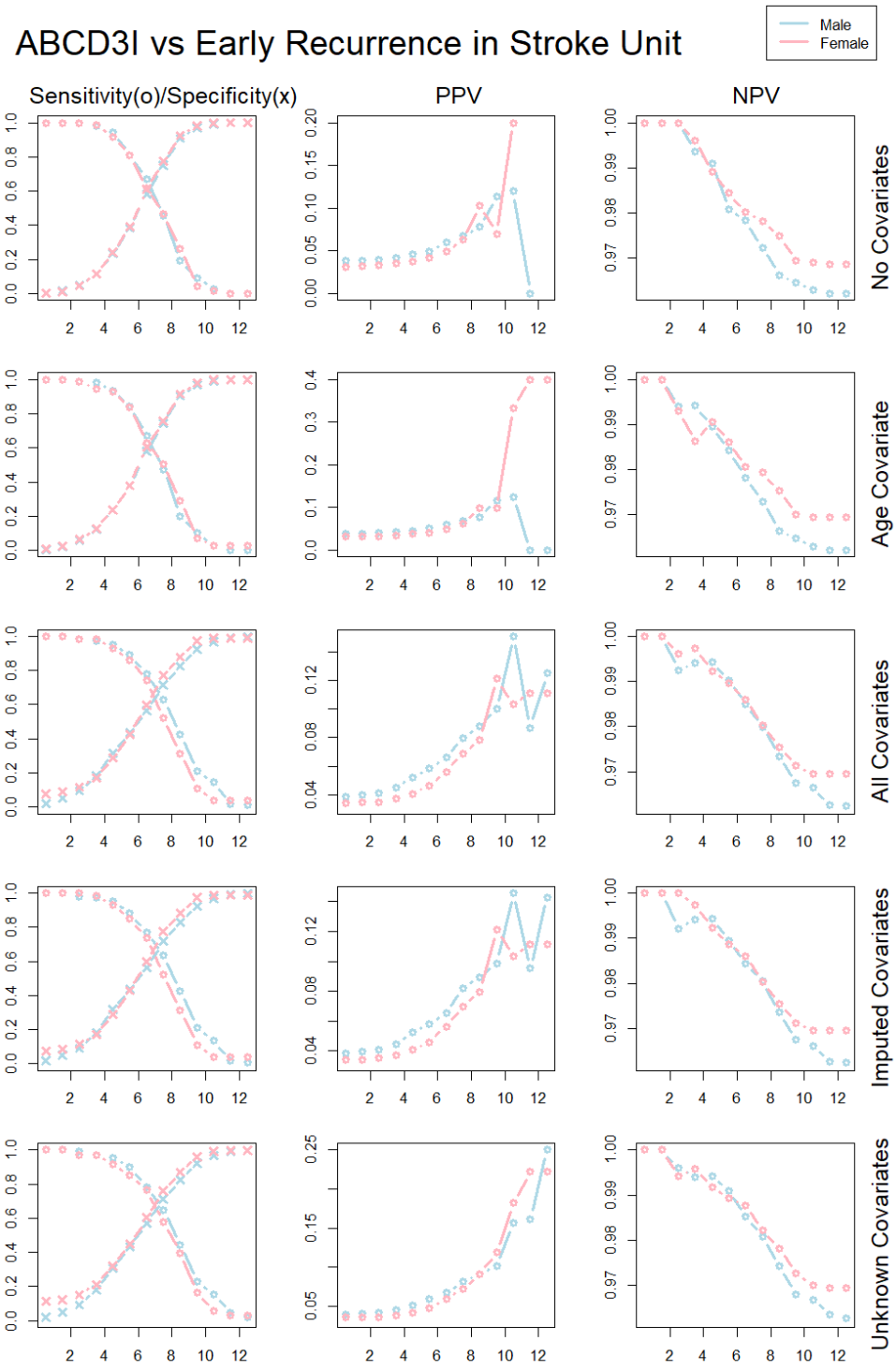
Figure 4: 'Calculated' Risk Scores Performance Metrics



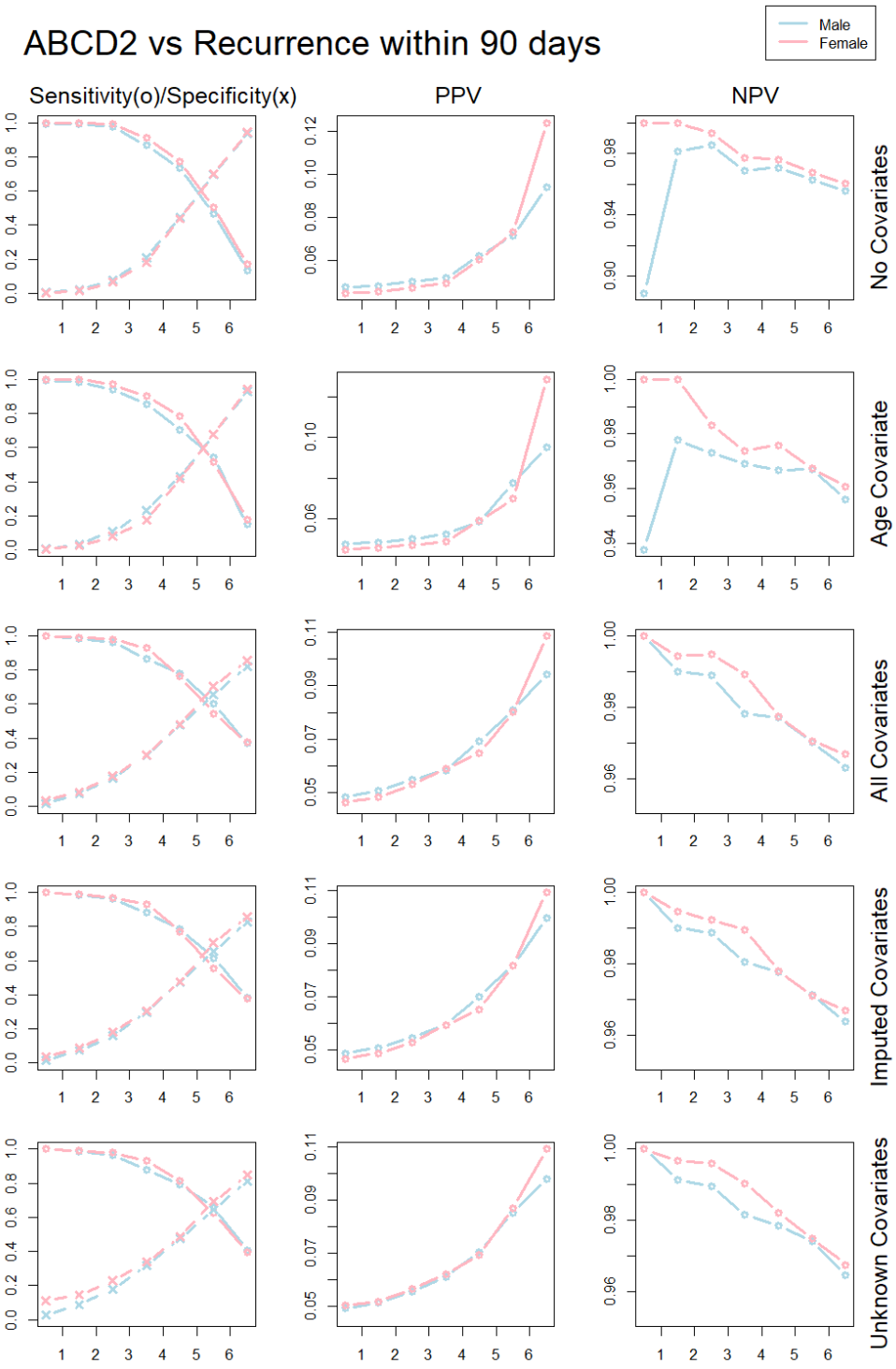
(b) ABCD3-I - Early Worsening



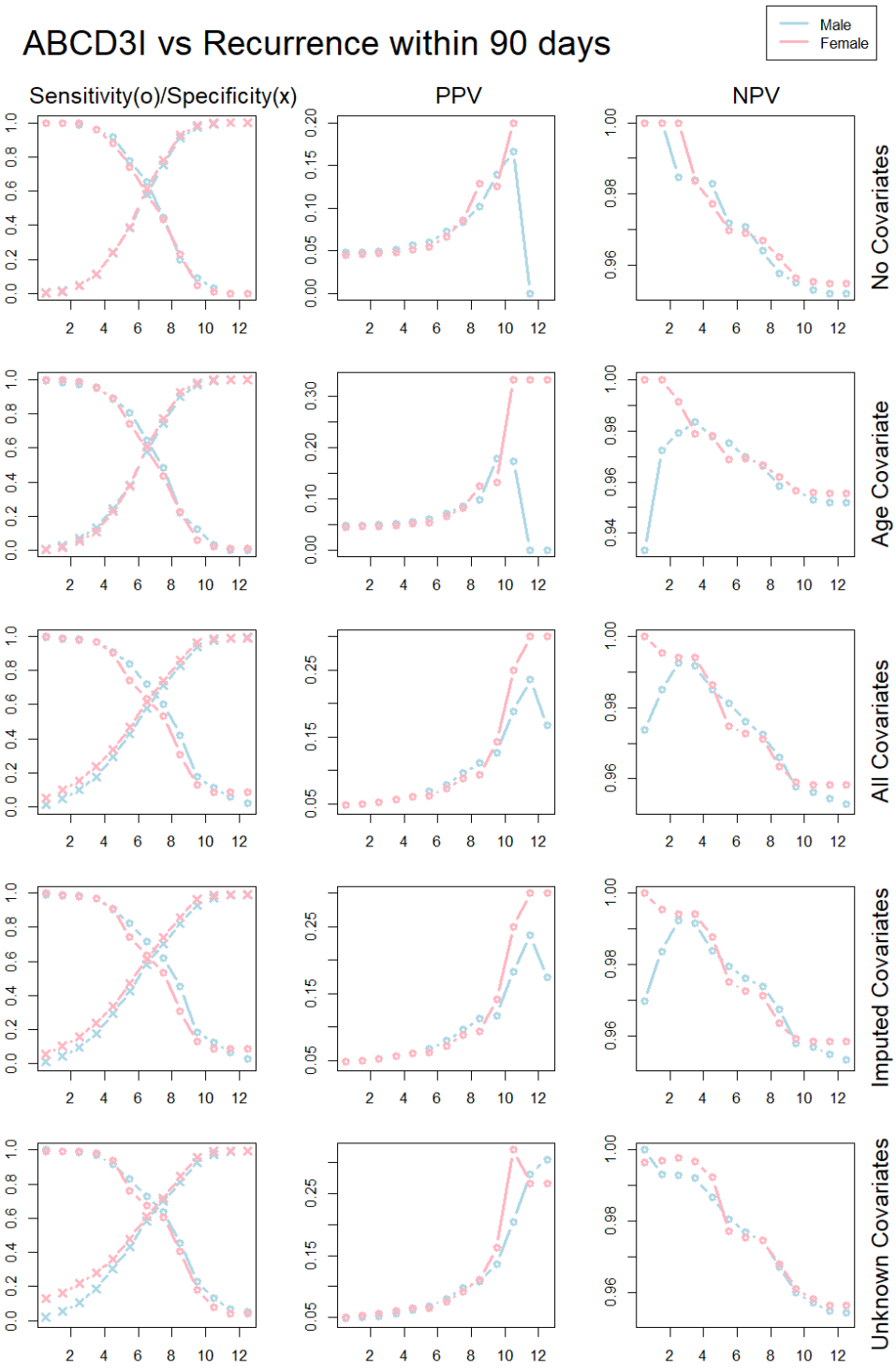
(c) ABCD2 - Early Recurrence in Stroke Unit



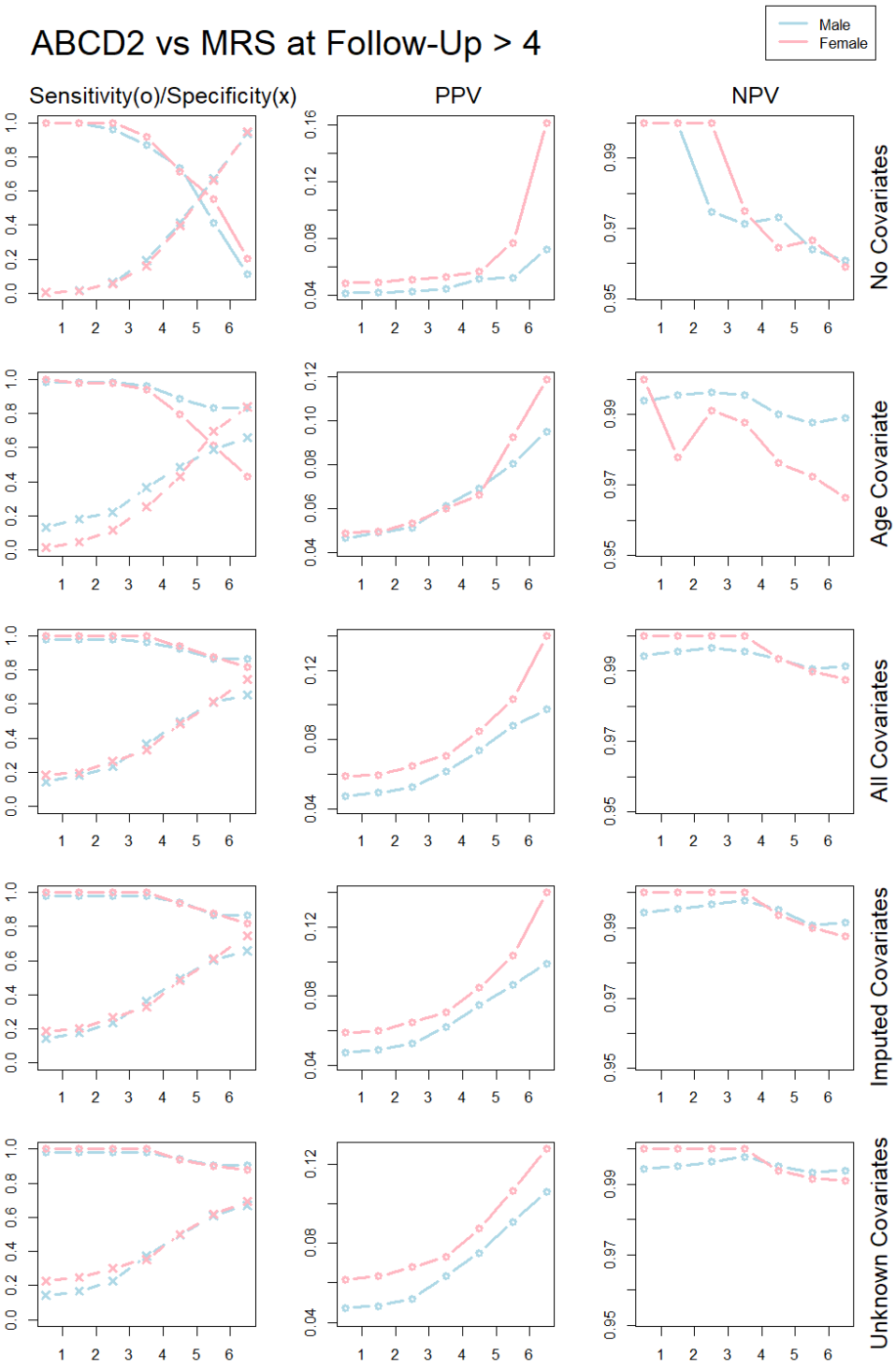
(d) ABCD3-I - Early Recurrence in Stroke Unit



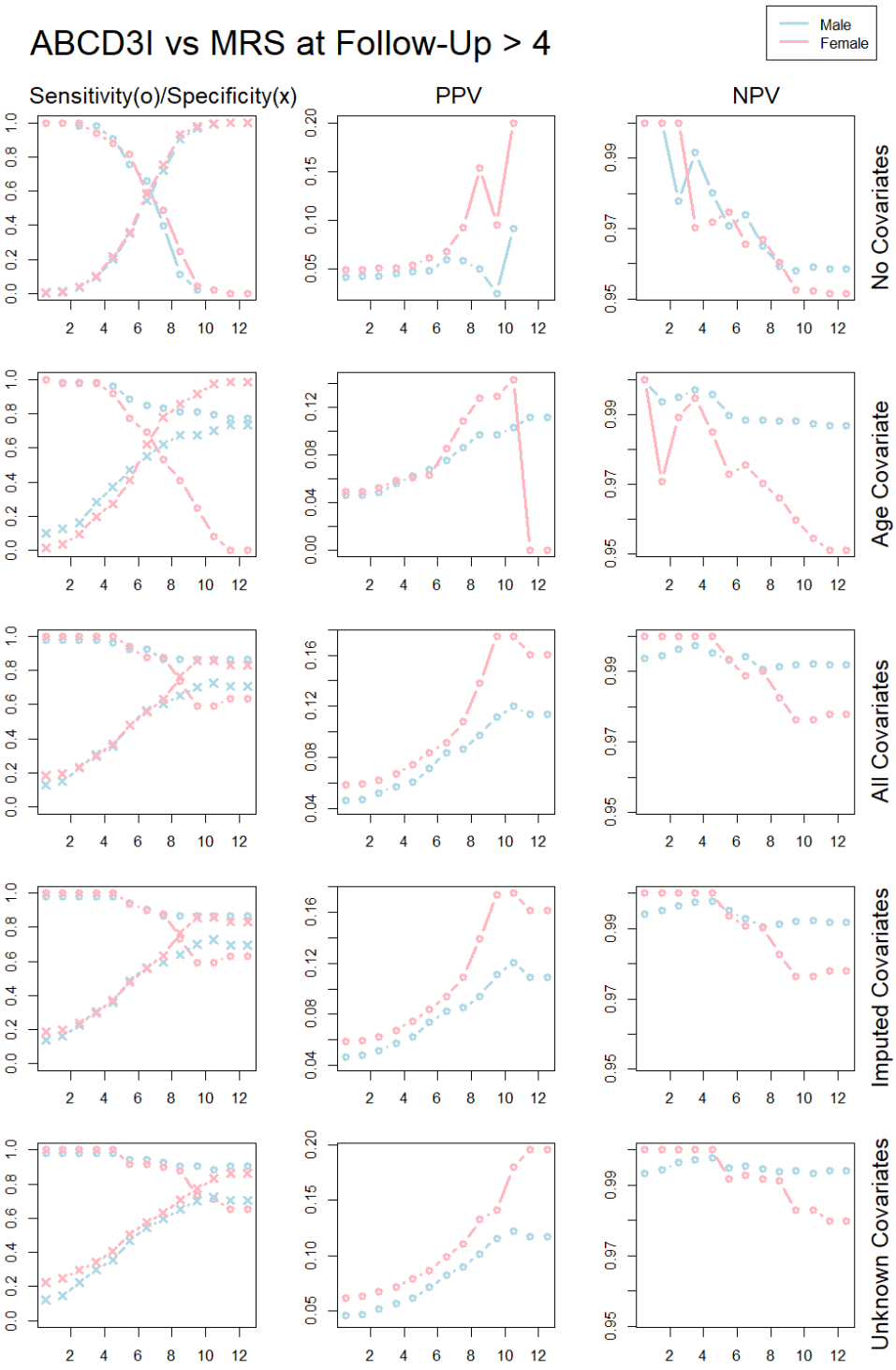
(e) ABCD2 - Recurrence within 90 days



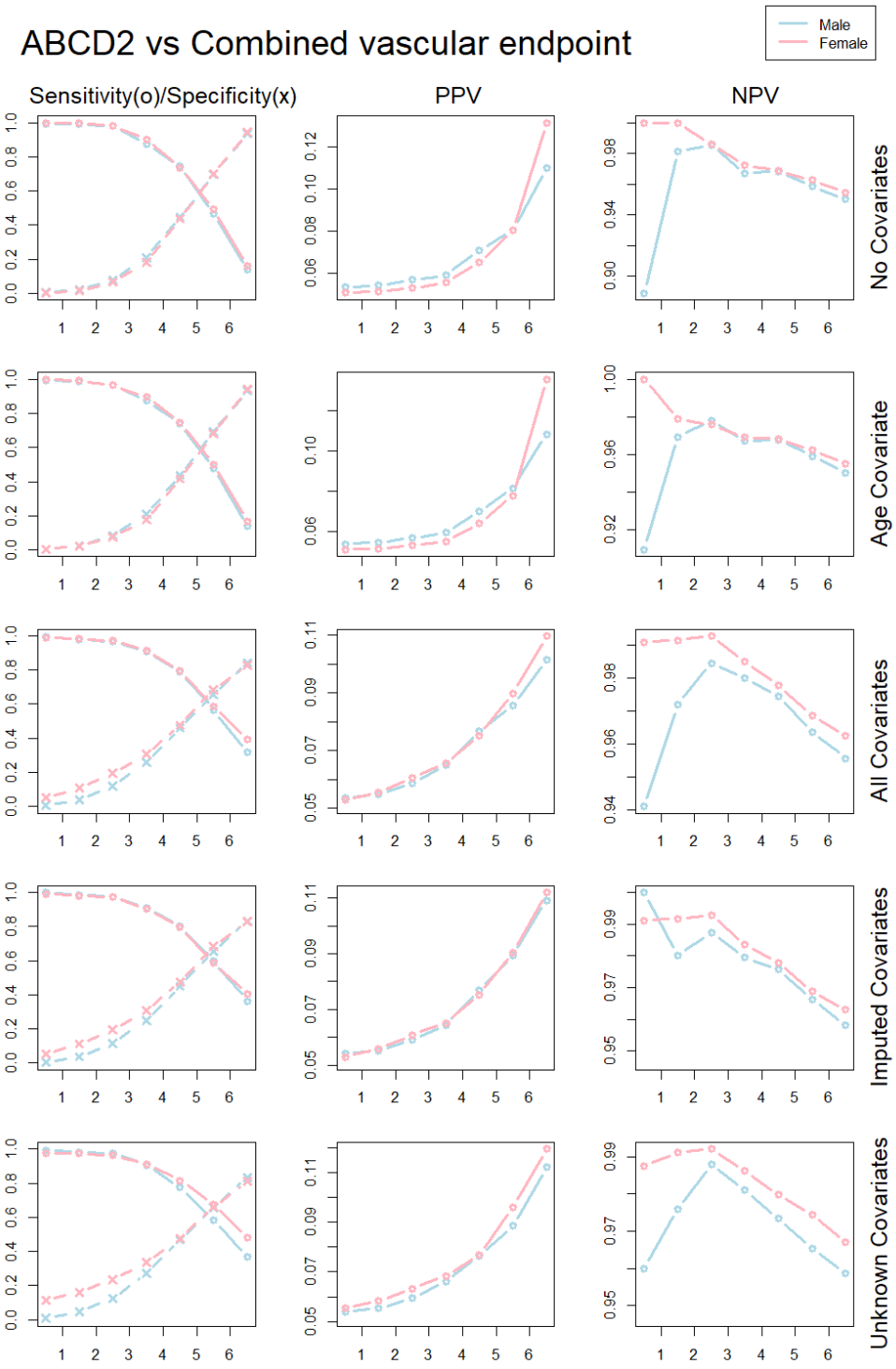
(f) ABCD3-I - Recurrence within 90 days



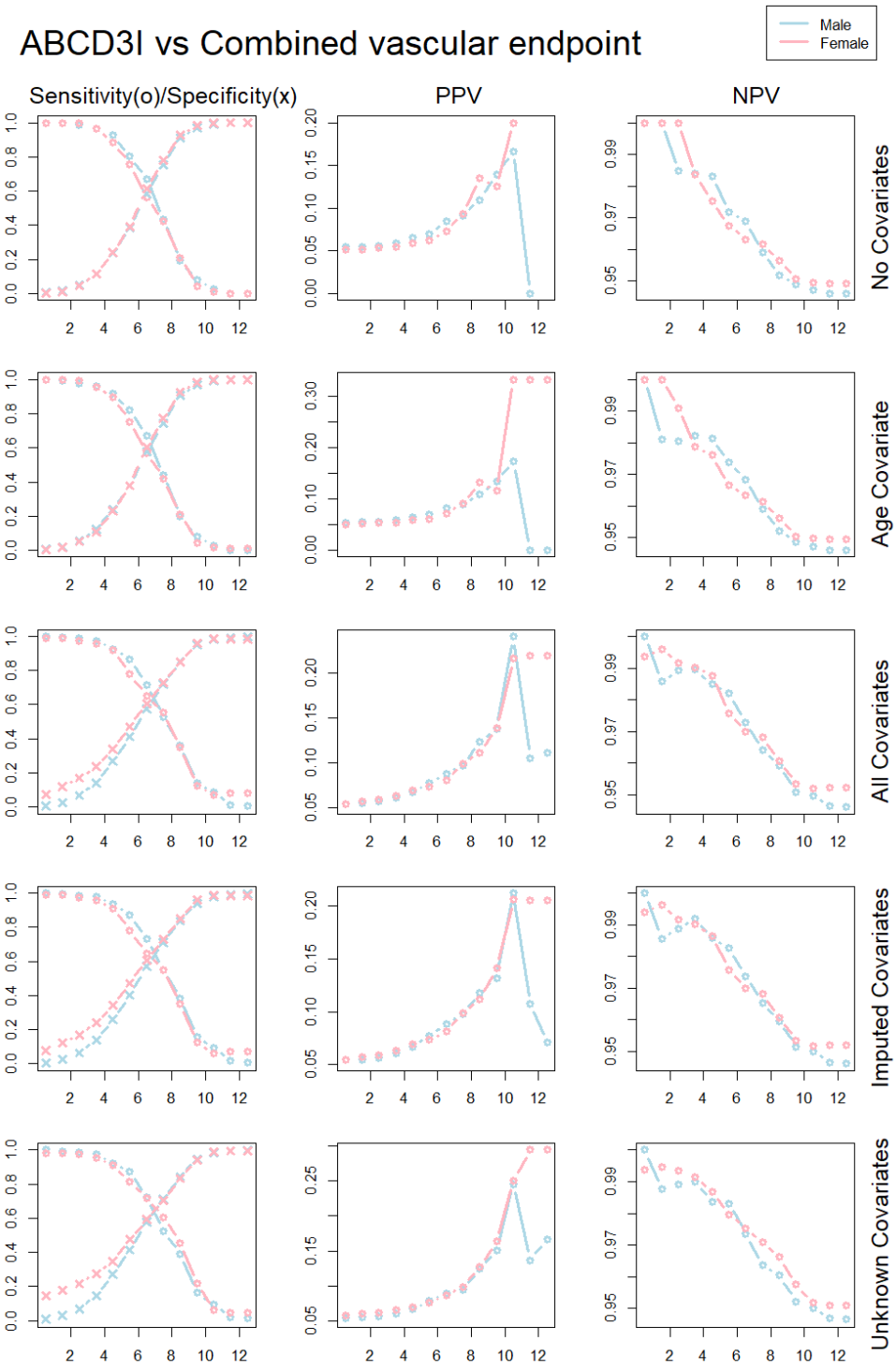
(g) ABCD2 - MRS at Follow-Up > 4



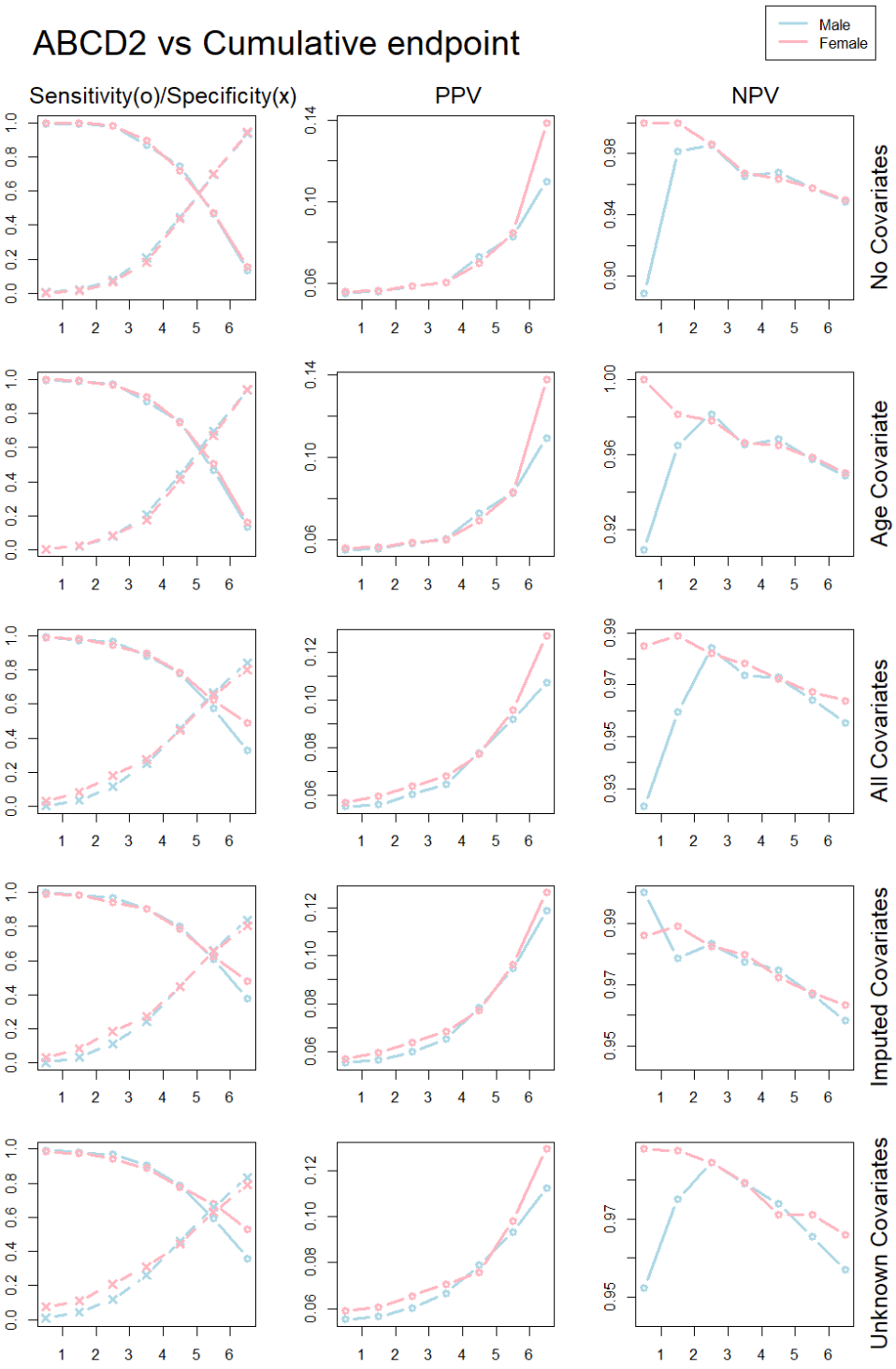
(h) ABCD3-I - MRS at Follow-Up > 4



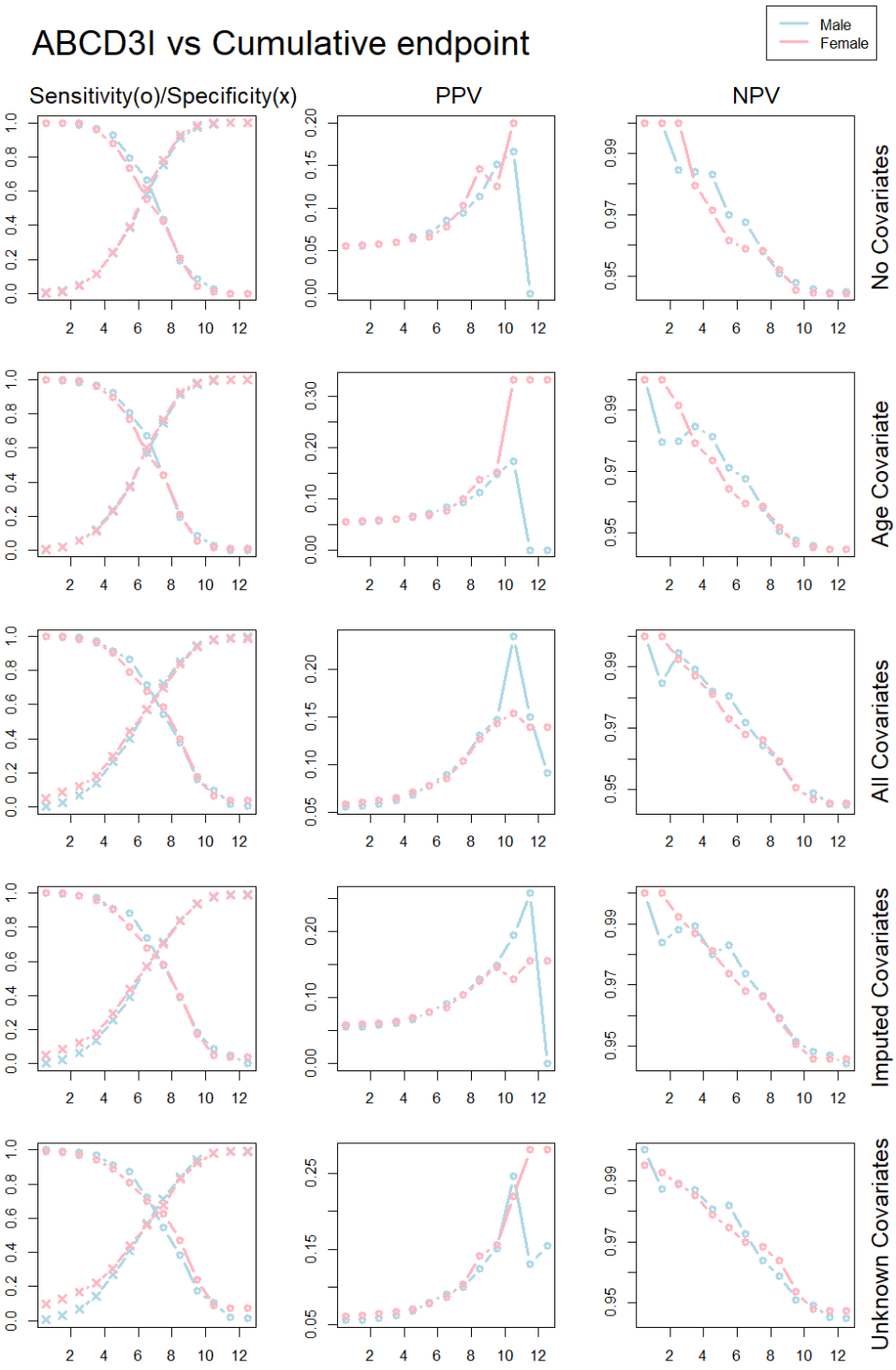
(i) ABCD2 - Combined vascular endpoint



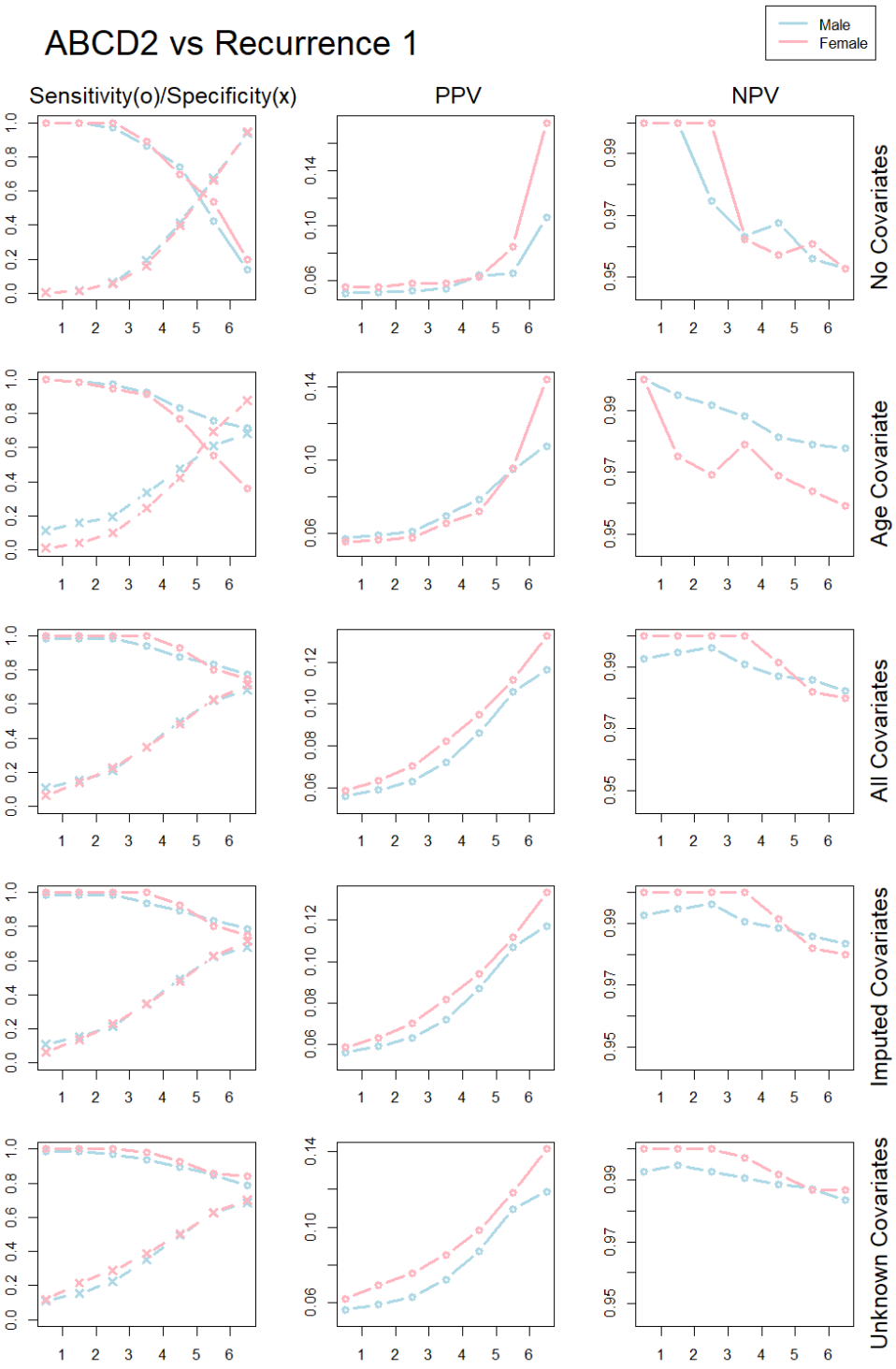
(j) ABCD3-I - Combined vascular endpoint



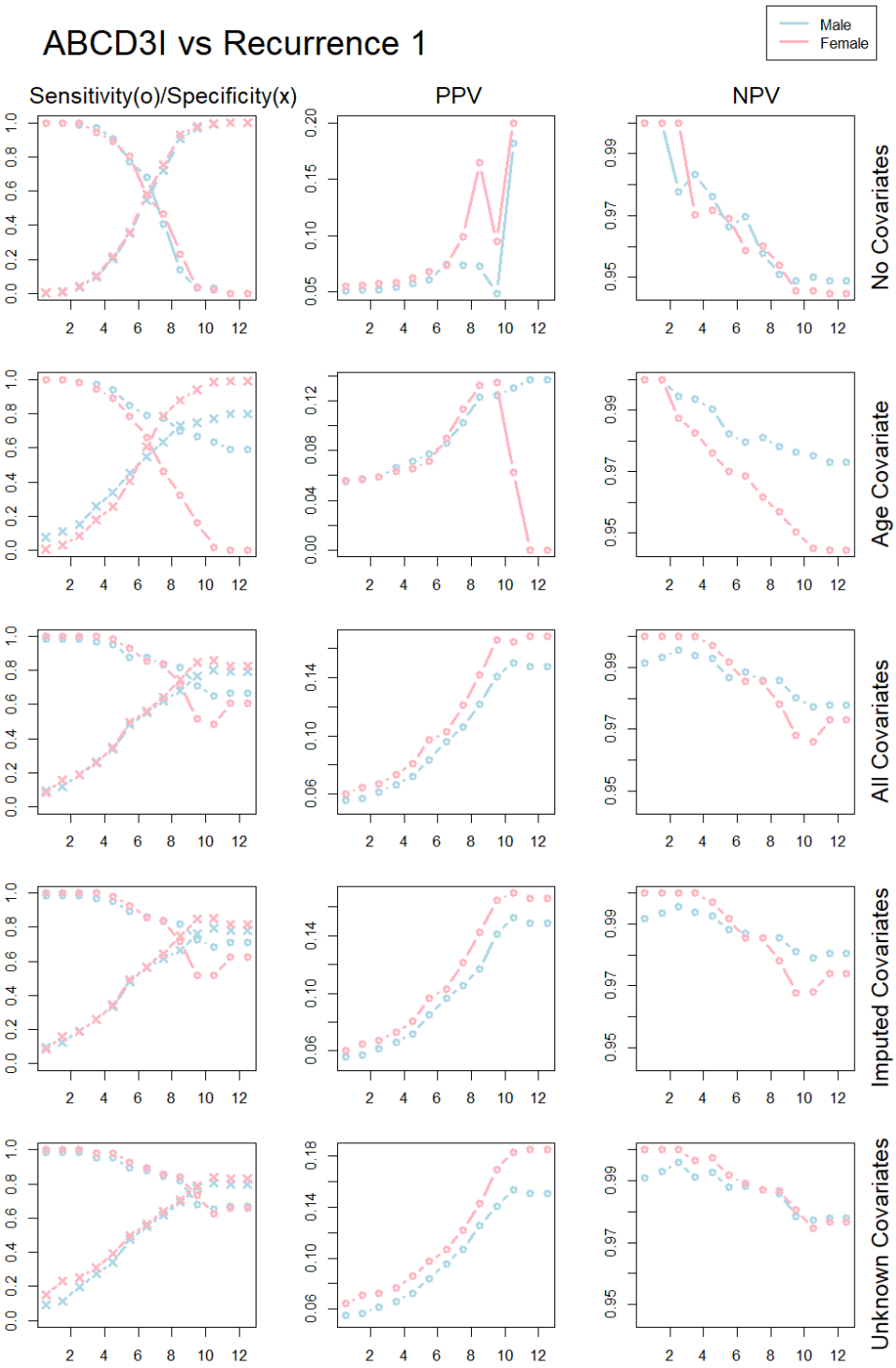
(k) ABCD2 - Cumulative endpoint



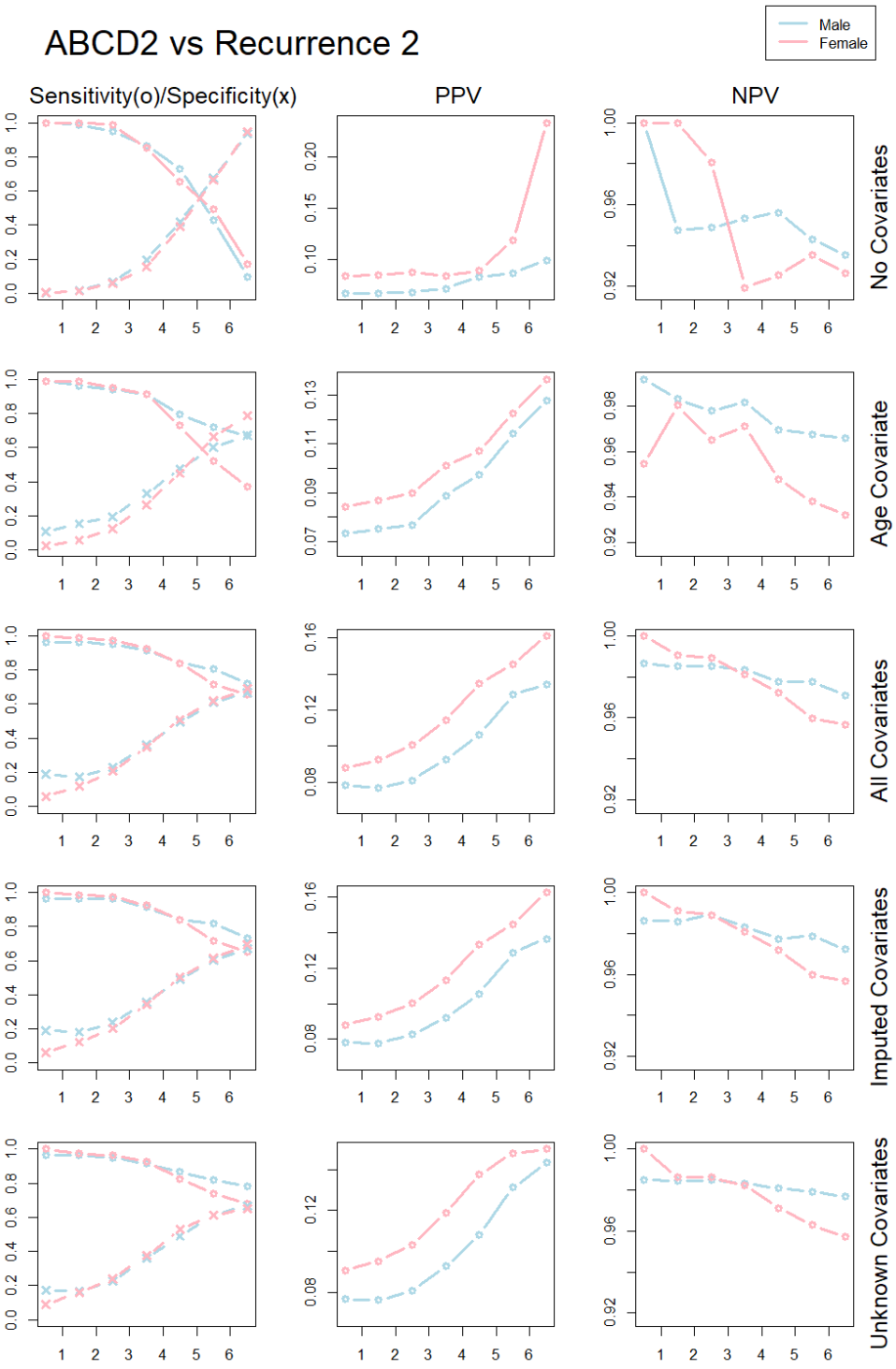
(l) ABCD3-I - Cumulative endpoint



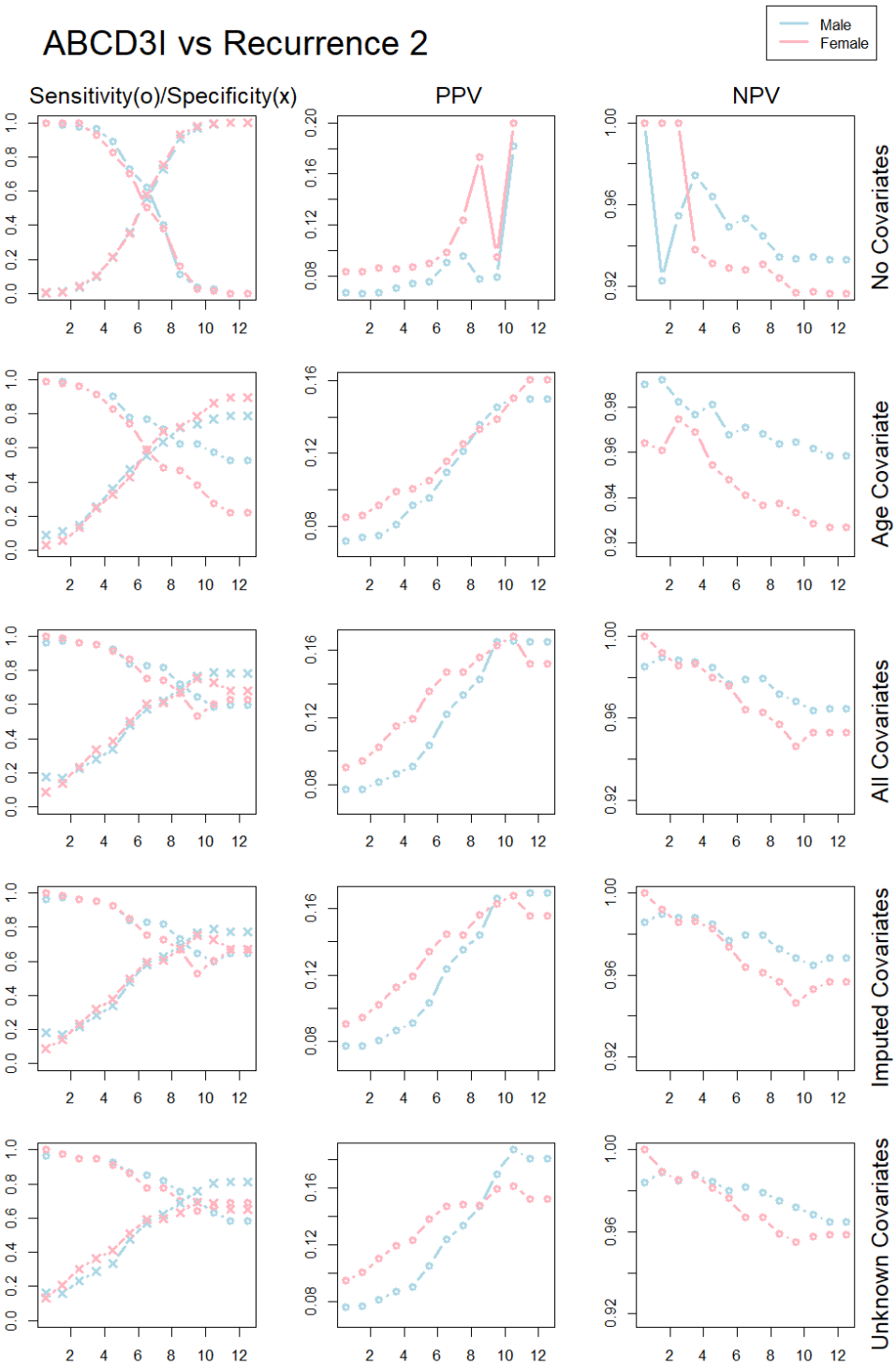
(m) ABCD2 - Recurrence 1



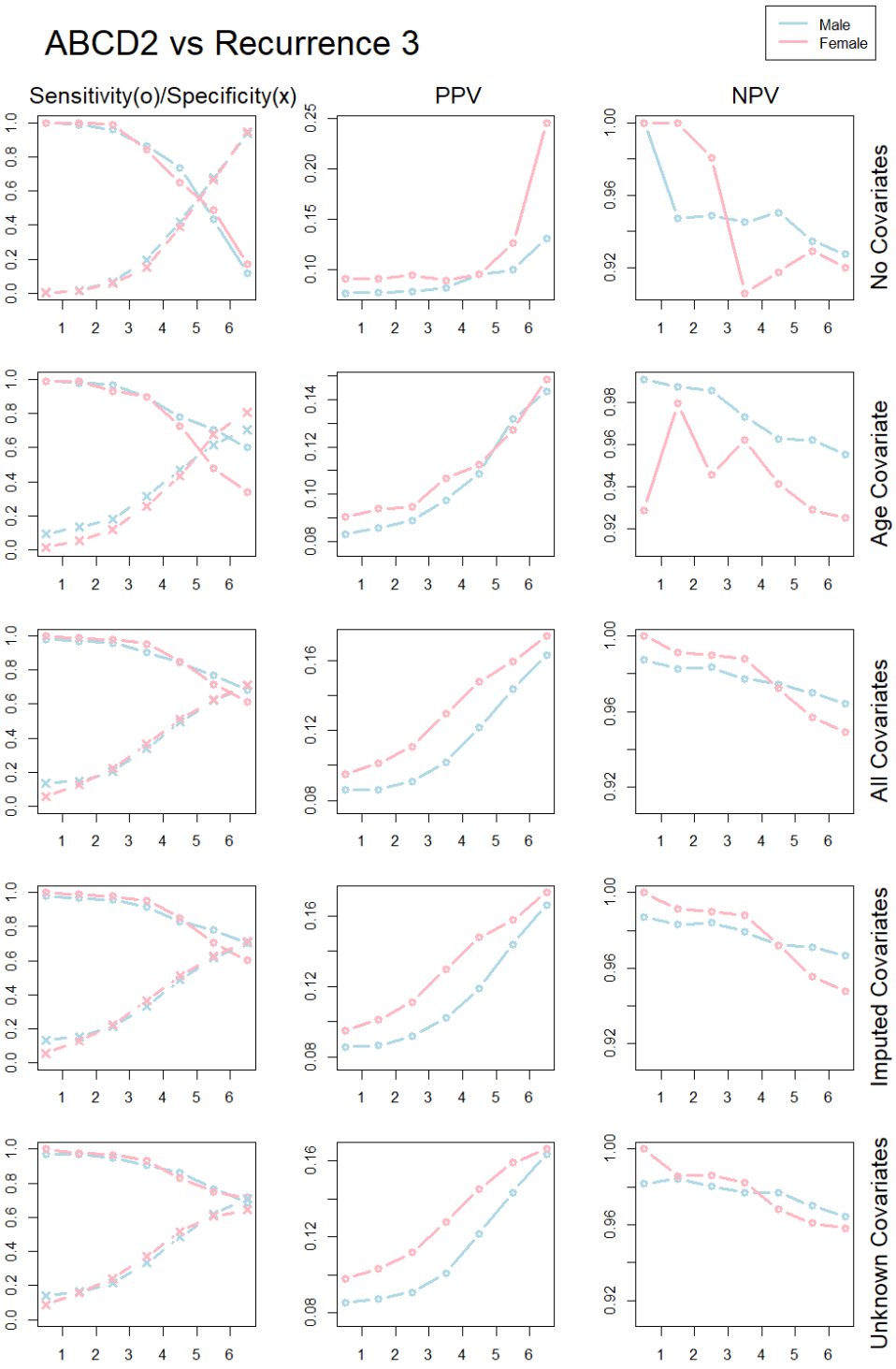
(n) ABCD3-I - Recurrence 1



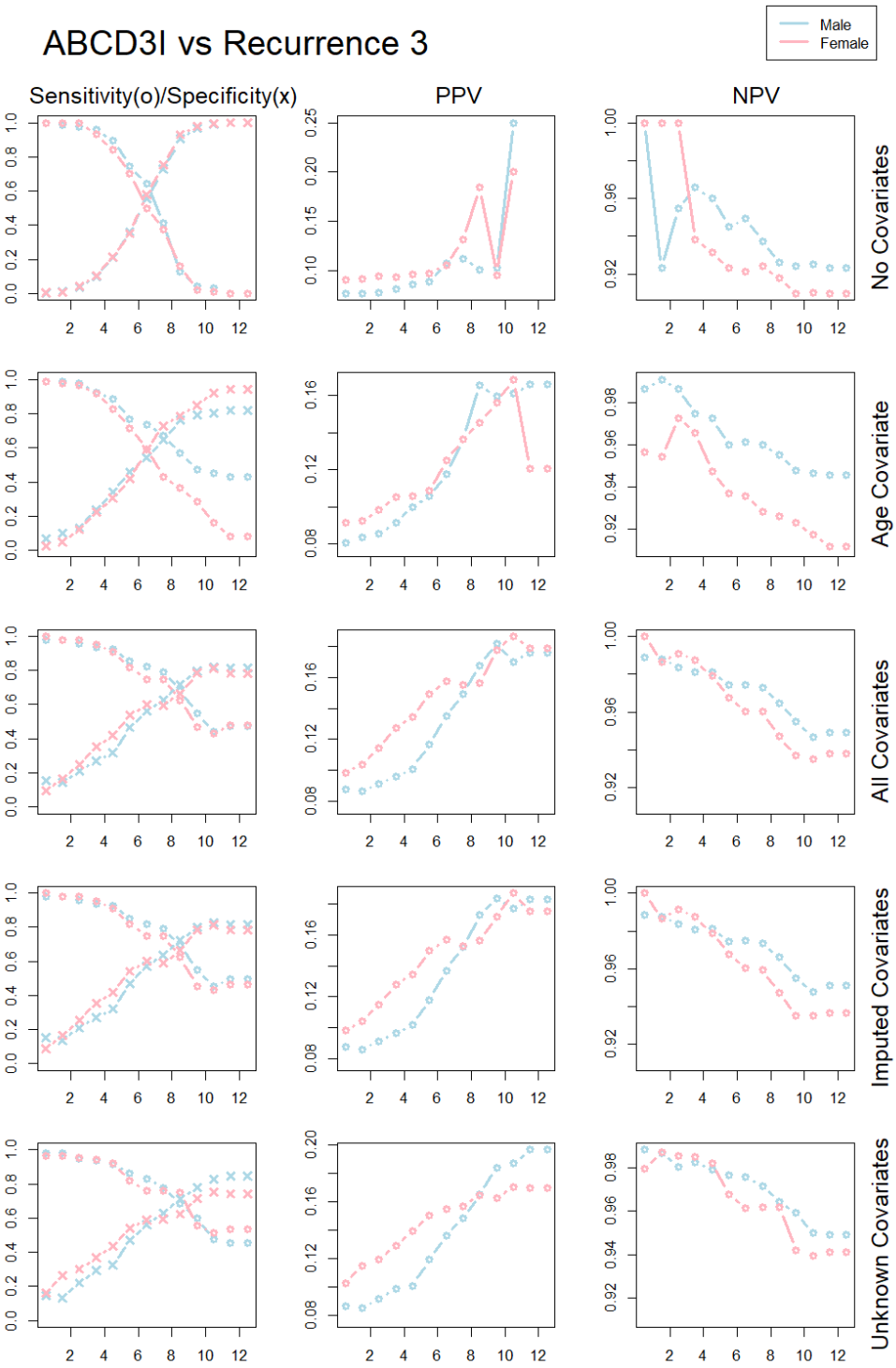
(o) ABCD2 - Recurrence 2



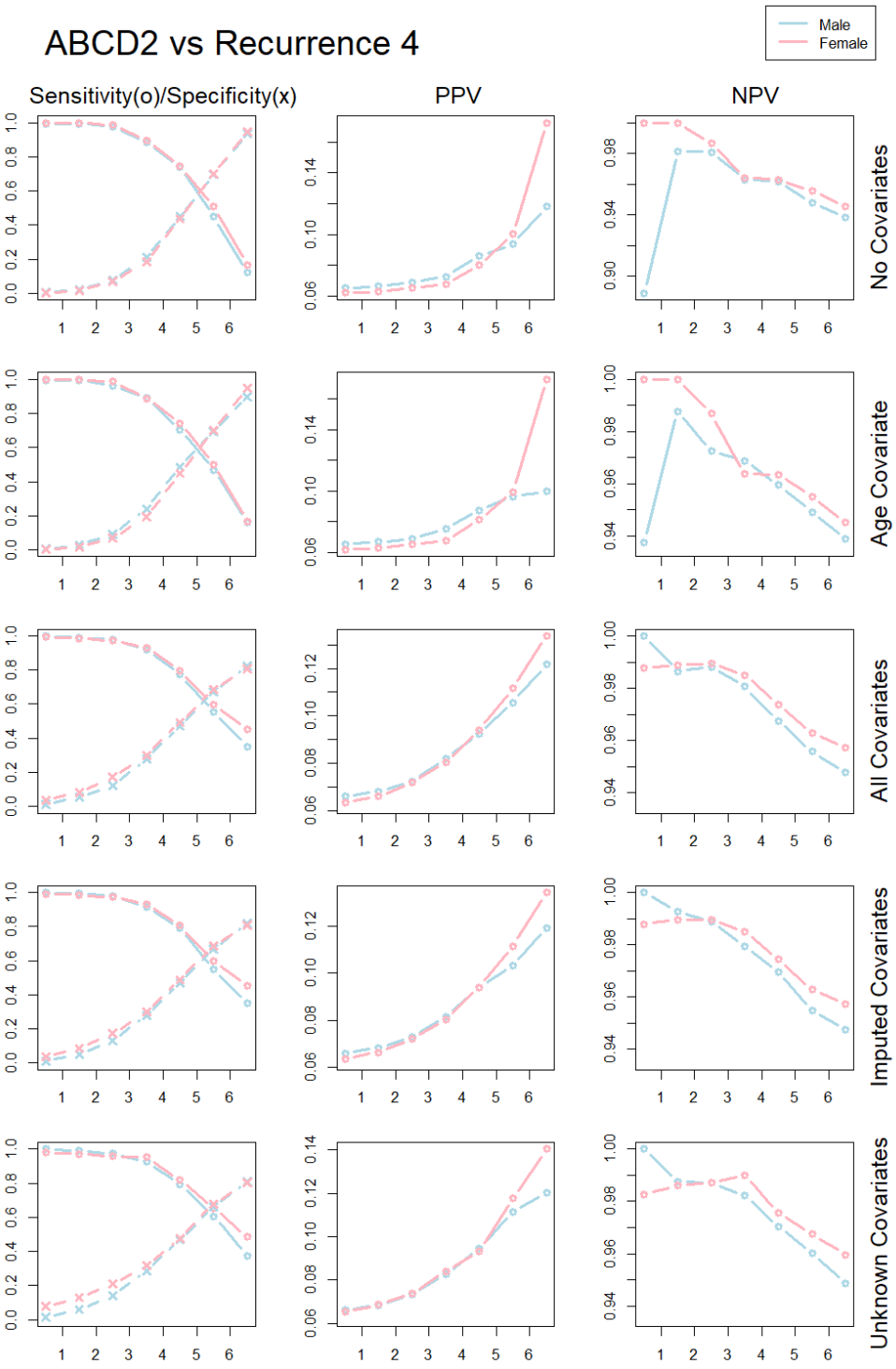
(p) ABCD3-I - Recurrence 2



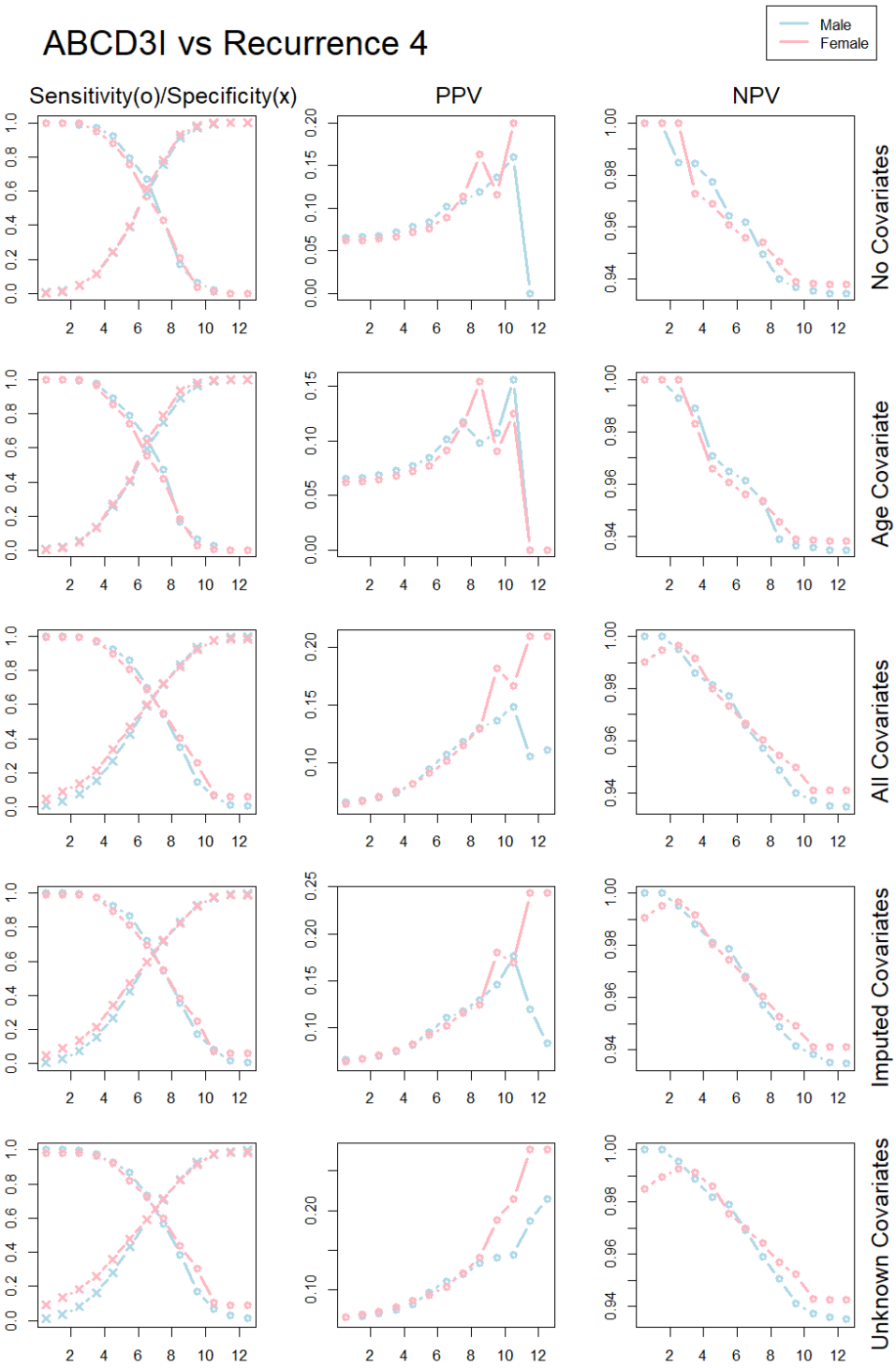
(q) ABCD2 - Recurrence 3



(r) ABCD3-I - Recurrence 3



(s) ABCD2 - Recurrence 4



(t) ABCD3-I - Recurrence 4

Population 1 - Diagnostic methods and secondary prophylactic treatments Odds Ratios

Variable	OR [95% CI]	P-Value	Interpretation
Length of hospital stay = 1-2	0.94 [0.79-1.11]	0.452	-
Length of hospital stay = 3-7	0.92 [0.78-1.09]	0.350	-
Length of hospital stay = 8-14	0.77 [0.6-0.98]	0.031	Women had lower odds.
Length of hospital stay = 14+	0.69 [0.48-1.01]	0.055	-
CCT	1.18 [1.07-1.29]	< 0.001	Women had higher odds.
CCT = Already on hand	1.11 [0.92-1.35]	0.281	-
CCT = Yes	1.18 [1.07-1.29]	< 0.001	Women had higher odds.
MRI - i8004	0.83 [0.76-0.9]	< 0.001	Women had lower odds.
MRI - i27022	0.74 [0.69-0.8]	< 0.001	Women had lower odds.
MRI - i27022 = Planned	0.77 [0.69-0.86]	< 0.001	Women had lower odds.
MRI - i27022 = Yes	0.74 [0.68-0.8]	< 0.001	Women had lower odds.
TTE	0.94 [0.88-1]	0.048	-
TTE = Planned	0.97 [0.89-1.06]	0.493	-
TTE = Yes	0.92 [0.86-0.99]	0.021	Women had lower odds.
TEE	0.75 [0.69-0.81]	< 0.001	Women had lower odds.
TEE = Planned	0.79 [0.7-0.89]	< 0.001	Women had lower odds.
TEE = Yes	0.71 [0.64-0.79]	< 0.001	Women had lower odds.
Antiplatelet agents	0.98 [0.9-1.06]	0.621	-
Acetylsalicylic acid	0.99 [0.93-1.06]	0.861	-
Clopidogrel	0.85 [0.79-0.92]	< 0.001	Women had lower odds.
Heparin subcutaneous = Medium	1.07 [0.94-1.21]	0.325	-
Heparin subcutaneous = Low	1.05 [0.96-1.16]	0.292	-
Heparin subcutaneous = High	0.95 [0.78-1.17]	0.654	-
PTA	0.62 [0.43-0.89]	0.010	Women had lower odds.
CEA	0.45 [0.36-0.55]	< 0.001	Women had lower odds.
Carotid Revascularization	0.47 [0.39-0.56]	< 0.001	Women had lower odds.
Rehabilitation	0.91 [0.81-1.02]	0.118	-
Inpatient rehabilitation	0.85 [0.76-0.96]	0.007	Women had lower odds.
Outpatient rehabilitation	1.03 [0.67-1.58]	0.893	-
OAC at Follow-Up	1.01 [0.44-2.37]	0.982	-

Table 36: Population 1 - Unadjusted treatment odds ratios by gender

Variable	OR [95% CI]	P-Value	Interpretation
Length of hospital stay = 1-2	0.9 [0.76-1.07]	0.241	-
Length of hospital stay = 3-7	0.88 [0.74-1.05]	0.154	-
Length of hospital stay = 8-14	0.76 [0.59-0.96]	0.024	Women had lower odds.
Length of hospital stay = 14+	0.65 [0.44-0.95]	0.025	Women had lower odds.
CCT	1.06 [0.96-1.16]	0.258	-
CCT = Already on hand	1.12 [0.92-1.36]	0.264	-
CCT = Yes	1.05 [0.96-1.16]	0.298	-
MRI - i8004	0.93 [0.85-1]	0.066	-
MRI - i27022	0.94 [0.87-1.01]	0.093	-
MRI - i27022 = Planned	0.95 [0.85-1.07]	0.398	-
MRI - i27022 = Yes	0.93 [0.86-1.01]	0.085	-
TTE	0.98 [0.92-1.05]	0.638	-
TTE = Planned	1.03 [0.94-1.12]	0.544	-
TTE = Yes	0.96 [0.9-1.03]	0.293	-
TEE	0.89 [0.82-0.98]	0.013	Women had lower odds.
TEE = Planned	0.96 [0.85-1.08]	0.495	-
TEE = Yes	0.85 [0.76-0.95]	0.004	Women had lower odds.
Antiplatelet agents	1.08 [0.99-1.17]	0.094	-
Acetylsalicylic acid	1.17 [1.09-1.25]	< 0.001	Women had higher odds.
Clopidogrel	0.78 [0.73-0.85]	< 0.001	Women had lower odds.
Heparin subcutaneous = Medium	1.04 [0.92-1.19]	0.518	-
Heparin subcutaneous = Low	1.05 [0.96-1.16]	0.296	-
Heparin subcutaneous = High	0.87 [0.71-1.08]	0.203	-
PTA	0.62 [0.42-0.89]	0.010	Women had lower odds.
CEA	0.41 [0.34-0.51]	< 0.001	Women had lower odds.
Carotid Revascularization	0.44 [0.36-0.53]	< 0.001	Women had lower odds.
Rehabilitation	0.93 [0.82-1.04]	0.208	-
Inpatient rehabilitation	0.88 [0.78-0.99]	0.033	Women had lower odds.
Outpatient rehabilitation	1.04 [0.67-1.61]	0.851	-
OAC at Follow-Up	0.93 [0.41-2.19]	0.857	-

Table 37: Population 1 - Treatment odds ratios by gender controlled for Age

Variable	OR [95% CI]	P-Value	Interpretation
Length of hospital stay = 1-2	0.96 [0.7-1.31]	0.787	-
Length of hospital stay = 3-7	1.02 [0.74-1.4]	0.909	-
Length of hospital stay = 8-14	0.88 [0.57-1.37]	0.575	-
Length of hospital stay = 14+	0.83 [0.43-1.61]	0.583	-
CCT	1.1 [0.94-1.28]	0.233	-
CCT = Already on hand	1.11 [0.82-1.51]	0.490	-
CCT = Yes	1.1 [0.94-1.28]	0.246	-
MRI - i8004	0.88 [0.77-1.01]	0.061	-
MRI - i27022	0.94 [0.81-1.07]	0.346	-
MRI - i27022 = Planned	0.88 [0.71-1.09]	0.249	-
MRI - i27022 = Yes	0.94 [0.82-1.09]	0.424	-
TTE	0.95 [0.85-1.07]	0.405	-
TTE = Planned	1 [0.85-1.16]	0.957	-
TTE = Yes	0.93 [0.82-1.06]	0.264	-
TEE	0.93 [0.8-1.09]	0.371	-
TEE = Planned	1.02 [0.83-1.25]	0.868	-
TEE = Yes	0.87 [0.71-1.06]	0.171	-
Antiplatelet agents	1.03 [0.87-1.23]	0.731	-
Acetylsalicylic acid	1.09 [0.96-1.23]	0.167	-
Clopidogrel	0.8 [0.7-0.93]	0.003	Women had lower odds.
Heparin subcutaneous = Low	0.99 [0.82-1.18]	0.899	-
Heparin subcutaneous = Medium	1.05 [0.82-1.34]	0.716	-
Heparin subcutaneous = High	0.89 [0.61-1.32]	0.572	-
PTA	1.22 [0.68-2.15]	0.492	-
CEA	0.56 [0.38-0.81]	0.002	Women had lower odds.
Carotid Revascularization	0.68 [0.48-0.94]	0.020	Women had lower odds.
Rehabilitation	0.9 [0.74-1.09]	0.294	-
Inpatient rehabilitation	0.91 [0.74-1.1]	0.323	-
Outpatient rehabilitation	1.01 [0.47-2.17]	0.970	-
OAC at Follow-Up	0.73 [0.27-1.99]	0.525	-

Table 38: Population 1 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors

Variable	OR [95% CI]	P-Value	Interpretation
Length of hospital stay = 1-2	0.97 [0.71-1.33]	0.865	-
Length of hospital stay = 3-7	1.02 [0.74-1.41]	0.884	-
Length of hospital stay = 8-14	0.9 [0.58-1.4]	0.639	-
Length of hospital stay = 14+	0.85 [0.44-1.67]	0.643	-
CCT	1.12 [0.96-1.3]	0.158	-
CCT = Already on hand	1.14 [0.84-1.55]	0.411	-
CCT = Yes	1.11 [0.96-1.3]	0.168	-
MRI - i8004	0.87 [0.76-1]	0.046	-
MRI - i27022	0.93 [0.81-1.06]	0.276	-
MRI - i27022 = Planned	0.88 [0.71-1.1]	0.257	-
MRI - i27022 = Yes	0.93 [0.81-1.07]	0.334	-
TTE	0.95 [0.85-1.07]	0.413	-
TTE = Planned	0.99 [0.85-1.16]	0.935	-
TTE = Yes	0.93 [0.82-1.06]	0.281	-
TEE	0.92 [0.79-1.08]	0.310	-
TEE = Planned	1.01 [0.82-1.25]	0.923	-
TEE = Yes	0.86 [0.7-1.05]	0.146	-
Antiplatelet agents	1.01 [0.85-1.21]	0.869	-
Acetylsalicylic acid	1.09 [0.96-1.23]	0.179	-
Clopidogrel	0.8 [0.7-0.92]	0.002	Women had lower odds.
Heparin subcutaneous = Low	0.98 [0.81-1.17]	0.795	-
Heparin subcutaneous = High	0.88 [0.6-1.3]	0.527	-
Heparin subcutaneous = Medium	1.04 [0.81-1.33]	0.777	-
PTA	1.23 [0.69-2.18]	0.476	-
CEA	0.54 [0.37-0.78]	0.001	Women had lower odds.
Carotid Revascularization	0.66 [0.47-0.92]	0.016	Women had lower odds.
Rehabilitation	0.92 [0.76-1.12]	0.403	-
Inpatient rehabilitation	0.9 [0.74-1.1]	0.310	-
Outpatient rehabilitation	1.02 [0.47-2.18]	0.967	-
OAC at Follow-Up	0.76 [0.28-2.09]	0.591	-

Table 39: Population 1 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors with Unknown included

Variable	OR [95% CI]	P-Value	Interpretation
Length of hospital stay = 1-2	0.95 [0.7-1.3]	0.763	-
Length of hospital stay = 3-7	1.02 [0.74-1.4]	0.923	-
Length of hospital stay = 8-14	0.87 [0.56-1.35]	0.542	-
Length of hospital stay = 14+	0.84 [0.43-1.64]	0.615	-
CCT	1.1 [0.94-1.28]	0.216	-
CCT = Already on hand	1.11 [0.82-1.51]	0.502	-
CCT = Yes	1.1 [0.94-1.28]	0.225	-
MRI - i8004	0.88 [0.77-1.01]	0.063	-
MRI - i27022	0.94 [0.81-1.07]	0.343	-
MRI - i27022 = Planned	0.88 [0.71-1.1]	0.273	-
MRI - i27022 = Yes	0.94 [0.82-1.08]	0.412	-
TTE	0.95 [0.85-1.07]	0.410	-
TTE= Planned	1 [0.85-1.16]	0.960	-
TTE = Yes	0.93 [0.82-1.06]	0.268	-
TEE	0.93 [0.8-1.08]	0.357	-
TEE = Planned	1.02 [0.83-1.25]	0.863	-
TEE = Yes	0.86 [0.71-1.06]	0.157	-
Antiplatelet agents	1.03 [0.86-1.22]	0.772	-
Acetylsalicylic acid	1.09 [0.97-1.24]	0.157	-
Clopidogrel	0.81 [0.7-0.93]	0.003	Women had lower odds.
Heparin subcutaneous = Low	0.98 [0.82-1.18]	0.854	-
Heparin subcutaneous = Medium	1.05 [0.82-1.34]	0.725	-
Heparin subcutaneous = High	0.89 [0.6-1.31]	0.540	-
PTA	1.22 [0.68-2.15]	0.498	-
CEA	0.56 [0.38-0.81]	0.003	Women had lower odds.
Carotid Revascularization	0.68 [0.48-0.94]	0.021	Women had lower odds.
Rehabilitation	0.91 [0.74-1.1]	0.318	-
Inpatient rehabilitation	0.91 [0.74-1.11]	0.342	-
Outpatient rehabilitation	1.01 [0.46-2.16]	0.987	-
OAC at Follow-Up	0.73 [0.27-2.01]	0.535	-

Table 40: Population 1 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors with Smoking and Alcohol Abuse imputed

Population 1 - Treatments by age group and gender

Age	Treatment: No Female	Treatment: Yes Female	Treatment: No Male	Treatment: Yes Male
18-49	15 (23.81%)	48 (76.19%)	34 (29.06%)	83 (70.94%)
50-59	27 (22.14%)	95 (77.87%)	59 (19.54%)	243 (80.46%)
60-69	57 (24.36%)	177 (75.64%)	118 (23.74%)	379 (76.26%)
70-79	126 (26.03%)	358 (73.97%)	134 (20.30%)	526 (79.70%)
80-89	102 (24.34%)	317 (75.66%)	86 (26.96%)	233 (73.04%)
90+	16 (25.00%)	48 (75.00%)	12 (34.29%)	23 (65.71%)

Table 41: Population 1 - Regular lipid-lowering drugs for patients with Hypercholesterolemia

Age	Treatment: No Female	Treatment: Yes Female	Treatment: No Male	Treatment: Yes Male
18-49	11 (16.67%)	55 (83.33%)	26 (25.74%)	75 (74.26%)
50-59	30 (21.13%)	112 (78.87%)	61 (18.89%)	262 (81.11%)
60-69	72 (23.76%)	231 (76.24%)	133 (21.91%)	474 (78.09%)
70-79	122 (18.51%)	537 (81.49%)	170 (19.41%)	706 (80.59%)
80-89	130 (20.12%)	516 (79.88%)	116 (22.92%)	390 (77.08%)
90+	19 (13.97%)	117 (86.03%)	18 (30.00%)	42 (70.00%)

Table 42: Population 1 - Regular blood pressure checks for patients with Hypertension

Age	Treatment: No Female	Treatment: Yes Female	Treatment: No Male	Treatment: Yes Male
18-49	16 (24.24%)	50 (75.76%)	28 (27.45%)	74 (72.55%)
50-59	41 (29.08%)	100 (70.92%)	84 (25.61%)	244 (74.39%)
60-69	83 (27.30%)	221 (72.70%)	163 (26.85%)	444 (73.18%)
70-79	207 (31.46%)	451 (68.54%)	232 (26.82%)	633 (73.18%)
80-89	216 (33.59%)	427 (66.41%)	174 (34.94%)	324 (65.06%)
90+	59 (44.03%)	75 (55.97%)	27 (45.00%)	33 (55.00%)

Table 43: Population 1 - Regular antihypertensives for patients with Hypertension

Age	Treatment: No Female	Treatment: Yes Female	Treatment: No Male	Treatment: Yes Male
18-49	4 (26.67%)	11 (73.33%)	8 (50.00%)	8 (50.00%)
50-59	13 (33.33%)	26 (66.67%)	39 (40.62%)	57 (59.38%)
60-69	35 (47.95%)	38 (52.05%)	65 (35.33%)	119 (64.67%)
70-79	70 (38.46%)	112 (61.54%)	110 (41.98%)	152 (58.02%)
80-89	74 (42.05%)	102 (57.95%)	58 (42.03%)	80 (57.97%)
90+	12 (50.00%)	12 (50.00%)	10 (58.82%)	7 (41.18%)

Table 44: Population 1 - Regular antidiabetica for patients with Diabetes mellitus

Age	Treatment: No Female	Treatment: Yes Female	Treatment: No Male	Treatment: Yes Male
18-49	0 (0.00%)	2 (100%)	0 (0.00%)	5 (100%)
50-59	0 (0.00%)	11 (100%)	0 (0.00%)	28 (100%)
60-69	1 (27.30%)	47 (97.92%)	0 (0.00%)	108 (100%)
70-79	1 (31.46%)	156 (99.36%)	0 (0.00%)	222 (100%)
80-89	0 (0.00%)	221 (100%)	0 (0.00%)	180 (100%)
90+	0 (0.00%)	57 (100%)	0 (0.00%)	35 (100%)

Table 45: Population 1 - Oral Anticoagulants (OAC) for patients with Atrial fibrillation

Population 2 - Diagnostic methods and secondary prophylactic treatments Odds Ratios

Variable	OR [95% CI]	P-Value	Interpretation
Length of hospital stay = 1-2	1.02 [0.87-1.2]	0.821	-
Length of hospital stay = 3-7	0.98 [0.83-1.17]	0.857	-
Length of hospital stay = 8-14	1.1 [0.79-1.53]	0.569	-
Length of hospital stay = 14+	1.3 [0.7-2.42]	0.399	-
CCT	1.01 [0.91-1.11]	0.900	-
MRI	1 [0.91-1.09]	0.994	-
TTE	0.89 [0.83-0.96]	0.004	Women had lower odds.
TTE = Planned	0.95 [0.87-1.04]	0.261	-
TTE = Yes	0.85 [0.78-0.93]	< 0.001	Women had lower odds.
TEE	0.64 [0.56-0.72]	< 0.001	Women had lower odds.
TEE = Planned	0.66 [0.57-0.78]	< 0.001	Women had lower odds.
TEE = Yes	0.6 [0.49-0.72]	< 0.001	Women had lower odds.
Antiplatelet agents = Mono	1.05 [0.96-1.14]	0.296	-
Antiplatelet agents = Dual	0.82 [0.73-0.91]	< 0.001	Women had lower odds.
Heparin = Low	0.97 [0.9-1.05]	0.446	-
Heparin = Medium	1.05 [0.9-1.23]	0.506	-
Heparin = High	0.69 [0.51-0.93]	0.014	Women had lower odds.
PTA	0.68 [0.45-1.01]	0.058	-
CEA	0.53 [0.42-0.66]	< 0.001	Women had lower odds.
Carotid Revascularization	0.55 [0.44-0.67]	< 0.001	Women had lower odds.
Rehabilitation	0.8 [0.65-0.98]	0.034	Women had lower odds.
Inpatient rehabilitation	0.71 [0.56-0.9]	0.004	Women had lower odds.
Outpatient rehabilitation	0.57 [0.36-0.87]	0.011	Women had lower odds.
Platelet inhibitors - i8070 = Mono	0.82 [0.76-0.88]	< 0.001	Women had lower odds.
Platelet inhibitors - i8070 = Dual	0.64 [0.49-0.82]	< 0.001	Women had lower odds.
Platelet inhibitors - i26019 = Mono	1.02 [0.89-1.18]	0.741	-
Platelet inhibitors - i26019 = Dual	0.68 [0.5-0.93]	0.015	Women had lower odds.
OAC at Follow-Up	0.95 [0.81-1.11]	0.523	-

Table 46: Population 2 - Unadjusted treatment odds ratios by gender

Variable	OR [95% CI]	P-Value	Interpretation
Length of hospital stay = 1-2	1.01 [0.86-1.19]	0.917	-
Length of hospital stay = 3-7	0.97 [0.81-1.14]	0.687	-
Length of hospital stay = 8-14	1.06 [0.76-1.47]	0.744	-
Length of hospital stay = 14+	1.3 [0.69-2.42]	0.414	-
CCT	0.94 [0.85-1.04]	0.222	-
MRI	1.06 [0.97-1.17]	0.180	-
TTE	0.94 [0.87-1.02]	0.155	-
TTE = Planned	1.01 [0.92-1.1]	0.887	-
TTE = Yes	0.9 [0.83-0.98]	0.018	Women had lower odds.
TEE	0.77 [0.67-0.87]	< 0.001	Women had lower odds.
TEE = Planned	0.8 [0.68-0.94]	0.006	Women had lower odds.
TEE = Yes	0.73 [0.6-0.88]	0.001	Women had lower odds.
Antiplatelet agents = Mono	1.24 [1.14-1.35]	< 0.001	Women had higher odds.
Antiplatelet agents = Dual	0.91 [0.82-1.02]	0.099	-
Heparin = Low	1.02 [0.95-1.1]	0.520	-
Heparin = Medium	1.12 [0.96-1.31]	0.163	-
Heparin = High	0.7 [0.52-0.94]	0.020	Women had lower odds.
PTA	0.67 [0.44-1]	0.056	-
CEA	0.51 [0.41-0.64]	< 0.001	Women had lower odds.
Carotid Revascularization	0.53 [0.43-0.65]	< 0.001	Women had lower odds.
Rehabilitation	0.81 [0.66-1]	0.049	-
Inpatient rehabilitation	0.74 [0.58-0.93]	0.011	Women had lower odds.
Outpatient rehabilitation	0.57 [0.37-0.88]	0.013	Women had lower odds.
Platelet inhibitors - i8070 = Mono	0.74 [0.68-0.79]	< 0.001	Women had lower odds.
Platelet inhibitors - i8070 = Dual	0.63 [0.49-0.81]	< 0.001	Women had lower odds.
Platelet inhibitors - i26019 = Mono	1.12 [0.97-1.3]	0.124	-
Platelet inhibitors - i26019 = Dual	0.75 [0.55-1.02]	0.068	-
OAC at Follow-Up	0.8 [0.68-0.94]	0.008	Women had lower odds.

Table 47: Population 2 - Treatment odds ratios by gender controlled for Age

Variable	OR [95% CI]	P-Value	Interpretation
Length of hospital stay = 1-2	1 [0.84-1.17]	0.953	-
Length of hospital stay = 3-7	0.97 [0.82-1.15]	0.737	-
Length of hospital stay = 8-14	1.09 [0.78-1.52]	0.630	-
Length of hospital stay = 14+	1.31 [0.69-2.46]	0.408	-
CCT	0.93 [0.84-1.03]	0.188	-
MRI	1.07 [0.97-1.17]	0.159	-
TTE	0.94 [0.87-1.02]	0.151	-
TTE = Planned	1.01 [0.92-1.11]	0.863	-
TTE = Yes	0.9 [0.82-0.98]	0.015	Women had lower odds.
TEE	0.74 [0.65-0.85]	< 0.001	Women had lower odds.
TEE = Planned	0.77 [0.65-0.91]	0.002	Women had lower odds.
TEE = Yes	0.71 [0.58-0.86]	< 0.001	Women had lower odds.
Antiplatelet agents = Mono	1.07 [0.97-1.19]	0.190	-
Antiplatelet agents = Dual	0.83 [0.73-0.94]	0.004	Women had lower odds.
Heparin = Low	0.98 [0.9-1.06]	0.610	-
Heparin = Medium	1.06 [0.9-1.25]	0.482	-
Heparin = High	0.71 [0.52-0.96]	0.029	Women had lower odds.
PTA	0.74 [0.48-1.12]	0.160	-
CEA	0.53 [0.42-0.66]	< 0.001	Women had lower odds.
Carotid Revascularization	0.56 [0.45-0.69]	< 0.001	Women had lower odds.
Rehabilitation	0.85 [0.69-1.04]	0.122	-
Inpatient rehabilitation	0.78 [0.61-0.98]	0.038	Women had lower odds.
Outpatient rehabilitation	0.57 [0.36-0.87]	0.011	Women had lower odds.
Platelet inhibitors - i8070 = Mono	0.76 [0.69-0.82]	< 0.001	Women had lower odds.
Platelet inhibitors - i8070 = Dual	0.72 [0.55-0.94]	0.014	Women had lower odds.
Platelet inhibitors - i26019 = Mono	1.06 [0.88-1.27]	0.549	-
Platelet inhibitors - i26019 = Dual	0.78 [0.56-1.08]	0.139	-
OAC at Follow-Up	0.84 [0.66-1.06]	0.142	-

Table 48: Population 2 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors

Variable	OR [95% CI]	P-Value	Interpretation
Length of hospital stay = 1-2	0.98 [0.83-1.15]	0.789	-
Length of hospital stay = 3-7	0.95 [0.8-1.14]	0.596	-
Length of hospital stay = 8-14	1.06 [0.76-1.49]	0.725	-
Length of hospital stay = 14+	1.33 [0.7-2.51]	0.382	-
CCT	0.93 [0.84-1.03]	0.175	-
MRI	1.1 [1-1.2]	0.056	-
TTE	0.94 [0.86-1.02]	0.115	-
TTE = Planned	1 [0.91-1.09]	0.939	-
TTE = Yes	0.89 [0.82-0.98]	0.012	Women had lower odds.
TEE	0.75 [0.65-0.85]	< 0.001	Women had lower odds.
TEE = Planned	0.77 [0.65-0.91]	0.002	Women had lower odds.
TEE = Yes	0.72 [0.59-0.88]	0.001	Women had lower odds.
Antiplatelet agents = Mono	1.07 [0.96-1.19]	0.233	-
Antiplatelet agents = Dual	0.83 [0.73-0.94]	0.004	Women had lower odds.
Heparin = Low	0.97 [0.9-1.05]	0.494	-
Heparin = Medium	1.03 [0.88-1.21]	0.713	-
Heparin = High	0.7 [0.52-0.95]	0.024	Women had lower odds.
PTA	0.76 [0.5-1.15]	0.197	-
CEA	0.53 [0.42-0.67]	< 0.001	Women had lower odds.
Carotid Revascularization	0.56 [0.46-0.69]	< 0.001	Women had lower odds.
Rehabilitation	0.84 [0.68-1.04]	0.116	-
Inpatient rehabilitation	0.77 [0.61-0.98]	0.037	Women had lower odds.
Outpatient rehabilitation	0.56 [0.36-0.86]	0.010	Women had lower odds.
Platelet inhibitors - i8070 = Mono	0.76 [0.7-0.83]	< 0.001	Women had lower odds.
Platelet inhibitors - i8070 = Dual	0.72 [0.56-0.94]	0.017	Women had lower odds.
Platelet inhibitors - i26019 = Mono	1.05 [0.88-1.26]	0.591	-
Platelet inhibitors - i26019 = Dual	0.77 [0.55-1.08]	0.125	-
OAC at Follow-Up	0.84 [0.67-1.06]	0.153	-

Table 49: Population 2 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors with Unknown included

Variable	OR [95% CI]	P-Value	Interpretation
Length of hospital stay = 1-2	0.97 [0.82-1.14]	0.685	-
Length of hospital stay = 3-7	0.94 [0.79-1.12]	0.497	-
Length of hospital stay = 8-14	1.06 [0.75-1.48]	0.743	-
Length of hospital stay = 14+	1.28 [0.68-2.4]	0.452	-
CCT	0.95 [0.86-1.05]	0.288	-
MRI	1.09 [0.99-1.2]	0.069	-
TTE	0.95 [0.87-1.03]	0.182	-
TTE = Planned	1 [0.91-1.1]	0.993	-
TTE = Yes	0.91 [0.83-0.99]	0.032	Women had lower odds.
TEE	0.75 [0.66-0.86]	< 0.001	Women had lower odds.
TEE = Planned	0.77 [0.65-0.9]	0.002	Women had lower odds.
TEE = Yes	0.73 [0.6-0.89]	0.002	Women had lower odds.
Antiplatelet agents = Mono	1.08 [0.97-1.2]	0.168	-
Antiplatelet agents = Dual	0.84 [0.74-0.95]	0.007	Women had lower odds.
Heparin = Low	0.97 [0.89-1.05]	0.440	-
Heparin = Medium	1.04 [0.88-1.22]	0.646	-
Heparin = High	0.7 [0.51-0.95]	0.021	Women had lower odds.
PTA	0.81 [0.53-1.23]	0.321	-
CEA	0.53 [0.42-0.66]	< 0.001	Women had lower odds.
Carotid Revascularization	0.57 [0.46-0.7]	< 0.001	Women had lower odds.
Rehabilitation	0.85 [0.69-1.04]	0.123	-
Inpatient rehabilitation	0.79 [0.62-1]	0.055	-
Outpatient rehabilitation	0.56 [0.35-0.86]	0.009	Women had lower odds.
Platelet inhibitors - i8070 = Mono	0.76 [0.7-0.83]	< 0.001	Women had lower odds.
Platelet inhibitors - i8070 = Dual	0.71 [0.55-0.93]	0.012	Women had lower odds.
Platelet inhibitors - i26019 = Mono	1.05 [0.88-1.26]	0.595	-
Platelet inhibitors - i26019 = Dual	0.75 [0.54-1.05]	0.091	-
OAC at Follow-Up	0.85 [0.67-1.07]	0.157	-

Table 50: Population 2 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors with Smoking and Alcohol Abuse imputed

Population 2 - Treatments by age group and gender

Age	Treatment: No Female	Treatment: Yes Female	Treatment: No Male	Treatment: Yes Male
18-49	10 (23.81%)	32 (76.19%)	14 (18.92%)	60 (81.08%)
50-59	18 (16.67%)	90 (83.33%)	29 (15.18%)	162 (84.82%)
60-69	34 (16.04%)	178 (83.96%)	48 (12.53%)	335 (87.47%)
70-79	63 (14.93%)	359 (85.07%)	50 (10.44%)	429 (89.56%)
80-89	83 (19.26%)	348 (80.74%)	48 (11.54%)	368 (88.46%)
90+	19 (34.55%)	36 (65.45%)	11 (30.56%)	25 (69.44%)

Table 51: Population 2 - Regular lipid-lowering drugs for patients with Hypercholesterolemia

Age	Treatment: No Female	Treatment: Yes Female	Treatment: No Male	Treatment: Yes Male
18-49	3 (13.04%)	20 (86.96%)	15 (32.61%)	31 (67.39%)
50-59	15 (17.65%)	70 (82.35%)	28 (18.42%)	124 (81.58%)
60-69	18 (9.05%)	181 (90.95%)	40 (11.20%)	317 (88.80%)
70-79	56 (12.53%)	391 (87.47%)	53 (9.89%)	483 (90.11%)
80-89	67 (12.48%)	470 (87.52%)	59 (12.50%)	413 (87.50%)
90+	10 (10.31%)	87 (89.69%)	5 (9.09%)	50 (90.91%)

Table 52: Population 2 - Regular blood pressure checks for patients with Hypertension

Age	Treatment: No Female	Treatment: Yes Female	Treatment: No Male	Treatment: Yes Male
18-49	8 (34.78%)	15 (65.22%)	12 (25.53%)	35 (74.47%)
50-59	13 (15.29%)	72 (84.71%)	28 (17.83%)	129 (82.17%)
60-69	31 (15.50%)	169 (84.50%)	53 (14.52%)	312 (85.48%)
70-79	74 (16.23%)	382 (83.77%)	80 (14.68%)	465 (85.32%)
80-89	121 (22.45%)	418 (77.55%)	73 (15.47%)	399 (84.53%)
90+	40 (43.48%)	52 (56.52%)	18 (32.73%)	37 (67.27%)

Table 53: Population 2 - Regular antihypertensives for patients with Hypertension

Age	Treatment: No Female	Treatment: Yes Female	Treatment: No Male	Treatment: Yes Male
18-49	3 (30.00%)	7 (70.00%)	4 (30.77%)	9 (69.23%)
50-59	7 (31.82%)	15 (68.18%)	17 (38.64%)	27 (61.36%)
60-69	18 (33.33%)	36 (66.67%)	37 (33.33%)	74 (66.67%)
70-79	40 (34.19%)	77 (65.81%)	36 (18.09%)	163 (81.91%)
80-89	41 (28.67%)	102 (71.33%)	50 (33.33%)	100 (66.67%)
90+	3 (18.75%)	13 (81.25%)	5 (41.67%)	7 (58.33%)

Table 54: Population 2 - Regular antidiabetica for patients with Diabetes mellitus

Age	Treatment: No Female	Treatment: Yes Female	Treatment: No Male	Treatment: Yes Male
18-49	0 (0.00%)	2 (100%)	2 (33.33%)	4 (66.67%)
50-59	2 (66.67%)	1 (33.33%)	3 (18.75%)	13 (81.25%)
60-69	1 (6.25%)	15 (93.75%)	11 (21.57%)	40 (78.43%)
70-79	28 (24.35%)	87 (75.65%)	24 (16.33%)	123 (83.67%)
80-89	37 (21.39%)	136 (78.61%)	41 (22.40%)	142 (77.60%)
90+	18 (38.30%)	29 (61.70%)	8 (30.77%)	18 (69.23%)

Table 55: Population 2 - Oral Anticoagulants (OAC) for patients with Atrial fibrillation

Overview of Generative AI Tools Used

Tool: OpenAI ChatGPT - ChatGPT (GPT-4), as of April 2025

- Used as assistance in the search for literature sources for the methodology and discussion sections. All suggested sources were read and verified by myself.
- Used to draft BibLaTeX entries, which I then edited and adapted as needed.
- Used to help with phrasing and stylistic issues of individual paragraphs. Particularly to avoid redundancy and improve clarity of my writing. The generated text suggestions were used as inspiration, but were revised and rephrased by myself.

List of Figures

3.1 Example ROC Curves 22

4.1 Filter Steps Flow Charts for Population 1 and 2 26
4.2 *calculated* vs *database* ABCD2 Score 29
4.3 Risk Score Scatter-plots 31
4.4 Distribution of Length of Hospital Stay 48

5.1 AUC Values - ABCD2 75
5.2 'Database' Scores - mRS at Follow-Up > 4: Performance Metrics 78
5.3 'Database' Scores - Recurrence 1: Performance Metrics 79
5.4 'Database' Scores - Recurrence 2: Performance Metrics 81
5.5 'Database' Scores - Early Worsening: Performance Metrics 82
5.6 'Database' Scores - Recurrence within 90 Days: Performance Metrics . . . 83
5.7 AUC Values 85
5.8 'Calculated' Scores - mRS at Follow-Up > 4: Performance Metrics 88
5.9 'Calculated' Scores - Recurrence 3: Performance Metrics 89
5.10 'Calculated' Scores - Early Worsening: Performance Metrics 90
5.11 'Calculated' Scores - Recurrence within 90 days: Performance Metrics . . 91
5.12 Distribution of conditions by age group and gender - Population 1 101
5.13 Population 1 - Treatment Mosaic Plots 103
5.14 Distribution of conditions by age group and gender - Population 2 113
5.15 Population 2 - Treatment Mosaic Plots 115

1 'Database' Risk Scores ROC Curves 153
2 'Calculated' Risk Scores ROC Curves 158
3 'Database' Risk Scores Performance Metrics 163
4 'Calculated' Risk Scores Performance Metrics 183

List of Tables

2.1 ABCD2 and ABCD3-I Scores 10

3.1 Confusion Matrix 22

4.1 Population 1. Risk Scores Subset 32
4.2 Populations. All Variables 38
4.3 Health Outcomes 45
4.4 Recurrence Variables 45
4.5 Diagnostic methods and secondary prophylactic treatments 46
4.6 Illnesses and Treatments 46
4.7 Populations. Encoded diagnostic methods and treatments 49
4.8 Covariates: Patient Characteristics 52
4.9 Populations. Risk-Factors 53

5.1 Population 1. Health Outcomes 59
5.2 Population 1 - Unadjusted Outcome Odds Ratios by Gender (Significant Results Only) 60
5.3 Population 1 - Outcome odds ratios by gender controlled for Age (Significant Results Only) 61
5.4 Population 1 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors (Significant Results Only) 62
5.5 Population 1 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors with Unknown included (Significant Results Only) 62
5.6 Population 1 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors with Smoking and Alcohol Abuse imputed (Significant Results Only) 63
5.7 Population 2. Health Outcomes 65
5.8 Population 2 - Unadjusted outcome odds ratios by gender (Significant Results Only) 66
5.9 Population 2 - Outcome odds ratios by gender controlled for Age (Significant Results Only) 66
5.10 Population 2 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors (Significant Results Only) 67

5.11 Population 2 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors with Unknown included (Significant Results Only) . . . 67
5.12 Population 2 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors with Smoking and Alcohol Abuse imputed (Significant Results Only) . . . 67
5.13 Population 1. Risk Stratification Scores . . . 70
5.14 Population 1. Recurrence Variables . . . 72
5.15 Population 1. Diagnostic methods and secondary prophylactic treatments 94
5.16 Population 1 - Unadjusted treatment odds ratios by gender (Significant Results Only) . . . 97
5.17 Population 1 - Treatment odds ratios by gender controlled for Age (Significant Results Only) . . . 98
5.18 Population 1 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors (Significant Results Only) . . . 99
5.19 Population 1 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors with Unknown included (Significant Results Only) . . . 100
5.20 Population 1 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors with Smoking and Alcohol Abuse imputed (Significant Results Only) . . . 100
5.21 Population 1 - Treatment Rates per gender . . . 100
5.22 Population 2. Diagnostic methods and secondary prophylactic treatments 105
5.23 Population 2 - Unadjusted treatment odds ratios by gender (Significant Results Only) . . . 108
5.24 Population 2 - Treatment odds ratios by gender controlled for Age (Significant Results Only) . . . 109
5.25 Population 2 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors (Significant Results Only) . . . 110
5.26 Population 2 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors with Unknown included (Significant Results Only) . . . 111
5.27 Population 2 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors with Smoking and Alcohol Abuse imputed (Significant Results Only) . . . 112
5.28 Population 2 - Treatment Rates per gender . . . 112

1 Population 1. Patient Characteristics and Risk Factors . . . 127
2 Population 2. Patient Characteristics and Risk Factors . . . 129
3 Population 1 Risk-Scores Subset. Patient Characteristics and Risk Factors 130
4 Population 1 - Unadjusted outcome odds ratios by gender . . . 132
5 Population 1 - Outcome odds ratios by gender controlled for Age . . . 132

6 Population 1 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors . . . 133
7 Population 1 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors with Unknown included . . . 133
8 Population 1 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors with Smoking and Alcohol Abuse imputed . . . 134
9 Population 2 - Unadjusted outcome odds ratios by gender . . . 134
10 Population 2 - Outcome odds ratios by gender controlled for Age . . . 135
11 Population 2 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors . . . 135
12 Population 2 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors with Unknown included . . . 135
13 Population 2 - Outcome odds ratios by gender controlled for Patient Characteristics and Risk Factors with Smoking and Alcohol Abuse imputed . . . 136
14 'Database' Scores Distribution for Early Worsening . . . 136
15 'Database' Scores Distribution for Early Recurrence in Stroke Unit . . . 137
16 'Database' Scores Distribution for Recurrence within 90 days . . . 137
17 'Database' Scores Distribution for Combined vascular endpoint . . . 138
18 'Database' Scores Distribution for Cumulative Endpoint . . . 139
19 'Database' Scores Distribution for Recurrence 1 . . . 139
20 'Database' Scores Distribution for Recurrence 2 . . . 140
21 'Database' Scores Distribution for Recurrence 3 . . . 141
22 'Database' Scores Distribution for Recurrence 4 . . . 141
23 'Calculated' Scores Distribution for Early Worsening . . . 142
24 'Calculated' Scores Distribution for Early Recurrence in Stroke Unit . . . 143
25 'Calculated' Scores Distribution for Recurrence within 90 days . . . 143
26 'Calculated' Scores Distribution for Combined vascular endpoint . . . 144
27 'Calculated' Scores Distribution for Cumulative Endpoint . . . 145
28 'Calculated' Scores Distribution for Recurrence 1 . . . 145
29 'Calculated' Scores Distribution for Recurrence 2 . . . 146
30 'Calculated' Scores Distribution for Recurrence 3 . . . 147
31 'Calculated' Scores Distribution for Recurrence 4 . . . 147
32 'Database' ABCD2 AUC Values . . . 148
33 'Database' ABCD3-I AUC Values . . . 149
34 'Calculated' ABCD2 AUC Values . . . 150
35 'Calculated' ABCD3-I AUC Values . . . 151
36 Population 1 - Unadjusted treatment odds ratios by gender . . . 203
37 Population 1 - Treatment odds ratios by gender controlled for Age . . . 204
38 Population 1 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors . . . 205
39 Population 1 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors with Unknown included . . . 206

40 Population 1 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors with Smoking and Alcohol Abuse imputed . . 207
41 Population 1 - Regular lipid-lowering drugs for patients with Hypercholesterolemia 207
42 Population 1 - Regular blood pressure checks for patients with Hypertension 208
43 Population 1 - Regular antihypertensives for patients with Hypertension . 208
44 Population 1 - Regular antidiabetica for patients with Diabetes mellitus . 208
45 Population 1 - Oral Anticoagulants (OAC) for patients with Atrial fibrillation 208
46 Population 2 - Unadjusted treatment odds ratios by gender 209
47 Population 2 - Treatment odds ratios by gender controlled for Age 210
48 Population 2 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors 211
49 Population 2 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors with Unknown included 212
50 Population 2 - Treatment odds ratios by gender controlled for Patient Characteristics and Risk Factors with Smoking and Alcohol Abuse imputed . . 213
51 Population 2 - Regular lipid-lowering drugs for patients with Hypercholesterolemia 213
52 Population 2 - Regular blood pressure checks for patients with Hypertension 214
53 Population 2 - Regular antihypertensives for patients with Hypertension . 214
54 Population 2 - Regular antidiabetica for patients with Diabetes mellitus . 214
55 Population 2 - Oral Anticoagulants (OAC) for patients with Atrial fibrillation 214

Bibliography

[Alb+02] Gregory W. Albers et al. “Transient Ischemic Attack — Proposal for a New Definition”. en. In: *New England Journal of Medicine* 347.21 (2002), pp. 1713–1716. ISSN: 0028-4793. DOI: 10.1056/nejmsb020987. URL: http://dx.doi.org/10.1056/NEJMsb020987.

[Arr+08] Jasmin Arrich et al. “Influence of Socioeconomic Status and Gender on Stroke Treatment and Diagnostics”. In: *Stroke* 39.7 (2008), pp. 2066–2072. DOI: 10.1161/STROKEAHA.107.506147. URL: https://doi.org/10.1161/STROKEAHA.107.506147.

[BA02] Kenneth P. Burnham and David R. Anderson. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Second. New York, NY: Springer, 2002. ISBN: 0-387-95364-7.

[Bio13] National Center for Biotechnology Information (NCBI). *The NCBI Handbook*. Online. "Accessed: 2025-03-18". 2013. URL: https://www.ncbi.nlm.nih.gov/books/NBK126194/.

[Bon+21] Anna K. Bonkhoff et al. “Female Stroke: Sex Differences in Acute Treatment and Early Outcomes of Acute Ischemic Stroke”. en. In: *Stroke* 52.2 (2021), pp. 406–415. ISSN: 0039-2499. DOI: 10.1161/strokeaha.120.032850. URL: http://dx.doi.org/10.1161/STROKEAHA.120.032850.

[Buu12] Stef van Buuren. *Flexible Imputation of Missing Data*. Interdisciplinary Statistics Series. Chapman & Hall/CRC, 2012. ISBN: 978-1-4398-6825-6. URL: https://stefvanbuuren.name/fimd/.

[Cha+17] Mustapha Chamekh et al. “Differential Susceptibility to Infectious Respiratory Diseases between Males and Females Linked to Sex-Specific Innate Immune Inflammatory Response”. In: *Frontiers in Immunology* 8 (2017), p. 1806. DOI: 10.3389/fimmu.2017.01806. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC5733536/.

[Eas+09] J. Donald Easton et al. “Definition and Evaluation of Transient Ischemic Attack”. en. In: *Stroke* 40.6 (2009), pp. 2276–2293. ISSN: 0039-2499. DOI: 10.1161/strokeaha.108.192218. URL: http://dx.doi.org/10.1161/STROKEAHA.108.192218.

[Faw06] Tom Fawcett. “An Introduction to ROC Analysis”. In: *Pattern Recognition Letters* 27.8 (2006), pp. 861–874. DOI: 10.1016/j.patrec.2005.10.010. URL: https://doi.org/10.1016/j.patrec.2005.10.010.

[Fon+21] Ana Catarina Fonseca et al. “European Stroke Organisation (ESO) guidelines on management of transient ischaemic attack”. en. In: *European Stroke Journal* 6.2 (2021), pp. CLXIII–CLXXXVI. ISSN: 2396-9873. DOI: 10.1177/2396987321992905. URL: http://dx.doi.org/10.1177/2396987321992905.

[Gal+10] S. L. Gall et al. “Sex differences in presentation, severity, and management of stroke in a population-based study”. In: *Neurology* 74.12 (2010), pp. 975–981. DOI: 10.1212/WNL.0b013e3181d5a48f. URL: https://doi.org/10.1212/WNL.0b013e3181d5a48f.

[Ges24] Gesundheit Österreich GmbH (GÖG). *Koordinationsstelle Qualitätsregister*. "Accessed: 2025-03-16". 2024. URL: https://goeg.at/Koordinationsstelle_Qualitaetsregister.

[Goc+20] Sophia Gocan et al. “Diagnosis of Transient Ischemic Attack”. en. In: *Stroke* 51.11 (2020), pp. 3371–3374. ISSN: 0039-2499. DOI: 10.1161/strokeaha.120.031510. URL: http://dx.doi.org/10.1161/STROKEAHA.120.031510.

[Goo99] Steven N. Goodman. “Toward Evidence-Based Medical Statistics. 1: The P Value Fallacy”. In: *Annals of Internal Medicine* 130.12 (1999), pp. 995–1004. DOI: 10.7326/0003-4819-130-12-199906150-00008. URL: https://www.acpjournals.org/doi/10.7326/0003-4819-130-12-199906150-00008.

[HR09] Peter J. Huber and Elvezio M. Ronchetti. *Robust Statistics*. Second. Hoboken, NJ: Wiley, 2009. ISBN: 978-0-470-12990-6.

[Hur+19] Robert Hurford et al. “Prognostic value of “tissue-based” definitions of TIA and minor stroke”. en. In: *Neurology* 92.21 (2019). ISSN: 0028-3878. DOI: 10.1212/wnl.0000000000007531. URL: http://dx.doi.org/10.1212/WNL.0000000000007531.

[Jam+21] Gareth James et al. *An Introduction to Statistical Learning with Applications in R*. Second. Springer, 2021. ISBN: 978-1-0716-1417-4. URL: https://www.statlearning.com/.

[Joha] Johns Hopkins Medicine. *Carotid Endarterectomy*. "Accessed: 2025-03-18". URL: https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/carotid-endarterectomy.

[Johb] Johns Hopkins Medicine. *Percutaneous Transluminal Angioplasty*. "Accessed: 2025-03-18". URL: https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/percutaneous-transluminal-angioplasty.

[Johc] Johns Hopkins Medicine. *Stroke*. "Accessed: 2025-03-26". Johns Hopkins Medicine. URL: https://www.hopkinsmedicine.org/health/conditions-and-diseases/stroke.

[Joh+07] S Claiborne Johnston et al. "Validation and refinement of scores to predict very early stroke risk after transient ischaemic attack". en. In: *The Lancet* 369.9558 (2007), pp. 283–292. ISSN: 0140-6736. DOI: 10.1016/s0140-6736(07)60150-0. URL: http://dx.doi.org/10.1016/S0140-6736(07)60150-0.

[Kim+10] Ji-Sun Kim et al. "Gender Differences in the Functional Recovery after Acute Stroke". In: *Journal of Clinical Neurology* 6.4 (2010), pp. 183–188. DOI: 10.3988/jcn.2010.6.4.183. URL: https://pubmed.ncbi.nlm.nih.gov/21264198/.

[KK10] David G. Kleinbaum and Mitchel Klein. *Logistic Regression: A Self-Learning Text*. 3rd. Statistics for Biology and Health. Springer, 2010. ISBN: 978-1-4419-1741-6. DOI: 10.1007/978-1-4419-1742-3. URL: https://doi.org/10.1007/978-1-4419-1742-3.

[Kno+16] Michael Knoflach et al. "Predictive value of ABCD2 and ABCD3-I scores in TIA and minor stroke in the stroke unit setting". en. In: *Neurology* 87.9 (2016), pp. 861–869. ISSN: 0028-3878. DOI: 10.1212/wnl.0000000000003033. URL: http://dx.doi.org/10.1212/WNL.0000000000003033.

[Kre+23] Christine Kremer et al. "Sex differences in outcome after carotid revascularization in symptomatic and asymptomatic carotid artery stenosis". In: *Journal of Vascular Surgery* 78.3 (2023), pp. 817–827. DOI: 10.1016/j.jvs.2023.03.502. URL: https://doi.org/10.1016/j.jvs.2023.03.502.

[Les+07] Thabele M. Leslie-Mazwi et al. "Sex Differences in Stroke Evaluations in the Ischemic Stroke Genetics Study". In: *Journal of Stroke and Cerebrovascular Diseases* 16.5 (2007), pp. 187–193. DOI: 10.1016/j.jstrokecerebrovasdis.2007.03.004. URL: https://pubmed.ncbi.nlm.nih.gov/17845914/.

[Mad+19] Tracy E. Madsen et al. "Temporal Trends of Sex Differences in Transient Ischemic Attack Incidence Within a Population". en. In: *Journal of Stroke and Cerebrovascular Diseases* 28.9 (2019), pp. 2468–2474. ISSN: 1052-3057. DOI: 10.1016/j.jstrokecerebrovasdis.2019.06.020. URL: http://dx.doi.org/10.1016/j.jstrokecerebrovasdis.2019.06.020.

[Mer+10] Áine Merwick et al. "Addition of brain and carotid imaging to the ABCD2 score to identify patients at early risk of stroke after transient ischaemic attack: a multicentre observational study". en. In: *The Lancet Neurology* 9.11 (2010), pp. 1060–1069. ISSN: 1474-4422. DOI: 10.1016/s1474-4422(10)70240-4. URL: http://dx.doi.org/10.1016/S1474-4422(10)70240-4.

[Öst] Österreichische Gesellschaft für Schlaganfall-Forschung (ÖGSF). *Stroke Units in Österreich*. "Accessed: 2025-03-16". URL: https://www.xn--gsf-rna.at/stroke-units/.

[Ott+22] Jessica Ott et al. "Scoping 'sex' and 'gender' in rehabilitation: (mis)representations and effects". In: *International Journal for Equity in Health* 21.179 (2022). DOI: 10.1186/s12939-022-01787-1. URL: https://doi.org/10.1186/s12939-022-01787-1.

[Pan12] Nikolaos Pandis. "The effect size". In: *American Journal of Orthodontics and Dentofacial Orthopedics* 142.5 (2012), pp. 739–740. DOI: 10.1016/j.ajodo.2012.06.011. URL: http://dx.doi.org/10.1016/j.ajodo.2012.06.011.

[Pan13a] Nikolaos Pandis. "Confidence intervals rather than P values". In: *American Journal of Orthodontics and Dentofacial Orthopedics* 143.2 (2013), pp. 293–294. DOI: 10.1016/j.ajodo.2012.11.012. URL: http://dx.doi.org/10.1016/j.ajodo.2012.11.012.

[Pan13b] Nikolaos Pandis. "The P value problem". In: *American Journal of Orthodontics and Dentofacial Orthopedics* 143.1 (2013), pp. 150–151. DOI: 10.1016/j.ajodo.2012.10.005. URL: http://dx.doi.org/10.1016/j.ajodo.2012.10.005.

[Pur+21] Francisco Purroy et al. "Sex-Related Differences in Clinical Features, Neuroimaging, and Long-Term Prognosis After Transient Ischemic Attack". en. In: *Stroke* 52.2 (2021), pp. 424–433. ISSN: 0039-2499. DOI: 10.1161/strokeaha.120.032814. URL: http://dx.doi.org/10.1161/STROKEAHA.120.032814.

[Ram+22] Niveditta Ramkumar et al. "Association of Sex With the Use and Outcomes of Carotid Revascularization: A Cohort Study". In: *Stroke: Vascular and Interventional Neurology* 2.5 (2022), e000380. DOI: 10.1161/SVIN.122.000380. URL: https://doi.org/10.1161/SVIN.122.000380.

[Rip19] Brian Ripley. *Multinomial Log-Linear Models (multinom)*. "Accessed: 2025-03-23". 2019. URL: https://www.rdocumentation.org/packages/nnet/versions/7.3-20/topics/multinom.

[Rob+23] Xavier Robin et al. *pROC: Display and Analyze ROC Curves*. "Accessed: 2025-03-23". 2023. URL: https://www.rdocumentation.org/packages/pROC/versions/1.18.5.

[Roo+22] Liliane C. Roosendaal et al. "Editor's Choice – Sex Differences in Response to Administration of Heparin During Non-Cardiac Arterial Procedures". In: *European Journal of Vascular and Endovascular Surgery* 64.5 (Nov. 2022), pp. 557–565. DOI: 10.1016/j.ejvs.2022.08.005. URL: https://doi.org/10.1016/j.ejvs.2022.08.005.

[Sav+21] Jeffrey L. Saver et al. “Standardized Nomenclature for Modified Rankin Scale Global Disability Outcomes: Consensus Recommendations From Stroke Therapy Academic Industry Roundtable XI”. In: *Stroke* 52.9 (2021), pp. 3054–3062. DOI: 10.1161/STROKEAHA.121.034480. URL: https://www.ahajournals.org/doi/abs/10.1161/STROKEAHA.121.034480.

[Sol+23] Jonathan Solomonow et al. “Is Dual Antiplatelet Therapy Underutilized Following TIA and Minor Stroke? (S3.006)”. In: *Neurology* 100.17_supplement_2 (2023), S3.006. DOI: 10.1212/WNL.0000000000203373. URL: https://www.neurology.org/doi/10.1212/WNL.0000000000203373.

[Son+13] Bo Song et al. “Validation of the ABCD3-I Score to Predict Stroke Risk After Transient Ischemic Attack”. In: *Stroke* 44.5 (2013), pp. 1244–1248. DOI: 10.1161/STROKEAHA.113.000969. URL: https://www.ahajournals.org/doi/abs/10.1161/STROKEAHA.113.000969.

[Syn+16] Nathalie E. Synhaeve et al. “Women have a poorer very long-term functional outcome after stroke among adults aged 18–50 years: the FUTURE study”. In: *Journal of Neurology* 263.6 (2016), pp. 1095–1102. DOI: 10.1007/s00415-016-8042-2. URL: https://doi.org/10.1007/s00415-016-8042-2.

[Szu10] Magdalena Szumilas. “Explaining Odds Ratios”. In: *Journal of the Canadian Academy of Child and Adolescent Psychiatry* 19.3 (2010), pp. 227–229. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC2938757/.

[Tea19] R Core Team. *Generalized Linear Models (glm)*. "Accessed: 2025-03-23". 2019. URL: https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/glm.

[Tea25] R Core Team. *mice: Multivariate Imputation by Chained Equations*. "Accessed: 2025-03-23". 2025. URL: https://www.rdocumentation.org/packages/mice/versions/3.17.0/topics/mice.

[Win+13] M. Wintermark et al. “Imaging Recommendations for Acute Stroke and Transient Ischemic Attack Patients: A Joint Statement by the American Society of Neuroradiology, the American College of Radiology, and the Society of NeuroInterventional Surgery”. en. In: *American Journal of Neuroradiology* 34.11 (2013), E117–E127. ISSN: 0195-6108. DOI: 10.3174/ajnr.a3690. URL: http://dx.doi.org/10.3174/ajnr.A3690.

[Yan+01] Anthony L. Yancey et al. “Gender Is a Major Factor in Determining the Severity of Mycoplasma Respiratory Disease in Mice”. In: *Infection and Immunity* 69.5 (May 2001), pp. 2865–2871. DOI: 10.1128/IAI.69.5.2865-2871.2001. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC98236/.

[Yan+10] Jie Yang et al. “Validation of the ABCD2 Score to Identify the Patients With High Risk of Late Stroke After a Transient Ischemic Attack or Minor Ischemic Stroke”. In: *Stroke* 41.6 (2010), pp. 1298–1300. DOI: 10.1161/STROKEAHA.110.578757. URL: https://www.ahajournals.org/doi/abs/10.1161/STROKEAHA.110.578757.