Université de Montréal

Methodological challenges in the comparative assessment of effectiveness and safety of oral anticoagulants in individuals with atrial fibrillation using administrative healthcare data.

Par

Liliya Gubaidullina

Faculté de médecine

Thèse présentée en vue de l'obtention du grade de Philosophae Doctor (Ph.D.)
en sciences biomédicales, option médecine expérimentale
Août, 2021

Cette thèse intitulée: **Methodological challenges in the comparative assessment of effectiveness and safety of oral anticoagulants in individuals with atrial fibrillation using administrative healthcare data.**

Présentée par

Liliya Gubaidullina

a été évaluée par un jury composé des personnes suivantes:

Marc Jolicoeur

président-rapporteur

Madeleine Durand

directeur de recherche

Christel Renoux

co-directeur de recherche

Sylvie Perreault

membre du jury

Caroline Sirois

examinateur externe

Rémi Goupil

représentant du doyen

Août, 2021

Résumé

La fibrillation auriculaire (FA), l'arythmie cardiaque la plus courante est un facteur de risque majeur pour le développement de l'accident vasculaire cérébral ischémique (AVC). Les anticoagulants oraux directs (AOD) ont largement remplacé la warfarine en usage clinique pour la prévention des AVC dans la FA. Cette recherche a examiné deux défis méthodologiques importants qui peuvent survenir dans les études observationnelles sur l'efficacité et l'innocuité comparatives des AOD et de la warfarine. Premièrement : un biais d'information résultant d'une classification erronée de l'exposition au traitement à la warfarine suite aux ajustements de doses fréquentes qui ne sont pas adéquatement consignés dans les données de dispensations pharmacologiques. Deuxièmement : un biais de sélection, en raison de la censure informative, généré par des mécanismes de censure différentiels, chez les patients exposés aux AOD, ou à la warfarine.

À l'aide des données administratives du Québec, j'ai mené trois études de cohortes rétrospectives qui ont portées sur toutes les personnes ayant initié un anticoagulant oral de 2010 à 2016. Ces études étaient restreintes aux résidents du Québec couverts par le régime public d'assurance médicaments (environ 40% de la population au Québec), c'est-à-dire : des personnes âgées de 65 ans et plus; des bénéficiaires de l'aide sociale; des personnes qui n'ont pas accès à une assurance-maladie privée; et les personnes à leur charge.

Dans la première étude, nous avons émis l'hypothèse que les données sur les réclamations en pharmacie ne reflètent pas correctement la durée de la dispensation de la warfarine. Les écarts entre les renouvellements consécutifs étaient plus grands pour la warfarine que les AOD. Dans

cette étude, on a trouvé que l'écart moyen pour les usagers de la warfarine était de 9.3 jours (avec un intervalle de confiance de 95% [IC]: 8.97-9.59), l'apixaban de 3.08 jours (IC de 95%: 2.96--3.20), et de 3.15 jours pour le rivaroxaban (IC de 95%: 3.03-3.27). Les écarts entre les renouvellements consécutifs présentaient une plus grande variabilité chez les personnes qui prenaient de la warfarine comparativement à celles qui prenaient des AOD. Cette variation peut refléter les changements de posologie de la warfarine lorsque la dose quotidienne est ajustée par le professionnel de la santé en fonction des résultats du rapport normalisé international (INR). L'ajustement de la dose peut prolonger (ou raccourcir) la période couverte par le nombre de comprimés délivrés.

Dans la deuxième étude, nous avons émis l'hypothèse que la définition de la durée d'exposition basée sur la variable des « jours fournis », disponible dans la base de données, et le délai de grâce fixe, entraîneront une erreur de classification différentielle de l'exposition à la warfarine par rapport aux AOD. Dans cette étude, on a utilisé deux approches pour définir la durée des dispensations : la variable des « jours fournis » disponible dans la base de données ainsi qu'une approche axée sur les données pour la définition de la durée de dispensation qui tient compte des antécédents de distribution précédents. La deuxième étude a révélé qu'en utilisant la variable des « jours fournis », la durée moyenne (et l'écart type) des durées des dispensations pour le dabigatran, le rivaroxaban, et la warfarine étaient de 19 (15), 19 (14), et de 13 (12) jours, respectivement. En utilisant l'approche fondée sur des données, les durées étaient de 20 (16), 19 (15), et de 15 (16) jours, respectivement. Ainsi, l'approche fondée sur les données s'est rapprochée de la variable des « jours fournis » pour les thérapies à dose standard telles que le

dabigatran et le rivaroxaban. Une approche axée sur les données pour la définition de la durée de dispensation, qui tient compte des antécédents de distribution précédents, permet de mieux saisir la variabilité de la durée de dispensation de la warfarine par rapport à la méthode basée sur la variable des « jours fournis ». Toutefois, cela n'a pas eu d'impact sur les estimations du rapport de risque sur la sécurité comparative des AOD par rapport à la warfarine.

Dans la troisième étude, nous avons émis l'hypothèse que lors de l'évaluation de l'effet d'un traitement continu avec des anticoagulants oraux (l'analyse per-protocole), la censure élimine les patients les plus malades du groupe des AOD et des patients en meilleure santé du groupe de warfarine. Cela peut baisser l'estimation de l'efficacité et de l'innocuité comparative en faveur des AOD. L'étude a démontré que les mécanismes de censure chez les initiateurs d'AOD et de warfarine étaient différents. Ainsi, certaines covariables pronostiquement significatives, telles que l'insuffisance rénale chronique et l'insuffisance cardiaque congestive, étaient associées avec une augmentation de la probabilité de censure chez les initiateurs d'AOD, et une diminution de la probabilité de censure chez les initiateurs de warfarine. Pour corriger le biais de sélection introduit par la censure, nous avons appliqué la méthode de pondération par la probabilité inverse de censure. Deux stratégies de spécification du modèle pour l'estimation des poids de censure ont été explorées : le modèle non stratifié, et le modèle stratifié en fonction de l'exposition. L'étude a démontré que lorsque les poids de censure sont générés sans tenir compte des dynamiques de censure spécifiques, les estimés ponctuels sont biaisés de 15% en faveur des AOD par rapport à l'ajustement des estimés ponctuels avec des poids de censure stratifiée selon l'exposition (rapport

de risque: 1.41; IC de 95%: 1.34, 1.48 et rapport de risque: 1.26; IC de 95%: 1.20, 1.33, respectivement).

Dans l'ensemble, les résultats de cette thèse ont d'importantes implications méthodologiques pour les futures études pharmacoépidémiologiques. À la lumière de ceux-ci, les résultats des études observationnelles précédentes peuvent être revus et une certaine hétérogénéité peut être expliquée. Les résultats pourraient également être extrapolés à d'autres questions cliniques.

Mots-clés: fibrillation auriculaire, anticoagulants oraux, biais d'information, biais de sélection, données administratives, erreur de classification, biais de sélection introduit par la censure, pondération par la probabilité inverse de censure, pharmacoépidémiologie, études observationnelles, l'efficacité et l'innocuité comparatives, inférence causale, durée d'exposition, méthodologie

Abstract

Atrial fibrillation (AF), the most common cardiac arrhythmia is a major risk factor for the development of ischemic stroke. Direct oral anticoagulants (DOACs) replaced warfarin in clinical use for stroke prevention in AF. This research investigated two important methodological challenges that may arise in observational studies on the comparative effectiveness and safety of DOACs and warfarin. First, an information bias resulting from misclassification of exposure to dose-varying warfarin therapy when using days supplied value recorded in pharmacy claims data. Second, a selection bias due to informative censoring with differential censoring mechanisms in the DOACs- and the warfarin exposure groups.

Using the Québec administrative databases, I conducted three retrospective cohort studies that included patients initiating an oral anticoagulant between 2010 and 2016. The studies were restricted to Québec residents covered by the public drug insurance plan (about 40% of Québec's population), including those aged 65 years and older, welfare recipients, those not covered by private medical insurance, and their dependents.

In the first study, we hypothesized that pharmacy claims data inadequately captured the duration of the dispensation of warfarin. Gaps between subsequent dispensations (refill gaps) and their variation are larger for warfarin than for DOACs. In this study, we found that the average refill gap for the users of warfarin was 9.3 days (95% confidence interval [CI]:8.97-9.59), apixaban 3.08 days (95%CI: 2.96--3.20), dabigatran 3.70 days (95%CI: 3.56-3.84) and rivaroxaban 3.15 days (95%CI: 3.03-3.27). The variance of refill gaps was greater among warfarin users than among DOAC users. This variation may reflect the changes in warfarin posology when the daily dose is

adjusted by a physician or a pharmacist based on previously observed international normalized ratio (INR) results. The dose adjustment may lead to prolongation of the period covered by the number of dispensed pills.

In the second study, we hypothesized that the definition of duration of dispensation based on the days supplied value and a fixed grace period will lead to differential misclassification of exposure to warfarin and DOACs. This may bias the estimate of comparative safety in favor of DOACs. In this study, we used two approaches to define the duration of dispensations: the recorded days supplied value, and the longitudinal coverage approximation (data-driven) that may account for individual variation in drug usage patterns. The second study found that using the days supplied, the mean (and standard deviation) dispensation durations for dabigatran, rivaroxaban, and warfarin were 19 (15), 19 (14), and 13 (12) days, respectively. Using the data-driven approach, the durations were 20 (16), 19 (15), and 15 (16) days, respectively. Thus, the data-driven approach closely approximated the recorded days supplied value for the standard dose therapies such as dabigatran and rivaroxaban. For warfarin, the data-driven approach captured more variability in the duration of dispensations compared to the days supplied value, which may better reflect the true drug-taking behavior of warfarin. However, this did not impact the hazard ratio estimates on the comparative safety of DOACs vs. warfarin.

In the third study, we hypothesized that when assessing the effect of continuous treatment with oral anticoagulants (per-protocol effect), censoring removes sicker patients from the DOACs group and healthier patients from the warfarin group. This may bias the estimate of comparative effectiveness and safety in favor of DOACs. The study showed that the mechanisms of censoring in the DOAC

and the warfarin exposure groups were different. Thus, prognostically meaningful covariates, such as chronic renal failure and congestive heart failure, had an opposite direction of association with the probability of censoring in the DOACs and warfarin groups. To correct the selection bias introduced by censoring, we applied the inverse probability of censoring weights. Two strategies for the specification of the model for estimation of censoring weights were explored: exposure-unstratified and exposure-stratified. The study found that exposure-unstratified censoring weights did not account for the differential mechanism of censoring across the treatment group and failed to eliminate the selection bias. The hazard ratio associated with continuous treatment with warfarin versus DOACs adjusted with exposure unstratified censoring weights was 15% biased in favor of DOACs compared to the hazard ratio adjusted with exposure-stratified censoring weights (hazard ratio: 1.41; 95% CI: 1.34, 1.48 and hazard ratio: 1.26; 95%CI: 1.20, 1.33, respectively).

Overall, the findings of this thesis have important methodological implications for future pharmacoepidemiologic studies. Moreover, the results of the previous observational studies can be reappraised, and some heterogeneity can be explained. The findings can be extrapolated to other clinical questions.

Keywords: Atrial fibrillation, Comparative effectiveness and safety, Observational study, Administrative claims data, Exposure misclassification, Selection bias, Censoring weights, Causal inference, Oral anticoagulants, Exposure definition, Information bias, Pharmacoepidemiology, Study methods

Acknowledgments

I would like to express my gratitude to all those whose support has helped me in my professional growth in the field of pharmacoepidemiology.

I would like to express my gratitude to my supervisor, Dr. Madeleine Durand, for making my Ph.D. experience possible. Thank you for facilitating my research and for your creativity, patience, and motivation.

To my co-supervisor, Dr. Christel Renoux, thank you for all your thoughtful advice, insight, and encouragement throughout my Ph.D. adventure.

I want to express my gratitude to the jury members Dr. Marc Jolicoeur, Dr. Sylvie Perreau, Dr. Caroline Sirois, and Dr. Rémi Goupil for their time and thorough review of my dissertation and for giving me valuable advice that helped me improve my work.

I would like to thank the Fonds de recherche du Québec – Santé (FRQS), and the Canadian network for observational drug effect studies (CNODES) for their financial assistance without which the completion of my research would not have been possible. Also, I would like to express my gratitude to CNODES for the outstanding training and networking opportunities through regular contact with the leading researchers in pharmacoepidemiology in Canada.

A very special thank you to Dr. Jacques LeLorier and his research group for sharing their expertise and recommendations that enhanced my research experience. I would like to thank Sophie Dellaniello, Sarasa Johnson, and Yu Lan Jin for their prompt help with programming and data extraction.

Table of contents

List of tables

List of figures

List of abbreviations

ACC	American College of Cardiology
AF	Atrial fibrillation
AHA	American Heart Association
AIDS	Acquired immunodeficiency syndrome
AOD	*Anticoagulants oraux directs*
ASHP	American Society of Health-System Pharmacists
AVC	*Accident vasculaire cérébral*
BID	Two times a day
CAB	Coronary artery bypass surgery
CAI	*Commission d'accès à l'information*
CCS	Canadian Cardiovascular Society
CHF	Congestive heart failure
CHSLD	*Centres d'hébergement de soins de longue durée*
CI	Confidence interval
CIHI	Canadian Institute for Health Information
COPD	Chronic obstructive pulmonary disease
DAG	Directed acyclic graphs
DALY	Disability adjusted life years
DIN	Drug identification number
DOAC	Direct oral anticoagulant
ECG	Electrocardiogram
ESC	European Society of Cardiology
FDA	Food and Drug Administration
HMO	Health maintenance organization
HR	Hazard ratio
HRS	Heart Rhythm Society
ICC	Intraclass correlation
ICD	International classification of diseases
INR	International normalized ratio
IPCW	Inverse probability of censoring weights
ISQ	*Institut de la statistique du Québec*
IQR	Interquartile range
IRR	Incidence rate ratios
ITT	Intention-to-treat

MED-ÉCHO	*Maintenance et exploitation des données pour l'étude de la clientèle hospitalière*
MI	Myocardial infarction
MSSS	*Ministère de la Santé et des services sociaux*
NSAID	Nonsteroidal anti-inflammatory drug
NVAF	Non-valvular atrial fibrillation
OAC	Oral anticoagulants
OR	Odds ratios
OTC	Over-the-counter
PCI	Percutaneous coronary intervention
PDC	Proportion of days covered
PS	Propensity scores
RAMQ	*Régie de l'assurance-maladie du Québec*
RCT	Randomized controlled trials
RR	Risk ratio
SD	Standard deviation
SE	Standard error
SSRI	Selective serotonin reuptake inhibitor
TIA	Transient ischemic attack
US	United States
VTE	Venous thromboembolism

CHAPTER 1: Background and significance

1.1 Pathophysiology and classification of atrial fibrillation

Atrial fibrillation (AF) is the most common type of sustained cardiac arrhythmia. It is associated with reduced quality of life, increased risk of disability, and all-cause mortality[1].

In AF, the cardiac electrical conduction system fails to effectively transmit electrical signals to cause the contraction of the myocardium (Figure 1.1). In a healthy heart, the ordered electrical signal is fired at the sinoatrial node and triggers atria contraction. Then, the signal travels through the atrioventricular node, and after a slight delay propagates to the bundle of His and Purkinje fibers to trigger ventricular contraction. In AF, multiple spontaneous signals are fired arising either from local ectopic firing, a single localized reentry circuit, or multiple functional reentry circuits[2]. These result in inadequate atrial and ventricular contractions and decreased general blood circulation[3]. Continuous circulatory disturbance leads to the development of clinical manifestations, e.g. lightheadedness, shortness of breath, fatigue, and chest pain. However, most individuals with AF are asymptomatic[4].

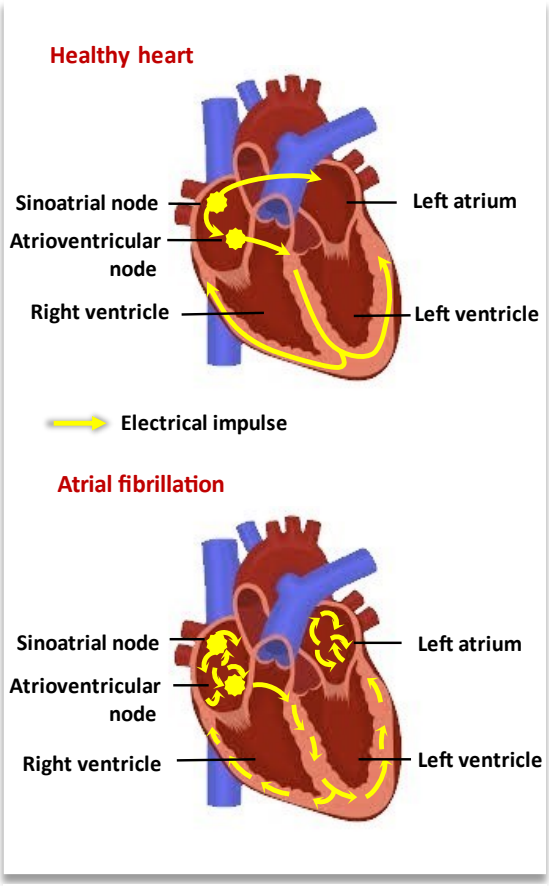


Figure 1.1. Healthy heart and atrial fibrillation.

The American Heart Association/American College of Cardiology/Heart Rhythm Society (AHA/ACC/HRS) 2014 Guideline for the Management of Patients with Atrial Fibrillation provides a clinically relevant classification that is based on the duration of episodes[5]. The Guidelines distinguish *paroxysmal* AF when the episodes terminate spontaneously within seven days, however, may recur in the future. *Persistent* AF occurs when the arrhythmia lasts seven days or longer without self-termination. A *long-standing* AF is a persistent AF that lasts more than one year. A *permanent* AF is a long-standing AF in which the normal heart rhythm cannot be restored.

A *non-valvular* AF refers to the presence of AF in individuals without a rheumatic mitral valve disease, prosthetic valve, or mitral valve repair[5]. The 2020 update of the American Heart Association/American College of Cardiology/Heart Rhythm Society (AHA/ACC/HRS) 2014 Guideline for the Management of Patients with Atrial Fibrillation clarifies that "valvular atrial fibrillation" refers to patients with either moderate or severe mitral stenosis or a mechanical heart valve[6].

1.2 Epidemiology of atrial fibrillation

The worldwide estimated prevalence of AF was 37, 574 million individuals (492 cases per 100,000 inhabitants) in 2017. It was 28,533 million individuals (422 cases per 100,000 inhabitants) in 2007, thus, increasing by 14% over the 10 years. Older individuals have a higher prevalence rate compared to younger individuals: in 2017, the prevalence rate was 5,062 per 100,000 in those aged ≥70 years, 996 per 100,000 in those aged 50-69 years, and 65 per 100,000 in those aged 15-49 years. Males have a higher age-standardized prevalence compared to females (in 2017, 561 and 416 per 100,000, respectively)[7].

The incidence of AF was 3.046 million new cases (40 cases per 100,000 inhabitants) in 2017, a 15% increase from 2.315 million new cases in 2007 (34 new cases per 100,000 inhabitants)[7]. By age group, the incidence was 321 in those aged ≥70 years, 110 in those 50-69 years, and five in those 15-49 years. In 2017, the incidence of AF in males was higher compared to females (41 and 38 per 100,000, respectively)[7].

The total estimated burden of AF was 5.976 million Disability Adjusted Life Years (DALYs) in 2017, increasing by 25% from 2007 when it was 4.397 million DALYs. The number of deaths for AF also considerably increased over 10 years. In 2017 there were an estimated 0.29 million deaths, an increase from 0.19 million deaths in 2007.

Overall, developed countries have a higher prevalence of AF as compared to developing countries[8]. In 2017, countries with the highest AF prevalence were Sweden (1,735 cases per 100,000 inhabitants), New Zealand (1,612 cases per 100,000 inhabitants), and the United States, US (1,565 cases per 100,000 inhabitants). In Canada, in 2017, the prevalence was also above the global

average with 1,243 cases per 100,000 inhabitants[7]. The higher prevalence of AF in developed countries is explained by a greater proportion of older individuals, enhanced detection, and longer survival of patients with AF[9].

In Québec, a study based on the analysis of the Québec administrative data (the *Régie de l'assurance-maladie du Québec*, RAMQ) reported that during the period from 2000 to 2009, the age- and sex-standardized incidence rate was as high as 324 (95% Confidence Interval, CI: 323-325) per 100,000 per year[10]. Men had a higher risk of incident AF compared to women (age-adjusted incidence rate ratio 1.51; 95% CI 1.50 to 1.52)[10].

1.3 Risk factors for atrial fibrillation

The factors associated with the development of AF are summarised in Figure 1.2.

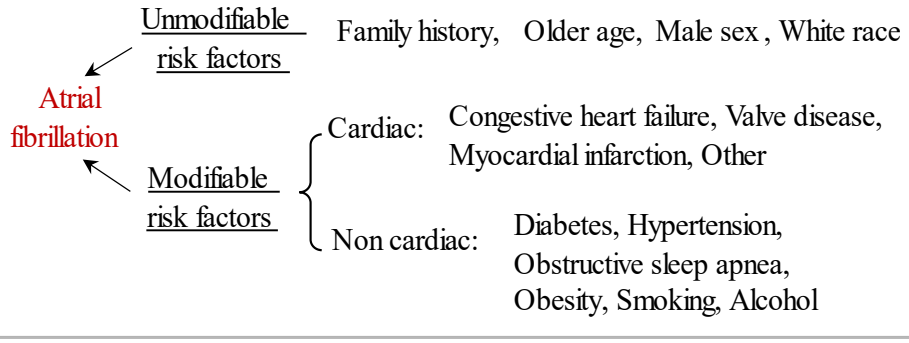


Figure 1.2. Risk factors associated with the development of AF.

Risk factors for AF can be grouped as modifiable and unmodifiable. The latter include family history, age, sex, and race (Figure 1.2). The Framingham Heart Study found that age was the strongest risk factor among other risk factors such as sex, body mass, diabetes, smoking, alcohol consumption, hypertension, and heart disease[11]. The study reported the incidence of AF (diagnosed with ECG) per 1000 person-years (py) was 0.5 for those aged 45–54 years, 1.1 for aged 55–64, 3.2 for those aged 65–74, 6.2 for those aged 75–84, and 7.7 for individuals aged ≥85 years[12]. Age was also associated with a more severe clinical course of AF: those older than 65 years were more often hospitalized and had higher mortality compared with individuals aged 65 years and younger[13]. The risk of AF is higher in men compared to women, in Whites compared to African Americans, Hispanics, or Asians[11,14,15]. Family history was associated with a 40% increased risk of developing AF in first-degree relatives[16].

Modifiable risk factors for atrial fibrillation include cardiac and non-cardiac conditions. Congestive heart failure (CHF) may increase the risk of AF by 4.5 and 5.9 times in men and women, respectively. Myocardial infarction was associated with the risk of AF only in men (40% increased risk). Other cardiac factors that may increase the risk of AF include coronary artery disease, hypertrophic cardiomyopathy, dilated cardiomyopathy, congenital heart disease, restrictive cardiomyopathies (e.g., amyloidosis, hemochromatosis, and endomyocardial fibrosis), cardiac tumors, constrictive pericarditis, calcification of the mitral annulus, right-sided heart failure, and idiopathic dilatation of the right atrium[17].

Diabetes mellitus may increase the risk of AF by 40% in men and 60% in women[18]. Duration of diabetes was associated with a 3% increased risk of AF for every additional year[19]. Hypertension increases the risk of AF by 50% and 40% in men and women, respectively[20]. Studies reported that treatment of diabetes or hypertension may modify the risk of AF. Thus, diabetic patients treated with thiazolidinediones (TZD) had a 31% decreased risk of AF compared to untreated patients[21]. Treatment with renin-angiotensin-aldosterone system blockers losartan[22] and valsartan[23] decreased the risk of AF by 33% and 20%, respectively. Telmisartan may reduce the risk of AF by 46%[24].

Obstructive sleep apnea may increase the risk of AF[25]. Therapy with continuous positive airway pressure may decrease the risk of recurrence of AF[26]. Obesity increases the risk of diseases associated with the development of AF, such as hypertension, diabetes, heart disease, and obstructive sleep apnea. Moreover, obesity was found to be independently associated with 49% higher AF risk[27]. Weight management programs may modify the risk of AF development[28]. Current smokers had a 200%, and ever-smokers a 58% higher risk of AF, compared to non-smokers[29].

Heavy alcohol consumption (more than 14 drinks per week) was associated with a 39% higher risk of AF[30].

Once, an individual develops AF, he/she has an increased risk of thromboembolism. That is outlined in the next section.

1.4 Risk of systemic thromboembolism in AF

Systemic thromboembolism is a serious complication in AF leading to disability and increased risk of mortality[5]. Systemic thromboembolism occurs when an arterial blood clot or thrombus forms in the heart chambers (mostly in the left auricular appendage) because of abnormal blood flow, damage to the heart endothelium, and abnormal blood coagulation[5]. Systemic thromboembolism includes ischemic stroke and extracranial systemic thromboembolism. A study, based on the analysis of all cases of thromboembolism from randomized clinical trials of anticoagulation in AF reported that among all participants, the incidence rate of stroke was 1.92, and extracranial systemic thromboembolism was 0.24 per 100 person years[31].

The Framingham Heart Study found that patients with valvular AF have an 18-fold, and patients with non-valvular AF have up to 6-fold higher risk of stroke compared to age- and blood pressure-matched patients without AF[32].

The risk of thromboembolism in AF is predicted using stratification schemes such as the CHA_2DS_2-VASc (AHA/ACC/HRS)[5], CHA_2DS_2-VASc (ESC)[33], and CHADS-65[34] (CCS) scores (Figure 1.3).

		CHA_2DS_2-VASc (AHA/ACC/HRS)	CHA_2DS_2-VASc (ESC)	CHADS-65 (CCS)
1	Age 65-74 years	1	1	1
2	Age ≥75 years	2	2	1*
3	Diabetes mellitus	1	1	1
4	Hypertension	1	1	1
5	Congestive heart failure	1	1	1
6	Vascular disease	1	1	N/A
7	Stroke, TIA, systemic embolism	2	2	1
8	Female Sex	1	1**	N/A

AHA, American Heart Association; ACC, American College of Cardiology ; HRS, Heart Rhythm Society; ESC, European Society of Cardiology; CCS, Canadian Cardiovascular Society; TIA, transient ischemic attack; N/A, not applicable.
* Included as part of age 65 years and older.
**Not counted in the absence of other risk factors

Figure 1.3. Stratification schemes for predicting thromboembolism in individuals with AF. Source: adapted from Andrade et al., 2017.

With these schemes, the risk of thromboembolism is calculated for each patient by summing up the scores for the manifested risk factors. The higher the calculated scores are, the higher the risk of thromboembolism. Figure 1.4 shows the annual rate of thromboembolism in patients with AF based on the value of the CHA_2DS_2 -VASc (ESC) score[35].

CHA_2DS_2 –VASc (ESC) Score	0	1	2	3	4	5	6	7	8	9
Thromboembolism, % per year	0	1.3	2.2	3.2	4.0	6.7	9.8	9.8	6.7	15.2

Figure 1.4. Annual rate of thromboembolism based on CHA2DS2 -VASc scoring. Source: adapted from Martin et al., 2017.

The antithrombotic treatment for the prevention of thromboembolism is recommended based on the total scores[36]. Thus, the CCS strongly recommends oral anticoagulants (OAC) for all patients with scores of ≥1[34]. The ESC considers that for men with a score of one, the treatment with OACs is reasonable, and for all patients with scores of ≥2, the treatment with OACs is strongly recommended[33]. The AHA/ACC/HRS strongly recommends OACs for those with scores of ≥2[5].

Other factors increasing the risk of thromboembolism in AF are not included in the risk stratification schemes. Thus, among AF patients, those with valvular heart disease have a higher risk of stroke compared to those without valvular heart disease[32]. Patients with a mechanical valve may have a higher risk of thrombosis compared to patients with a biologic valve[37]. Patients with persistent AF had higher adjusted rates of stroke or systemic embolism compared to those with paroxysmal AF (2.18 vs. 1.73 events per 100 patient-years, p=0.048)[38]. Next, patients with atrial fibrillation and a renal disease requiring renal replacement therapy were found to have an 80% (HR 1.83, 95%CI: 1.57–2.14) higher risk of stroke compared with those without such disease[39]. Furthermore, AF patients with renal failure requiring hemodialysis, are not only at a higher risk of thromboembolism but also have an increased risk of bleeding. Anticoagulation in these patients remains a challenge and is strictly individualized[40].

Smoking was found to be significantly associated with an increased risk of thromboembolism, however, the addition of this factor to CHA_2DS_2-VASc did not improve prediction[41].

For the prevention of thromboembolism in individuals with AF, lifelong therapy with oral anticoagulants is recommended[5].

1.5 Prevention of thromboembolism in non-valvular AF with oral anticoagulants

Anticoagulation therapy for stroke prevention is a part of the integrated care of patients with AF aiming at improving life expectancy, quality of life, autonomy, and social functioning[42].

There are two classes of oral anticoagulants: vitamin K antagonists (e.g., warfarin) and direct oral anticoagulants (DOACs: dabigatran, rivaroxaban, apixaban, edoxaban). Figure 1.5 shows the part of the coagulation cascade and the points of action of oral anticoagulants[43,44].

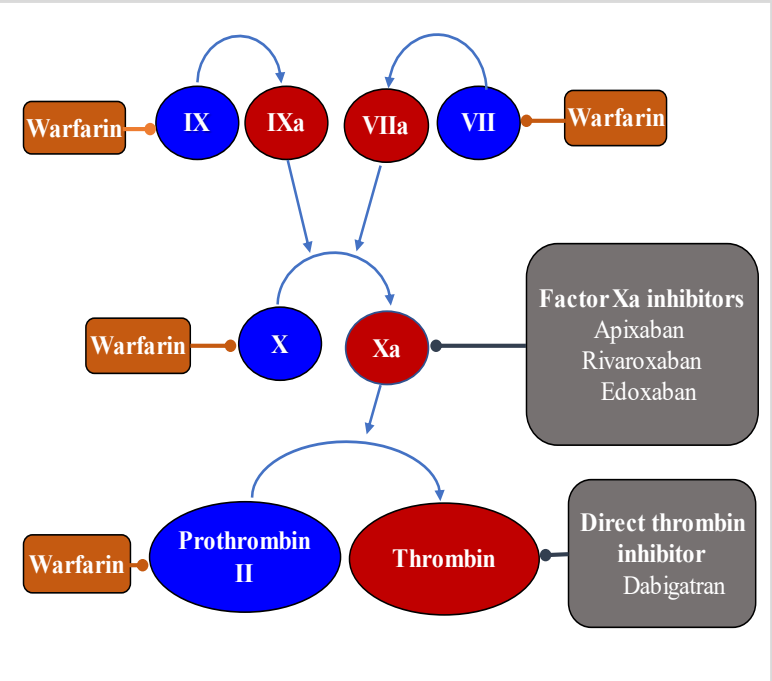


Figure 1.5. Coagulation cascade and points of action of oral anticoagulants. Source: adapted from Makaryus et al., 2013 and Dempfle et al., 2014.

Vitamin K antagonist, warfarin was approved by the United States Food and Drug Administration (FDA) in 1954. The earliest marketed date of warfarin recorded in the Drug Product Database of Health Canada is December 31, 1957[45].

Warfarin is inexpensive and safe to use in patients with renal failure. A meta-analysis of randomized controlled trials (RCT) found that compared to placebo, warfarin reduced the risk of stroke by 62% (pooled odds ratio 0.38; 95%CI: 0.28-0.52), and all-cause mortality by 26% (pooled odds ratio 0.74; 95%CI:0.57-0.96)[46]. However, bleeding is the major side-effect of warfarin. Compared to placebo,

warfarin may be associated with a 90% higher risk of major bleeding (pooled odds ratio 1.9; 95%CI:0.89-4.04)[47]. Furthermore, there was a significantly increased risk of intracranial bleeding in those taking versus those not taking warfarin (hazard ratio 1.97; 95%CI:1.24-3.13)[48]. The daily dose of warfarin is not standard and can vary from 2 to 10 mg[49]. Moreover, the required dosage for each individual is unpredictable and is adjusted during the treatment based on international normalized ratio (INR) measurements. Warfarin has a narrow therapeutic index: it is effective for the prevention of thromboembolism only if the anticoagulation level is maintained within the INR range of 2 to 3[50-52]. Higher INR values increase the risk of bleeding: the incidence of major bleeding was 50% higher in patients on warfarin with INR>3.0 compared to those with INR 2.0-3.0[53].

Time in therapeutic range (TTR) measures the proportion of days that a patient on warfarin was within INR of 2.0 to 3.0 over the total number of days on treatment[54]. A greater TTR is associated with improved outcomes[55]. Thus, a study reported that oral anticoagulation with vitamin K antagonists was beneficial compared to dual antiplatelet therapy only in patients with TTR maintained above 58% to 65%[56]. In real-world clinical settings, the reported TTRs varied from 55%[57] to 76%[58]. In clinical trial settings, the TTR was found to be higher compared to real-world settings (absolute difference of 12.2%; 95% CI, from 4.8% to 19.5%; $p < 0.0001$)[59].

In patients treated with warfarin, the risk of major bleeding was higher during the first 90 days of therapy compared with the later period (11.0 vs 6.3 per 100 person-years, risk ratio (RR):1.75, 95%CI: 1.27-2.44, $p < 0.001$)[60]. Furthermore, poor anticoagulation control due to laboratory error (now rare), or inexperience of health care personnel with dose adjustments may also lead to bleeding in individuals on warfarin[53]. The risk of bleeding is higher in patients with co-existing

comorbidities such as cancer[61], hypertension, cerebrovascular disease, previous ischemic stroke, abnormal renal or liver function, serious heart disease[62], history of peptic ulcer[63], or in individuals of older age[60]. Bleeding risk is also increased with concomitant use of cytochrome P-450 inhibitors (e.g. selective serotonin reuptake inhibitor [SSRIs], clopidogrel, azole antifungals, cimetidine, amiodarone), loop diuretics and sodium valproate, antibiotics (through modification of gut microbiome), aspirin, anti-inflammatory drugs, herbal supplements (garlic and ginkgo Biloba)[64]. Conversely, food or herbal supplement rich in vitamin K vitamins (coenzyme Q_{10}, St. John's wort, ginseng) may decrease the effect of warfarin and increase the risk of thromboembolism[64].

As an alternative for warfarin, Health Canada approved four DOACs for the prevention of stroke in AF: dabigatran in 2010, rivaroxaban and apixaban in 2012, and edoxaban in 2016. Compared to warfarin, DOACs have a more predictable anticoagulation response and standard daily doses (apixaban 5 mg two times a day [BID], dabigatran 150 mg BID, rivaroxaban 20 mg once daily, and edoxaban 60 mg once daily)[65]. Reduced doses of DOACs are indicated for patients who meet prespecified criteria. In Canada, the CCS recommends using apixaban 2.5 mg BID if at least two of the following criteria are met: serum creatinine ≥133 μmol/L, age≥80 years, body weight ≤60 kg. Dabigatran 110 mg BID is recommended if age≥80 years or age≥75 years and other risk factors for bleeding are present. Rivaroxaban 15 mg daily should be taken when creatinine clearance is 15-49 mL/min, and edoxaban 30 mg daily is recommended if creatinine clearance is 30-49 mL/min, body weight ≤60 kg, or if concomitant use of P-glycoprotein inhibitors, except amiodarone or verapamil[65]. In Quebec, similar fixed criteria for the reduced doses of DOACs for use in clinical practice are recommended by the *Institut national d'excellence en santé et en services sociaux*

(INESSS)[66]. Overall, DOACs have lower occasions of interaction with food and other drugs, and no dietary restrictions are needed. However, DOACs are susceptible to drug-drug interactions resulting in changes in their serum concentration that may alter their efficacy and safety. Thus, P-glycoprotein or cytochrome P450 3A4/5 inhibitors (e.g., ketoconazole, itraconazole, ritonavir, fluconazole, boceprevir, amiodarone, diltiazem, quinidine) are most likely to increase the serum concentration of DOACs and increase the risk of bleeding. Strong inducers of cytochrome P450 3A4/5 (e.g., carbamazepine, phenobarbital, phenytoin, rifampin, apalutamide, enzalutamide) may decrease the effect of DOACs and increase the risk of thromboembolism[67].

In patients on DOACs, there is no need for regular blood monitoring. However, DOACs are contraindicated in end-stage or severe chronic renal failure. Furthermore, treatment with DOACs is more expensive compared to warfarin[68]. Thus, the Canadian Agency for Drugs and Technologies in Health (CADTH) reported that for the fiscal year 2011/2012, the annual cost of warfarin 5 mg daily was 54$, whereas this cost was 1,289.44$ for dabigatran 150 mg BID, 1,147.53$ for rivaroxaban 20 mg daily, and 1,289.44$ for apixaban 5 mg BID[69]. Noviyani et al.[70] conducted a systematic review and meta-analysis of economic evaluation studies on DOACs and warfarin by pooling incremental net benefits. The study found that in high-income countries, despite the higher drug acquisition costs, the overall cost of DOACs for the healthcare system may be lower compared to warfarin. This may be explained by fewer hospitalizations related to bleeding complications of warfarin treatment[71,72].

At the time of the study, no specific antidotes were available for DOACs. To date, idarucizumab (reverses the direct thrombin inhibitor dabigatran)[73], andexanet alfa (reverses effects of factor Xa

inhibitors)[74], and ciraparantag (reverses effects of dabigatran and factor Xa inhibitors)[75] have been developed as reversal agents for the DOACs. Idarucizumab is the only approved by Health Canada[73].

The benefits and risks of each DOAC versus warfarin have been assessed by randomized clinical trials and observational studies.

1.6 Randomized controlled trials on efficacy and safety of oral anticoagulants

1.6.1 Pivotal trials and their findings

Four large phase-III randomized clinical trials assessed the efficacy and safety of each DOAC compared to warfarin for stroke prevention in non-valvular AF (Table 1.1). The trials' objective was a test for non-inferiority of DOACs against warfarin, and the secondary objective was a test for superiority.

Table 1.1. Selected characteristics of pivotal randomized controlled trials on efficacy and safety of DOACs vs. warfarin

RCT	Blinding	Number of participants	Follow-up duration (median, years)
RE-LY[76] Dabigatran 150 mg or 110 mg BID	Dabigatran- blinded to the dose received Warfarin- unblinded fashion	18,113	2
ROCKET-AF[77] Rivaroxaban 20 mg or 15 mg QD.	Double-blinded, double-dummy	14,264	1.9
ARISTOTLE[78] Apixaban 5 mg or 2.5 mg BID	Double-blinded, double-dummy	18,201	1.8
ENGAGE AF-TIMI trial"[79] Edoxaban 60 mg or 30 mg QD	Double-blinded, double-dummy	21,105	2.8

Only RE-LY[76] and ENGAGE AF-TIMI[79] trials assessed comparative efficacy and safety in participants randomly assigned to higher and lower doses of dabigatran or edoxaban versus

warfarin. ROCKET-AF[77] and ARISTOTLE[78] trials did not provide comparative estimates for lower doses of rivaroxaban and apixaban versus warfarin since the number of participants receiving lower doses was less than 5%[80].

Dabigatran 150 mg was superior to warfarin, and dabigatran 110 mg was non-inferior to warfarin for the prevention of stroke or systemic embolism (hazard ratio, RR 0.66; 95% CI; 0.53-0.82 and RR 0.91; 95% CI: 0.74-1.11, respectively). The hazards of major bleeding were similar between patients in the dabigatran 150 mg arm and the warfarin arm (RR 0.93; 95% CI; 0.81-1.07). Both, 150 mg and 110 mg dabigatran were associated with a lower risk of hemorrhagic stroke compared to warfarin (RR 0.26; 95% CI: 0.14-0.49, and RR 0.31; 95%CI: 0.17-0.56, respectively). However, higher hazards of gastrointestinal bleeding were observed in the 150 mg dabigatran arm compared to the warfarin arm (RR 1.50; 95% CI: 1.19-1.89).

Rivaroxaban was non-inferior to warfarin in the prevention of stroke or systemic embolism (HR: 0.79; 95% CI; 0.66-0.96). There were similar hazards of major and clinically relevant nonmajor bleeding in the rivaroxaban vs the warfarin group (HR 1.03; 95%CI:0.96-1.11). The hazards of intracranial hemorrhage were lower in the rivaroxaban vs warfarin group (HR 0.67; 95%CI:0.47-0.93). The rate of major gastrointestinal bleeding was higher in the rivaroxaban vs warfarin group (3.2% and 2.2%, respectively, $p<0.001$, HR was not reported).

Apixaban was superior to warfarin in the prevention of stroke or systemic embolism (HR 0.79; 95% CI; 0.66-0.95, $P<0.001$ for noninferiority; $P = 0.01$ for superiority). The hazards of major bleeding were also lower in the apixaban vs the warfarin group (HR 0.69; 95%CI:0.60-0.80, $P<0.001$), including intracranial bleeding (HR 0.42, 95%CI:0.30–0.58).

Edoxaban 60 mg or 30 mg was non-inferior to warfarin for prevention of stroke and systemic embolism (HR 0.79; 95% CI; 0.63-0.99, and HR 1.07; 95% CI; 0.87-1.31, respectively). The hazards of major bleeding were lower in both, the 60 mg or 30 mg edoxaban group vs the warfarin group (HR 0.80; 95%CI:0.71-0.91 and HR 0.47; 95%CI:0.41-0.55). Both, 60 mg and 30 mg of edoxaban were associated with lower hazards of intracranial bleeding (HR 0.47, 95%CI:0.34–0.63, and HR: 0.30, 95%CI:0.21–0.43, respectively). Participants from the edoxaban 60 mg group had higher hazards of GI bleeding as compared to the warfarin group (HR 1.23; 95% CI; 1.02-1.50).

A recent meta-analysis of patient-level data from the four pivotal RCTs[80] reported that compared to warfarin, standard-dose DOACs when analyzed collectively, were more effective in the prevention of stroke and systemic embolism when used in standard doses (HR 0.81; 95% CI; 0.74-0.89), and had lower hazards of intracranial bleedings (HR 0.45; 95% CI; 0.37-0.56), however, had higher hazards of gastrointestinal bleeding (HR 1.31; 95% CI; 1.08-1.57). Compared to warfarin, lower-dose DOACs (i.e., dabigatran 110 mg BID or edoxaban 15 mg QD), were not statistically different in the prevention of stroke and systemic embolism (HR 1.06; 95% CI; 0.95-1.19), had lower hazards of intracranial bleedings (HR 0.28; 95% CI; 0.21-0.37), and statistically not different hazards of gastrointestinal bleeding (HR 0.85; 95% CI; 0.62-1.18).

Overall, the pivotal RCTs and the meta-analysis of patient-level data from the four pivotal RCTs showed that DOACs are no-inferior or superior to warfarin in the prevention of thromboembolism, and are at least as safe as warfarin. However, RCTs have certain limitations that may compromise their external validity. These limitations are discussed in the next section.

1.6.2 Limitations of randomized clinical trials

RTCs are the gold standard in clinical research, they provide the highest level of evidence to guide clinical decisions or adopt new therapies[81]. However, the external validity or generalizability of RTCs may be limited by strict inclusion/exclusion criteria for study participation, omission or underrepresentation of important groups of patients, highly controlled study settings, limited sample size, or limited study follow-up[82]. External validity is defined as the applicability of the findings from RCTs "to a definable group of patients in a particular clinical setting in routine practice"[83].

The results on the efficacy and safety of DOACs were extrapolated to some vulnerable subgroups of patients in real clinical practice. However, the real benefit-risk balance of DOACs in these subgroups is not known. Thus, a study found that due to the strict inclusion criteria of the ARISTOTLE trial, up to 58% of patients in need of anticoagulation in a Swedish clinical practice would not be included in this trial[84]. The reasons for non-eligibility were not meeting ECG criteria (54%); coexisting psychosocial problems, including not uncommon alcohol abuse or dementia (30%); having both these criteria (8%); and other reasons (8%). On the other hand, the indications and contraindications for DOACs in the Summary of Product Characteristics (SmPC), a document that is a part of marketing authorization, are less restrictive. A study found that among AF patients with suspected stroke, a significantly higher proportion was eligible for DOAC therapy based on SmPC than based on the inclusion/exclusion criteria of the RE-LY, ROCKET-AF, and ARISTOTLE trials[85]. These proportions were respectively 72.9 versus 47.6 % for dabigatran ($p <$

0.001), 75.6 versus 39.3 % for rivaroxaban ($p < 0.001$), and 62.0 versus 45.5 % for apixaban ($p < 0.001$).

Some important subgroups of AF patients requiring anticoagulation and excluded from the pivotal trials were those with dementia, a history of intracranial bleeding, patients with anemia, or chronic kidney disease with severe renal impairment (creatinine clearance <30 ml/min)[85,86]. The ARISTOTLE trial allowed participation for those with creatinine clearance <25 ml/min. Yet, in this trial, individuals with severe renal impairment were underrepresented and comprised only 1.5% of the total study population. Underrepresentation is another issue that is common in RCTs[87]. A low number of trial participants with a certain prognostic characteristic makes it difficult to provide a subgroup analysis for any possible differences in the benefit-risk balance of DOACs. For example, in clinical practice, individuals of very advanced ages (80 years or over) may represent about 30% of those receiving oral anticoagulants in AF[88]. Yet, these individuals at high risk of stroke and bleeding, including gastrointestinal hemorrhage[32,89,90]were not separately stratified in all pivotal DOACs trials, instead, they were part of the age group of ≥75 years. Other groups of patients that may have a higher risk of bleeding, however, underrepresented in the pivotal RCTs are patients with multimorbidity (defined as a presence of multiple diseases that are associated with frailty, polypharmacy, and adverse events), obese individuals, and those with co-administration of cardiac P-glycoprotein inhibitors (amiodarone, carvedilol and verapamil).

The next point of concern is that, in general, in the highly controlled settings of RCTs, the adherence to the treatment is higher than in real-world settings. For example, a secondary analysis of data from the ARISTOTLE trial found that six months after trial initiation, 6.5% of patients had

low adherence to the therapy (<80% by pill count). In 8.7% of patients, the adherence was moderate (80%-90% by pill count), and in 84.8% it was high (>90% by pill count)[91]. Conversely, a study in the US that used real-world medical and pharmacy healthcare claims reported lower adherence rates based on the proportion of days covered (PDC). Thus, at six months, the proportion of patients with high adherence (PDC>80%) was 63% for apixaban, 59% for rivaroxaban, and 53% for dabigatran[92]. Such differences in adherence between RCTs and real-world clinical data are expected to affect the effectiveness and safety profiles of DOACs.

Four pivotal RCTs comparing efficacy and safety of DOACs and warfarin were global multicernter trials involving thousands of patients (Table 1.1.) with the median follow-up ranging from 1.8 years in the ARISTOTLE trial to 2.8 years in the ENGAGE AF-TIMI. However, they can be underpowered to provide measures of events that are rare (between 1 in 1,000 and 1 in 10,000 people may be affected) and very rare (fewer than 1 in 10,000 may be affected)[93-95].

Due to the limitations of RCTs exposed here, observational studies are an important tool to supplement the existing knowledge on the effectiveness and safety of DOACs compared to warfarin.

1.7 Observational studies on the effectiveness and safety of oral anticoagulants

1.7.1 Results from observational studies

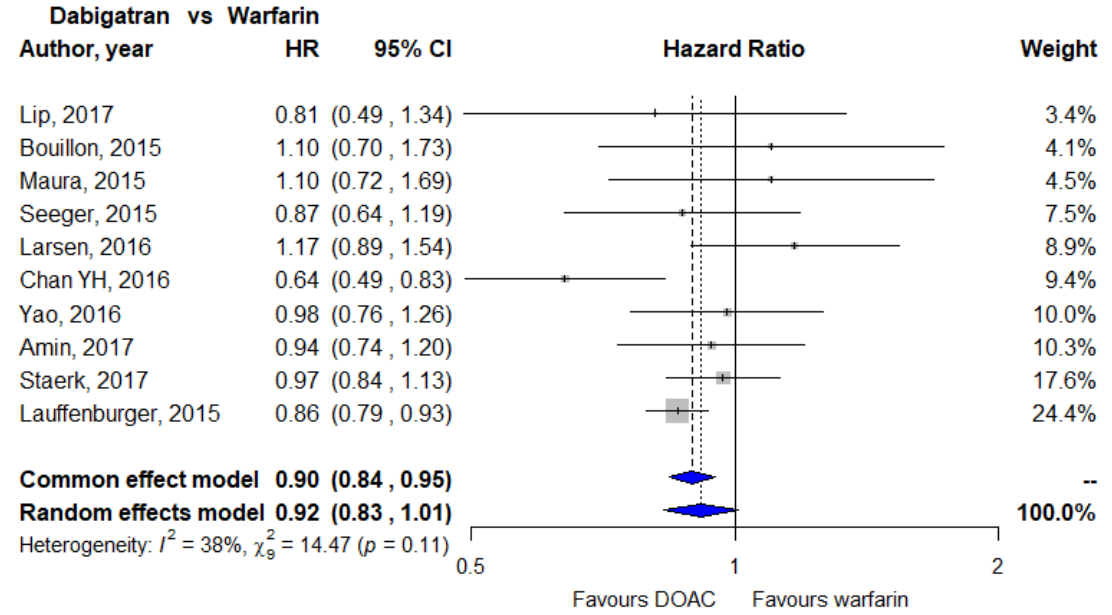
Observational studies in pharmacoepidemiology are usually based on the analysis of data from electronic healthcare databases containing information on large groups of individuals (population-based observational studies). These include administrative healthcare databases[96] and electronic health records[97]. Post-marketing observational studies are essential to supplement and bridge the findings from the RCTs to the daily clinical practice[98]. Unlike RCTs, observational studies have higher generalizability and may address the effectiveness and safety of medications in populations of patients that were excluded or underrepresented in the clinical trials[94]. For example, several studies assessed the effectiveness and safety of DOACs versus warfarin in patients with chronic kidney disease[39], the elderly[99], or frail patients[100].

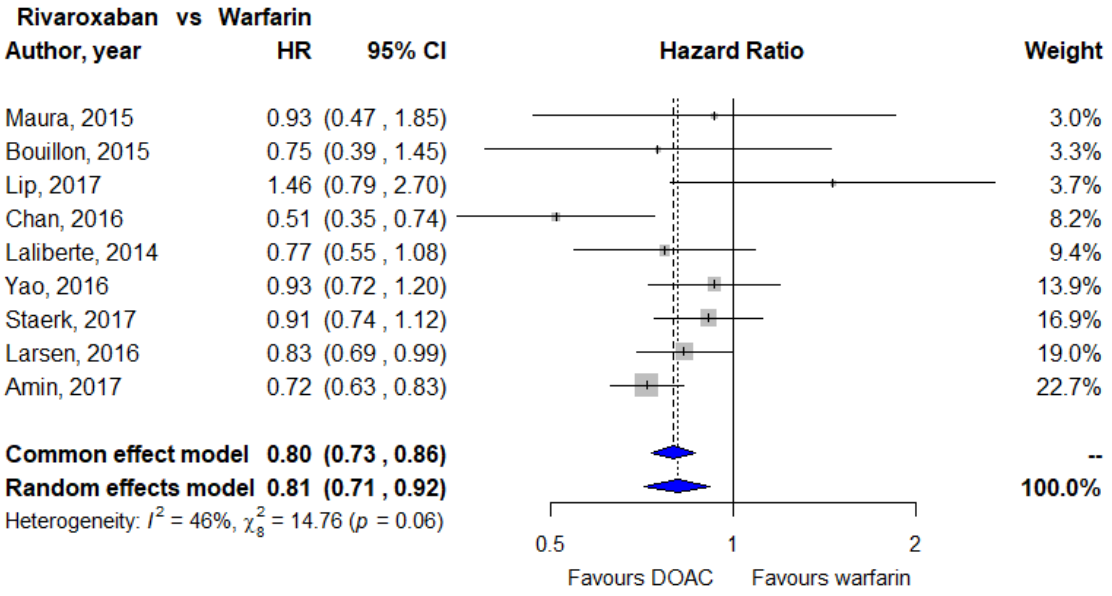
Four meta-analyses[101-104] and one systematic review[105] of observational studies comparing DOACs to warfarin were published during 2016 – 2019. The findings from these meta-analyses were similar to those reported by RCTs. Thus, all DOACs were superior or similar to warfarin for the prevention of stroke or systemic embolism. For example, Hirschl et al.[103] reported that rivaroxaban was superior and dabigatran and apixaban were similar compared to warfarin (pooled HR 0.80; 95%CI: 0.69-0.93, HR 0.92; 95%CI: 0.76-1.11, and HR 0.88; 95%CI: 0.64-1.21 respectively). Escobar et al.[101] found that all DOACs were superior to warfarin for the prevention of stroke or systemic embolism (dabigatran, pooled HR 0.87; 95%CI: 0.81-0.93, rivaroxaban, pooled HR 0.78; 95%CI: 0.71-0.85, and apixaban pooled HR 0.79; 95%CI: 0.67-0.93). Compared to warfarin, the risk of major bleeding was lower for dabigatran and apixaban and similar for rivaroxaban versus

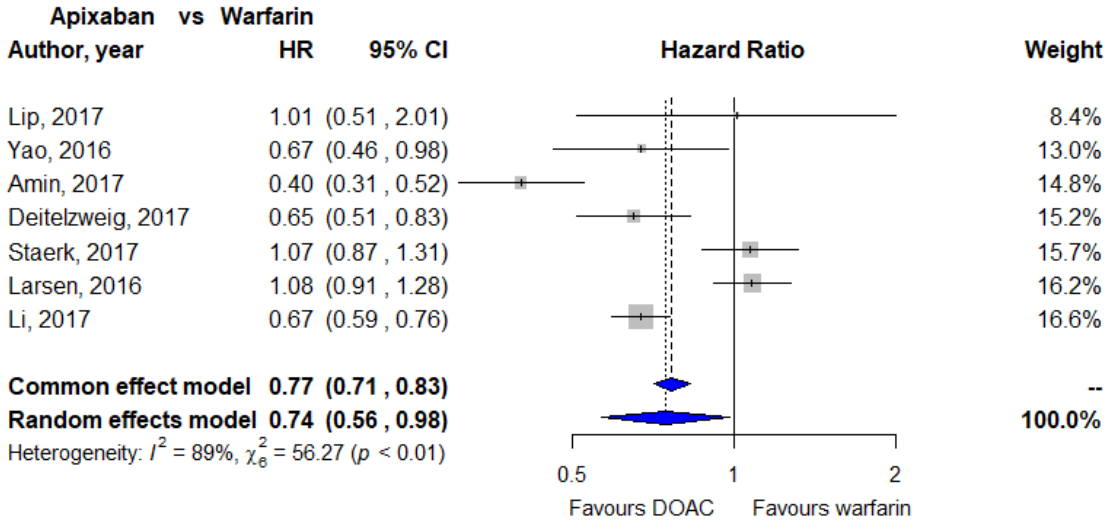
warfarin (pooled HR 0.80; 95%CI: 0.73-0.88, pooled HR 0.60; 95%CI: 0.56-0.66, and HR 1.02; 95%CI: 0.93-1.12, respectively)[104]. Compared to warfarin, dabigatran and rivaroxaban were associated with a higher risk of gastrointestinal bleeding (pooled HR 1.11; 95%CI: 1.004-1.23 and HR 1.25; 95%CI: 1.09-1.42, respectively), whereas apixaban had a lower risk of gastrointestinal bleeding (pooled HR 0.64; 95%CI: 0.55-0.75). All-cause mortality was lower in dabigatran and similar in rivaroxaban/apixaban (pooled HR 0.78; 95%CI: 0.65-0.97 and HR 0.99; 95%CI: 0.78-1.25, respectively)[104].

In the scope of this thesis, for illustration of results from observational studies, I extracted some findings from studies included in the abovementioned meta-analyses[101-104] and the systematic review[105] and created forest plots displaying HRs for stroke and systemic embolism, major bleeding, and gastrointestinal bleeding (Figure 1.6, Figure 1.7, Figure 1.8). The methods that I used for this analysis are provided in Annex 1. Test for heterogeneity was done with the Cochran Q test (X^2 test)[106]. Additionally, the I^2 test measured the percentage of the total variation in the effect estimates that was not due to chance[107]. The reference list of individual studies used to create the forest plots is in Annex 1.

Figure 1.6. Forest plots displaying hazard ratios of stroke and systemic embolism reported in observational studies comparing DOACs and warfarin [a,b].



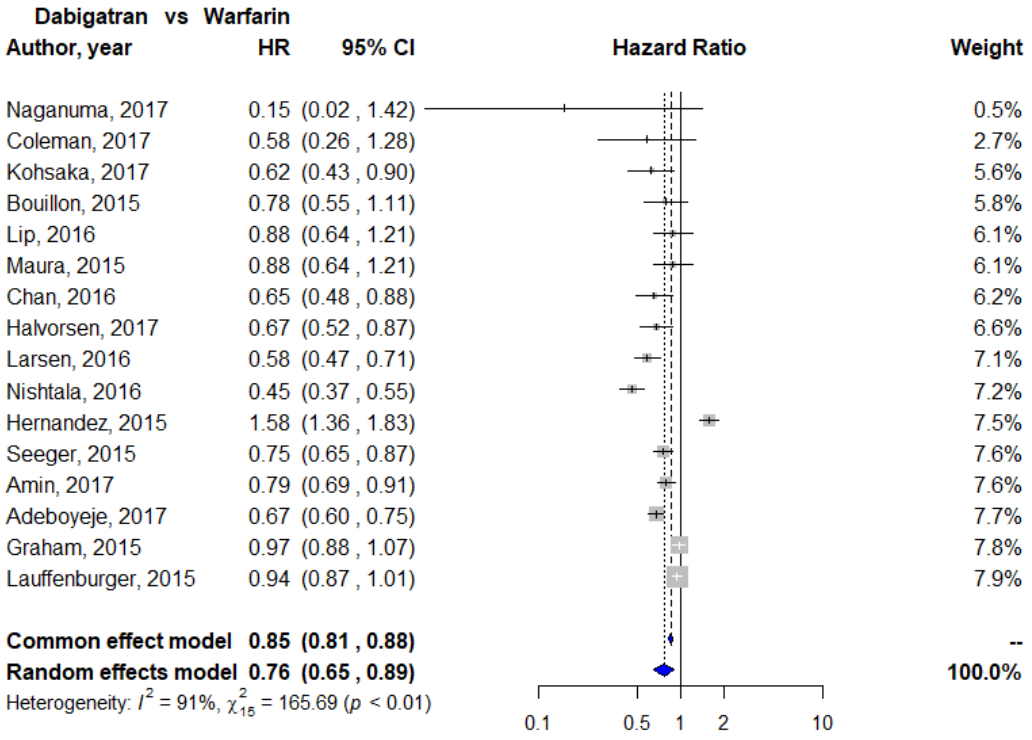


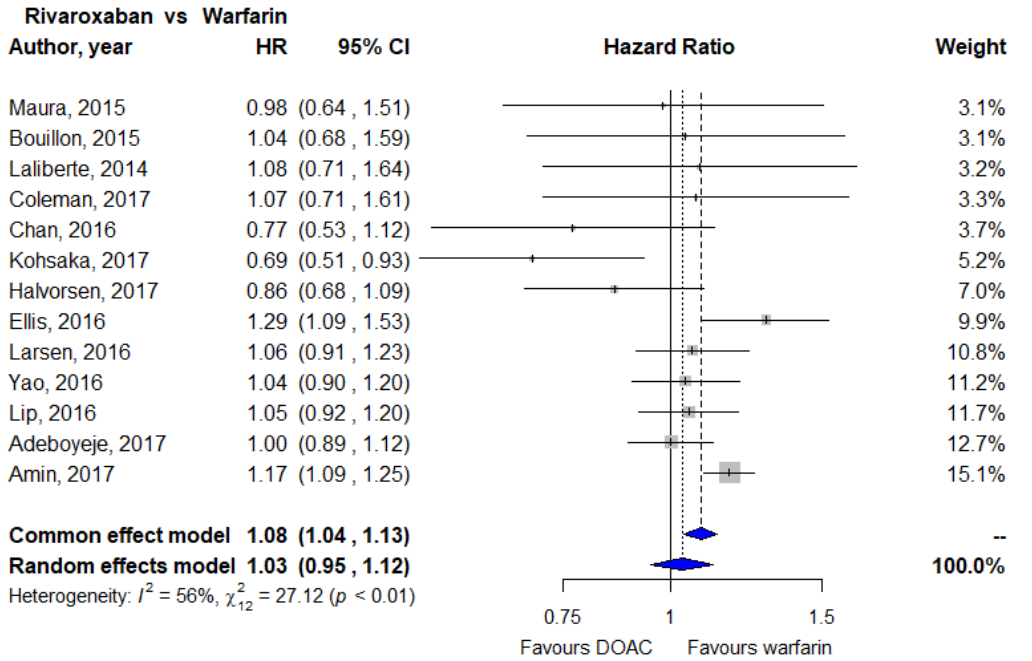


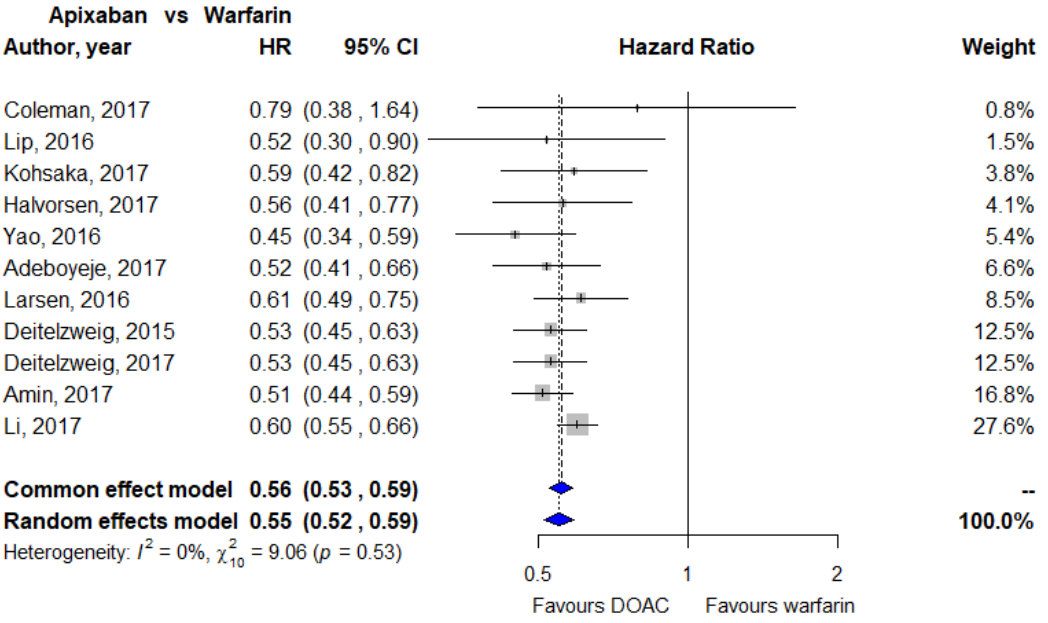
[a] Test for heterogeneity was done with the Cochran Q test (X^2 test)[106]. Additionally, the I^2 test measured the percentage of the total variation in the effect estimates that was not due to chance[107].

[b] The reference list of individual studies used to create the forest plots is in Annex 1.

Figure 1.7. Forest plots displaying hazard ratios of major bleeding reported in observational studies comparing DOACs and warfarin [a,b].



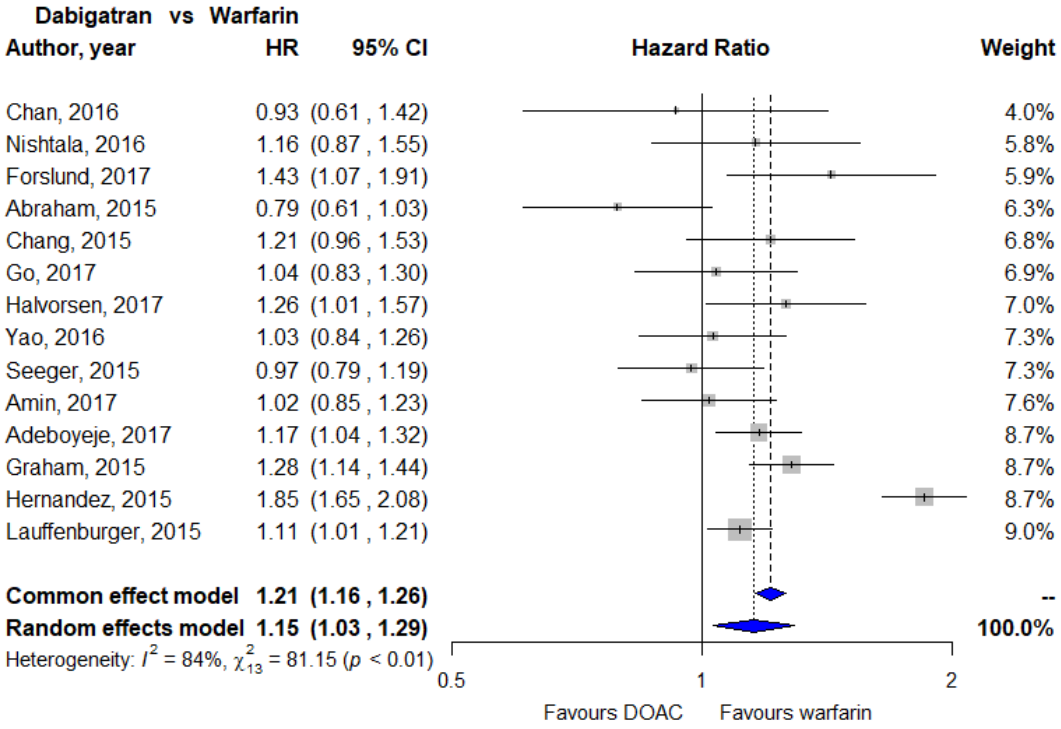


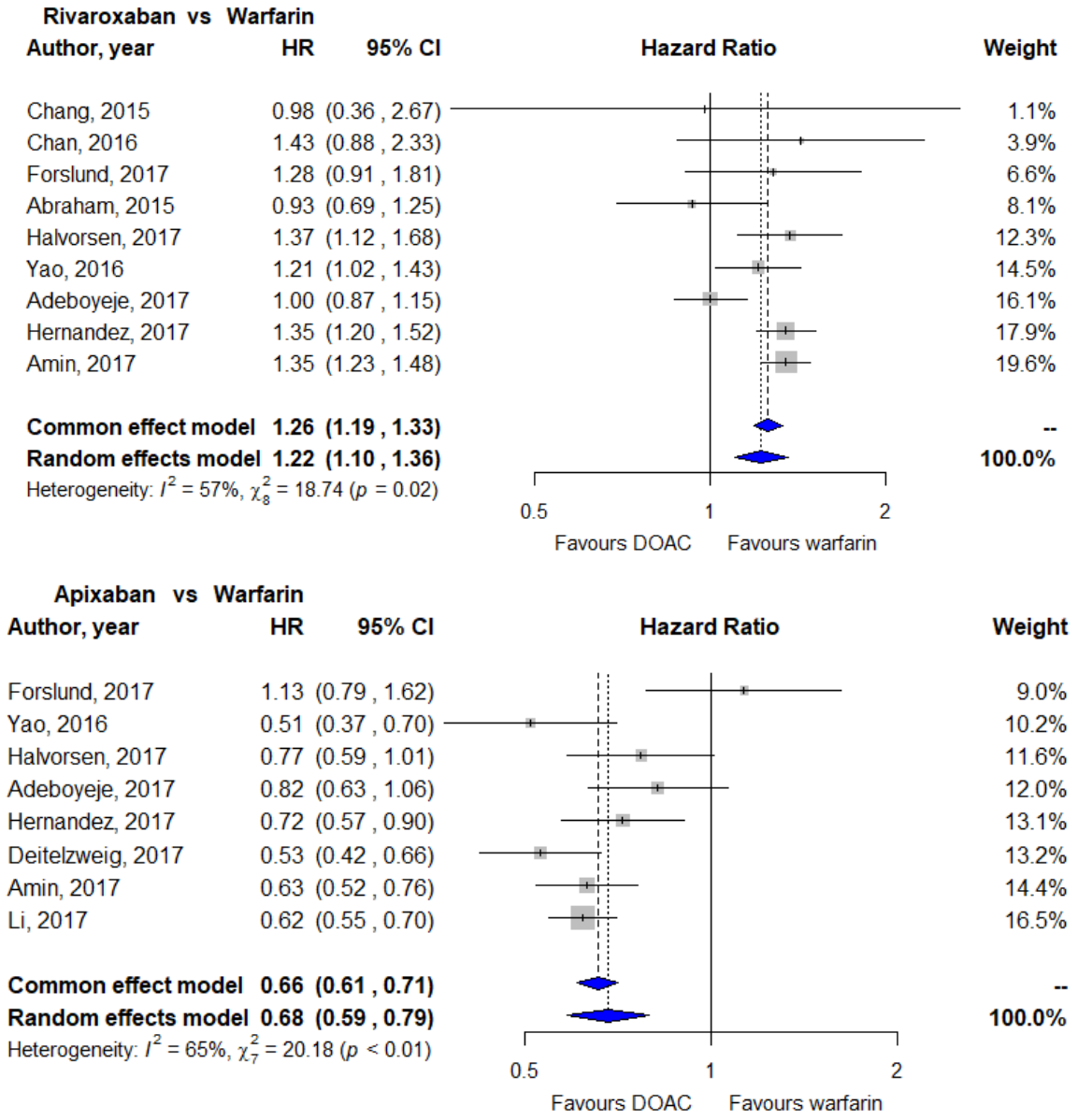


[a] Test for heterogeneity was done with the Cochran Q test (X^2 test)[106]. Additionally, the I^2 test measured the percentage of the total variation in the effect estimates that was not due to chance[107].

[b] The reference list of individual studies used to create the forest plots is in Annex 1.

Figure 1.8. Forest plots displaying hazard ratios of gastrointestinal bleeding reported in observational studies comparing DOACs and warfarin [a,b].





[a] Test for heterogeneity was done with the Cochran Q test (X^2 test)[106]. Additionally, the I^2 test measured the percentage of the total variation in the effect estimates that was not due to chance[107].

[b] The reference list of individual studies used to create the forest plots is in Annex 1.

Overall, across the observational studies, there was a statistically significant heterogeneity (X^2 p<0.1 or I^2>50%) between the reported HR estimates for the comparative risk of thromboembolism (Figure 1.6), major bleeding (Figure 1.7) or gastrointestinal bleeding (Figure 1.8). For example, for major bleeding (Figure 1.7), some studies reported benefits for DOACs, whereas others reported benefits for warfarin. Heterogeneity between the reported estimates can be explained by the differences in the methodological approaches, studies' population[108], or differences in the proportion of patients with advanced age and higher comorbidity burden who may use lower doses of oral anticoagulant[80]. However, the heterogeneity may partially be due to biases in some studies.

Observational studies may evaluate the effectiveness and safety in broad groups of the population, and they have the statistical power to detect the association of exposure even with a rare or very rare health condition. However, observational studies are prone to biases that limit their validity. Two of these biases are the particular topic of this thesis.

1.7.2 Limitations of observational studies

Observational studies are susceptible to bias, including confounding, information bias, and selection bias. Bias is a systematic error and a threat to the internal validity of the study's estimates. Internal validity refers to the extent to which the study findings reflect the true effect of the drugs under investigation on the outcome, and are not the result of a systematic error[109]. Selection bias can affect the external validity as well when the results from a study may not be generalizable to the whole population[110]. This chapter gives a brief description of the most prevalent types of bias in pharmacoepidemiology and a more detailed description of two types of bias that are the focus of this thesis.

Confounding arises when "all or part of the apparent association between the exposure and the outcome is in fact accounted for by other variables that affect the outcome and are not themselves affected by exposure"[110]. One of the advantages of randomized controlled trials is the random treatment allocation that makes the distribution of measured and unmeasured prognostic characteristics balanced between individuals in the exposure groups. In observational studies, treatment selection is not random, and, at baseline, the exposure groups differ in demographic and behavioral factors, comorbid clinical conditions, concomitant treatments, etc. *Confounding by indication* is most common in observational studies on comparative effectiveness with non-treated individuals as a reference group. In such studies, the indication (e.g., a disease) for the drug of interest determines both the exposure to the drug and the intended outcome. However, studies with an active comparator used for the same indication are lucking the confounding by indication[111]. C*hanneling bias* may happen in observational studies comparing drugs with the

same therapeutic indications when a newer drug and an older comparator are prescribed to patients with different prognostic factors[112]. Sicker patients may be more prescribed the newer drug due to its expected advantages or intolerance/poor response to the older drug. Conversely, healthier patients may be more prescribed the new drug due to safety concerns. For example, after approval to use in patients with AF, dabigatran and rivaroxaban were prescribed to individuals with a lower risk of stroke or bleeding[113-116] compared to patients prescribed warfarin, likely due to safety concerns. In this case, even when controlling for the baseline characteristics, the residual confounding would make warfarin appear a poorer drug than dabigatran or rivaroxaban. Generally, because of such inequalities in patients' characteristics, the comparative effectiveness research done soon after the market entry may be biased. However, over time, the difference in prognostic characteristics of patients prescribes older and newer drugs become more balanced[117].

Information bias occurs as a result of inaccuracy when measuring exposure, covariate, or outcome variables. Populations-based observational studies rely on information collected for purposes other than research, and information bias may be introduced at multiple points in the course of a study. For example, exposure to medication is defined based on the dispensing information contained in the pharmacy claims database. However, having the medication dispensed does not necessarily indicate exposure to this medication, as the drug that is dispensed may not be ingested by the patients[118]. Moreover, dispensation claims do not contain information on medications received during hospitalizations, free medication samples given by physicians, over-the-counter (OTC) medications, or drugs priced less than the co-payment that is required by

some insurance plan[96]. Furthermore, information bias may arise when using diagnostic or procedural codes for the ascertainment of study outcomes or covariates. In administrative data, these codes are assigned to all patient encounters in the healthcare system. However, diagnostic or procedural codes may inaccurately present the true disease status, for example, due to the inaccuracy of coding, or lack of granularity and clinical details when recording diagnoses[118].

Selection bias comes from an error in the procedures used to select individuals in the study or the analysis. In pharmacoepidemiological research, a *protopathic bias* may arise when selecting patients for whom treatment was prescribed for early manifestations of the undiagnosed health condition under the study[119]. The *depletion of susceptible bias* (prevalent user bias) may affect studies that select both prevalent and incident users of the study drug[120]. This bias occurs when the risk of the study outcome is not stable over time, thus, the risk is higher right after initiation of the drug and, then it decreases with the longer duration of exposure.

The time-related bias occurs when the follow-up or exposure time is incorrectly accounted for in the study design or analysis[121]. Time-related bias may fall into the category of selection, information bias, or confounding. *Immortal-time bias* may arise when the period during which the study outcome cannot occur is either misclassified or excluded from the analysis[121,122]. *Immeasurable time bias* arises when the information on exposure cannot be assessed based on data availability[121]. For example, in RAMQ prescription claims, information on exposure during hospitalization is missing, therefore, the exposure status cannot be ascertained. *Time-lag bias* occurs when patients at different disease stages are compared. This is a form of confounding by

disease severity[123]. *Time-window bias* may arise in case-control studies when cases and control have differential follow-up times during which the outcome of interest occurs[124].

This thesis focuses on two types of bias that are important in studies on the comparative effectiveness and safety of oral anticoagulants in individuals with AF. They are the ***information bias due to misclassification of exposure*** and ***selection bias due to informative censoring***. These biases and analytical strategies to mitigate them are described in the following sections.

1.7.3 Exposure misclassification

Definition

Misclassification of exposure to a drug is an error in the measurement of exposure to that drug that does not reflect its real use[125]. Exposure misclassification is differential if its probability depends on the outcome of interest[126]. Differential misclassification may result in a biased study estimate either away or toward the null. Non-differential misclassification usually leads to biased study estimates toward the null[118,126].

In general, automated pharmaceutical dispensation claims provide more accurate information on exposure to medications as compared to self-reported data or prescription information from medical records[127-129]. Data in the automated dispensation claims are not subject to recall, participation, or reporting biases[130]. The drawbacks of the prescription dispensation claims include not recording the information on medications dispensed during hospitalizations or stays in long-term care centers (*Centres d'hébergement de soins de longue durée*, CHSLD in Québec), free medication samples given by physicians, OTC medications, or drugs priced less than the co-payment that is required by some insurance plans[96]. These lead to missing information on exposure and misclassification of exposed individuals as unexposed. On the contrary, some unexposed individuals may be classified as exposed if the prescription was filled but never taken, or in case of low adherence to the treatment[118]. When defining the timing or duration of the exposure to a drug from dispensing events, the following information may be available in the claims data: dispensing dates, prescribed dosage, unit strength, number of dispensed pills, and duration of

supply (days supplied)[131]. The days supplied variable (either recorded or estimated) may be extended by a specified number of days (grace-period)[132].

A grace period is a prespecified time interval used to extend the *duration of dispensation* variable to compensate for small irregularities in refills, changes in dosage, stockpiling, non-compliance, or to account for residual effects of some drugs[132]. For example, for patients dispensed warfarin with a 30-day supply, a grace period of 60 days may be considered appropriate to account for possible changes in the dosing or non-compliance (Figure 1.9). The days supplied variable and a grace period is used when constructing the periods of continuous drug exposure, or the time interval when a patient is assumed to take a medication before a switch to another medication or permanent treatment discontinuation. Because there is no convincing argument behind the choice of the grace period, it is recommended to investigate the robustness of the comparative estimates with several varying grace periods[133].

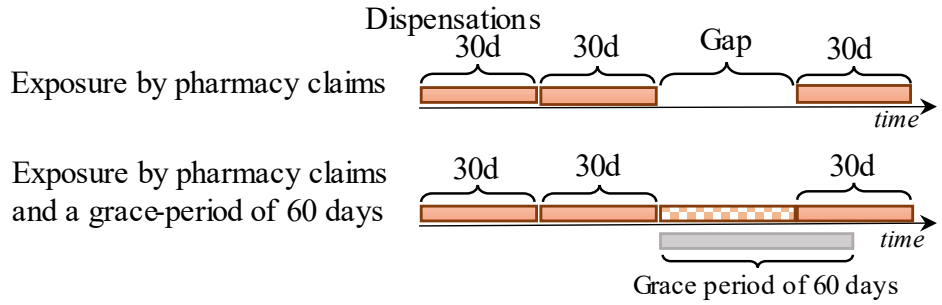


Figure 1.9. Defining period of continuous drug exposure using a variable denoting the duration of dispensation and a grace period.

Two types of exposure misclassification scenarios may arise when defining exposure periods (Figure 1.10).

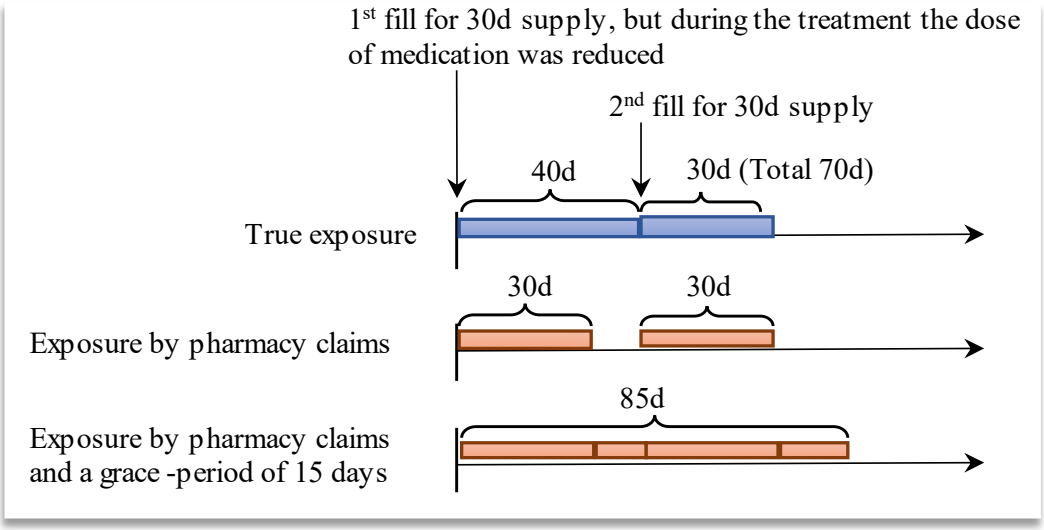


Figure 1.10. Exposure misclassification scenarios when using prescription claims data. Source: adapted from Schneeweiss and Avorn.

First, when a patient or his/her physician decides to lower the dose of medication, the real supply would last longer than the recorded days supplied variable leading to misclassification of some exposed time as unexposed, i.e. exposure definition is lacking sensitivity. Second, when a patient discontinues the drug before the supply ends, or when a lengthy grace period is used, some unexposed time may be misclassified as exposed, i.e. exposure definition is lacking specificity[97].

The risk of substantial misclassification of exposure exists when studying an acute effect of symptomatic treatments or therapies with time-varying patterns of use. In these instances, information recorded in pharmacy claims may not be reliable for defining the timing of exposure. For example, benzodiazepines, psychotropics, salicylates, analgetics[134], or nonsteroidal anti-

inflammatory drugs (NSAID)[135] are symptomatic treatments taken intermittently which makes it difficult to determine changes in exposure status between consecutive dispensations. Another example is inhaled corticosteroids and inhaled beta-agonists for the treatment of asthma and COPD. These are drugs with high between-individual dose variability. These medications are commercialized in canisters that have a fixed number of puffs, and the lifespan of the dispensed canister depends on the prescribed dosage and may vary substantially from patient to patient. Therefore, the information on the recorded days supplied in pharmacy claims data may not be reliable[136]. The next example is insulin, a drug whose doses may change day to day depending on diet, physical activity, or target blood glucose level. Thus, the dispensed quantity of insulin may last longer, or be wasted due to the expiration of the medications or during the priming of the insulin pen[137]. Therefore, the dispensing information for insulin is not likely to be compatible with its real exposure.

DOACs and warfarin are medications with short induction periods and different pharmacokinetics and pharmacodynamics[43] that impact the outpatient management of their users. The goal for the outpatient management of patients on warfarin is to maintain an optimal therapeutic INR range[138]. In clinical practice, warfarin dosage is frequently adjusted by physicians, and pharmacy dispensing data may inadequately reflect the flexible warfarin regimen. For example, the true duration of warfarin supply may last longer than recorded due to pill splitting. This inconsistency may be more substantial for patients with a large quantity of warfarin supply, such as 30 or 60 days. Moreover, when dispencing warfarin to patients, pharmacists tend to record a shorter duration of treatment than the planned one to have some

flexibility in a situation when the dosage would be increased and the drug supply would run out early. This is done to avoid issues with the prescription renewal, as the early claim will not be accepted by the pharmacy system (C. Sirois, personal communication, June 20, 2022).

On the contrary, DOACs have fixed doses, and the goal of outpatient management is to ensure strict adherence and persistence[138]. For these drugs, the information in pharmacy claims should be nearer to the true use.

Analytical strategies to address exposure misclassification

The assumptions on exposure definition may be explored with sensitivity analyses. The objective of these analyses is to reveal any substantial changes in the association between exposure and the outcome using alternative definitions of exposure[139]. An example is a study by Tang et al. on the risk of gastrointestinal bleeding among individuals treated with dabigatran versus those treated with warfarin[140]. For the main analysis, the exposure to anticoagulants was defined based on the days supplied value and a grace period of 14 days. In the sensitivity analysis, the grace period was changed to 7 days, however, the incidence rate ratios (IRR) estimates were robust with both grace periods. Thus, in the main analysis, the adjusted IRR was 0.99 (95%CI: 0.75-1.31). In the sensitivity analysis with a grace period of 7 days, the adjusted IRR was 0.97 (95%CI: 0.74-1.28)[140]. Conversely, the change of assumptions on exposure definition affected some estimates in the study of Go et al. that was based on the cohort of matched dabigatran and warfarin patients[141]. In the main analysis, using the days supplied and a 7-day grace period, the HR estimate indicated that compared to warfarin, the use of dabigatran was associated with a statistically significant higher risk of myocardial infarction (HR:1.88, 95%CI: 1.22-2.90). The HR estimates for myocardial

infarction from sensitivity analyses were not consistent with the main analysis. Thus, the change in the grace period from 7-day to 14-day resulted in a 45% decrease in the HR estimate from 1.88 (95%CI: 1.22-2.90) to 1.43 (CI:0.99-2.08), respectively. Next, the authors use the "expanded exposure algorithm", e.i. stockpiling was added to any refill that occurred within 7 days before the estimated end of the previous refill. The HR estimate with the "expanded exposure algorithm" decreased to 1.38 (95%CI:1.00-1.92)[141]. Thus, in the study of Go et al., altering the definition of exposure affected the magnitude of the effect and its statistical significance, however, did not change the direction of the HR estimate associated with the risk of myocardial infarction.

Some authors encourage using methods for quantifying the impact of information bias on a study estimates[118,142]. These methods may show the magnitude and direction of the bias[70]. However, to use these methods in observational studies on comparative effectiveness and safety of oral anticoagulants, one needs some parameters from previous validation studies (e.g. sensitivity and specificity of the days supplied variable for oral anticoagulants)[97], which are currently not available.

1.7.4 Selection bias due to censoring

Definition

In longitudinal studies, censoring is a termination of follow-up of a study participant for reasons other than the event of interest[143]. Overall, censoring results in missing information on survival under the exposure of interest.

Censoring was first described in the scope of the standard survival analysis, for example, by Cox and Oakes[144], Kalbfleisch and Prentice[145], and Klein and Moeschberger[146]. Thus, for the survival analysis to produce unbiased results, the assumptions of censoring at random[144,146], independent censoring[145], or non-informative censoring[146] must held. Kleinbaum and Klein[147] explain that although these assumptions are used interchangeably in the literature they have differences.

The assumption of censoring at random implies that an individual whose follow-up is censored at time t has the same prognosis for the event ("survival experience") as other individuals in the same risk set at time t[147].

The assumption of independent censoring is censoring at random within a subgroup of individuals with the same value of a covariate. For example, censoring is at random by sex groups, however, it may not be at random in the whole cohort[147].

The assumption of non-informative censoring indicates that the distribution of the survival times does not affect the distribution of censorship times. The example of informative censoring that Kleinbaum and Klein[147] give is a person leaving the study because his family member who is also in the study developed the study event, i.e. censoring is related to the survival distribution.

From the analytical perspective, the following types of censoring are described:

1. Administrative censoring[148] or end-of-study censoring[149] occurs at the prespecified end of the study date. It is usually at random and does not lead to bias in absolute or relative estimates (under the assumption that the time of cohort entry is not associated with the risk of the outcome event)[144].

2. The lost-to-follow-up[149] censoring occurs when a study participant dropped out of the study and there is no further information on his/her exposure and survival. In a study based on the analysis of administrative data, the lost-to-follow-up censoring may be applied in case of disenrollment from pharmaceutical coverage for any reason. In commercial insurance claims databases, disenrollment may potentially lead to selection bias, as it is a common event (up to 40%) and may vary depending on patients' characteristics (e.g. age, number of comorbidities, frailty, hospitalization, or emergency room visit, use of prescription medication), health plan characteristics (health maintenance organization [HMO] or other), and geographic characteristics (urban versus rural)[150]. In Quebec, disengagement from the RAMQ prescription drug insurance plan may lead to selection bias, for example, in case a previously unemployed individual aged < 65 years would join a new employer's private prescription drug plan. However, likely, the number of such individuals and hence the risk of selection bias is low.

3. Artificial censoring was described by Robins and Finkelstein[151] and Joffe[148]. This censoring is done when information on survival is available, however, for analytical purposes, the follow-up is censored in case of discontinuation of the treatment assigned at baseline. Thus, in RCTs and observational studies, the follow-up may be censored at the time when a patient switches from treatment assigned at baseline or permanently discontinues the baseline treatment. In the context

of RCTs, this analytical approach is known as "per-protocol"[152], in observational studies, this approach sometimes is referred to as "as-treated"[153,154], or as "per-protocol"[155]. This type of censoring may induce bias because reasons for non-compliance with the treatment assigned at baseline can be associated with time-evolving prognostic factors. For example, the AIDS Clinical Trial Group (ACTG) randomized trial 021 compared the effect of Bactrim versus aerosolized pentamidine on survival (secondary outcome) in people with AIDS and pneumocystis pneumonia[151]. Study participants from the aerosolized pentamidine group were more likely to stop the assigned therapy and switch to more potent therapy in case of worsening of their clinical condition. Conversely, in the Bactrim arm, study participants were more likely to stop because of side effects (such as allergic skin rash) which were not important prognostic factors for survival[151].

4. Competing-risk censoring[156] is censoring at the time of occurrence of some other event that precludes the occurrence of the primary study event. For example, in a study on the risk of gastrointestinal bleeding in individuals exposed to dabigatran, death is a competing event. In the presence of competing risk censoring, the estimated crude incidence of the outcome is biased upward[156]. The relative measure (hazard ratio) may be biased in either direction.

Analytical strategies to address selection bias due to censoring

Hernan et al. suggest using causal diagrams (directed acyclic graphs, DAG) to assess the possible bias from censoring, define variables that may influence censoring, and choose an appropriate analytical method to reduce or avoid a bias that may be introduced by censoring[157].

Overall, there are three analytical strategies to address selection bias due to censoring: intention-to-treat (ITT) analysis without any censoring to avoid the introduction of selection bias; per-protocol/as-treated analysis with the assumption that censoring is at random; and per-protocol/as-treated analysis with the assumption that censoring is not at random and requiring an adjustment for selection bias. Of note, the term "per-protocol analysis" is used with RCTs, whereas "as-treated analysis" is used in observational studies. However, in the "causal inference" literature, the term "per-protocol analysis" is used with both, RCTs and observational studies. Thereafter in this thesis, the term "per-protocol" analysis is used with observational studies as well.

The ITT analysis assesses the effect of the assigned treatment[158]. The exposure status of individuals is defined based on the drug that was initiated at cohort entry. Individuals are followed until the defined end of the study, regardless of the baseline treatment discontinuation during the follow-up (no censoring is applied). The benefit of the ITT analysis is that it preserves the balance of the prognostic characteristics that was achieved at baseline by randomization in RCTs, or by matching, covariate adjustment, or other methods in observational studies[159]. However, the estimated ITT effect depends on the magnitude of adherence to the assigned treatment in each study group. Two identical studies with the same effect of a new treatment, but with different patterns of adherence may report different ITT estimates[160]. Usually, the observational studies of the acute outcome event limit their follow-up time because adherence to the drug of interest is expected to decline over time.

The per-protocol analysis assesses the effect of treatment that would be observed if all individuals had adhered to the assigned treatment[155]. In the literature, this treatment effect is called *the effect of sustained treatment*[161], or *the effect of continuous treatment*[155]. Under the per-protocol analysis,

individuals are followed until any discontinuation of treatment that was initiated at baseline. At the time of discontinuation, the follow-up is censored. This artificial censoring creates a population that is continuously treated with the drug initiated at baseline. However, the validity of the estimates from the per-protocol analysis is based on expert knowledge and the correctness of the assumption on the mechanism of censoring[155]. Thus, the per-protocol analysis without adjustment for artificial censoring may estimate the causal effect of treatment only under the assumption that censoring was at random. However, in the presence of time-varying confounding that is a predictor of the study outcome and the subsequent treatment strategy, this analytical approach may not provide a valid estimate of the treatment effect[157]. When the time-varying confounding is the reason for censoring, this may introduce a selection bias. Robins et all.[151] proposed using the inverse probability of censoring weights (IPCWs) to correct for induced selection bias and estimate the causal effect of exposure. The underlying idea is to "compensate" for those censored by assigning extra weight to uncensored patients having a similar distribution of prognostic factors. Thus, weighting creates a pseudo population of patients that would have been observed if the censoring hadn't occurred[151].

Thus, for an individual (i) in the study, the probability (P) of not being censored ($C=0$) is estimated for each observation representing updated information on this individual at a given time (t) (e.g. months of follow-up or visits to the clinic). This updated information may include the baseline (X) and time-updated (L) characteristics. The non-stabilized weights are the inverse of the time-specific probability of not being censored up until this time.

$$\text{Non} - \text{stabilized censoring weight } i(t) = \prod_{t=1}^{t} \frac{1}{P[Ci(t) = 0 | Ci(t-1) = 0, Xi, Li(t-1)]} \quad (1)$$

The values of non-stabilized weights may be very large in individuals with a small probability of not being censored. Large weights increase the variance and may lead to a biased estimate of the treatment effect. Hernán et al[162] recommend stabilizing the weights to reduce the variability. Stabilization is done by substituting 1 in the numerator in the IPCWs. Thus, a new numerator is estimated as the probability of remaining in the study given the baseline characteristics that were also used when estimating non-stabilized weights[162].

The stabilized censoring weight for individual *i* at time *t* is thus given by:

$$\text{Stabilized censoring weight } i(t) = \prod_{t=1}^{t} \frac{P[Ci(t) = 0 | Ci(t-1) = 0, Xi]}{P[Ci(t) = 0 | Ci(t-1) = 0, Xi, Li(t-1)]} \quad (2)$$

The numerator estimates the probability of not being censored at time *t* given that this individual was not censored at the time (*t-1*), and baseline covariates *X*.

The denominator estimates the probability of not being censored at time *t* given that this individual was not censored at the time (*t-1*), baseline covariates *X*, and time-updated covariates *L* measured at the time (*t-1*).

Three assumptions must be met for the IPCWs to create a pseudo-population free of selection bias introduced by censoring[163]. First, the non-testable ignorability assumption (no unmeasured/residual confounders), i.e. for each time *t*, censoring occurs at random given the probability of remaining

uncensored up to time *t*, baseline covariates, and the time-updated covariates at the time (*t-1*)[164]. Second, the positivity assumption: at any time *t*, the study individuals with any combination of the observed baseline covariates and the time-updated covariates at the time (*t-1*) have a greater than zero probability of not being censored[165]. This assumption may be tested by examining the distributions of the IPCWs. Any extreme values of the IPCWs would signify the violation of the positivity assumption. The third assumption is the correct specification of the censoring model, which means the right choice of the model form, the predictors of censoring, and the functional form between the censoring and the predictors of censoring. This assumption may also be verified with the distributions of the estimated IPCWs. Thus, very extreme values of IPCWs may indicate misspecification of the censoring model (e.g., too many predictors or multicollinearity may result in a large standard error)[166]. Furthermore, the mean of stabilized weights should be close to 1.00 (proof in Hernán, M. A., & Robins, J. M., 2006[167]), the mean that is far from 1.00 indicates misspecification of the model for estimation of censoring weights.

The IPCWs were used in previous studies to adjust for censoring, e.g. smoking and cognitive decline[168]; postmenopausal hormone therapy and coronary heart disease[159]; statins and primary prevention of coronary heart disease[169]; antiretroviral treatment and AIDS-free survival[170]. Thus, the study by Weuve et al.[168] used the data from the observational Chicago Health and Aging Project to analyze the effect of smoking on cognitive decline in older individuals. However, smoking was also strongly associated with mortality and attrition in the study. This created a selection bias leading to underestimation of the effect of smoking on cognitive decline. Indeed, the analysis weighed with IPCWs showed that cognitive decline in smokers was up to 0.20 standard units per

decade faster than in never smokers (95% CI= −0.36 to −0.04). These estimates were 56%–86% larger compared to the estimates obtained using the unweighted analysis (0.11 standard units per decade, 95% CI= −0.20 to −0.02)[168].

CHAPTER 2: Rationale for research, research aims, objectives, and hypotheses

The two aims of this research were to investigate two important methodological challenges that may compromise the internal validity of pharmacoepidemiological studies on the comparative effectiveness and safety of oral anticoagulants in AF. First, an information bias resulting from misclassification of exposure to dose-varying warfarin therapy when using days supplied value recorded in pharmacy claims data. Second, a selection bias due to censoring with differential censoring mechanisms in the DOACs- and the warfarin-exposed individuals.

2.1 The rationale for the present research

2.1.1 Exposure misclassification using pharmacy data

Vitamin K antagonists, including warfarin, are a class of drugs with high variability of pharmacodynamics and pharmacokinetics between individuals and within an individual over time[171,172]. A recent meta-analysis of observational studies on adherence to oral anticoagulants in individuals with AF was not able to provide a pooled estimate on adherence to warfarin due to lacking measures on adherence to warfarin in the included studies[173]. The authors of this meta-analysis stated that, overall, observational studies did not report on adherence to warfarin due to some difficulties in the ascertainment of the true duration of warfarin supply because of variation of the dosage during the treatment course[173].

With administrative data, exposure to warfarin is commonly defined based on dispensing information recorded by community pharmacists. However, pharmacists dispensing the drug may

be unaware of any adjustments in the warfarin dosage by the treating physicians. On the other hand, knowing that the warfarin dose may be adjusted, pharmacists tend to record a shorter duration of treatment than the planned one to have some flexibility if the dosage is increased, and the drug supply runs out early. This is done to avoid issues with the prescription renewal, as the early claim will not be accepted by the pharmacy computer system (C. Sirois, personal communication, June 20, 2022). Therefore, the dispensed supply may last well beyond the duration recorded by the dispensing pharmacist. The resulting gap between observed and expected dates of warfarin refill may be explained by the discrepancy between the posology recorded in the pharmacy claim and that truly taken by the patient. These may lead to exposure misclassification which is likely to be a differential between warfarin and DOACs, as DOACs do not require dose adjustments or monitoring.

To our knowledge, no previous studies assessed the potential for misclassification of exposure to warfarin when using pharmacy claims. However, some recognized that the true duration of warfarin dispensation may last longer than the recorded duration, and considered this issue in the study design.

In the settings with available laboratory data, the intervening INR measurements were used to bridge the gaps between warfarin prescriptions. Go et al.[48] conducted a study on the effectiveness of warfarin in stroke prevention using the US commercial administrative claims databases (the Kaiser Permanente of Northern California). Individuals were defined as continuously exposed to warfarin if gaps between two consecutive prescriptions did not exceed 60 days (this choice may be explained by the fact that at the time of the study, in the Kaiser Permanente pharmacy claims, the

most common duration of dispensation for long-term treatments was 90 days, thereby, the authors suggested that the supply may last another 60 days because of dose changes and some nonadherence). In this study, individuals with longer gaps were defined as exposed in the presence of intervening INR measurements at least every 6 weeks (this choice may be explained by the existing recommendation to perform INR measurements every four weeks[174], and two additional weeks were allowed by the authors to likely to account for medical practice- or patient-induced delays or longer INR recall intervals in patients with long-term INR stability). The authors validated this algorithm in 1,207 patients with an incident thromboembolic or hemorrhagic event during the follow-up. At the time of the event, the exposure status defined with the algorithm was verified with the exposure status documented in the medical records, with the agreement of k=0.84. Other studies used this algorithm. An example is a study by Casciano et al.[175] on the economic burden of nonadherence to warfarin therapy among patients with NVAF using the US MarketScan Research Database[175]. Furthermore, Azoulay et al.[176,177] adapted the algorithm developed by Go et al.[48] in their studies using the General Practice Research Database (CPRD), a primary care database from the United Kingdom. In the setting of CPRD, the algorithm may not only account for changes in warfarin dosages, but also for those warfarin prescriptions that inherently are not captured in the CPRD (e.g. prescriptions given at hospital discharge or in an anticoagulation clinic)[176]. Next, Webster-Clark et al.[178,179] in their studies based on the analysis of the US Medicare data, used a definition of exposure time that was tailored to the specific classes of oral anticoagulants. Thus, the allowed gap between consecutive dispensations of dabigatran was 30 days, whereas, for warfarin, it was 45 days (this may be explained by the fact that for long-term treatments, Medicare commonly supplies a drug for 30 days). Additionally, the Medicare *current procedural*

terminology (CPT) codes for anticoagulation management were used to "refresh" the warfarin supply for another 30 days[178]. Information on the validation of this method is not available in the published papers.

The abovementioned algorithms may account for the dose-varying warfarin regimen and minimize misclassification of truly exposed time as unexposed (improved sensitivity). This would result in increasing the follow-up time of patients on warfarin and detection of some additional events of both bleeding and thromboembolism. The method used by Webster-Clark et al.[178,179] also allows to minimize misclassification of truly unexposed time as exposed (improving specificity) for patients on DOACs, as these medications have standard doses, and long gaps between DOACs dispensation are most likely to signify treatment discontinuation. Therefore, the follow-up time for DOACs patients would decrease. The number of thromboembolism events should also decrease, as these events would likely be due to lacking anticoagulation. The number of bleeding events should not change (no bleeding event when a patient was misclassified as exposed to a DOAC).

Missing information on INR testing is common in the laboratory datasets and this may limit the implementation of the algorithm described by Go et al.[48]. Moreover, laboratory files or procedural codes for anticoagulation management are not available in some population-based administrative datasets, including the RAMQ database.

2.1.2 Selection bias due to censoring

In clinical practice, therapy with oral anticoagulants is dynamic, involving switching and treatment discontinuation. A switch from warfarin to a DOAC is likely to happen for convenience purposes, or due to non-compliance with warfarin therapy and dietary restrictions driving poor TTR.

A switch from a DOAC to warfarin may happen because of evolving new health conditions when the use of DOACs is considered unsafe. These conditions are associated with an increased risk of stroke and bleeding and include new-onset renal failure, valvular disease or surgery, and coronary vascularization[180]. The per-protocol analysis estimates the effect of continuous treatment with a drug that was assigned at the time of cohort entry, and the follow-up of patients is censored at the time of discontinuation of this drug. Thus, censoring the follow-up of individuals who initiated a DOAC and then switched to warfarin removes from the study those patients who are at higher risk for stroke or bleeding. Indeed, previous population-based studies reported that patients who switched from a DOAC to warfarin during the treatment were older and had a higher risk of bleeding compared to those who did not switch[181,182]. At the same time, individuals who initiated warfarin and developed similar risk profiles would stay in the study. In such a situation, the measure of comparative effectiveness and safety would be biased in favor of DOACs. To correct selection bias, one needs to apply the IPCWs. However, a model of censoring probabilities that is adjusted for treatment value, time of follow-up, and all prognostic factors, would estimate the average coefficients for prognostic variables across treatment groups. Thus, renal failure is expected to be positively associated with censoring in DOAC users and negatively associated with censoring among warfarin users. The coefficient for the probability of censoring given renal failure would be

an average value of the treatment-specific censoring probabilities. As a result, applying the IPCWs would fail to eliminate the selection bias associated with renal failure. One potential solution for the correct specification of the censoring model in this clinical situation is to stratify by exposure.

Previously, Webster-Clark et al.[179] in their study comparing dabigatran and warfarin-treated individuals, recognized this differential mechanism of censoring across the treatment groups and adjusted for the resulting selection bias with the exposure-stratified censoring weights. However, to our knowledge, no previous literature described the conceptual framework for this bias and its effect on the study conclusion. Moreover, no studies demonstrated the effect of misspecification of the censoring model on the estimates of comparative effectiveness and safety of DOACs and warfarin in AF.

2.2 Hypotheses

1. Prescription claims data inadequately capture the duration of the dispensation of warfarin. Gaps between subsequent dispensations and their variation are larger for warfarin than for DOACs.

2. Definition of exposure duration based on the days supplied value and a fixed grace period will lead to differential misclassification of exposure to warfarin and DOACs. This may bias the estimate of comparative safety in favor of DOACs.

3. When assessing the effect of continuous treatment with oral anticoagulants (per-protocol effect), censoring removes the sicker individuals from the DOACs group and healthier individuals from the warfarin group. This biases the estimate of comparative effectiveness and safety in favor of DOACs.

2.3 Aims and objectives

Aim 1. To identify the potential for information bias resulting from misclassification of exposure to dose-varying warfarin therapy when using days supplied value recorded in pharmacy claims data.

Objectives:

1.1. To characterize the dispensation patterns of warfarin, apixaban, dabigatran, and rivaroxaban.

1.2. To compare the variation of refill gaps between users of different oral anticoagulants, as well as within users of the same oral anticoagulant during the treatment course.

1.3. To evaluate whether a data-driven method to define the exposure duration captures the variability in warfarin dispensations better than the days supplied method.

1.4. To investigate the impact of the method for definition of exposure duration on the hazard ratios of major bleedings associated with the use of dabigatran or rivaroxaban versus warfarin.

Aim 2. To describe the selection bias due to censoring with differential censoring mechanisms in the DOACs- and the warfarin-exposed individuals.

Objectives:

2.1. To describe the associations between the prognostic variables and censoring in the DOACs and warfarin groups.

2.2. To describe the direction of selection bias due to censoring in the DOACs and warfarin groups.

2.3. To compare two strategies for estimation of censoring weights: exposure-stratified and exposure-unstratified.

2.4. To estimate and compare the hazard ratios of the effect of continuous treatment with warfarin versus DOACs on the composite of stroke, major bleeding, myocardial infarction, and all-cause mortality using exposure-stratified and exposure-unstratified censoring weights.

CHAPTER 3: Research methods

This chapter provides an overview of the data source and the methods used for the design and analyses of three projects.

3.1 Source of data

This research used the computerized health care databases of the *Régie de l'assurance-maladie du Québec* (RAMQ). RAMQ administers the government health insurance plan for all Québec residents, and the prescription drug insurance plan for individuals of 65 years and older, welfare recipients, and residents not covered by private medical insurance and their dependents. In total, about 40% of Québec's population is covered by the RAMQ prescription drug insurance plan[183,184]. Those not eligible for the RAMQ prescription drug insurance plan are Québec residents under the age of 65 years who are covered by private plans (group insurance or employee benefit plans)[185]. As such, information on drugs dispensed to these individuals is not available in the RAMQ prescription claims data.

RAMQ routinely collects diagnostic and medical information for health statistics and reimbursement purposes from healthcare providers all over the province.

The RAMQ maintains four medical databases. The *registry of beneficiaries' database* contains socio-demographic information on all recipients of the health insurance plan. This information includes the month and year of birth, sex, the first three digits of the postal code, and the year and month of death (if applicable). The *medical services database* contains fee-for-service physician

claims for outpatient, inpatients, or emergency room services with the date, location, and type of the rendered service, the physician's medical specialty, medical diagnosis, type of medical procedures, and the date of procedures. Diagnoses are coded using the International Classification of Diseases (ICD) 9th revision system, and medical procedures are coded using the Canadian classification of diagnostic, therapeutic, and surgical procedures (CCP). The *prescription claims database* contains information on filled prescriptions from all healthcare providers in Québec. The recorded information includes the date of prescription fill, drug identification number (DIN), class of drug (coded by the American Society of Health-System Pharmacists [ASHP] Pharmacologic-Therapeutic Classification System), drug generic name, unit strength, dispensed quantity, intended duration of treatment in days, drug's gross cost and contribution of the insured person, the specialty of the prescribing physician. RAMQ proceeds the payment for the dispensed drug only if all these fields are completed and within range. A study found that in a random sample of prescription claims, the proportion of missing or out-of-range data varied from 0% to 0.7%[131]. The *database with eligibility periods for public drug insurance plans* includes the insurance plan details and the period of coverage.The Maintenance et exploitation des données pour l'étude de la clientèle hospitalière (MED-ÉCHO) is managed by the Québec Ministry of Health and Social Services (*Ministère de la Santé et des services sociaux*, MSSS). The database collects discharge summaries following any hospitalization, including intensive care. Information is available on the hospitalization dates, types of admission, primary diagnosis on admission, the principal diagnosis, the cause of in-hospital death, and up to 19 secondary diagnoses, as well as the type of medical procedures. Since 2006, the ICD 10th

revision system, and the Canadian Classification of Health Interventions (CCI) is used for coding diagnoses and medical procedures[186].

All databases may be linked using the unique patient identifier. The RAMQ data have been extensively used for pharmacoepidemiologic research, including dispensation patterns, and the effectiveness and safety of oral anticoagulants[187-189].

However, there are limitations of the RAMQ databases when using them for research purposes. Thus, the RAMQ databases were designed by the Government to collect information for administrative and reimbursement purposes, not for research. Moreover, the public drug insurance plan is restricted to those aged 65 years and older, welfare recipients, those not covered by private medical insurance, and their dependents (about 40% of the Québec population). A previous study reported that the prescription claims database is over-represented by individuals of lower socio-economic status[190]. Because of that, studies based on the prescription database may have limited external validity (generalizability) as they are restricted to specific segments of the Québec population.

The internal validity of studies on comparative effectiveness and safety may be compromised, and information bias or residual and unmeasured confounding may arise in several ways. Thus, in the prescription claims database, the information on drug exposure may be incomplete since it does not include data on medications dispensed during hospitalizations or stay in long-term care facilities, over-the-counter medications, and free medication samples that are given by physicians. The therapeutic indications for drug use are not recorded. There may be incomplete longitudinal data on dispensed drugs for those individuals who migrate between the RAMQ and private health

insurance, as well as for those who travel out of the province. Furthermore, the information recorded in the physician claims and hospital discharge summaries databases is lacking clinical details, including disease severity. RAMQ databases do not provide information on referrals for laboratory tests or their results (e.g., kidney function or INR measurements). These limits the identification of patients in whom DOACs are contraindicated, i.e., dabigatran is contraindicated when the estimated glomerular filtration rate (eGFR) is <30 mL/min, rivaroxaban when it <30 mL/min (at time of the study) or 15 mL/min (at present), and apixaban when it is <15 mL/min. Lack of information on INR measurements limits defining the anticoagulation levels in patients treated with warfarin. Furthermore, RAMQ data missing some important determinants of health such as lifestyle factors (e.g., physical activity, smoking, use of alcohol, etc.), income, education, gender (not biological sex), and morphometric characteristics (body mass index).

The recording of diagnoses in the *medical services database* may not be accurate as it is not mandatory for reimbursement. Furthermore, physicians can only enter one diagnosis for each visit, thus multiple diagnoses are not included. For example, for patients with diabetes and hypertension, the latter would not be in the outpatient claims[191].

Several studies assessed the validity of diagnoses in the *medical services database* and found that the recorded information was highly specific, however, had varying sensitivity[192-194]. Thus, a study by Wilchesky et al.[192] validated the diagnostic information for 14, 980 patients using medical charts as a gold standard. Diagnoses were obtained from the medical services claims data using ICD-9 codes e.g., ICD-9 codes 250.0–250.9 for diabetes, 428.0–428.9 and 429.3 for congestive heart failure, 410-410.9 for myocardial infarction, 401.0–401.9 hypertension, 585.0–586.9 for renal

failure, 443 and 441 for peripheral vascular disease, 430-438 for cerebrovascular disease. The study reported high specificity of recorded diagnosis. For example, specificity for diabetes was 96.8% (95% CI: 96.5-97.1), congestive heart failure 96.1% (95%CI: 95.7-96.4), myocardial infarction 96.8% (95% CI: 96.5-97.1), hypertension 81.9% (95%CI: 81.0-82.8), renal failure 99.1 % (95%CI: 98.9-99.2), peripheral vascular disease 95.0% (95%CI: 94.6-95.4), cerebrovascular disease 95.3% (95%CI: 95.0-95.7). However, the reported sensitivity was substantially lower, and ranged from 68.7% (95%CI:67.7-69.7) for hypertension and 64.4% (95%CI:62.6-66.2) for diabetes to 25.4% (95%CI: 22.35 -28.66) for myocardial infarction, 18.6% (96%CI: 15.7-21.8) for chronic renal failure, and 19.7% (95%CI: 15.4-24.7) for dementia.

Studies assessing the validity of diagnoses in the hospital discharge dataset *MED-ÉCHO* concluded that information is sufficiently reliable[195-197]. Thus, Lambert et al.[195] investigated the reliability and predictability of data in MED-ÉCHO compared to data from medical records in a sample of 1,989 randomly selected patients hospitalized with myocardial infarction or those who underwent angioplasty or bypass surgery. The study evaluated the conditions included in the Charlson Comorbidity Index and some other predictors of mortality in cardiac patients and found that information recorded in hospital discharge data, in general, was reliable. Thus, for hypertension, the reported sensitivity was 87.5% (95%CI: 85.7-89.4), specificity was 97.4% (95%CI: 96.1-98.3), positive predictive value (PPV) was 97.7% (95%CI: 96.5-98.5), and negative predictive value (NPV) was 86.3% (95%CI: 84.3-88.2). For chronic renal disease these measures were 88.2% (95%CI: 84.0-91.6), 99.3% (95%CI: 98.8-99.7), 96.1% (95%CI: 93.1-98.0), and 97.9% (95%CI: 97.1-98.5), respectively. For acute renal disease, the measures were 66.2% (95%CI: 59.8-72.6),

99.0% (95%CI: 98.5-99.4), 83.6% (95%CI: 75.4-90.0), and 97.5% (95%CI: 96.7-98.1), respectively. The lower sensitivity was reported for the history of previous myocardial infarction (54.6%; 95%CI: 51.1-58.2), however, with high sensitivity (99.1%, 95%CI: 98.5-99.5) PPV (94.1%, 95%CI: 90.4-96.8), and NPV (89.3%, 95%CI: 87.9-90.6).

The next limitation of RAMQ data is missing information on individuals seen by physicians who are paid by the Alternative payment plan (APP) and not fee for service, hence the services under the APP are not included in the RAMQ. A study found that in 2016, about 20% of physicians were covered by the APP in Québec[198]. RAMQ does not collect information on services provided by medical officers for the Canadian Armed Forces, as well as services provided in the Canadian Forces Health Services Clinic. Furthermore, the RAMQ information collection system is passive: data is collected when an individual presents to a doctor. This may cause an overrepresentation of individuals with certain underlying characteristics and reduce the external validity of the study. Thus, those with better follow-up are more likely to have their diagnosis ascertainment and get the treatment. For example, individuals with asymptomatic conditions such as hypertension or diabetes are more likely to be diagnosed and treated if they reside in urban versus rural areas, have high socioeconomic status, and had a family physician. Moreover, healthcare access bias (a form of selection bias) may arise when individuals from different study groups have different access to treatment, diagnostic tests, etc. due to economic, geographic, cultural, or other reasons. For example, in the US, this bias may partially explain the finding from cohort studies that demonstrated that incidence and prevalence of atrial fibrillation were lower in African Americans compared to Caucasians[199].

Overall, because of strict data protection regulations in Québec, it is difficult to link the information from RAMQ to individual outpatient or inpatient medical charts. These restrict the collection of more detailed clinical information (to address the unmeasured confounding) or the conduct of studies for validation of recorded diagnoses or drug use patterns (to address information bias)[200].

To have an access to linked, patient-level data contained in RAMQ or MSSS databases, a request must be submitted to the Quebec government agency *Commission d'accès à l'information* (CAI) for approval[201]. Following submission, the turnaround time for approval may take up to six-eight months[202]. Since 2019, an online submission process has been available via the *Institut de la statistique du Québec* (ISQ) Research Data Access Point that reduces the data processing time[203]. The CAI grants the authorization to receive data for research or statistical purposes without the patient's consent, provided the obligations of patients' confidentiality and inability to identify personal information. These conditions are set in the *Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels* and *Loi sur la protection des renseignements personnels dans le secteur privé*[201].

3.2 Study population

The study population was new users of DOACs or warfarin in atrial fibrillation. Inclusion and exclusion criteria for the study population are outlined in Table 3.1. A list of DIN and ICD codes used to define the study population is in Annex 2 and Annex 3, respectively.

Table 3.1. Eligibility criteria for the study population

Inclusion criteria
1. Dispensation of an oral anticoagulant (warfarin, dabigatran, rivaroxaban, or apixaban) between January 1, 2010, and December 31, 2016[ab]. The date of the first dispensation is assigned as the date of cohort entry
2. Covered by the RAMQ public drug insurance plan at least one year before the date of cohort entry[b]
3. Aged at least 18 years at the date of cohort entry[b]
4. Having a diagnostic code for atrial fibrillation (ICD-9 427.3, ICD-10 I48.x) in the year before the date of cohort entry
Exclusion criteria
1. Having a dispensation for any oral anticoagulant in the year before the date of cohort entry
2. Diagnosis of valvular disease (including rheumatic heart disease) or prior cardiac valve surgery in the year before the date of cohort entry
3. Diagnosis of venous thromboembolic disease in the 3 months before and including the date of cohort entry
4. Hemodialysis 3 months before cohort entry

[a]For the first study the period for patient identification was from 1 January 2010 to 15 March 2015 because of data availability
[b]Preselected by RAMQ

The study population was patients who initiated oral anticoagulation (warfarin, dabigatran, rivaroxaban, or apixaban) between January 1, 2010 (the year when Health Canada approved the first DOAC, dabigatran for use in AF), and December 31, 2016, were aged ≥18 years and covered by the RAMQ public drug insurance plan at least one year before treatment initiation (note, for the first study the period for patient identification was from 1 January 2010 to 15 March 2015 because of data availability). Edoxaban was not included because it was not approved for stroke prevention in AF during the study period. The DINs used to identify oral anticoagulants are listed in Annex 2.

Treatment initiation was defined as having no dispensation for any oral anticoagulant in the year before. This was required to ensure that prevalent users are not included in the study to avoid the depletion of the susceptibles bias[204]. This bias may be introduced by including in the study patients who were newly dispensed a DOAC, but previously were treated with warfarin. Such individuals may have a lower risk for stroke and bleeding compared to patients without previous treatment with oral anticoagulants[120].

The date of the first dispensation was set as the date of cohort entry (time zero, t0). Further, those eligible for the cohort must be aged at least 18 years at the date of cohort entry and had a diagnostic code for AF or atrial flutter (ICD-9 427.3, ICD-10 I48.x) in the hospitalization discharge summaries at any position, or in the physician claims. The diagnosis of AF or atrial flutter had to be within a year before the date of cohort entry. A systematic review of previous validation studies reported that the PPV of ICD-9 codes for AF ranged from 70% to 96%, with a median of 89%, and the sensitivity ranged from 57% to 95%, with a median of 79%[205].

We excluded those patients who had a diagnosis of valvular disease (including rheumatic heart disease) in the year before cohort entry (see Annex 3 for ICD codes). This period was chosen considering the disease severity and existing recommendations for routine follow-up. Thus, according to CCS recommendations, patients with diagnosed valvular heart disease should undergo regular follow-up with periodic examinations at least every year[65]. If these patients develop an AF, their risk for thromboembolism is considered high, and they should receive oral anticoagulation with warfarin. Furthermore, in clinical practice, all patients newly diagnosed with AF are routinely examined for the presents of valvular heart disease. Because RAMQ contains information on all

medical visits, we assumed that a one-year look-back period is sufficient to ascertain patients with valvular heart disease. Another category of patients that were excluded from the study were those who underwent cardiac valve surgery in the year before cohort entry because anticoagulation is recommended temporarily for these patients for the prevention of postoperative thromboembolic events[206].

To improve the precision in identifying patients with AF, we excluded those with a diagnosis of venous thromboembolic disease in the three months before cohort entry (see Annex 3 for ICD codes). Three months is the minimum duration of treatment with oral anticoagulants in patients with thromboembolic disease[207], thus any anticoagulation within that time window was assumed to be prescribed for venous thromboembolism. Next, we excluded patients who underwent hemodialysis three months before cohort entry. Hemodialysis is indicated for patients with renal failure to whom the use of DOACs is contraindicated. Three-month period was chosen to allow the inclusion of patients with a reversible form of renal failure (for example, acute tubular necrosis) that may require hemodialysis temporarily[208].

Individuals were followed into the cohort until the earliest of the following: end of health or drug insurance, death, or end of the study period (March 31, 2017). Thus, patients who entered the cohort close to December 31, 2016, were allowed to have at least three months of follow-up. Note, that for the first study, the end of the study was March 31, 2015, because of data availability.

3.3 First study

3.3.1 Aim and objectives

The first and the second studies of this thesis aimed to explore the potential for information bias resulting from misclassification of exposure to dose-varying warfarin therapy when using days supplied value recorded in pharmacy claims data. The first study was descriptive and had two objectives: (1) to characterize the dispensation patterns of warfarin, apixaban, dabigatran, and rivaroxaban, and (2) to compare the variation of refill gaps between users of different oral anticoagulants, as well as within users of the same oral anticoagulant over the treatment course.

3.3.2 Base cohort definition

Individuals with NVAF newly dispensed an oral anticoagulant were identified as described in section *3.2 Study population*. Study participants were followed from the date of the first dispensation until the last dispensation, end of pharmaceutical coverage, or end of the study period (March 31, 2015), whichever occurred first. Because the study was descriptive, we collected the entire dispensation history regardless of switching between oral anticoagulants.

Further, patients were grouped into four cohorts based on the type of dispensed oral anticoagulant (warfarin, dabigatran, apixaban, or rivaroxaban). When a patient switched between anticoagulants during the follow-up, this was considered a new treatment episode, and this patient was reallocated to the corresponding cohort at the date of the switch.

3.3.3 Matched cohort definition

To estimate the variation of refill gaps, only those having at least three consecutive dispensations of the same oral anticoagulant were eligible. At baseline, compared to patients on DOACs, those on warfarin had more comorbidities (particularly prior hemorrhage) (Table 4.2) and thus were more likely to stop treatment because of the high risk of bleeding or have temporary treatment interruptions for medical procedures. In our calculation of refill gaps, we tried to account for possible interruptions in the therapy by subtracting the days spent in a hospital from the refill gaps. Furthermore, to make patients more comparable we matched patients from four cohorts.

Patients receiving warfarin and each of the three DOACs were matched through a common-referent group[209]. For this, first, I chose the apixaban cohort as the referent cohort because it had the lowest number of participants. Those from the apixaban cohort were consecutively matched in a 1:1 ratio with those from dabigatran, rivaroxaban, and warfarin cohorts. Finally, only patients with an available match for all drugs were retained for the analysis of variation of refill gaps. Matching was done on age, sex, and propensity scores (PS) estimated from the prognostic characteristics measured in the year before the base cohort entry date (see below section *3.3.7 Propensity scores*).

Next, among matched individuals from four cohorts (dabigatran, rivaroxaban, apixaban, and warfarin), I identified the subgroup of weekly pillbox users. In Québec, to support patients in their adherence to the treatment, community pharmacies offer a 7-day supply of prescribed medications either in sealed pillboxes or in a plastic vial if the dosage of medication may be changed[210] (further in the text, I will refer to all 7-day supplies as “pillboxes”). Pillboxes are automatically prepared and often delivered at home without the need for the patient to request dispensation. The weekly

pillbox users were defined as those having had at least four successive dispensations with seven days duration. The use of pharmacy-prepared weekly pillboxes may influence treatment adherence, therefore, in secondary analysis, I stratified the matched individuals on the users and non-users of weekly pillboxes.

The flowchart of patient selection in the First study is in Table 4.1, Chapter 4.

3.3.4 Exposure definition

I identified all dispensations of warfarin, dabigatran, rivaroxaban, and apixaban by their DINs (Annex 2. List of Drug Identification Numbers (DIN) used to define exposure to anticoagulants in all studies of this thesis.). For each dispensation, I extracted the drug generic name, dispensing date, and days supplied. For a given individual, all dispensations were sorted by their dates in ascending order.

3.3.5 Refill gaps

To estimate the refill gaps, for each patient, all dispensations for a given oral anticoagulant were ordered chronologically to replicate the refill history. A refill gap was defined as the absolute number of days between the expected date of refill of the current dispensation (estimated based on days supplied) and the observed date of refill of the following dispensation. For example, Figure 3.1 shows an extract from the pharmacy claims dataset for an individual who has four consecutive dispensations of dabigatran. The first dispensation with a 30-day supply was on July 4th, 2012. Thus, the expected date of the next dispensation would be July 4th, 2012 + 30 days=August 3d, 2012. Yet, the actual date of refill recorded in pharmacy claims was August 12, 2012. Therefore, a gap of nine days occurred between the expected and actual dates of refill.

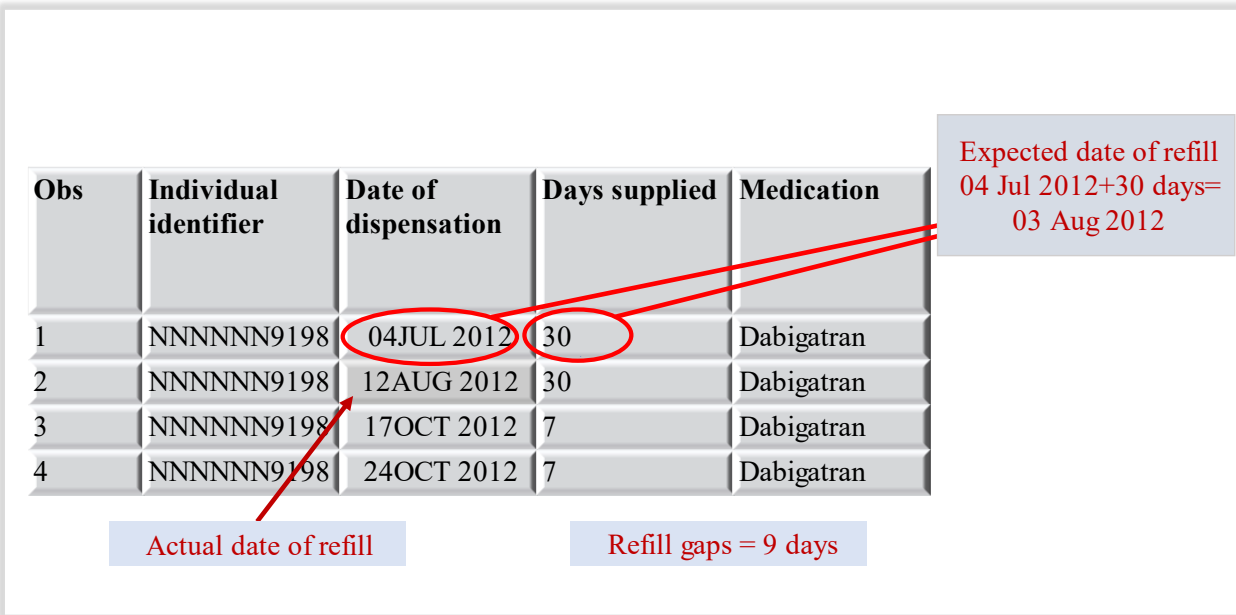


Figure 3.1 Estimation of refill gaps using pharmacy claims data.

Pharmacy claim data does not provide information on medications provided to a patient in the hospital. Furthermore, in case of hospitalization, a patient would have a surplus of drugs at home and may refill the prescription at the community pharmacy later. This would result in a refill gap. To account for this, I identified all hospitalization periods and adjusted accordingly the refill dates (e.g.: if a patient was hospitalized for eight days following a dispensation with 30 days supplied, the expected day of refill was calculated as days supplied (30) + days in the hospital (8) = 38 days following the dispensation).

Refill gaps are commonly used to describe treatment adherence[211]. These studies assume that the number of dispensed pills is adequate to cover the period recorded by the dispensing pharmacist (days supplied variable). In a situation with complete adherence, when, for example, a patient has been dispensed 30 pills for 30 days and returns for a refill on the 31st day, the refill gap is zero. Alternatively, in a situation of poor adherence, the patient may occasionally miss some daily pills, and return for a refill on day 39^{th,} resulting in a refill gap of 9 days.

However, refill gaps may also capture deviations from prescribed posology, such as in the case of dosage adjustments of warfarin. For example, a patient with a one-pill-daily regimen is given 30 pills by his pharmacist who enters 30 days supplied in the dispensation database. The following day (for simplicity of calculations), the physician instructs the patient to take half a pill every day. This patient, even completely adherent, has pills for 60 days and returns on day 61 resulting in a refill gap of 30 days.

Thus, the refill gaps capture information on both adherence and adequacy of the days supplied data. Under the assumption of no (or negligible) adherence variation between warfarin and DOACs at a

population level, any significant increase in refill gaps for a given medication would indicate that days supplied data is an imprecise measure of the actual number of days of availability of that drug.

Other situations may result in gaps between consecutive refills among patients with optimal adherence. For example, (a) in the case of the first prescription, the patient may have a free medication sample and return for the refill later, and (b) the stockpiling of a drug may occur if a patient returns for a refill a few days before completing his previous supply, (c) a change in dosage may require a new prescription for tablets with a different dose (e.g., a patient had 5 mg tablets of warfarin, then he/she is prescribed 2.5 mg tablets), but this patient may still use the tablets from the old supply and return later for the refill.

3.3.6 Statistical analysis

Patients from the base cohort (before matching) were characterized by age, sex, comorbidity profile, and use of medications measured at the time of cohort entry. Descriptive statistics were used to summarize the distribution of the recorded days supplied variable, the dispensed quantity, strength of the dispensed drug, the number of days between two consecutive dispensations, and the refill gaps for each oral anticoagulant. For these variables, in a separate table, I described the mean with standard deviation (SD), median with interquartile range (IQR), modes, minimum and maximum values, or the number of dispensations with percentage.

Matching individuals from four cohorts. Logistic regression was used to calculate the PS at baseline, or each time a patient switched oral anticoagulants. Thus, I fitted three PS models, and in each of these models, those treated with apixaban were the reference group. The first PS model

estimated the probabilities of receiving warfarin versus apixaban, the second PS model was for dabigatran versus apixaban, and the third model was for rivaroxaban versus apixaban. All three PS models were adjusted for covariates that were assumed to be associated with treatment assignment and health events leading to treatment interruptions.

Those individuals with the PSs <= the 5th percentile or >= the 95th percentile of the overall propensity score distribution were truncated (their PSs were increased or reduced to the values of the 5th or 95th percentile, respectively).

For the analyses of the variation of the refill gap, those from the apixaban cohort were consecutively matched in a 1:1 ratio with those from the dabigatran, rivaroxaban, and warfarin cohorts. Greedy nearest neighbor matching[212] was performed on age (±1year), sex, and PS (using a caliper of 0.2 standard deviations of the logit of PS). This caliper was chosen based on recommendations from the study by P. Austin[213] that used Monte-Carlo simulations to determine the optimal caliper widths for PS matching. The study showed that the caliper of 0.2 standard deviations of the logit of PS was the best trade-off between the precision of the treatment effect estimate and the reduction of bias due to systematic differences between treatment groups.

The balance in the distribution of covariates between matched patients from cohorts of users of oral anticoagulants was examined using the median (IQR) for the continuous variables, and numbers and percentages for dichotomous variables.

Next analyses on refill gaps were done with the matched individuals from four cohorts. To account for the clustered structure of the data (multiple prescriptions per patient) and to estimate the variation of refill gaps, I used multilevel linear regression models (see description in Section

3.3.8)[214]. From the variance components produced by the multilevel model, I calculated the intraclass correlation (ICC) that assesses the clustering, or similarity of refill gap values in the same patient[215]. I repeated the multi-level analysis, stratified according to the use of pharmacist-prepared weekly pillboxes.

3.3.7 Propensity scores

Randomized controlled trials are the gold standard to estimate the effect of treatment on an outcome. Because of random treatment allocation, there is a balanced distribution of measure and unmeasured prognostic baseline characteristics between treatment groups. In observational studies, treatment selection is not random. This gives rise to a special methodological problem, confounding. Confounding occurs when the factors that prompt the use of a certain drug also modify the probability of the study outcome, and these factors are not on the causal pathway between the drug use and the outcome[216]. Methods for minimizing the effect of confounding include those applied during the design phase and the analysis phase of an observational study. In the design phase, restriction or matching may be used to exclude or control for confounding. In the analysis phase, methods of confounding control include stratification, standardization, regression adjustment, or propensity score (PS) methods[217].

PSs are defined as the balancing scores that make treatment assignment to be ignorable on the potential outcome conditional on measured baseline prognostic factors[218]. PS predicts the probability of exposure to the study drug, given pre-existing patient characteristics. For example, in our study cohort, a higher proportion of individuals initiating warfarin had a diagnosis of congestive heart failure at baseline compared to those initiating DOACs (Table 6.2, Chapter 6:).

Thus, the baseline risk of stroke was higher for those on warfarin. After matching using propensity scores (0.2 of the standard deviation of logit of propensity scores), age (±1 year), sex, and cohort entry date (±2 years), the proportions of patients with congestive heart failure were close in the DOACs and warfarin treatment groups (Table 6.2, Chapter 6:). Furthermore, an advantage of the PS method over other methods of confounding control is flexibility with the number of baseline covariates to control as it allows combining information on many potential confounders into a single score[219]. Patients from two treatment groups with the same PS value have a similar distribution of prognostic characteristics, thus, the estimated treatment effect will not be affected by confounding. However, the PS methods as well as other methods of confounding control (except randomization) cannot account for unmeasured confounding.

In all three studies of this thesis, I applied PSs to control for confounding. It is worth noticing that in such a large observational study using traditional covariate adjustment of the outcome model may yield estimates with the same precision[220]. The traditional covariate adjustment allows inferring the effect of each covariate on the outcome of interest. However, this was not our research interest. Instead, all prognostic factors that may be associated with the outcome of interest we identified from the previously published papers. In this research, we used the PS methods because of the advantage of assessment of correct specification for the PS model as opposed to the outcome model with traditional covariate adjustment[219]. To assess the correct specification of the PS model, we compared the distribution of measured baseline covariates between DOACs and warfarin-treated patients in a matched sample. For example, tables 4.4 (Chapter 4) and 6.2 (Chapter 6) show that after matching, the distributions of all covariates were

similar across the two treatment groups. Furthermore, PS methods allow separation of the design and the analysis phases of the study. This precludes the modification of the outcome model until the desired outcome estimate is achieved[221].

The PSs were estimated with logistic regression. Variables included in the PS model were those considered to be the true confounders, i.e., baseline patient characteristics that may be associated with the exposure assignment and with the outcome. Moreover, those characteristics that are strongly associated with the outcome regardless of their association with the exposure assignment (potential confounders)[222] were included as well. Based on the simulation study, Brookhart et al. concluded that the inclusion of variables associated only with the outcome but not with the exposure may minimize bias and reduce the variance of the estimated effect of exposure[223]. The possible explanation is as follows: "if a covariate is theoretically unassociated with exposure, there can be some slight chance relation between the covariate and the exposure for any given realization of a data set. If that covariate is also related to the outcome, then it is an empirical confounder for that particular data set"[223]. Thus, in a given study population, a prognostic variable thought to be unassociated with exposure may, in fact, be associated with exposure in a subset of patients with a specific combination of the baseline characteristics. Conversely, the inclusion of variables that are only predictors of exposure may fail to eliminate bias and lead to a less precise estimate of the exposure and outcome relashionships[223]. Austin et al. give an example of such a situation using a cohort of patients assigned to a newer and an older drug[224]. Patients entering the cohort earlier have a higher probability to receive an older drug, whereas patients entering the cohort later have a higher probability to receive a newer drug. A variable denoting

time is associated with treatment assignment but is not associated with the outcome. Thus, the inclusion of this time variable in the PS model may result in estimated PSs that are substantially different between the treatment groups (e.g., the extreme case would be when one treatment group has very high PSs, and the other has very low PSs). Therefore, this may lead to the formation of fewer matched pairs, less statistical power, and an imprecise treatment effect estimate or introduction of Type II error. Furthermore, P. Austin notes that it may be difficult to categorize the baseline patient characteristics into true confounders, potential confounders, or variables associated only with exposure[219]. However, variables denoting dates or time, or policy-related variables (e.g., prescription drug coverage may differ in public versus private drug insurance plans) warrant more investigation.

The variables that I used to build the PS were those that may be associated with treatment assignment and/or may modify the risk of thromboembolism or/and bleeding. These variables were identified based on expert opinion and a review of observational studies on the comparative effectiveness and safety of oral anticoagulants.

Several variables were identified as confounders and included in the PS model

- Age and sex
- Medical diagnoses (congestive heart failure, hypertension, diabetes, any stroke, transient ischemic attack, chronic kidney disease, acute kidney injury, liver disease, cancer, chronic obstructive pulmonary disease, coronary atherosclerosis, myocardial infarction, peripheral vascular disease, prior major bleeding, and dementia)

- Medical procedures (implantation of pacemaker or pacemaker-defibrillator, prior coronary artery bypass surgery, prior percutaneous coronary intervention, atrial fibrillation ablation, or auricular appendage closure)
- Use of drugs (aspirin, other antiplatelet agents, non-steroidal anti-inflammatory drugs, H2 receptor blockers, proton pump inhibitors, medications used for the treatment of hypertension, antidiabetic drugs, insulin, statins, selective serotonin reuptake inhibitors, benzodiazepines, typical and atypical antipsychotics, cholinesterase inhibitors, opioids, systemic corticosteroids).

Soon after DOACs entered the market, warfarin was more likely to be prescribed to individuals of older age, women, those with a history of stroke or bleeding, myocardial infarction, CHF, diabetes, renal disease, hypertension, or peripheral vascular disease[113-116]. Furthermore, age, sex, congestive heart failure, hypertension (or use of medications recommended for treatment of hypertension as a proxy), diabetes (or use of antidiabetic drugs and insulin as a proxy), any stroke, transient ischemic attack, chronic kidney disease, acute kidney injury, liver disease, cancer (or use of opioids as a proxy), chronic obstructive pulmonary disease, coronary atherosclerosis, myocardial infarction, peripheral vascular disease, and history of major prior major bleeding are the predictors of either stroke or bleeding, or both stroke and bleeding[36,39,62,225-227]. Surgical procedures for the treatment of coronary atherosclerosis (coronary artery bypass surgery or percutaneous coronary intervention) increase the risk of stroke[228]. Surgical treatment of cardiac arrhythmias (implantation of a pacemaker or implantable cardioverter-defibrillator, atrial fibrillation ablation or left auricular appendage closure) may modify the risk of stroke by either reducing it[229], or increasing it[230].

Aspirin and other antiplatelet agents, non-steroidal anti-inflammatory drugs, H2 receptor blockers, and proton pump inhibitors are drugs that increase bleeding risk[231]. Patients with dementia (or use of typical and atypical antipsychotics, and cholinesterase inhibitors as proxies) have a high risk of fall[232] and a low level of adherence to medication[233] which in turn increases the risk for stroke and bleeding[234]. Moreover, patients with vascular dementia may have a higher risk of stroke[235].

Comorbidities were identified using inpatient (hospital discharge dataset *MED-ÉCHO*) and outpatient data (*medical services database*). The use of medication during the baseline period was identified from the pharmacy claims data. Patients were considered exposed to a concomitant medication if the date of cohort entry was within the time frame estimated from the date of dispensation plus the duration of the prescription and grace period of 30 days.

In all three studies, the lookback period for ascertainment of comorbidities was one year. This lookback period was chosen based on findings from previously published studies. These studies found that with administrative health care data, the increase of the lookback period beyond one year did not result in significant improvement in the identification of comorbidities[236,237]. Moreover, the combination of health data from outpatient, inpatient records, or pharmacy claims allows for improving sensitivity in identifying comorbid conditions[238].

The ICD codes for medical diagnoses are provided in Annex 3. List of the International Classification of Diseases (ICD) codes used in all studies of the thesis.The use of drugs was identified with corresponding DINs converted from the Anatomical Therapeutic Chemical (ATC) classification codes for medications listed in Annex 4.

Overall, there are two strong assumptions for PS to effectively eliminate confounding: (1) within a group of patients with the same PS, treatment assignment is independent of the potential outcome (i.e., treatment assignment is not associated with prognostic variables), and (2) each study patient has a nonzero probability to receive either treatment. These two assumptions imply other conditions: no unmeasured confounders and correct specification of statistical model for PS estimation. In my studies, the correct model specification issue was resolved by recommended approaches to variable selection[224], testing several statistical models, including interaction and quadratic terms, and finally testing the covariate balance between study treatment groups when matching on PS was effectuated (first and third studies)[219]. To determine the best-fitted model, I used Akaike's Information Criteria (AIC) and Bayesian Information Criteria (BIC). However, the condition of "no unmeasured confounders" is not testable. Thus, the effectiveness of the estimated PS in eliminating confounding depends on expert knowledge, availability and completeness of information in the database, including information on surrogate variables[239].

3.3.8 Multilevel or mixed-effect model

Traditional statistical methods such as regression methods or T-tests treat all observations in the dataset as if they are independent. Independence of observations from each other means that the measurements are unrelated[240]. However, some data have a hierarchical structure that implies some degree of dependence or correlation between observations[241]. An example is clustered data when the outcome variable is measured once for each individual, but individuals are grouped by some shared activity or characteristics (e.g., patients of the same physician who has certain treatment preferences). Another example of non-independent observations is the repeated measures data

when the outcome variable is measured more than once for each individual (e.g., repeated measurement of blood pressure for patients over the follow-up).

Violating the assumption of independence leads to the underestimation of the standard error (SE) and the overestimation of statistical significance (type-1 error)[242].

Thus, the SE of the mean is estimated as:

$$Standard\ error = S/\sqrt{N} \quad (3)$$

where S = standard deviation, N = the total number of observations.

In the case of non-independent observations, the effective number of observations (i.e., the number of independent observations that should be used in computing the SE)[243] is lower than the total number of observations. Thus, when accounting for non-independence, the estimate of SE would be larger.

The first study of this thesis used repeated measurements of refill gaps for each patient in the study. The frequency and regularity of dispensations received by one individual reflect characteristics that are inherent to that person. These include adherence to medication, dosage, tendency to have stable or unstable warfarin dosage, stockpiling preferences, traveling habits, etc. Therefore, refill gap patterns are more likely to be similar within an individual than between two different individuals. Thus, any individual patient will have a smaller variation of refill gaps compared to the variation of refill gaps between two distinct patients.

The data with refill-gaps measurements is structured in two levels: refill-gaps (level 1) measured within individuals (level 2) (Figure 3.2). With two levels, there are two sources of variation of refill gaps: within-individual and between-individual.

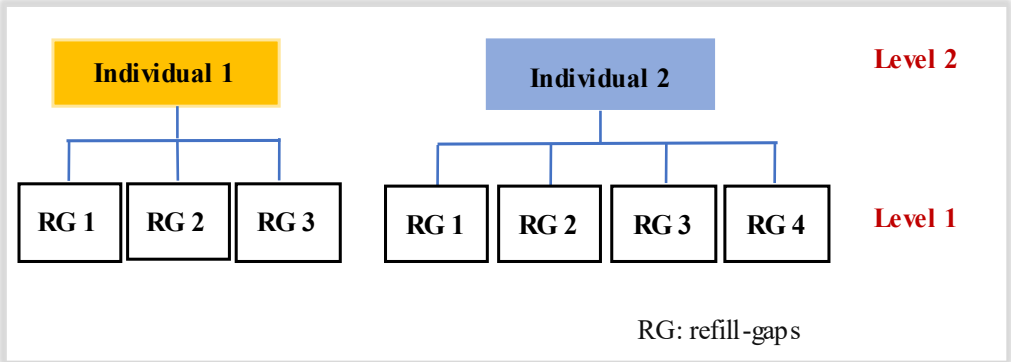


Figure 3.2. Two-level structure of data for analysis of refill-gaps.

Let Y_{ij} is an outcome variable (a refill-gap) for measurement i within individual j.

Then the models for two levels of variation of refill gaps are:

Level 1 (estimation of an individual mean and the within-individual variance of refill gaps):

$$Y_{ij}=a_{0j}+e_{ij}, \quad e_{ij} \sim N(0, \sigma^2) \tag{4}$$

Where a_{0j} is the mean refill-gap of individual j, e_{ij} is the residual individual differences from the mean of individual j, and σ^2 is the within-individual variance.

Level 2 (estimation of the grand mean and the between-individual variance):

$$a_{0j}=a+u_{0j}, \quad u_{0j} \sim N(0, \sigma_{00}^2) \tag{5}$$

Where a is the grand mean for all patients, u_{0j} is the residual differences between the individual means and the grand mean, σ_{00}^2 is the between-individual variance

Composite equation:

$$Y_{ij}=a+(u_{0j}+e_{ij}), \quad u_{0j} \sim N(0, \sigma_{00}^2)\ e_{ij} \sim N(0, \sigma^2) \tag{6}$$

Where a is the grand mean for all patients, u_{0j} is the between-individual residuals, e_{ij} is the within-individual residuals, σ^2 is the within-individual variance, σ_{00}^2 is the between-individual variance

In the first study, I fitted two models – the unconditional mean model and the unconditional growth model. The unconditional mean model was fitted for each of the four groups of matched patients (patients on warfarin, apixaban, dabigatran, or rivaroxaban). The goal of this analysis was to describe the variation of refill gaps in each matched cohort. The unconditional mean model or intercept-only model with no predictors provided the estimate of the grand mean of refill gaps and quantified the within-individual and between-individual variation of refill gaps across patients without considering the time. An example of the output of the SAS MIXED with the estimates from the unconditional mean model for the matched patients treated with apixaban is provided in Figure 3.3. The output contains the estimation of the grand mean (a) as a fixed effect, the within-individual variance (σ^2), and the between-individual variance (σ_{00}^2) as random effects[244]. The grand mean is the true mean of refill gaps across all patients in the cohort. The T-test for the estimate of the grand shows that the mean refill gap was non-zero. The estimate of the within-individual variation shows how a patient's measurements of refill gaps deviate from his/her individual-specific mean (default SAS output for proc MIXED did not provide the estimates of individual-specific means). The between-individual variation summarised how the patient-specific means of refill gaps deviated from the grand mean.

To describe these parameters in the users and non-users of weekly pillboxes, I stratified patients by pillbox use and run the unconditional mean model in each stratum separately.

Figure 3.3. Output from SAS MIXED procedure.

The SAS System 11:22 Tuesday, August 3, 2021 1

The Mixed Procedure

Model Information

Data Set	WORK.SIGMA_1
Dependent Variable	Gap
Covariance Structure	Variance Components
Subject Effect	NAM
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Name of the dataset
Response variable (refill gap)
Covariance structure
Group variable (which is an individual in my analysis)
Estimation method (REML=restricted maximum likelihood) default in SAS

Class Level Information

Class	Levels	Values
nam	7376	not printed

Dimensions

Covariance Parameters	2
Columns in X	1
Columns in Z per Subject	1
Subjects	7376
Max Obs per Subject	508

Number of groups (individuals in my analysis)
Maximum number of observations (refill gaps) per individual

Number of Observations

Number of Observations Read	121407
Number of Observations Used	121407
Number of Observations Not Used	0

Number of observations (refill gaps) in the dataset
Number of observations (refill gaps) used in the analysis

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	767865.13718338	
1	1	746679.62613679	0.00000046
2	1	746679.50472241	0.00000000

Convergence criteria met.

Estimates of two variances (random effects)
SAS uses Wald z -test to test where the covariances are significantly different from 0

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr > Z
Intercept	nam	21.8336	0.4585	47.62	<.0001
Residual		23.9582	0.1001	236.34	<.0001

Between -individual variance (σ_{00}^2)
Within -individual variance (σ^2)

Fit Statistics

-2 Res Log Likelihood	746679.5
AIC (Smaller is Better)	746683.5
AICC (Smaller is Better)	746683.5
BIC (Smaller is Better)	746697.3

Estimates of fixed effect (intercept and slope)
In my analysis with no predictors, it shows only intercept.
SAS uses t -test to test where the coefficients are significantly different from 0

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	3.0774	0.05747	7376	53.55	<.0001

Estimate of the Grand mean (a)

The next step was to explore if the pharmacy claims data may capture the dose adjustment patterns at the initiation of warfarin. This was done with the unconditional growth model with time as the only predictor of changes in refill gaps. As a time variable, I used an ordered dispensation number. This model quantified the variation of refill gaps across patients over time under the assumption that the trajectory of change was linear. The latter was investigated in a sample of 200 randomly selected patients treated with warfarin by constructing nonparametric growth trajectories and OLS (ordinary least squares) fitted trajectories (Figure 3.4. Growth trajectories representing the change of refill gaps over time in a sample of 200 randomly selected patients from the matched cohort of warfarin users.).

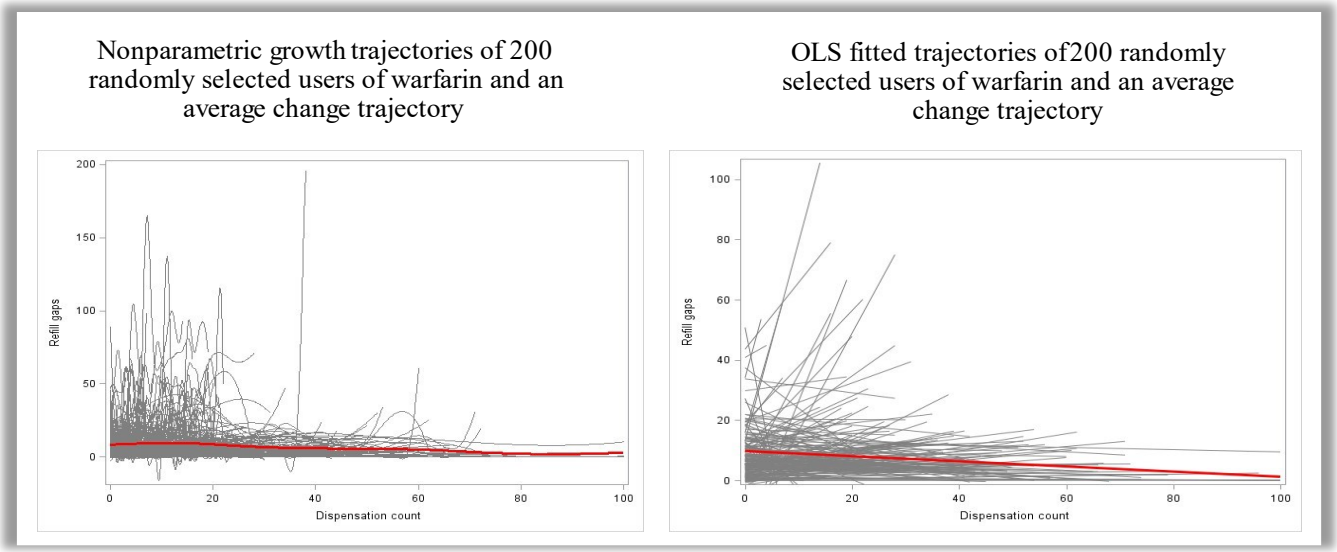


Figure 3.4. Growth trajectories representing the change of refill gaps over time in a sample of 200 randomly selected patients from the matched cohort of warfarin users.

The unconditional growth model provided the estimate of the within-individual variance that indicated how person-specific measurement of refill gaps deviated from the person-specific

change trajectory. The between-person variation was partitioned on the variation in the initial status (first refill gap measurements) and the variation in the rate of change over time (Table 4.6, Chapter 4:).

The two models that I used for my first study provided the basic estimates that allow the detection of systematic variation of refill gaps and the source of this variation (within- or between individuals).

Multi-level has several assumptions:

- Residuals at level 1 and level 2 are normally distributed;
- Relationships between the outcome variable and the predictor are linear;
- The variance of the level 1 and level 2 residuals is equal at each level of every predictor (homoscedasticity).

Normal distribution assumption

The distribution of refill gaps in our study population was skewed on the right for all oral anticoagulants. Most patients either did not have any gaps between consecutive dispensations or had small gaps, however, some patients had extremely large gaps (Table 4.3, Chapter 4:). However, we proceeded with the linear multilevel model and did not perform any data transformation because, the goal of this analysis was to quantify and describe the variation of refill gaps for each DOAC and warfarin, thus, the model did not include any predictor to explain the variation of refill gaps. Moreover, mixed-effects models are robust to some violations of the distributional assumptions including the skewed distribution[245] given a sufficiently large study sample[246].

The normal distribution assumption was verified based on the visual inspection using the Q-Q (quantiles-quantiles) plots of the distributions of level 1 (within-individual, e_{ij}) and level 2 (between-individual, u_{0j}) raw residuals (Annex 7). The Q-Q plot shows the ordered values of residuals against the quantiles of the normal distribution. When the distribution is normal, the points should form a roughly straight line[247]. In our study, the distributions of residuals for all oral anticoagulants were "heavy-tailed" with the points falling along the middle line and curves on the extremities.

Linearity and Homoscedasticity assumptions

For the unconditional mean model, these assumptions were not tested because of no predictors in the models.

For the unconditional growth model, the linearity and homoscedasticity assumptions were verified in the warfarin cohort by using a plot with the distribution of level-2 residuals versus the dispensation number (Annex 8). Based on this plot, the assumption of linear individual change of refill gaps seemed reasonable. The distribution of residuals has approximately equal range and variability for all dispensation counts. However, some patients have an extremely large variability of refill gaps mainly at treatment initiation.

3.4 Second study

3.4.1 Aim and objectives

In the second study, I continued exploring the potential for misclassification of exposure when using pharmacy claims. We investigated if the way to define exposure to oral anticoagulants may affect the estimate of the comparative safety of rivaroxaban versus warfarin and dabigatran versus warfarin. We decided to explore this after published controversial results from RCT and observational studies. Thus, pivotal RCTs reported a similar risk of major bleeding in patients treated with dabigatran and rivaroxaban compared to those treated with warfarin (RR 0.93; 95% CI; 0.81-1.07 and HR 1.03; 95%CI:0.96-1.11, respectively)[76,77]. However, the results from observational studies raised concern that in real-world patents, the use of dabigatran and rivaroxaban may be associated with an increased risk of major bleeding. Thus, a paper published in the Journal of the American Medical Association (JAMA) in 2015 reported that patients treated with dabigatran were almost 60% likely to develop major bleeding (HR 1.58; 95% CI, 1.36 to 1.83)[248]. Another paper published in the Chest journal found that rivaroxaban was associated with a higher risk of major bleeding compared to dabigatran (HR 1.30; 95% CI, 1.10-1.53)[249]. In the second study, we investigated if the way to define exposure to oral anticoagulants may affect the estimate of the comparative safety of rivaroxaban versus warfarin and dabigatran versus warfarin.

The objectives of the second study were: to evaluate whether a data-driven method to define the exposure duration captures the variability in warfarin dispensations better than the days supplied method, and to investigate the impact of the method for definition of dispensation duration on the

hazard ratios of major bleeding associated with the use of dabigatran or rivaroxaban versus warfarin.

3.4.2 Study cohort

Individuals with NVAF newly dispensed an oral anticoagulant were identified as described in section *3.2 Study population*. The cohort entry date was defined as the date of the first dispensation for an anticoagulant. The follow-up of individuals ended at the earliest of the following: occurrence of major bleeding, end of pharmaceutical coverage, death, discontinuation of the treatment allocated at cohort entry, or end of the study period (March 31st, 2017).

The flowchart of patient selection in the Second study is in

Annex **9**.

3.4.3 Exposure definitions

I identified all dispensations for dabigatran, rivaroxaban, and warfarin and extracted information on the dispensing date, days supplied, unit strength, and dispensed quantity. For a given individual, all dispensations were sorted in chronological order.

The duration of dispensations was defined using two approaches:

1. The duration of dispensation was equal to the recorded days supplied value (this approach is used most often in pharmacoepidemiological studies on comparative effectiveness and safety of oral anticoagulants);

2. The duration of dispensation was estimated for an individual based on his previous dispensation history (i.e. data-driven approach).

The data-driven approach or the longitudinal coverage approximation (COV)[250], is an algorithm that was described and evaluated previously[251]. Unlike the approach that relies on the recorded days supplied value, the longitudinal coverage estimates the duration of each dispensation given the observed time interval since drug initiation, the dispensed quantities, and the unit strength. Thus, for the prescription dispensed at time t_i, the duration of dispensation is estimated from the total milligrams of drug dispensed at time t_i divided by the average daily dose used from drug initiation (t_0) to the time of current dispensation (t_i)[97]. The average daily dose is a quotient of the total mg dispensed up until the current dispensation and the number of days passed from drug initiation to current dispensation.

Then the duration of dispensation is estimated using the formula.

$$Duration\ of\ dispensation = \frac{Total\ mg\ dispensed\ at\ time\ t_i}{\frac{Total\ mg\ dispensed\ between\ t_0\ and\ t_{i-1}}{Number\ of\ days\ between\ t_0\ and\ t_i}} \qquad (7)$$

Figure 3.5 shows an extract from the pharmacy claims dataset for an individual who has three consecutive dispensations of warfarin. First dispensation of 30 pills of warfarin 5 mg on day zero (01 July 2012); the second dispensation of 30 pills of warfarin 2mg on day 60 (30 August 2012); and the third dispensation of 30 pills of warfarin 2mg on day 90 (29 September 2012).

Obs	Individual identifier	Date of dispensation	Dispensed quantity	Unit strength, mg	Drug
1	NNNNNN9190	01 JUL 2012	30	5	Warfarin
2	NNNNNN9190	30 AUG2012	30	2	Warfarin
3	NNNNNN9190	29 SEP 2012	30	2	Warfarin

t_0

t_{t-1}

t_t

Figure 3.5 An extract from pharmacy claims to show the estimation of the duration of dispansation using the data-driven approach.

In this example, the estimated duration of the current dispensation (third dispensation) of warfarin is:

*Duration of dispensation =[(30pills*2mg)/(30pills*5mg+30pills*2mg)/90days]=26 days*

Moreover, with this approach, the maximum allowed length of the estimated duration of dispensation was two and a half of the dispensed quantity at t_i. For the first dispensation or when the preceding estimated duration exceeded the maximum allowed length, the average daily dose was replaced by the daily doses recommended or commonly prescribed to older individuals: 220 mg for dabigatran[252,253], 20 mg for rivaroxaban[252], and 3 mg for warfarin[254]. A previous study showed that the longitudinal covering accounted for individual variation in usage patterns and accurately approximated the duration of dispensation compared to the simulated duration, including medications with large dosage variability, such as vitamin K antagonist phenprocoumon and nonsteroidal anti-inflammatory drugs (NSAIDs)[251].

Duration of dispensations was adjusted for identified hospitalization periods, e.g.: if a patient was hospitalized for eight days following a dispensation with the duration of 30 days, the expected duration of dispensation was extended by eight days (new duration of dispensation =38 days).

For each individual in the study, I constructed periods of continuous drug exposure using the duration of dispensation obtained with both approaches and one of three grace periods (15, 30, and 60 days) that were used in the previous studies[90,114,141,255]. The grace period was added to account for irregular refill patterns. Therefore, we disregarded any drug stockpiling due to overlaps of the subsequent dispensations of the same oral anticoagulant. The follow-up was censored if the gap between dispensations exceeded the allowed grace period. In the case of switching to a different oral anticoagulant, the follow-up was censored at the date of this new dispensation.

3.4.4 Outcome definition

The outcome of interest was time to a major bleeding event. Major bleeding was defined as gastrointestinal, ocular, intracranial, or other bleeding requiring hospitalization. All outcome events were identified using ICD-10 codes from hospital discharge summaries using primary, principal diagnosis, or the cause of in-hospital death. The admission date to the hospital was assigned as the date of the outcome event. We did not use physician claims for identifying the outcome events to avoid a detection bias, since regular outpatient visits of individuals treated with warfarin may increase the likelihood of less important bleeding being recorded in the physician claims database.

Several studies investigated the validity of reporting major bleeding events in the administrative hospital discharge data. Thus, Joos et al. assessed the accuracy of ICD-10 codes for acute anticoagulation therapy-related bleeding in patients admitted to the University of Utah hospital using patients' medical records as a reference standard. ICD-10 codes were extracted for diagnoses at any position[256]. In total, the study identified 71 bleeding events from the medical records. The reported sensitivity, specificity, PPV, and NPV for intracranial hemorrhage were 100% (95%CI: 78%–100%), 98% (95%CI: 97%–99%), 58% (95%CI: 37%–77%), and 100% (95%CI: 99.%–100%), respectively. For gastrointestinal bleeding, reported sensitivity, specificity, PPV, and NPV for intracranial hemorrhage were 96% (95%CI: 78%–100%), 97% (95%CI: 96%–98%), 55% (95%CI: 38%–79%), and 100% (95%CI: 99.%–100%), respectively. Arnason et al. investigated the accuracy of coding (based on ICD-9 codes) of possible warfarin complications in the chart abstractions submitted by a university-associated teaching hospital in Ottawa for billing

purposes[257]. The abstracts for 1,964 hospitalizations were validated against hospital medical records as a reference standard. In this study, the possible warfarin complications were major bleeding or thromboembolism. Major bleeding was defined as an event of intracranial bleeds, GI bleeds, hematuria, vaginal bleeding, epistaxis, hemoptysis, hemorrhage not otherwise specified, hemarthrosis, or hemopericardium. The study reported that for major bleeding, the sensitivity, specificity, PPV, and NPV were 94% (95%CI: 91%–96%), 83% (95%CI: 78%–87%), 87% (95%CI: 83%–90%), and 92% (95%CI: 88%–95%), respectively.

The list of ICD-10 codes that we used for the identification of major bleeding is reported in Annex 3. ICD-10 codes for major bleeding were identified from previous publications on the comparative effectiveness and safety of oral anticoagulants and the set of codes was endorsed by our research group.

We did not use procedural codes for transfusions because procedures are generally underreported in administrative claims data. Thus a study by Lucyk et al. found the sensitivity of blood transfusion procedure code was 26%[258].

3.4.5 Statistical analysis

We used descriptive statistics to summarize the patients' demographics, comorbidity profile, and use of medications measured at the time of cohort entry. For each anticoagulant, the duration of dispensations was estimated using the days supplied and the data-driven approaches. At this stage, no grace period was added, and each individual had two alternative durations of dispensation. Distributions of the estimated durations in days were summarized with means and standard deviations (SD).

The mead duration of follow-up was reported. The incidence rates of major bleeding were estimated using Poisson regression.

Logistic regression models were used to estimate the PS of each patient being treated either with dabigatran vs. warfarin, or rivaroxaban vs. warfarin given age, sex, and clinical characteristics measured at the time of cohort entry and listed earlier (see Section *3.3.7 Propensity scores*).

Next, I assessed the impact of the approach for defining the duration of dispensation on the strength of association between treatment with dabigatran or rivaroxaban versus warfarin and major bleeding using Cox proportional hazards regression models adjusted for the year of cohort entry and deciles of PS. The estimate from this analysis can be interpreted as the relative hazard of major bleeding in individuals taking dabigatran or rivaroxaban versus that in individuals taking warfarin, adjusted for the year of cohort entry and PSs. The analyses were repeated for each grace period. The model building statistics are shown in tables in Annex 10. These tables contain the information on model fit statistics (log-likelihood ratio [-2 LOG L], Akaike Information Criterion [AIC], and Schwarz information criterion [SBC]), Chi-square test of whether the model with the covariate/covariates is better than the model with no variables, and type 3 test of whether there are any differences in event rate across any of the levels of the covariate.

3.4.6 Cox proportional hazard regression

The Cox proportional hazards model or the proportional hazards regression is the commonly used model for censored longitudinal data. It establishes the difference in survival due to treatment options or other prognostic factors[259].

This model was chosen because, in the second and third studies, the outcome of interest was the time between the date of cohort entry (time zero) and the study event. Furthermore, the follow-up was censored in those patients who did not experience the study events, died, or discontinued the oral anticoagulants received at time zero.

The Cox regression gives the estimate of the hazard ratio (HR) and its confidence interval. The hazard is an instantaneous rate of the event of interest at time *t*. For example, in the second study, for each exposure group, the hazard is the instantaneous rate of major bleeding at time *t* in individuals treated with a given anticoagulant. The Cox regression examines the risk sets of individuals that are still being followed at the time of occurrence of each event of interest. At each time of the event, the exposure values are compared in those who experienced the event, and those without the event.

Mathematically, it is expressed as

$$h(t) = h_0(t) * exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p) \quad (8)$$

Where *h(t)* is the hazard at time *t*, $h_0(t)$ is the baseline hazard (the hazard rate for those in whom all covariates have the reference values), at time *t*, and x_1 to x_p are the covariates, including the treatment.

If a study has only one covariate which is a treatment with two exposure categories (treated and untreated), the hazard ratio comparing treated with untreated at time *t* is

$$HR(t) = \frac{h_0(\mathrm{t})\exp\,(\beta_1)}{h_0(\mathrm{t})} = \exp\,(\beta_1) \qquad (9)$$

Cox regression uses the Wald chi-square statistics to test if each explanatory covariate in the model has a significant effect on the outcome variable.

The model assumes that the hazard ratio remains constant over time (proportional hazard assumption). In the second and third studies, this assumption was examined graphically by plotting the cumulative hazard functions of each treatment group against time since the start of follow-up[259]. The cumulative hazard function is the sum of the hazards experienced up to time *t*. The lines of the cumulative hazard functions of treatment groups were reasonably parallel, which signified that the proportional hazard assumption was held.

3.5 Third study

3.5.1 Aim and objectives

The third study aimed to explore the selection bias due to censoring in the DOACs versus warfarin groups, and its impact on the hazard ratio estimate of comparative effectiveness and safety. The objectives were: to describe the associations between the prognostic variables and censoring in the DOACs and warfarin groups; to describe the direction of selection bias due to censoring in the DOACs and warfarin groups; to compare two strategies for estimation of censoring weights: exposure-stratified and exposure-unstratified; and to estimate and compare the hazard ratios of the effect of continuous treatment with warfarin versus DOACs on the composite of stroke, major bleeding, myocardial infarction, and all-cause mortality using exposure-stratified and exposure-unstratified censoring weights.

3.5.2 Case presentation

When comparing the effectiveness and safety of DOACs and warfarin using the per-protocol analytical approach, the follow-up of individuals is censored at the time of deviation from the baseline treatment[260]. However, reasons to deviate from baseline treatment allocation are distinct for DOACs and warfarin initiators. For example, a switch from a DOAC to warfarin may happen as a precaution, in the case of a new-onset renal failure, since DOACs are contraindicated in end-stage renal disease. Renal failure also increases the risks of stroke and bleeding[3,39,261,262]. On the contrary, a switch from warfarin to a DOAC may happen for treatment simplification in individuals free of renal failure. Moreover, at the time of the study, the RAMQ drug insurance plan classified DOACs as exception drugs for patients with AF. DOACs would be covered if patients met one of

two conditions: (1) patients having a history of prior treatment with warfarin but who had failed to maintain the targeted therapeutic range, (2) patients for whom the monitoring of anticoagulation levels with warfarin was not possible or not available. Thus, the intent for clinicians was to attempt warfarin in a short term before switching to DOAC. Because of this differential switching dynamics between treatment groups, during the follow-up, sicker individuals with higher outcome risks are removed from the DOACs group, whereas healthier individuals with lower outcome risks are removed from the warfarin group, resulting in selection bias.

Figure 3.6 presents the different mechanisms of censoring in the DOACs versus the warfarin treatment group when the censoring is applied at the time of the switch from treatment assigned at baseline. In this figure, bleeding is an outcome event, and a new-onset renal failure is a time-varying prognostic factor.

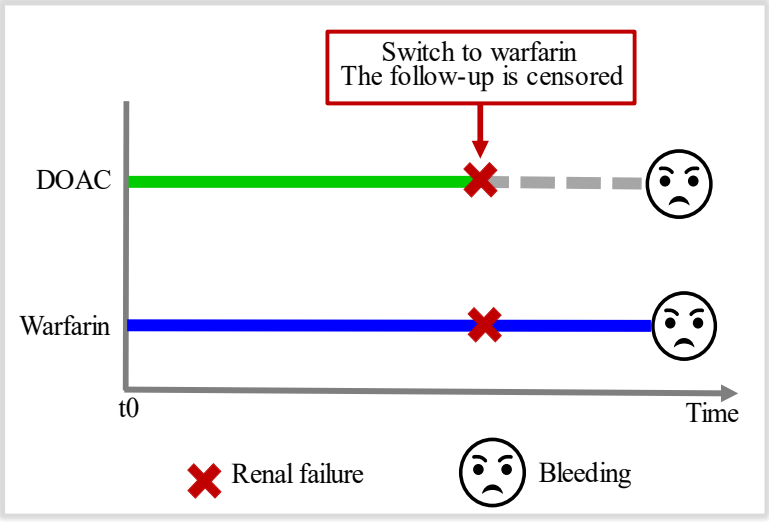


Figure 3.6 Different mechanisms of censoring in the DOACs versus the warfarin treatment group.

The IPCWs may be used to correct this selection bias. In this situation, a model estimating the probabilities of censoring that is adjusted for treatment value, time of follow-up, and all prognostic factors would estimate the average coefficients for prognostic variables across treatment groups. For example, a patient with renal failure would more likely be censored in the DOAC treatment group and less likely be censored in the warfarin treatment group. However, this model would estimate the coefficient for the probability of censoring given renal failure as an average value across two treatment groups. This average would obscure the treatment-specific censoring probabilities. As a result, applying the IPCWs would fail to eliminate the selection bias associated with renal failure. One potential solution for the correct specification of the censoring model in this clinical situation is to stratify it by exposure.

3.5.3 Base cohort definition

Individuals with NVAF newly dispensed an oral anticoagulant were identified as described in Section *3.2 Study population*. The cohort entry date was defined as the date of the first dispensation for an anticoagulant. We followed individuals until the earliest of the following: occurrence of the study outcome, end of pharmaceutical coverage, five years since the initiation of the oral anticoagulant, or end of the study period (March 31st, 2017).

3.5.4 Matched cohort definition

To balance the observed demographic and clinical characteristics at baseline, patients dispensed warfarin were matched 1:1 without replacement to patients dispensed a DOAC on age (±1 year), sex, cohort entry date (±2 years), and PSs (0.2 of the standard deviation of logit of propensity scores) estimated from baseline clinical predictors of treatment censoring listed in Section *3.3.7*

Propensity scores. Those individuals without a match were excluded from the study. The balance of the baseline characteristics in both exposure groups was assessed using standardized mean differences.

The flowchart of patient selection in the Third study is in

Annex **12**.

3.5.5 Definition of exposure

Exposure to DOACs and warfarin was defined based on the pharmaceutical dispensation records. The periods of continuous exposure were identified using dispensation dates, reported days supplied, and a grace period of 30 days.

3.5.6 Outcome definition

The outcome of interest was time to a composite of ischemic or hemorrhagic stroke, systemic embolism, major bleeding (intracranial, gastrointestinal, ocular, and any other bleeding necessitating hospitalization), myocardial infarction, and all-cause mortality.

The composite effectiveness and safety outcome corresponds to our research question. This study aimed to improve the methodology of observational studies comparing DOACs and warfarin. The composite effectiveness and safety outcome was used to make our argument more compelling in making a stronger case for the application of exposure-stratified censoring weights. Having more outcome events improved the precision of the HR estimate. Furthermore, in this study, we hypothesized that because of differential selection bias across the treatment groups, each component of the composite outcome may be biased in favor of DOACs.

However, there are some disadvantages when using the composite outcome. In general, in studies on comparative effectiveness and safety, the study conclusions may be misleading, and demonstrated treatment benefits may be assumed to relate to all components when, in fact, only a few events of an individual component may contribute to the composite outcome. This may be more starkly if the components include the most important clinical events (e.g., myocardial infarction, a serious bleeding event requiring hospitalization, or death) and less serious components (e.g., bleeding not requiring hospitalization). To communicate the study findings in a publication more clearly, it is recommended to present data for all individual components[263].

In our research, each component of the composite outcome was assessed in a previous study conducted by our research group (Durand, *CMAJ*, 2020). This study on the comparative effectiveness and safety of DOACs versus warfarin used exposure-stratified censoring weights. The study found no difference in the incidence of ischemic stroke or systemic embolization in patients with NVAF (HR: 0.98, 95%CI: 0.77-1.25), however, the use of DOACs was associated with fewer major bleeding events in patients treated with DOACs versus those treated with warfarin (HR: 0.71, 95%CI: 0.64-0.80)[264].

Stroke, systemic embolism, and major bleeding events were identified using the International Classification of Diseases (10th revision) codes from hospital discharge summaries using primary or principal diagnoses, or the cause of in-hospital death. The admission date to the hospital was assigned as the date of the outcome event. The date of death was obtained from the RAMQ *registry of beneficiaries' database* . The list of ICD codes is in Annex 3.

3.5.7 Covariates

We collected information on the potential prognostic factors that are listed in Section *3.3.7 Propensity scores*. Our team assumed that these factors may have an impact on initiating or stopping warfarin or a DOAC and may increase the risk of the components of the composite outcome. These prognostic factors were identified using expert knowledge and information from previously published studies on the comparative effectiveness and safety of warfarin and DOACs.

3.5.8 Analytical approach

Our goal was to emulate a target trial where study participants who would be prescribed warfarin under usual care, and were eligible for DOACs and warfarin treatment, would be randomized to receive either warfarin or a DOAC. All subjects would continue their initial treatment for the entire follow-up (per-protocol analysis[169,265]). To balance the observed demographic and clinical characteristics at baseline, patients dispensed warfarin were matched in a 1:1 ratio without replacement to patients dispensed a DOAC on age (±1 year), sex, cohort entry date (±2 years), and propensity scores (0.2 of the standard deviation of logit of propensity scores) estimated from baseline clinical predictors of treatment/censoring defined above[213]. The follow-up was right censored at the time of deviation from the baseline treatment allocation, either due to a switch to an alternative anticoagulant or treatment cessation for any reason.

I conducted three analyses. In the first analysis, censoring due to deviation from baseline treatment was considered non-informative or completely at random and was not adjusted for. In the second analysis, this censoring was considered informative with similar associations between the prognostic factors and the probability of censoring in individuals treated with warfarin and DOACs.

We adjusted for this censoring with the IPCWs calculated with a common censoring model, unstratified by exposure[165]. In the third analysis, we still considered censoring due to deviation from baseline treatment to be informative, however, the dynamic of censoring differed between treatment groups (i.e., different associations between the prognostic factors and the probability of censoring in individuals treated with warfarin and DOACs). Accordingly, for the third analysis, we built IPCWs using exposure-stratified censoring models (i.e.: weights were built separately for patients on warfarin and those on DOACs).

In a sensitivity analysis, I further explored if the mechanism of censoring is different between those who switched between treatment groups and those who permanently discontinued treatment with oral anticoagulants. For this analysis, censoring due to deviation from baseline treatment was considered informative, with a different dynamic of censoring between treatment groups. IPCWs were estimated using exposure-stratified censoring models. Furthermore, for each treatment group, two sets of IPCWs were estimated, (1) to account for censoring due to switching between treatment groups, and (2) to account for censoring due to permanent discontinuation of treatment.

We did not include in the censoring model variables for health conditions when DOACs became absolutely contraindicated (i.e., initiation of hemodialysis in end-stage renal disease, heart valve surgery). This was done to avoid the violation of the positivity assumption when estimating censoring weights. The positivity assumption states that at any time t, the study individuals with any combination of covariates have a greater than zero probability of not being censored. However, for patients in the DOACs treatment group who developed end-stage renal disease or underwent heart valve surgery, the probability of not being censored (i.e. the probability of continuing

treatment with a DOAC) would be zero. Instead, in both treatment groups, the follow-up was right-censored at the date when these health conditions were recorded. This censoring was assumed as being at random conditional on covariates as both the warfarin- and the DOACs-treated individuals were expected to have equal probabilities of developing these health conditions after the baseline matching. Table 6.2 (Chapter 6) shows the distribution of baseline patient characteristics in warfarin and DOACs users before and after matching. Before matching, compared to patients on DOACs, patients initiating warfarin were frailer and had a worse prognosis for stroke and bleeding. Matching allowed us to balance the prognostic characteristics between the two treatment groups and reduce the baseline confounding. Among others, we balanced two treatment groups for hypertension, diabetes, and chronic renal disease, the most common causes of end-stage renal disease. Furthermore, in the design phase, we exclude patients with recorded valvular heart disease. For these reasons, we assumed that the risk of development of the conditions when DOACs become contraindicated during the follow-up is similar between the treatment groups.

Another variable that we did not include in the model for the estimation of censoring weights is non-major bleeding defined as bleeding not requiring hospitalization. Treatment discontinuation and hence censoring due to events of non-major bleeding may be considered informative, however, this censoring should be non-differential across the treatment groups, i.e., patients in both treatment groups (DOACs and warfarin) would discontinue treatment after non-major bleeding. Furthermore, identifying non-major bleeding from physician claims may lead to a detection bias. Regular outpatient visits of individuals treated with warfarin may increase the likelihood of non-major bleeding being recorded in the physician claims database.

3.5.9 Statistical analysis

We used descriptive statistics to summarize the patients' demographics, comorbidity profile, and use of medications measured at the time of cohort entry. The incidence rate of composite effectiveness and safety outcome was estimated using Poisson regression.

Propensity score. A logistic regression model was used to estimate the baseline probabilities of each patient being treated with warfarin vs DOACs, given all covariates stated above (see Section 3.3.7 for the description of PS). Propensity scores were used to create the matched study cohort. Descriptive statistics were used to summarize distributions of patients' characteristics at cohort entry. A characteristic was considered as balanced across exposure groups with ≤10 % absolute standardized difference[266].

Probability of being censored due to deviation from baseline treatment and IPCWs. To estimate the probabilities of being censored, I structured the data in the "counting process" style with multiple rows of data per subject. Thus, for each study individual, starting from the time of cohort entry (*t0*), I split the follow-up into 30-day intervals (months, hereafter), and measured the changes in the values of covariates (defined in Section 3.3.7) as follows: covariates were measured within the year before the cohort entry date, and then updated at the beginning of every month of follow-up. Once a medical condition occurred, it remained present for the remainder of the follow-up time. For exposure to medication, current use was determined, and patients were considered as exposed starting from the date of medication dispensation until the end of the dispensation duration, plus a 30-day grace period. In addition, for each month, I put an indicator of censoring: i.e., where the follow-up of an individual was censored because of deviation from the baseline treatment or

this individual continued the treatment assigned at baseline (censored=1 or censored=0, respectively). An individual was considered to be "at-risk" at a given month (time t) if he hadn't yet had the outcome event nor been censored in the previous month (*time t-1*). Every at-risk patient at each month was described as a patient-month.

Figure 3.7 shows an example of the data format. For each individual, the number of rows was equal to the number of months of follow-up, i.e., one row corresponding to each month (t) with the updated information on the time-dependent covariates, censoring, and outcome. Thus, for a patient with an individual identifier 1 (ID=1), the follow-up time is split into four months. The months of follow-up are labeled t_1, t_2, t_3, t_4. Next, *t-1* refers to the time point a month before t. For example, if $t=t_2$, then $t\text{-}1=t_1$.

Obs	ID	t	Exposure	Year of cohort entry	Outcome	Censored	Covariate 1	Covariate 2	Covariate N
1	1	1	doac	2012	0	0	0	0	0
2	1	2	doac	2012	0	0	1	0	0
3	1	3	doac	2012	0	0	1	0	0
4	1	4	doac	2012	1	0	1	0	0
5	2	1	warfarin	2011	0	0	0	0	0
6	2	2	warfarin	2011	0	1	0	0	0
7	3	1	doac	2013	0	0	0	0	0
8	3	2	doac	2013	0	1	0	0	0
9	4	1	doac	2013	0	0	0	0	0
10	4	2	doac	2013	0	0	1	0	0
11	5	1	warfarin	2012	0	0	0	0	0
12	5	2	warfarin	2012	0	0	0	0	0
13	5	3	warfarin	2012	0	0	0	0	0
14	5	4	warfarin	2012	0	0	0	1	0
15	6	1	doac	2011	0	0	0	0	0
16	6	2	doac	2011	0	0	0	0	0
17	6	3	doac	2011	0	0	0	0	1
18	7	1	warfarin	2013	0	0	0	0	0
19	7	2	warfarin	2013	0	0	0	0	0
20	7	3	warfarin	2013	0	0	0	0	1

Figure 3.7 An example of data format for estimation of the probabilities of remaining uncensored.

For every month, I fitted a pooled logistic regression to estimate the probabilities of being censored given sex, age at baseline, month (*t*), and the vector of covariates with the values updated for the previous month (*t-1*). Thus, in Figure 3.7, the individual with ID=1would have four estimated probabilities of being censored for every *t*.

In this study, the assumption was that the probability of censoring given the covariates must be constant over the month. This assumption seemed to be reasonable as the covariates could only be updated when someone comes in contact with the healthcare system.

Two sets of probabilities of censoring were estimated:

1. With an unstratified model:

```
PROC LOGISTIC data=dataset;
MODEL censored=t exposure age sex covariate1 covariate2 … covariateN;
RUN;
```

2. With exposure-stratified models:

```
PROC LOGISTIC data=dataset (where=(exposure="warfarin"));
MODEL censored=t age sex covariate1 covariate2 … covariateN;
RUN;
PROC LOGISTIC data=dataset (where=(exposure="doac"));
MODEL censored=t age sex covariate1 covariate2 … covariateN;
RUN;
```

Thus, the unstratified model produced average estimates of probabilities of censoring for prognostic variables across treatment groups, whereas the exposure-stratified model produced exposure-specific estimates of probabilities of censoring. The output from these models also contained the estimates of the odds ratios (OR) with 95% confidence intervals (CI) measuring the association of each covariate with the probability of censoring. For each censoring model (unstratified and exposure-stratified), the ORs (95%CI) were presented in a table (Table 6.3,

Chapter 6:). The model fit statistics, test for the global null hypothesis, and Wald chi-square tests for each covariate are shown in Annex 14.

In a sensitivity analysis, for each treatment group, two probabilities of censoring were estimated, (1) the probability of censoring due to switching between treatment groups, and (2) the probability of censoring due to permanent treatment discontinuation:

```
PROC LOGISTIC data=dataset (where=(exposure="warfarin") and (censoring_discontin=0));
MODEL censored_switch=t age sex covariate1 covariate2 … covariateN;
RUN;

PROC LOGISTIC data=dataset (where=(exposure="warfarin") and (censoring_switch=0));
MODEL censored_ discontin =t age sex covariate1 covariate2 … covariateN;
RUN;

PROC LOGISTIC data=dataset (where=(exposure="doac") and (censoring_discontin=0));
MODEL censored_switch =t age sex covariate1 covariate2 … covariateN;
RUN;
PROC LOGISTIC data=dataset (where=(exposure="doac") and and (censoring_switch=0));
MODEL censored_ discontin =t age sex covariate1 covariate2 … covariateN;
RUN;
```

The ORs (and 95% CI) measuring the association of each covariate with the probability of censoring due to switch or treatment discontinuation are presented in Annex 15.

The next step was to estimate the probabilities of remaining uncensored which are:

1-probabilities of censoring

Finally, I estimated the IPCWs as the inverse of the product of the probabilities of remaining uncensored up to and including the month of follow-up[151].

Thus, in Figure 3.7, for the individual with ID=1, the IPCW for $t=3$ is estimated as

$$IPCW_{t=3} = \frac{1}{Pr\ uncensored_{t=1} * Pr\ uncensored_{t=2} * Pr\ uncensored_{t=3}} \quad (10)$$

Accordingly, I estimated unstratified and exposure-stratified IPCWs.

In a sensitivity analysis, IPCWs were estimated for each treatment group as a product of two censoring weights (censoring due to switching and censoring due to permanent treatment discontinuation)[264].

To reduce the influence of extreme values of the estimated IPCWs on the size of the standard error for the hazard ratio estimate, the weights were stabilized. Thus, in the equation above, 1 in the numerator was substituted by the conditional probability of not being censored given t and treatment group for exposure-unstratified IPCWs, and only t for unstratified IPCWs[267].

Thus, the numerator was estimated for unstratified and exposure-stratified models as follows:

1. Unstratified model:

```
PROC LOGISTIC data=dataset;
MODEL censored=t exposure;
RUN;
```

2. With exposure-stratified models:

```
PROC LOGISTIC data=dataset (where=(exposure="warfarin"));
MODEL censored=t;
RUN;
PROC LOGISTIC data=dataset (where=(exposure="doac"));
MODEL censored=t;
RUN;
```

For sensitivity analysis:

```
PROC LOGISTIC data=dataset (where=(exposure="warfarin") and (censoring_discontin=0));
MODEL censored_switch=t;
RUN;
PROC LOGISTIC data=dataset (where=(exposure="warfarin") and (censoring_switch=0));
MODEL censored_ discontin =t;
RUN;

PROC LOGISTIC data=dataset (where=(exposure="doac") and (censoring_discontin=0));
MODEL censored_switch =t;
```

```
RUN;
PROC LOGISTIC data=dataset (where=(exposure="doac") and and (censoring_switch=0));
MODEL censored_ discontin =t;
RUN;
```

To check if the selection of the model for the numerator was correct, I run the outcome model weighted with non-stabilized and then, with stabilized IPCWs. The hazard ratio estimates from the models adjusted with non-stabilized and stabilized weights provided identical point estimates aside from the larger standard errors for the analyses with non-stabilized weights.

The distribution of stabilized IPCWs estimated with unstratified and exposure-stratified models was presented using box plots. Truncation of censoring weight was not used because of no extremely large or small values (Figure 6.3, Chapter 6:). The smallest value was 0.37, and the largest was 29.5.

Applying the IPCWs creates a pseudo population resulting from re-weighting the study subjects based on their IPCW values. The re-weighting ensures that covariates are balanced across those censored and those who continue in the study. For example, if participants with a certain prognostic factor are more likely to be censored, remaining participants with the same prognostic factor receive greater weight in the analysis, thereby compensating for those censored. However, the pseudo-populations being created with exposure-stratified versus unstratified IPCWs would differ. Thus, the pseudo-population created by unstratified IPCWs ensures the balance of those predictors of censoring that have a similar effect on censoring in both DOACs and warfarin treatment groups, whereas, exposure-stratified IPCWs would create a pseudo-population for each treatment group. This creates a balance of both types of predictors of censoring, those with a similar effect on

censoring across treatment groups, as well as those with a different effect on censoring in DOACs versus warfarin-treated individuals.

Aside from the correct specification of the censoring model that I explored in the third study, other assumptions are required for the IPCWs to create a pseudo-population free of selection bias introduced by informative censoring[163]. First, the non-testable ignorability assumption, i.e. we assume that in the current month, censoring occurs at random given the probability of remaining uncensored up to the current month, sex, age, treatment assigned at baseline, and the observed values of covariates in the previous month. Second, the positivity assumption, that is we assume that at any given month, the study individuals with any combination of the observed covariates in the previous month, age, sex, and treatment assigned at baseline have a greater than zero probability of not being censored. I tested the positivity assumption by examining the distributions of the unstratified and exposure stratified IPCWs[166].

Outcome model. The outcome model estimated the risk of the composite outcome in warfarin vs DOACs-treated individuals. For this, I fitted three pooled logistic regression models that approximated Cox proportional hazards regression models[268] with the treatment as the only independent variable.

Thus, with the first unweighted model, I used all at-risk subjects at time t to fit a pooled logistic regression model. Thus, the logistic regression model was fitted for every month of follow-up, and then the results were pooled across all individuals and all months of follow-up. The only covariate in this model was the indicator for the treatment group. In my analysis, the pooled logistic regression approximated a Marginal Structural Cox model as the probability of the outcome event

at any given time point was low. The exponential of the coefficient of treatment with DOACs could be interpreted as the marginal hazards ratio relative to treatment with warfarin.

The second and third models were models weighted with the unstratified IPCWs and the exposure-stratified IPCWs, respectively. Again, the logistic regression model was fitted for every month of follow-up and weighted with the IPCWs estimated for this particular month. The results were pooled across all individuals and all months of follow-up. In the weighted models, the variance of the estimated coefficient for treatment was approximated using a robust variance estimator that takes into account the weights and repeated measures[269].

CHAPTER 4: Methodological challenges in assessment of current use of warfarin among patients with atrial fibrillation using dispensation data from administrative healthcare databases[270].

Sinyavskaya L, Matteau A, Johnson S, Durand M. *Pharmacoepidemiology and Drug Safety* 2018;27(9):979-986.

4.1 Abstract

Purpose: Algorithms to define current exposure to warfarin using administrative data may be imprecise. The study objectives were to characterize dispensation patterns and to measure gaps between expected and observed refill dates for warfarin and direct oral anticoagulants (DOACs).

Methods: Retrospective cohort study using administrative healthcare databases of the Régie de l'assurance-maladie du Québec. We identified every dispensation of warfarin, dabigatran, rivaroxaban, or apixaban for patients with AF initiating oral anticoagulants between 2010 and 2015. For each dispensation, we extracted the date and duration. Refill gaps were calculated as the difference between expected and observed dates of successive dispensation. Refill gaps were summarized using descriptive statistics. To account for repeated observations nested within patients and to assess the components of variance of refill gaps, we used unconditional multilevel linear models.

Results: We identified 61,516 new users. The majority were prescribed warfarin (60.3%), followed by rivaroxaban (16.4%), dabigatran (14.5%), apixaban (8.8%). The most frequent recorded duration of dispensation was seven days, suggesting the use of pharmacist-prepared weekly pillboxes. The average refill gap from the multilevel model was higher for warfarin (9.28 days, 95%CI:8.97-9.59) compared to DOACs (apixaban 3.08 days, 95%CI: 2.96--3.20, dabigatran 3.70, 95%CI: 3.56-3.84, rivaroxaban 3.15, 95%CI: 3.03-3.27). The variance of refill gaps was greater among warfarin users than among DOAC users.

Conclusions: The greater refill gaps for warfarin may reflect inadequate capture of the period covered by the number of dispensed pills recorded in administrative data. A time-dependent

definition of exposure using dispensation data would lead to greater misclassification of warfarin than DOACs use.

4.2 Introduction

Vitamin K inhibitor warfarin decreases the risk of stroke and other systemic thromboembolism in patients with atrial fibrillation (AF)[50-52]. Clinical use of warfarin is fraught with unpredictable individual dosage, the need for frequent monitoring of the International Normalized Ratio (INR), and numerous interactions with food and other drugs[271]. Therefore, the dosage of warfarin is frequently adjusted by physicians over the treatment course without communicating with community pharmacists dispensing the drug.

Administrative health databases containing drug dispensation data are commonly used to study population-level adherence, and to compare the safety and effectiveness of direct oral anticoagulants (DOACs) and warfarin in real-world settings[272]. Exposure is defined based on dispensing information recorded by community pharmacists. A period of continuous exposure is assumed when gaps between subsequent dispensations do not exceed an investigator-defined grace period, which accounts for irregular refill patterns[90,114,248,273,274]. However, in the case of warfarin, due to individually adaptive dosage adjustments over a single dispensation period, the dispensed drug supply may last well beyond the duration recorded by the dispensing pharmacist (e.g., because of pill-splitting or/and stockpiling). Thus, the resulting gap between observed and expected dates of refill is due to a discrepancy between the posology assumed by the dispensing pharmacist and that truly taken by the patient. Furthermore, refill gaps may vary between patients, as well as within

individual patients over the treatment course. These may lead to exposure misclassification which is likely to be a differential between warfarin and DOACs, as DOACs do not require dose adjustments or monitoring. Though there is an algorithm that uses linked data on INR measurement to bridge the gaps between warfarin prescriptions[48], it cannot be used in some settings, including the Québec administrative healthcare databases, in which information on laboratory test results or INR testing is lacking.

We hypothesized that in the Québec administrative healthcare databases, refill gaps and their variation are larger for warfarin than for DOACs. Relying on dispensation data to define current exposure to anticoagulants will lead to greater misclassification of exposure for warfarin than for DOACs. We sought to characterize the dispensation patterns of warfarin, apixaban, dabigatran, and rivaroxaban. We also aimed to compare the variation of refill gaps between users of different oral anticoagulants, as well as within users of the same oral anticoagulant over the treatment course.

4.3 Methods

Data source

The study was conducted using computerized healthcare databases of the *Régie de l'assurance-maladie du Québec* (RAMQ) and the *Maintenance et exploitation des données pour l'étude de la clientèle hospitalière* (MED-ÉCHO)[183]. The RAMQ administers the government health insurance plan for all Québec residents, and prescription drug insurance plan for individuals of 65 years and older, welfare recipients, and those not covered by private medical insurance (in total, about 40% of Québec's population)[183,184]. It routinely collects diagnostic and medical information for

physician and hospital reimbursement, and monitoring and management of healthcare programs. The RAMQ maintains four medical databases. The *registry of beneficiaries' database* contains socio-demographic information for all recipients of the health insurance plan. The *medical services database* with fee-for-service physician claims for ambulatory and hospitals services. The *MED-ÉCHO database* collects discharge summaries following any hospitalization. The *prescription database* includes insurance plan details, period of coverage, date of prescription, drug identification number (DIN), drug class (coded by ASHP, American Hospital Formulary Service), generic name, formulation, prescribed dosage, quantity, duration of treatment, gross cost and contribution of the insured person. Of note, in Québec, all prescription drugs are paid for by the public drug insurance plan. The plan imposes no fixed co-pays, and the contribution paid by patients (deductibles and co-insurance) is determined by drug cost and patient income. Therefore, data completeness for warfarin and DOACs dispensed at community pharmacies is close to 100%, with small exceptions being drugs dispensed while traveling outside of the province. However, the RAMQ prescription database does not contain information on treatments provided in hospital settings.

Base cohort definition

All new users of oral anticoagulants between 1 January 2010 and 15 March 2015 were identified as those with no dispensation of any oral anticoagulant in the preceding year. Patients had to be aged at least 18 years and have a diagnostic code of AF or atrial flutter (ICD-9 427.3, ICD-10 I48.x) within a year before initiation of the oral anticoagulant. We excluded those with a diagnosis of venous thromboembolism in the 90 days before the first dispensation. We followed patients

from the date of the first dispensation until the last dispensation, end of pharmaceutical coverage, or end of the study period (March 31, 2015), whichever occurred first. Patients were grouped into four cohorts based on the type of prescribed oral anticoagulant (warfarin, dabigatran, apixaban, or rivaroxaban). Those who switched between anticoagulants during the follow-up were considered as starting a new treatment episode and were re-assigned to the corresponding cohort at the time of each switch.

Matched cohort definition

To estimate the variation of refill gaps, patients who had at least three consecutive dispensations of the same oral anticoagulant were eligible. Patients receiving warfarin and each of the three DOACs were matched through a common-referent group, a methodology described elsewhere[209]. Briefly, patients taking apixaban were matched in a 1:1 ratio with users of dabigatran, rivaroxaban, and warfarin, and only those with an available match for all three drugs were retained. Matching was done based on age, sex, and propensity score (PS) calculated from the variables listed in Table 4.2.

Subgroups of weekly pillboxes users: In Québec, to support patients in their adherence to the treatment, community pharmacies offer sealed pillboxes of the prescribed medication for a seven-day period[210]. The 7-day supply of medications requiring dose adjustments is provided in a plastic vial[210]. In this text, the term “pillbox” will be used to refer to all 7-day drug supplies.

Pillboxes are automatically prepared and often delivered at home without the need for the patient to request dispensation. As this practice can influence treatment adherence, in a secondary analysis

we stratified the cohorts according to weekly pillbox use which was defined as having had at least four successive dispensations of seven-day duration.

Exposure definition

We identified all dispensations of warfarin, dabigatran, rivaroxaban, and apixaban. For each dispensation, we extracted the dispensing date, DIN, and the number of days supplied dispensed.

Refill gap definition

For each patient, all dispensations for a given oral anticoagulant were ordered chronologically. Refill gaps between two successive dispensations were calculated as the absolute number of days between the observed and expected refill date (identified from days supplied) of the next dispensation. Refill dates were adjusted accordingly for identified hospitalization periods (e.g.: if a patient was hospitalized for eight days following a dispensation with 30 days supplied, the expected day of refill was calculated as days supplied (30) + days in the hospital (8) = 38 days following the dispensation).

Interpretation of refill gaps. Refill gaps are often used to describe treatment adherence[211]. This assumes that the number of dispensed pills is adequate to cover the period indicated by the dispensing pharmacist using the day supplied field. In a situation with complete adherence, when, for example, a patient was dispensed 30 pills for 30 days and returns for a refill on the 31st day, the refill gap is zero. Alternatively, in a situation of non-adherence, the patient takes a pill every other day only, and returns for a refill on day 61st, the resulting refill gap is 30 days.

Additionally, refill gaps may capture deviations from prescribed posology that are unknown to the dispensing pharmacist, such as in the case of dosage adjustments of warfarin. For example, a patient with a one-pill-daily regimen is given 30 pills by his pharmacist who enters 30 days supplied in the dispensation database. The following day (for simplicity of calculations), the physician instructs the patient to take half a pill every day. This patient, even completely adherent, has pills for 60 days and returns on day 61 resulting in a refill gap of 30 days. Thus, the refill gaps capture information on both adherence and adequacy of the days supplied data. Under the assumption of similar adherence between warfarin and DOACs at a population level, any significant increase in refill gaps for a given medication would indicate that days supplied data is an imprecise measure of the actual number of days of availability of that drug.

Statistical analysis

Patients from the base cohort were characterized by age, sex, comorbidity profile, and use of medications measured at the time of cohort entry. For the days supplied values, the dispenced quantity, strength of the dispenced drug, the calculated number of days between two consecutive prescriptions, and the refill gaps, we estimated means with standard deviation (SD), medians with interquartile range (IQR), modes, and minimum and maximum values, or the number of dispensations with percentage.

Constitution of matched cohorts. Logistic regression was used to calculate the PS at baseline, or each time a patient switched oral anticoagulants. The distributions of PSs estimated for patients treated with apixaban and dabigatran; apixaban and rivaroxaban; and apixaban and warfarin are in

Annex **6**.

Greedy matching[212] was performed on age (±1year), sex, and PS (using a caliper of 0.2 standard deviations of the logit of PS).

To account for the clustered structure of the data (multiple prescriptions per patient) and to estimate the variation of refill-gaps, we used multilevel linear regression models for the matched cohorts to produce estimates of the cluster-adjusted average of refill gaps (fixed effect) and the variance of refill gaps partitioned on within- and between-patient components (random effects)[214]. From the variance components, we calculated the intraclass correlation (ICC) that assesses the clustering, or similarity of refill gap values in the same patient[215]. We repeated the multi-level analysis, stratified according to the use of pharmacist-prepared weekly pillboxes.

As warfarin dosage is expected to be adjusted more frequently during the initial dose-finding, we examined the longitudinal change of refill-gaps and their variation in the warfarin cohort with an unconditional growth model, using an ordered dispensation number as a temporal predictor.

4.4 Results

We included 61,516 new users of oral anticoagulants in the study (Table 4.1). At baseline, most patients were prescribed warfarin (60.3%), followed by rivaroxaban (16.4%), dabigatran (14.5%), and apixaban (8.8%) (Table 4.2). Median age varied from 75 years (IQR: 69-81) in rivaroxaban to 78 (IQR: 72-85) in apixaban and warfarin patients, with equal proportions of males and females among all types of oral anticoagulants (Table 4.2). A higher proportion of warfarin compared to DOACs patients had comorbidities, including hypertension, cardio-and cerebra-vascular diseases,

diabetes, kidney injury, and chronic renal failure. Warfarin patients more often used antiplatelet agents. Among new users of DOAC, a higher proportion of those dispensed apixaban had comorbidities than those dispensed dabigatran or rivaroxaban.

After the base cohort entry, 13,472 (22%) of 61,516 patients switched at least once to another type of oral anticoagulant (Annex 5). Overall, over the follow-up, there were 17,214 switches.

For more than half of all dispensations, the reported number of days supplied was seven days (Table 4.3). Apixaban 5 mg accounted for 66% of all apixaban dispensations, dabigatran 150 mg for 47% of all dabigatran dispensations, and rivaroxaban 20 mg for 26% of all rivaroxaban dispensations.

For most dispensations, there were no gaps in medication supply (mode=0). The mean refill gap value was longer for warfarin compared to DOACs.

After matching, users of different oral anticoagulants had similar distributions by age, sex, comorbidities, and use of other medications (Table 4.4). Median follow-up was 0.86 years for apixaban, 2.1 years for dabigatran, 1.4 years for rivaroxaban, and 1.5 years for warfarin cohorts.

Cluster-adjusted average refill gap in days (the true average, once clustering at the patient level is appropriately accounted for) was higher for warfarin (9.28, 95%CI:8.97-9.59) compared to apixaban (3.08, 95%CI:2.96--3.20), dabigatran (3.70, 95%CI:3.56-3.84), and rivaroxaban (3.15, 95%CI:3.03-3.27) (Table 4.5). Clustering of refill gaps within patients was higher for warfarin (ICC: 0.67) compared with apixaban (ICC:0.48), dabigatran (ICC:0.53), and rivaroxaban (ICC:0.38). This can be interpreted as follows: variation of refill gap values is greater between any two warfarin users than between any two DOAC users.

Refill gaps in patients on warfarin were longer at treatment initiation and decreased over the follow-up, but this effect was very modest: on average, the refill gap was 0.72 days shorter for every 10 dispensations (Table 4.6).

Stratification by use of weekly pillboxes: Average length of refill gaps considerably differed between users and non-users of pharmacist-prepared weekly pillboxes (Table 4.7). Pillbox users of all oral anticoagulants had the lowest averages of refill gaps: warfarin 1.80 (95%CI:1.68-1.92), apixaban 0.61 (95%CI:0.57-0.65), dabigatran 0.73 (95%CI:0.69-0.77), and rivaroxaban 0.71 (95%CI:0.65-0.77) days. On the other hand, in non-users of pharmacist-prepared weekly pillboxes, the difference in mean refill gaps was even more striking between warfarin (12.56, 95%CI:12.17-12.95) and apixaban (4.28, 95%CI:4.12-4.44), dabigatran (4.98, 95%CI:4.80-5.16), or rivaroxaban (4.25, 95%CI:4.09-4.41).

4.5 Discussion

In this large, population-based study based on the analysis of administrative pharmacy records, we found that the gaps between observed and expected dates of prescription refill were more than twice as long for warfarin than for DOACs. Moreover, compared to a DOAC user, a patient taking warfarin had a higher variation of refill gaps over the follow-up, likely, due to frequent dosage adjustments, particularly after treatment initiation. Furthermore, compared to DOACs, there was also a higher variation of refill gaps among patients taking warfarin, likely, indicating different dosing regimens between patients.

With administrative data, exposure to oral anticoagulants is often assessed using refill dates, days of medication supply, and allowing for a grace period between dispensations, which varies by studies from three[90] to 60[248] days. The relatively rapid onset of action of oral anticoagulants and the reversal of their effects upon discontinuation (both for effectiveness and safety outcomes) make time-dependent exposure definitions desirable for safety and effectiveness research. Ideally, a patient should be classified as exposed starting on the day the drug is dispensed, and for the continuity of the days supplied, plus a certain grace period of a few days that should be chosen based on the individual drug pharmacokinetics and possible dosage adjustment. Selection of the grace period requires careful clinical consideration to account for the pharmacologically pertinent period of observation or a precise time when a drug under the study is active in the body[275].

Refill gaps reflect not only adherence but also the discrepancy between the prescribed posology registered by the dispensing pharmacist and the true posology due to dosage adjustments by the treating physician. In clinical practice, there are frequent warfarin dosage adjustments between two dispensations that are based on INR monitoring. Our findings suggest that under the reasonable assumption of similar adherence between warfarin and DOACs users at the population level[276], the high variation in refill gaps for warfarin is due to inadequate capture in the dispensation data of the periods covered by the number of dispensed pills. This can be explained by the lack of communication between the treating physician, who adjusts warfarin, and the dispensing pharmacists. Moreover, at pharmacist-managed warfarin administration services, when the exact dose of warfarin is known, pharmacists tend to record a shorter duration of treatment than what is planned. This is done to avoid issues with the prescription renewal, as the early claim (if a dose of

warfarin is increased) will not be accepted by the pharmacy system (C. Sirois, personal communication, June 20, 2022).

Of importance to pharmacoepidemiological studies of oral anticoagulants, our findings indicate that universal application of the same grace period to warfarin and DOACs will lead to misclassification of exposure. As this degree of misclassification will be differential, it may bias estimates in unpredictable directions. For example, a grace period of three days may correctly approximate exposure to DOACs use, however, many patients taking warfarin would be erroneously considered unexposed. While a longer grace period (for example 30 days) may be more appropriate for warfarin, it would lead to greater misclassification for DOACs, which have short average refill gaps. To correctly assign a continuous exposure period of warfarin and DOACs, a grace period defined as a percentage of the duration of precedent dispensation could be more accurate as opposed to a fixed grace period. However, the data presented here does not allow to recommend any grace period in particular, whether relative or absolute. More research is needed to determine the appropriate strategy. When comparing the safety and effectiveness of DOACs vs warfarin, one should be aware of differential data limitations in dispensing data, and test several grace periods in sensitivity analyses to assess the impact of exposure misclassification.

We also found that patients classified as users of pharmacist-prepared weekly pillboxes had markedly shorter refill gaps and their variation within and between patients compared to non-users. This supports previous reports that pharmacist-prepared pillboxes may improve patients' adherence to a therapy[277]. The results seen in patients not using weekly-prepared pillboxes reinforce those seen in our entire population: warfarin users have longer refill gaps, and more

variation in refill gaps, than DOAC users. In a previous study, adherence to DOACs was found superior to that of warfarin, using a refill gap of >89 days. Those results may be explained by inadequate interpretation of refill gaps rather than by a true difference in adherence[278].

Our study has some limitations. First, reporting errors in pharmacy claims may result in inaccurate estimates of refill gaps. However, we believe that reporting errors are non-differential among users of different oral anticoagulants. Second, the use of pharmacy-prepared weekly pillboxes may explain the absence of gaps between most of the dispensations. However, this does not imply good adherence: pharmacies renew pillboxes automatically if the patient picks up them regularly even without consuming the previous drug supply. Third, we assumed that population-level adherence is similar between DOACs and warfarin, and that refill gaps variation is explained by inadequate capture of the true medication regimen, but similar results would be found if adherence to warfarin was lower than adherence to DOACs. However, it has been proposed that, on the contrary, routine measures of INR and monitoring visits may favor adherence to treatment[279]. Nevertheless, we acknowledge that differential adherence between oral anticoagulants would influence the results of our study, but believe it is unlikely to completely explain our findings. Fourth, a different definition of the use of weekly-prepared pillboxes may change the results for this population. Finally, in Québec, the public drug insurance program limits the possibility for low-cost drugs, like warfarin, to be filled outside of the RAMQ prescription database. However, this possibility exists for persons traveling out of the province. This may limit the completeness of our data, but its frequency is likely negligible.

Our study has several strengths. First, it is based on the prospectively collected longitudinal population-based data from the large provincial database, thus avoiding potential selection bias. Second, our source population was all individuals eligible for the Québec public drug program, including elderly and welfare recipients (40% of the Québec population[183]) making our results generalizable to a similar population. Third, because of the linkage capacities with other provincial databases, we characterized our patients in terms of comorbidities profile and medication use. The information on hospitalizations was available, allowing us to extract the periods when a patient was receiving medication from the hospital pharmacy.

In conclusion, we found that refill gaps and their variation were greater for warfarin than for other oral anticoagulants. This suggests an inadequate capture of the period covered by the number of dispensed pills for warfarin. An algorithm defining the current use of oral anticoagulants using dispensation data would lead to greater misclassification for warfarin than other oral anticoagulants and cause bias in unpredictable directions. This should also be considered when assessing adherence to warfarin.

4.6 Tables

Table 4.1. Flow of study participants for the first study. Québec, 2010-2015.

	Apixaban	Dabigatran	Rivaroxaban	Warfarin	Total
Total number of new users of oral anticoagulants between 01/01/2010 and 03/15/2015	6,320	10,894	33,670	74,272	125,156
Patients with AF or atrial flutter	5,531	9,093	10,633	41,769	67,026
Excluded those with VTE 90 days before the date of cohort entry	99	146	565	4,703	5,513
New users of oral anticoagulants in the base cohorts	5,432	8,947	10,071	37,066	61,516
Total number of oral anticoagulant treatment episodes in the base cohort over the follow-up[†]	9,042	14,135	15,618	39,575	78,730
Number of matched treatment episodes	7,376	7,376	7,376	7,376	29,504

Abbreviations: VTE, venous thromboembolism

[†]Each treatment switch between oral anticoagulants defines a new treatment episode

Table 4.2. Baseline characteristics of the new users of oral anticoagulants in the base cohort of the first study. N=61,516, Québec, 2010-2015.

	Apixaban	**Dabigatran**	**Rivaroxaban**	**Warfarin**
Cohort size, n (%)	5,432 (8.8)	8,947 (14.5)	10,071 (16.4)	37,066 (60.3)
Age in years, median (IQR)	78 (72-85)	76 (69-82)	75 (69-81)	78 (71-84)
Male sex, n (%)	2,675 (49.2)	4,671 (52.2)	5,333 (53.0)	18,856 (50.9)
Comorbidities*, n (%)				
Hypertension	3,117 (57.4)	4,690 (52.4)	5,352 (53.1)	24,896 (67.2)
Myocardial infarction	155 (2.9)	154 (1.7)	187 (1.9)	2,484 (6.7)
Congestive heart failure	1,138 (20.9)	1,430 (16.0)	1,560 (15.5)	11,428 (30.8)
Stroke or TIA	672 (12.4)	955 (10.7)	937 (9.3)	6,051 (16.3)
Hemorrhage	417 (7.7)	617 (6.9)	634 (6.3)	3,728 (10.1)
Diabetes	1,303 (24.0)	1,949 (21.8)	2,197 (21.8)	12,002 (32.4)
Peripheral vascular disease	643 (11.8)	864 (9.7)	970 (9.6)	7,297 (19.7)
Chronic renal failure	566 (10.4)	489 (5.5)	619 (6.1)	7,672 (20.7)
Cancer	945 (17.4)	1,286 (14.4)	1,497 (14.9)	5,846 (15.8)
Acute kidney injury	478 (8.8)	318 (3.6)	478 (4.7)	4,923 (13.3)
COPD	922 (17.0)	1,320 (14.8)	1,519 (15.1)	8,094 (21.8)
Liver disease	114 (2.1)	152 (1.7)	207 (2.1)	1,095 (3.0)
Dementia	367 (6.8)	396 (4.4)	414 (4.1)	2,736 (7.4)
Diseases related to alcohol abuse	154 (2.8)	272 (3.0)	362 (3.6)	1,469 (4.0)
Medications**, n (%)				
Medications used for the treatment of hypertension	5,144 (94.7)	8,448 (94.4)	9,292 (92.3)	35,366 (95.4)
Antiplatelets	3,508 (64.6)	5,933 (66.3)	6,287 (62.4)	27,357 (73.8)
Antipsychotics	399 (7.3)	525 (5.9)	628 (6.2)	2,706 (7.3)
NSAIDs	874 (16.1)	1,576 (17.6)	2,077 (20.6)	5,942 (16.0)
Statins	3,223 (59.3)	5,121 (57.2)	5,687 (56.5)	24,148 (65.1)

* Measured in the year before the base cohort entry date

**Measured at the time of cohort entry

Abbreviations: IQR, interquartile range; TIA, transient ischemic attack; COPD, chronic obstructive pulmonary disease, NSAIDs, nonsteroidal anti-inflammatory drugs

Table 4.3. Patterns of utilization of oral anticoagulants from dispensation level data of the base cohort. Québec, 2010-2015.

	Apixaban	**Dabigatran**	**Rivaroxaban**	**Warfarin**
Total dispensations, n (%)	147,008 (5.8)	500,922 (19.6)	333,542 (13.1)	1,571,275 (61.6)
Duration of the dispensing period (days)				
Median (IQR)	7 (7-30)	7 (7-30)	7 (7-30)	7 (7-22)
Mean ± SD	14.9 ± 12.8	17.4 ± 14.1	17.0 ± 13.8	13.3 ± 11.6
Mode	7	7	7	7
Min-Max	0 - 365	0 - 365	0 - 365	0 - 365
Strength of dispensed tablets	2.5 mg - 34% 5 mg – 66%	110 mg- 53% 150 mg – 47%	10 mg – 55% 15 mg- 19% 20 mg – 26%	1 mg – 7% 2 mg – 26% 2.5 mg – 5% 3 mg – 9% 4 mg – 10% 5 mg – 40% 6 mg – 2% 7.5 mg – 0.1% 10 mg –0.6%
Dispenced quantity				
Median (IQR)	14 (14-60)	14 (14-60)	7 (7-30)	7 (7-30)
Mean ± SD	16.7 ± 15.3	18.8 ± 17.1	17.7 ± 15.1	13.0± 10.3
Mode	14	14	7	7
Min-Max	1-360	1-360	1-363	1-360
Days between dispensations				
Median (IQR)	7 (7-28)	7 (7-29)	7 (7-29)	7 (7-21)
Mean ± SD	14.5 ± 13.2	17.8 ± 15.9	17.0 ± 15.4	14.6 ± 16.8
Mode	7	7	7	7
Min-Max	0 - 331	0 - 363	0 - 363	0 - 364
Gaps in medication supply (days)				
Median (IQR)	0 (0-1)	0 (0-2)	0 (0-2)	0 (0-3)
Mean ± SD	1.6 ± 5.8	2.0 ± 6.9	1.9 ± 7.7	3.8 ± 10.8
Mode	0	0	0	0
Min-Max	0 - 301	0 - 350	0 - 355	0 - 345

Abbreviations: IQR, interquartile range; SD, standard deviation, Min, minimum value, Max, maximum value

Table 4.4. Patients' baseline characteristics of matched apixaban, dabigatran, rivaroxaban, and warfarin treatment episodes. N=29,504, Québec, 2010-2015.

	Apixaban	**Dabigatran**	**Rivaroxaban**	**Warfarin**
Cohort size, n (%)	7,376 (25.0)	7,376 (25.0)	7,376 (25.0)	7,376 (25.0)
Age in years, median (IQR)	78 (72-84)	78 (72-84)	78 (72-84)	78 (72-84)
Male sex, n (%)	3,587 (48.6)	3,587 (48.6)	3,587 (48.6)	3,587 (48.6)
Comorbidities*, n (%)				
Hypertension	4,102 (55.6)	4,087 (55.4)	4,064 (55.1)	4,113 (55.8)
Myocardial infarction	178 (2.4)	190 (2.6)	181 (2.5)	214 (2.9)
Congestive heart failure	1,575 (21.4)	1,625 (22.0)	1,558 (21.1)	1,523 (20.6)
Stroke or TIA	891 (12.1)	889 (12.1)	862 (11.7)	877 (11.9)
Hemorrhage	693 (9.4)	733 (9.9)	706 (9.6)	620 (8.4)
Diabetes	1,812 (24.6)	1,846 (25.0)	1,808 (24.5)	1,806 (24.5)
Peripheral vascular disease	878 (11.9)	880 (11.9)	846 (11.5)	843 (11.4)
Chronic renal failure	682 (9.2)	730 (9.9)	669 (9.1)	653 (8.9)
Cancer	1,208 (16.4)	1,184 (16.1)	1,176 (15.9)	1,181 (16.0)
Acute kidney injury	469 (6.4)	506 (6.9)	445 (6.0)	418 (5.7)
COPD	1,242 (16.8)	1,221 (16.6)	1,219 (16.5)	1,141 (15.5)
Liver disease	165 (2.2)	162 (2.2)	156 (2.1)	119 (1.6)
Dementia	474 (6.4)	442 (6.0)	457 (6.2)	442 (6.0)
Diseases related to alcohol abuse	215 (2.9)	217 (2.9)	199 (2.7)	180 (2.4)
Medications**, n (%)				
Medications used for the treatment of hypertension	7,029 (95.3)	7,033 (95.3)	7,045 (95.5)	7,060 (95.7)
Antiplatelets	4,231 (57.4)	4,205 (57.0)	4,127 (56.0)	4,289 (58.1)
Antipsychotics	515 (7.0)	470 (6.4)	485 (6.6)	451 (6.1)
NSAIDs	1,050 (14.2)	1,010 (13.7)	1,014 (13.7)	1,016 (13.8)
Statins	4,589 (62.2)	4,572 (62.0)	4,628 (62.7)	4,577 (62.1)

* Measured in the year before the base cohort entry date

**Measured at the time of cohort entry

Abbreviations: IQR, interquartile range; TIA, transient ischemic attack; COPD, chronic obstructive pulmonary disease, NSAIDs, nonsteroidal anti-inflammatory drugs

Table 4.5. Estimates from multilevel models for the refill gaps in the matched cohorts of patients treated with apixaban, dabigatran, rivaroxaban, or warfarin. Québec, 2010-2015.

	Apixaban	**Dabigatran**	**Rivaroxaban**	**Warfarin**
N of patients in the analysis	7,376	7,376	7,376	7,376
Weighted average of refill gaps	3.08*	3.70 *	3.15 *	9.28*
(SE)	(0.06)	(0.07)	(0.06)	(0.16)
Variance Components				
Within-person	23.95*	31.78*	38.08*	85.89*
(SE)	(0.10)	(0.08)	(0.13)	(0.22)
Between-person	21.83*	35.22*	23.23*	173.44*
(SE)	(0.46)	(0.68)	(0.51)	(3.09)
Intraclass Correlation, ICC	0.48	0.53	0.38	0.67

*p<.001
Abbreviation: SE standard error

Table 4.6. Longitudinal change of refill gaps in the matched warfarin cohort for every 10 dispensations. N=7,376, Québec, 2010-2015.

	Estimate
Initial average refill-gap	9.98
(SE)	(0.16)
Rate of change per 10 dispensations	-0.72*
(SE)	(0.01)
Variance Components	
Within-person	85.96*
(SE)	(0.23)
Between-person at initial refill-gap	183.46*
(SE)	(3.27)
Between-person in the rate of change	1.04*
(SE)	(0.04)
Covariance between initial refill-gap and rate of change	-13.84*
(SE)	(0.34)

*p<.001
Abbreviation: SE standard error

Table 4.7. Estimates from multilevel models for the refill gaps by use of pharmacist prepared weekly pillboxes for the matched cohorts of patients treated with apixaban, dabigatran, rivaroxaban, or warfarin. Québec, 2010-2015.

	Use of weekly pillboxes				**Non-use of weekly pillboxes**			
	Apixaban	**Dabigatran**	**Rivaroxaban**	**Warfarin**	**Apixaban**	**Dabigatran**	**Rivaroxaban**	**Warfarin**
N of patients in the analysis	2,389	2,278	2,319	2,236	4,987	5,098	5,057	5,140
Weighted average of refill gaps (SE)	0.61* (0.02)	0.73* (0.02)	0.71* (0.03)	1.80* (0.06)	4.28* (0.08)	4.98* (0.09)	4.25* (0.08)	12.56* (0.20)
Variance Components								
Within-person (SE)	8.18* (0.04)	8.54* (0.03)	16.09* (0.07)	24.68* (0.08)	56.48* (0.41)	71.88* (0.32)	78.44* (0.45)	206.17* (0.93)
Between-person (SE)	0.94* (0.05)	1.01* (0.04)	1.26* (0.07)	6.89* (0.23)	19.53* (0.66)	36.44* (0.97)	21.25* (0.72)	189.76* (4.42)
Intraclass Correlation, ICC	0.10	0.11	0.07	0.22	0.26	0.34	0.21	0.48

*p<.001
Abbreviation: SE standard error

CHAPTER 5: Defining the duration of the dispensation of oral anticoagulants in administrative healthcare databases.

Sinyavskaya, L, Renoux, C, Durand, M. *Pharmacoepidemiology and Drug Safety* 2022; 31(1): 105- 109.

5.1 Abstract

Purpose: In clinical practice, warfarin therapy requires frequent dose adjustments. In pharmacy claims, the days supplied value may not reflect the true duration of warfarin dispensation. This may affect the measures of association comparing the safety of direct oral anticoagulants (DOACs) vs. warfarin.

Methods: Using Québec healthcare administrative databases, we formed a cohort of 55,230 patients newly treated with oral anticoagulants between 2010 and 2016. The duration of dispensations was defined using two approaches: the recorded days supplied value, and the longitudinal coverage approximation (data-driven) that may account for individual variation in drug usage patterns. Propensity scores adjusted Cox proportional hazards regression models were used to estimate the hazard ratio (HR) of major bleeding with dabigatran or rivaroxaban vs. warfarin.

Results: Using the days supplied, the mean (and standard deviation) dispensation durations for dabigatran, rivaroxaban, and warfarin were 19 (15), 19 (14), and 13 (12) days, respectively. Using the data-driven approach, the durations were 20 (16), 19 (15), and 15 (16) days, respectively. The choice of the approach had no impact on the HR estimates.

Conclusions: In our settings, the data-driven approach closely approximated the recorded days supplied value for the standard dose therapies such as dabigatran and rivaroxaban. For warfarin, the data-driven approach captured more variability in the duration of dispensations compared to the days supplied value, which may better reflect the true drug-taking behavior of warfarin. Both approaches may provide valid estimates when comparing the safety of DOACs vs. warfarin.

5.2 Introduction

Administrative healthcare databases are used for the post-marketing studies on the comparative effectiveness and safety of direct oral anticoagulants (DOACs) and vitamin K antagonist warfarin[114,255]. To define the duration of dispensations, most studies used the days supplied value. However, the days supplied value may be vulnerable to exposure misclassification when comparing therapies with different medication use patterns, such as warfarin and DOACs[97,270,280].

Individuals on warfarin need frequent dosing adjustments. The true duration of warfarin supply may last longer than the recorded days supplied due, for example, to pill-splitting. In such a case, the gaps between dispensations do not necessarily imply poor adherence. Conversely, gaps between dispensations for DOACs may mostly be attributable to the lack of adherence. A previous study showed that a data-driven approach to defining the duration of dispensation may account for past medication use patterns and accurately approximate the duration of dispensation for medications with large dosage variability, such as vitamin K antagonist phenprocoumon[251].

We evaluated whether the data-driven approach better captures the variability in the duration of warfarin dispensations compared to the days supplied approach. We further investigated how the approach used to define the duration of dispensation affects the strength of an association between major bleeding and treatment with dabigatran or rivaroxaban vs. warfarin.

5.3 Methods

Study cohort

Using the computerized administrative healthcare databases of the *Régie de l'assurance-maladie du Québec* (RAMQ), we identified all individuals ≥18 years initiating dabigatran, rivaroxaban, or warfarin for non-valvular atrial fibrillation (NVAF) between January 2010 and December 2016. An initiator was defined as a person with no dispensation of any oral anticoagulant in the preceding year. The date of the first dispensation was defined as the date of cohort entry. Individuals were followed until the occurrence of a major bleeding, end of pharmaceutical coverage, discontinuation of the treatment allocated at cohort entry, death, or end of study (March 31, 2017), whichever occurred first.

Definition of exposure

From all dispensations we extracted the dispensing date, days supplied, unit strength, and dispensed quantity. The duration of dispensations was defined using two approaches: based on the recorded days supplied value and the data-driven approach.

The data-driven approach or the longitudinal coverage approximation was described and evaluated previously[251]. Unlike the approach that relies on the recorded days supplied value, the data-driven approach estimates the duration of dispensation given the observed time intervals since drug initiation, the dispensed quantities, and the unit strength. Thus, for the prescription dispensed at time t_i, the duration of dispensation is estimated from the total milligrams of drug dispensed at time t_i divided by the average daily dose used from drug initiation (t_0) to current dispensation (t_i).

The average daily dose is a quotient of the total mg dispensed from the drug initiation to current dispensation and the number of days passed from drug initiation to current dispensation.

Then the duration of dispensation is estimated using the next formula.

$$Duration\ of\ dispensation = \frac{Total\ mg\ dispensed\ at\ time\ ti}{\frac{Total\ mg\ dispensed\ between\ t_0\ and\ t_{i-1}}{Number\ of\ days\ between\ t_0\ and\ t_i}}$$

For example, a patient receives the first dispensation of 30 pills of warfarin 5 mg on day zero; the second dispensation of 30 pills of warfarin 2mg on day 60; and the third dispensation of 30 pills of warfarin 2mg on day 90. The estimated duration of the current dispensation (third dispensation) of warfarin is:

*[(30pills*2mg)/(30pills*5mg+30pills*2mg)/90days]=26 days*

When applicable, the number of days that a patient spent at the hospital was subtracted from the estimated number of days between t_0 and t_i. This measure was taken to account for the medication supply provided by the hospital, which could increase the coverage of the supply obtained at the pharmacy.

The maximum allowed length of the estimated duration of dispensation was two and a half of the dispensed quantity at t_i[251]. For the first dispensation or when the preceding estimated duration exceeded the maximum allowed length, the average daily dose was replaced by the daily doses recommended or commonly prescribed to older individuals: 220 mg for dabigatran[252], 20 mg for rivaroxaban[252], and 3 mg for warfarin[254].

For each oral anticoagulant, we constructed the periods of continuous drug exposure using the duration of dispensation obtained with both approaches plus one of three grace periods: 15, 30, and 60 days[114,255]. When applicable, the gaps between two consecutive dispensations were adjusted for the number of days spent in the hospital. We censored the follow-up if the gap between dispensations exceeded the allowed grace period.

Outcome definition

The outcome was time to gastrointestinal, ocular, intracranial, or other bleeding requiring hospitalization. The admission date to the hospital was assigned as the date of the outcome event.

Statistical analysis

We used descriptive statistics to summarize the patients' demographics, comorbidity profile, and use of medications measured at the time of cohort entry.

The distributions of the estimated duration of dispensation in days were summarized with means and standard deviations (SD).

The incidence rates of major bleeding were calculated using Poisson regression.

We used logistic regression models to estimate the propensity scores of each patient being treated either with dabigatran vs. warfarin, or rivaroxaban vs. warfarin given age, sex, and clinical characteristics measured in the year before cohort entry. Cox proportional hazard regression models adjusted for the year of cohort entry and deciles of propensity scores were used to estimate the association between treatment with dabigatran or rivaroxaban versus warfarin and major bleeding. The analyses were repeated for each grace period. The model building statistics are in Annex 10. The study was approved by the research ethics board of the Centre Hospitalier de l'Université de Montréal (CHUM).

5.4 Results

The study cohort included 55,230 patients initiating an oral anticoagulant for NVAF, of whom 8,800 (16%) were dispensed dabigatran, 15,683 (28%) rivaroxaban, and 30,743 (56%) warfarin. At the time of cohort entry, the median age was 77 (IQR: 70-83) years, and 52% were males.

Compared to patients initiating dabigatran or rivaroxaban, patients initiating warfarin were older, and have a higher burden of comorbidities, including hypertension, cardio-and cerebra-vascular diseases, diabetes, kidney injury, and chronic renal failure. Warfarin patients more often used antiplatelet agents (Table 5.1). At least 50% of patients initiating dabigatran were dispensed a 110 mg dose (median dispensed strengths - 100 mg, IQR: 110-150 mg) and at least 25 % of patients initiating rivaroxaban were dispensed a 15 mg dose (median dispensed strengths - 20 mg, IQR: 15-20 mg). In patients initiating warfarin, the median estimated daily dose was 4 mg (IQR: 3-5 mg).

Using the days supplied and the data-driven approaches, the means and SDs of the defined durations were close for dabigatran and rivaroxaban (Figure 5.1). For warfarin, the mean and SD were, respectively 15% and 33% greater with the data-driven compared to the days supplied approaches.

In general, with both approaches, longer grace periods resulted in longer follow-up (Table 5.2), a greater number of major bleeding events, and lower incidence rates of major bleeding (Table 5.3). However, using the same grace period, the estimates on the duration of follow-up, number of major bleeding events, and incidence rates of major bleeding were close using both approaches. With both approaches and across all grace periods, we consistently found that the hazards of major bleeding were 30 to 40% significantly lower in dabigatran vs. warfarin-treated individuals. Overall, the hazards were 9 % to 20% lower in rivaroxaban vs. warfarin-treated individuals with borderline statistical significance (Table 5.3).

5.5 Discussion

In this study based on the analysis of administrative healthcare datasets, the data-driven approach to define the duration of dispensation captured more variability for warfarin compared to the days supplied value. However, the choice of the approach, as well as the grace period length, did not have a substantial impact on the strength of association between treatment with dabigatran or rivaroxaban vs. warfarin and the risk of major bleeding.

We found that the data-driven approach to define the duration of dispensation may approximate the days supplied value for standard-dose therapies such as dabigatran and rivaroxaban. For warfarin, a drug with a dose-varying regiment, the data-driven approach may better reflect the true drug-taking behavior of warfarin. These findings may have important implications for future studies using databases that do not contain information on the intended prescription duration[281] or databases where the missing information for days supplied is a substantial problem[282].

In our study, the choice of the approach for defining the duration of dispensation, as well as the length of the grace period, did not have a differential impact on the incidence rates of major bleeding across all oral anticoagulants. Also, both approaches consistently produced very close estimates of the HRs comparing the risk of major bleeding in dabigatran or rivaroxaban vs. warfarin. This suggests that in our study population, there were few individuals on warfarin having long gaps between dispensations. Indeed, we previously found that for up to 75% of warfarin dispensations, the gaps were no longer than three days[270]. This relatively high persistence in the therapy may be explained by the usage of the pharmacy-prepared pillboxes that are distributed weekly at the community pharmacies in Québec and are aimed to improve clinical outcomes in the population [277]. However, in settings with different community pharmacy interventions, the

approach used to define the duration of dispensations may have a different effect on the study estimates.

In our study, the incidence rates of major bleeding were among the lowest of those reported in previous observational studies of individuals with NVAF[105]. Thus, a systematic review of real-world observational studies comparing the risk of bleeding in patients with AF reported that the overall rate of major bleeding ranged between studies from 1.48 to 8.42 per 100 person-years with a median of 3.46 per 100 person-years[105]. The low incidence rates in our study may partially be explained by a more restrictive definition of major bleeding that included only those bleeding events requiring hospitalization. Our finding is consistent with that of a Canadian multicenter population-based observational cohort study on the effectiveness and safety of oral anticoagulants[264]. The study was based on the analysis of health care databases from seven Canadian provinces, and in each province, data were analyzed following a common research protocol. The study found some heterogeneity across provinces in the incidence of major bleeding[264]. The possible explanation for lower rates in Quebec was possibly different healthcare practices and management of patients on oral anticoagulants.

In our study, the estimates of HRs comparing the association between major bleeding and treatment with dabigatran vs warfarin are consistent with those reported in meta-analyses of observational studies[101,103]. Thus, a meta-analysis of observational studies reported the pooled HR of the association between major bleeding and treatment with dabigatran versus warfarin was 0.75 (95%CI: 0.62-0.86). However, the same meta-analysis reported that the pooled HR of the association between major bleeding and treatment with rivaroxaban versus warfarin was 1.02 (95%CI: 0.93–1.12). In our study, the direction of the estimates was in favor of rivaroxaban across all grace periods and methods for the definition of exposure duration. Furthermore, the RE-LY

trial reported the relative risk of major bleeding was 0.93 (95%CI:0.81–1.07) and 0.80 (95%CI: 0.69–0.93) in patients treated with dabigatran 150 mg versus warfarin and dabigatran 110 mg versus warfarin, respectively. The ROCKET AF clinical trial found the HR of major bleeding was 1.03 (95%CI: 0.96-1.11) in patients on rivaroxaban versus those on warfarin. Our findings were more in favor of dabigatran and rivaroxaban. This may be explained by a higher proportion of patients treated with reduced doses of dabigatran or rivaroxaban in our setting. Furthermore, in real-world clinical settings, patients on warfarin may have a higher risk of bleeding or thromboembolism because of poorer INR control compared to the clinical trial setting. Finally, at baseline, compared to patients on dabigatran or rivaroxaban, patients initiating warfarin were frailer and older, and a higher proportion of patients initiating warfarin had a history of previous stroke or bleeding. Although we adjusted for the baseline confounding using PS, the residual confounding may bias the estimates in favor of DOACs.

Our study has some limitations. First, the time of exposure to oral anticoagulants was approximated based on information from dispensation claims, and the true drug intake by patients was unknown. Second, in some individuals, low adherence to anticoagulants would lead to misclassification of unexposed person-time as exposed, especially with longer grace periods. This may be more prominent for dabigatran, which is taken twice daily[283]. However, in this study, we aimed to investigate how the operational definition of duration of dispensation may impact the effect estimates of the important clinical outcome under the general assumption of good treatment adherence at the population level.

Overall, in our settings, the data-driven approach closely approximated the recorded days supplied value for the standard dose therapies such as dabigatran and rivaroxaban. For warfarin, the data-driven approach captured more variability in the duration of dispensations compared to the days

supplied value, which may better reflect the true drug-taking behavior of warfarin. Both approaches may provide valid estimates when comparing the safety of DOACs vs. warfarin.

5.6 Table and figure

Table 5.1. Baseline characteristics of patients treated with dabigatran, rivaroxaban, and warfarin. Québec, 2010-2016.

Characteristic	Dabigatran	Rivaroxaban	Warfarin
Number of patients	8,800	15,683	30,743
Age in years, median (IQR)	75 (69-82)	74 (68-80)	78 (71-84)
Male sex, n(%)	4,663 (53.0)	8,543 (54.5)	15,706 (51.1)
Comorbidities, n(%)			
Hypertension	8,277 (94.1)	14,413 (91.9)	29,448 (95.8)
Congestive heart failure	1,214 (13.8)	2,047 (13.1)	7,966 (25.9)
Coronary atherosclerosis	2,810 (31.9)	4,714 (30.1)	14,695 (47.8)
Diabetes	2,183 (24.8)	3,744 (23.9)	10,237 (33.3)
Peripheral vascular disease	770 (8.8)	1,294 (8.3)	5,236 (17.0)
Chronic renal failure	423 (4.8)	828 (5.3)	5,792 (18.8)
Cancer	1,242 (14.1)	2,328 (14.8)	4,869 (15.8)
Acute kidney injury	289 (3.3)	675 (4.3)	3,774 (12.3)
COPD	1,235 (14.0)	2,245 (14.3)	6,371 (20.7)
Liver disease	126 (1.4)	261 (1.7)	780 (2.5)
Dementia	533 (6.1)	868 (5.5)	3,005 (9.8)
Diseases related to alcohol abuse	265 (3.0)	566 (3.6)	1,188 (3.9)
Bleeding	717 (8.1)	1,127 (7.2)	3,246 (10.6)
Stroke	898 (10.2)	1,256 (8.0)	4,819 (15.7)
Medical procedures, n(%)			
Pacemaker implantation/Catheter ablation	614 (7.0)	917 (5.8)	2,410 (7.8)
PCI/CAB	155 (1.8)	375 (2.4)	2,732 (8.9)
Medications, n(%)			
Antiplatelets	4,796 (54.5)	7,982 (50.9)	18,981 (61.7)
Antipsychotics	367 (4.2)	786 (5.0)	1,695 (5.5)
NSAIDS	406 (4.6)	943 (6.0)	1,233 (4.0)
Statins	4,729 (53.7)	8,142 (51.9)	18,459 (60.0)
Benzodiazepines	1,961 (22.3)	3,344 (21.3)	8,699 (28.3)
Systemic corticosteroids	351 (4.0)	695 (4.4)	1,720 (5.6)
H-2 receptor blockers	87 (1.0)	132 (0.8)	333 (1.1)
Proton pump inhibitors	3,417 (38.8)	6,200 (39.5)	15,240 (49.6)
Opioids	454 (5.2)	1,949 (12.4)	2,532 (8.2)
SSRI	612 (7.0)	1,155 (7.4)	2,487 (8.1)

Table 5.2. Duration of follow-up of the study individuals by applied grace period and approach for defining the duration of dispensation.

Grace period	Approach	Duration of follow-up (person-years)		
		Mean	Median	Total
15-day	Days supplied	1.2	0.6	65,561
	Data-driven	1.1	0.6	63,369
30-day	Days supplied	1.6	1.0	85,974
	Data-driven	1.6	1.0	89,429
60-day	Days supplied	1.9	1.3	103,063
	Data-driven	1.9	1.3	106,273

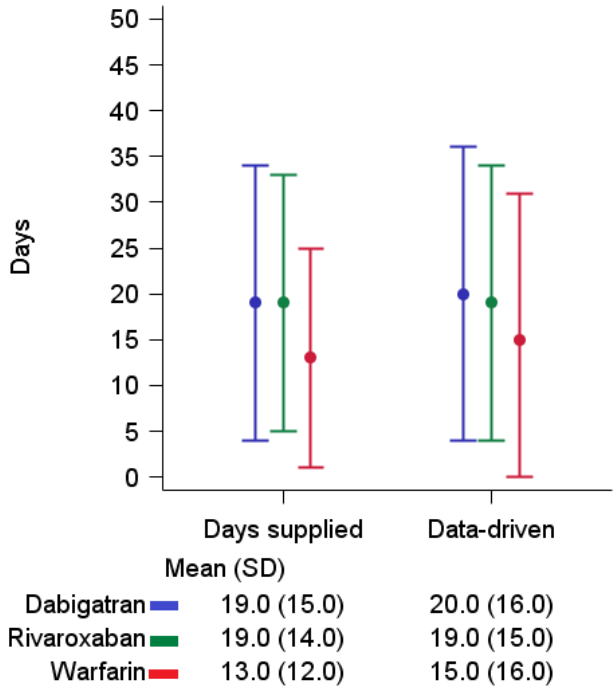
Table 5.3. Hazard ratios for the association between the use of dabigatran or rivaroxaban versus warfarin and the risk of major bleeding, using two approaches to define the duration of dispensations. N=55,230, Québec, 2010-2016.

Grace- period	Exposure	N of events	Person-years	Incidence rate (95% CI) *	Crude HR	Adjusted HR (95% CI)
15-day						
Days supplied	Dabigatran	167	14,087	1.18 (1.02-1.38)	0.47	0.63 (0.52-0.75)
	Rivaroxaban	347	19,740	1.76 (1.58-1.95)	0.66	0.91 (0.77-1.07)
	Warfarin	863	31,735	2.72 (2.54-2.91)	Ref	Ref
Data-driven						
	Dabigatran	164	13,497	1.20 (1.02-1.41)	0.49	0.67 (0.56-0.80)
	Rivaroxaban	328	18,480	1.77 (1.59-1.98)	0.67	0.91 (0.77-1.08)
	Warfarin	837	31,392	2.67 (2.49-2.85)	Ref	Ref
30-day						
Days supplied	Dabigatran	194	17,745	1.09 (0.95-1.26)	0.46	0.59 (0.50-0.69)
	Rivaroxaban	386	23,588	1.64 (1.48-1.81)	0.63	0.81 (0.70-0.95)
	Warfarin	1,132	44,642	2.54 (2.39-2.69)	Ref	Ref
Data-driven						
	Dabigatran	188	17,646	1.06 (0.92-1.23)	0.45	0.58 (0.49-0.68)
	Rivaroxaban	378	23,216	1.63 (1.47-1.80)	0.64	0.82 (0.70-0.95)
	Warfarin	1,192	48,568	2.45 (2.32-2.60)	Ref	Ref
60-day						
Days supplied	Dabigatran	211	20,013	1.05 (0.92-1.21)	0.46	0.58 (0.50-0.68)
	Rivaroxaban	410	26,073	1.57 (1.43-1.73)	0.63	0.79 (0.69-0.91)
	Warfarin	1,350	56,977	2.37 (2.25-2.50)	Ref	Ref
Data-driven						
	Dabigatran	213	20,039	1.06 (0.93-1.21)	0.47	0.60 (0.51-0.69)
	Rivaroxaban	410	26,000	1.58 (1.43-1.74)	0.64	0.81 (0.70-0.93)
	Warfarin	1,401	60,233	2.33 (2.21-2.45)	Ref	Ref

*per 100 person-years

Abbreviations: N, number; CI, confidence interval; HR, hazard ratio

Figure 5.1. Mean and standard deviation of the defined duration of dispensations for oral anticoagulants using the days supplied and the data-driven approaches. Québec, 2010-2016.



Abbreviations: SD, standard deviation

CHAPTER 6: Evidence of the different associations of prognostic factors with censoring across treatment groups and impact on censoring weight model specification: the example of anticoagulation in atrial fibrillation[284].

Sinyavskaya L, Schnitzer M, Renoux C, Guertin JR, Talbot D, Durand M. American Journal of Epidemiology. 2021; *90*(12), 2671–2679.

6.1 Abstract

Applying the inverse probability of censoring weights (IPCWs) may reduce selection bias due to informative censoring in longitudinal studies. However, in studies with an active comparator, the associations between predictors and censoring may differ across treatment groups, and failure to take this into account may result in biased estimates. We used the clinical example of anticoagulation treatment with warfarin or a direct oral anticoagulant (DOAC) in atrial fibrillation to illustrate this based on an analysis of the Québec administrative databases. The parameter of interest was the hazard ratio of the composite of stroke, major bleeding, myocardial infarction, or death associated with continuous use of warfarin versus DOACs. Two strategies for the specification of the model for estimation of censoring weights were explored: exposure-unstratified and exposure-stratified. The hazard ratio associated with continuous treatment with warfarin versus DOACs adjusted with exposure stratified IPCWs was 1.26 (95% confidence interval (CI): 1.20, 1.33). Using exposure-unstratified IPCW, the hazard ratio differed by 15 % in favor of DOACs (1.41; 95% CI: 1.34, 1.48). Not accounting for the different associations between the predictors and informative censoring across exposure groups may lead to the misspecification of censoring weights and a biased estimate of the treatment effect.

6.2 Introduction

Informative censoring introduces a selection bias in randomized clinical trials and observational studies[169,285,286]. Informative censoring may occur when study participants are lost to follow-up, in case of treatment crossover, or upon treatment cessation[151]. In addition, in observational studies using large population-based databases, informative censoring may happen at the time of disenrollment from the health plan[287,288]. Inverse-probability-of-censoring weights (IPCWs) can be used to correct selection bias introduced by informative censoring[151,165,168,289].

One of the assumptions necessary for IPCWs to create a pseudo-population free of selection bias is the correct specification of the model used to estimate the IPCWs[163]. We present a case where the association between covariates and censoring probability differs by treatment groups, using an observational study of individuals with atrial fibrillation initiating a direct oral anticoagulant (DOAC) or warfarin for stroke prevention. Atrial fibrillation is the most common cardiac arrhythmia affecting more than 2.7 million Americans[290]. Anticoagulation is the cornerstone of treatment to prevent the intracardiac formation of blood clots and secondary embolization of the systemic circulation causing a stroke. The use of warfarin, approved by the FDA in 1954 for stroke prevention in atrial fibrillation, has largely been replaced by DOACs (dabigatran, rivaroxaban, apixaban, and edoxaban) which began to enter the market in 2010. DOACs' market entry represents the major change in atrial fibrillation management in the last 10 years. Following the landmark randomized clinical trials that led to regulatory approval of DOACs, there has been much interest in estimating the real-life safety and effectiveness of DOACs vs warfarin. A 2018 systematic review found more than 70 observational studies aiming to estimate the safety and effectiveness of DOACs vs warfarin[104].

The first objective of our study was to describe the associations between the prognostic variables and censoring in DOACs and warfarin groups. The second objective was to compare two IPCW strategies to estimate the effect of continuous treatment with warfarin versus DOACs on the composite of stroke, major bleeding, myocardial infarction, and all-cause mortality. We compared three estimates of the hazard ratio obtained by using unweighted, weighted with exposure-unstratified IPCWs, and weighted with exposure-stratified IPCWs analyses.

6.3 Methods

Case presentation

We designed an observational study of new initiators of DOACs or warfarin in atrial fibrillation. To estimate the sustained treatment effect of warfarin vs DOACs on the occurrence of stroke or bleeding, follow-up may be censored at the time of deviation from baseline treatment (per-protocol effect)[260]. However, reasons to deviate from baseline treatment allocation are distinct for DOACs and warfarin initiators. For example, a switch from a DOAC to warfarin may happen as a precaution, in the case of a new-onset renal failure, since DOACs are contraindicated in end-stage renal disease. Renal failure also increases the risk of the outcome[3,39,261,262]. On the contrary, a switch from warfarin to a DOAC may happen for treatment simplification in individuals free of renal failure. Therefore, during follow-up, sicker individuals with higher outcome risks are removed from the DOACs group, whereas healthier individuals with lower outcome risks are removed from the warfarin group, resulting in selection bias. A causal diagram (Figure 6.1) presents the structure of selection bias due to censoring using the bleeding event as an outcome and a new-onset renal failure (henceforth, renal failure) as a time-varying prognostic factor[157,291].

In this situation, a model of censoring probabilities that is adjusted for treatment value, time of follow-up, and all prognostic factors, would estimate the average coefficients for prognostic variables across treatment groups. Thus, renal failure is expected to be positively associated with censoring in DOACs users and negatively associated with censoring among warfarin users. The coefficient for the probability of censoring given renal failure would be an average value of the treatment-specific censoring probabilities. As a result, applying the IPCWs would fail to eliminate

the selection bias associated with renal failure. One potential solution to correctly estimate censoring weights in this clinical situation is to use an exposure-stratified censoring model.

Data source

This retrospective cohort study was conducted by linking three computerized health care databases of the *Régie de l'assurance-maladie du Québec* (RAMQ). The RAMQ administers the government health insurance plan for all Québec residents, and prescription drug insurance plan for individuals aged 65 and older, welfare recipients, and those not covered by private medical insurance, totaling about 40 % of the Québec population [184]. The study data included outpatient and inpatient physician claims, hospitalization discharge summaries, and outpatient medication dispensations. The RAMQ data have been extensively used for pharmacoepidemiologic research, including dispensation patterns, and the effectiveness and safety of oral anticoagulants [187-189].

Study cohort

We included all individuals 18 years or older who initiated oral anticoagulation (warfarin or DOACs) between January 1, 2010, and December 31, 2016, and enrolled in the database at least one year before treatment initiation. DOACs included were dabigatran, rivaroxaban, or apixaban; edoxaban was not available during the study period. Treatment initiation was defined as having no dispensation for any oral anticoagulant in the year before. Individuals had to receive a diagnostic code for atrial fibrillation in the year before initiation of anticoagulation with no record of prior valvular heart disease, venous thromboembolism, cardiac valve surgery, or hemodialysis (see ICD9 and ICD10 codes in Table 6.1). The cohort entry date was defined as the date of the first dispensation for an anticoagulant. We followed individuals until the earliest of the following: occurrence of the study outcome, end of pharmaceutical coverage, five years since the initiation of the oral anticoagulant, or end of the study period (March 31st, 2017).

Definition of exposure

Exposure to DOACs and warfarin was defined based on the pharmaceutical dispensation records. The periods of continuous exposure were identified using dispensation dates, reported days supplied, and a grace period of 30 days.

Outcome definition

The outcome of interest was time to a composite of ischemic or hemorrhagic stroke, systemic embolism, major bleeding (intracranial, gastrointestinal, ocular, and any other bleeding necessitating hospitalization), myocardial infarction, and all-cause mortality. Stroke, systemic embolism, and major bleeding events were identified using the International Classification of Diseases (10th revision) codes from hospital discharge summaries using primary or principal diagnoses, or the cause of in-hospital death. The admission date to the hospital was assigned as the date of the outcome event. The date of death was obtained from Québec's health insurance database.

Covariates

We collected information on the potential prognostic factors that may both have an impact on initiating or stopping warfarin or a DOAC and may increase the risk of the components of the composite outcome. These prognostic factors were identified using expert knowledge and information from previously published studies on the comparative effectiveness and safety of warfarin and DOACs. We gathered information on age; sex; and the following medical conditions: hypertension, congestive heart failure, coronary artery disease, diabetes, peripheral vascular disease, stroke or transient ischemic attack, chronic kidney disease, acute kidney injury, liver disease, cancer (excluding non-melanoma skin cancer), chronic obstructive pulmonary disease,

prior major bleeding, dementia, use of revascularization for coronary atherosclerosis (coronary artery bypass surgery or percutaneous coronary intervention) or surgical treatment for cardiac arrhythmias (implantation of a pacemaker or implantable cardioverter-defibrillator, atrial fibrillation ablation or left auricular appendage closure). Information on comorbid medical conditions and surgical treatment was collected from outpatient and inpatient physician claims, and hospitalization discharge summaries. A list of ICD codes used to define comorbidities is in Annex 3.

The use of the following medications was also collected from the outpatient medication dispensations dataset: antiplatelets, antipsychotics, non-steroidal anti-inflammatory drugs, statins, benzodiazepines, systemic corticosteroids, histamine H2-receptor blockers, proton pump inhibitors, selective serotonin reuptake inhibitors, and opioids.

The data was structured in the "counting process" style with multiple rows of data per subject. For each study individual, we split the follow-up into 30-day intervals, and measured the changes in the values of covariates and the outcome events at every interval, as follows: covariates were measured within the year before the cohort entry date, and then updated at the beginning of every 30-day time-period of follow-up. Once medical conditions occurred, they were assumed to remain present for the remainder of the follow-up time. For exposure to medication, current use was determined with patients considered exposed from the date of medication dispensation until the end of the dispensation duration, plus a 30-day grace period. We assumed that these covariates were sufficient to achieve conditional exchangeability when adjusting the baseline confounding with matching on propensity scores (see next section) and selection bias due to informative censoring with the IPCWs[165,292].

Analytical approach

Our goal was to emulate a target trial where study participants who would be prescribed warfarin under usual care, and were eligible for DOACs and warfarin treatment, would be randomized to receive either warfarin or a DOAC. All subjects would continue their initial treatment for the entire follow-up (per protocol analysis[169,265]). To balance the observed demographic and clinical characteristics at baseline, patients dispensed warfarin were matched 1:1 without replacement to patients dispensed a DOAC on age (±1 year), sex, cohort entry date (±2 years), and propensity scores (0.2 of the standard deviation of logit of propensity scores) estimated from baseline clinical predictors of treatment/censoring defined above[213]. The follow-up was right-censored at the time of deviation from the baseline treatment, either due to a switch to an alternative anticoagulant or treatment cessation for any reason. Additionally, in both treatment groups, we right-censored the follow-up at the date of a recorded health condition when DOACs became absolutely contraindicated (i.e., initiation of hemodialysis in end-stage renal disease, heart valve surgery). We considered this latter censoring as being at random conditional on covariates as both the warfarin- and the DOACs-treated individuals were expected to have equal probabilities of developing these health conditions after the baseline matching.

We conducted three analyses. In the first analysis, the censoring due to deviation from baseline treatment was considered as non-informative or completely at random and was not adjusted for. In the second analysis, this censoring was considered informative with similar associations between the prognostic factors and the probability of censoring in individuals treated with warfarin and DOACs. We adjusted with the IPCWs calculated with a common censoring model, unstratified by exposure[165]. In the third analysis, we assumed that the associations between prognostic factors and the probability of censoring differ between DOAC- and warfarin-treated individuals.

Accordingly, we built IPCWs using exposure-stratified censoring models (i.e.: weights were built separately for patients on warfarin and those on DOACs).

Statistical analysis

Propensity score**.** A logistic regression model was used to estimate the baseline probabilities of each patient being treated with warfarin vs DOACs, given all covariates stated above. Propensity scores were used to create the matched study cohort. Descriptive statistics were used to summarize distributions of patients' characteristics at cohort entry (see Table 6.2). A characteristic was considered as balanced across exposure groups with ≤10 % absolute standardized difference[266].

Probability of being censored and IPCWs. To describe the predictors of censoring in each treatment group, a pooled logistic regression was used to estimate the probabilities of being censored[289] for each month of follow-up given sex, age at baseline, month, and the vector of covariates with the values updated for the previous months (*t*-1). Two sets of probabilities of censoring were estimated, one utilizing an unstratified model and the second with exposure-stratified models.

We presented the odds ratios (OR) with 95% confidence intervals (CI) as an association of each covariate with the probability of censoring. The output from logistic regression models to estimate probabilities of censoring from the unstratified and exposure-stratified models is in Annex 14.

In a sensitivity analysis, we further explored if the mechanism of censoring is different between those who switched between treatment groups and those who permanently discontinued treatment with oral anticoagulants. For this analysis, censoring due to deviation from baseline treatment was considered informative, with a different dynamic of censoring between treatment groups. Probabilities of censoring were estimated using exposure-stratified censoring models.

Furthermore, for each treatment group, two sets of probabilities of censoring were estimated, (1) to account for censoring due to switching between treatment groups, and (2) to account for censoring due to permanent discontinuation of treatment.

We estimated the IPCWs as the inverse of the product of the probabilities of remaining uncensored (1-probabilities of censoring) up to and including the month of follow-up[151]. Accordingly, we estimated unstratified and exposure-stratified IPCWs. For the sensitivity analysis, IPCWs were estimated for each treatment group as a product of two censoring weights (to account for censoring due to switching and censoring due to permanent treatment discontinuation)[264].

To reduce the influence of extreme values of the estimated IPCWs on the size of the standard error for the hazard ratio estimate, the weights were stabilized by the conditional probability of not being censored given the month of follow-up and treatment group for exposure-unstratified IPCWs, and month of follow-up only for unstratified IPCWs[267]. The hazard ratio estimates from the models adjusted with non-stabilized and stabilized weights provided identical point estimates aside from the larger standard errors for the analyses with non-stabilized weights. We presented the distribution of stabilized IPCWs estimated with unstratified and exposure-stratified models using box plots.

Aside from the correct specification of the censoring model that we explored in our case study, other assumptions are required for the IPCWs to create a pseudo-population free of selection bias introduced by informative censoring[163]. First, the non-testable ignorability assumption, i.e. we assume that in the current month, censoring occurs at random given the probability of remaining uncensored up to the current month, sex, age, treatment assigned at baseline, and the observed values of covariates in the previous month. Second, the positivity assumption, that is we assume that at any given month, the study individuals with any combination of the observed covariates in

the previous month, age, sex, and treatment assigned at baseline have greater than zero probability of not being censored. We tested the positivity assumption by examining the distributions of the unstratified and exposure stratified IPCWs[166].

Outcome model. To estimate the risk of the composite outcome in warfarin vs DOACs-treated individuals, we fitted three pooled logistic regression models that approximated Cox proportional hazards regression models[268] with the treatment as the only independent variable. We fitted unweighted, and two weighted outcome models with the weights changing every month of follow-up. One model was weighed with the unstratified IPCWs and the other with the exposure-stratified IPCWs. In the weighted models, the variance of the estimated coefficient for treatment was approximated using a robust variance estimator that takes into account the weights and repeated measures[269].

For the sensitivity analysis, a pooled logistic regression model was weighed with the exposure-stratified censoring weights that were the product of two IPCWs (IPCW for switching and IPCW for permanent discontinuation of treatment).

The study used the de-identified individuals' data. It was approved by the research ethics board of the Centre Hospitalier de l'Université de Montréal (CHUM).

6.4 Results

Characteristics of the matched cohort

The study cohort included 23,927 matched pairs of warfarin- and DOAC-treated individuals with balanced distributions of demographic and clinical characteristics (Table 6.2). The flow of study participants for the study is shown in

Annex 12. During the study period, the proportion of censoring due to treatment cessation for any reason, or switching from the baseline treatment, was twice as high in the warfarin compared to the DOACs group (69 and 36 %, respectively). Of those censored, 41 % of warfarin- and 15 % of DOAC-treated individuals were censored due to treatment switch. Furthermore, 390 (0.8%) patients were censored because of developing health conditions when the use of DOACs is contraindicated. Of those, 148 (0.3%) patients underwent hemodialysis (47 from the DOACs group and 101 from the warfarin group), and 242 (0.5%) patients had valve surgery (106 from the DOACs group and 136 from the warfarin group).

The total follow-up was 76,177 person-years. The Kaplan-Meier survival curves (Figure 6.2) display that compared to patients on warfarin, those on DOACs had a higher crude cumulative survivorship with respect to composite effectiveness and safety outcome.

The incidence of some major prognostic factors in the study cohort during the follow-up is shown in

Annex 13.

Probabilities of censoring

Table 6.3 provides the odds ratios estimated from the unstratified and exposure-stratified censoring models (used to calculate unstratified and exposure-stratified IPCWs, respectively). In the unstratified censoring model, individuals had higher odds of being censored if they were males and had comorbid congestive heart failure, cancer, dementia, or alcohol-related disorders, underwent surgical treatment for coronary atherosclerosis or cardiac arrhythmias, or were treated with concomitant antiplatelets, non-steroidal anti-inflammatory drugs, corticosteroids, or opioids. On the other hand, hypertension, concomitant treatment with statins, proton pump inhibitors, or serotonin reuptake inhibitors were associated with lower odds of being censored. In the exposure-stratified censoring model, being male or having comorbid congestive heart failure was associated with higher odds of censoring only in the DOACs group, whereas in the warfarin group the odds ratios of these variables were close to the null (1.01 and 0.99, respectively). Chronic renal failure was associated with higher odds of censoring in DOACs treated individuals (OR 1.34; 95%CI:1.25-1.43), lower odds of censoring in individuals on warfarin (OR 0.86; 95%CI: 0.82-0.91), and it was not found to be significantly associated with censoring in the unstratified model (OR 1.01; 95%CI: 0.97-1.06). Overall, individuals with more comorbidities and expected poorer prognosis had higher odds of being censored in the DOAC group compared to the warfarin group. As expected, the odds ratios from the unstratified analysis lay between the two odds ratios from the stratified analysis. The sensitivity analysis showed that these differences in characteristics were driven by censoring due to the switch between treatment groups (Annex 15).

Inverse probabilities of censoring weights

Figure Figure 6.3 presents the distribution of IPCWs derived from the unstratified and stratified models. For the exposure-unstratified estimated weights, the overall mean was 1.00. For the

exposure stratified IPCWs, the mean weights were 1.00 and 1.02 in the DOACs and warfarin groups, respectively.

Comparative effectiveness and safety estimates

Estimates of the hazard ratios of the composite outcome of stroke, major bleedings, myocardial infarction, and death in individuals treated with warfarin vs DOACs are presented in Table 6.4. The hazard ratio estimate adjusted with weights unstratified by exposure was 15 % more in favor of DOACs than the hazard ratio estimate adjusted with exposure-stratified weights (1.41; 95%CI:1.34,1.48 and 1.26; 95%CI:1.20,1.33, respectively). Finally, the hazard ratio estimate without any adjustments for informative censoring was close to the estimate adjusted with unstratified weights (1.38; 95%CI:1.32,1.45). The standard error for the estimated log-hazard ratio was only 1.1 % greater when using exposure-stratified versus unstratified weights (0.0264 and 0.0261, respectively).

Results from sensitivity analysis were consistent with the results obtained using the exposure-stratified weights (Annex 16).

6.5 Discussion

In this study comparing the continuous use of warfarin versus DOAC, we showed the different associations of prognostic variables with censoring in treatment groups. We found that prognostically meaningful covariates, such as renal and heart failure, had the opposite direction of association with the probability of censoring in the DOAC and warfarin groups. Applying the exposure-unstratified censoring weights did not result in any change in the HR estimate compared to the unadjusted analysis. Conversely, using exposure-stratified censoring weights changed the estimate by 15 % compared to unstratified censoring weights. Together, these results suggest that in this particular scenario, IPCWs estimated using the exposure-stratified model may better capture the true probabilities of censoring associated with each covariate, and, hence, better correct for informative censoring. Although the direction of the HR estimates did not change, the magnitude of the difference is clinically significant.

Randomized controlled trials showed non-inferiority and even superiority of DOACs over warfarin in stroke prevention[293-295]. A recent meta-analysis of observational studies confirmed similar or lower risks of major bleeding and stroke with DOACs versus warfarin[104]. However, in observational studies comparing an older and a newer drug, the channeling may result in baseline confounding between treatment groups[112]. Soon after entering the market, DOACs were prescribed to individuals with fewer comorbidities and concomitant medications, and those with a lower risk of stroke or bleeding[113-116]. This conservative DOAC prescribing was likely due to existing evidence of a higher risk of gastrointestinal bleeding and myocardial infarction in individuals on

dabigatran[294], as well as the absence of an antidote to reverse the action of DOACs. Further, studies reported that over the course of treatment, older patients and those with a higher risk of bleeding were switched from a DOAC to warfarin[181,182]. Consistently, in this study, we found that in the DOACs group, individuals who were censored at the deviation from baseline treatment were those with higher odds of a new-onset renal or heart failure. Thus, in observational studies comparing the effectiveness and safety of DAOCs and warfarin, worse outcomes in individuals on warfarin compared to those of DOACs may partially be due to informative censoring.

Other situations are possible when treatment discontinuation may result in informative censoring and selection bias, e.g., low-risk patients with transient atrial fibrillation requiring a few months of anticoagulation; or high-risk patients developing a new health condition that increases their risk of bleeding when no form of anticoagulation is allowed. However, most likely, the probability of censoring for these patients would be similar across treatment groups, hence, the direction of selection bias would be similar. Indeed, our sensitivity analysis showed that the differences in characteristics of those censored in the DOACs and the warfarin group were driven by censoring due to the switch between treatment groups.

Stratifying by exposure groups to derive the inverse probability weights was used in some previous studies. Cook et al. re-analyzed the Women's Health Study, a randomized trial of aspirin versus placebo on prevention of cardiovascular disease, to estimate the effect of continuous aspirin use on cardiovascular prevention. After identifying that predictors of aspirin use differed according to initial randomization and past aspirin use, the authors estimated the inverse probability of treatment weights separately for each group[166]. Webster-Clark et al., in their recent paper contrasting warfarin

and dabigatran initiators, used exposure-stratified censoring weights "to account for differential discontinuation and switching in the two groups"[179]. To our knowledge, our study is the first to formally present the differential association of prognostic factors with censoring for each treatment group and show it affects the HR estimates.

Our study has several limitations. First, the validity of our analyses adjusted with the IPCWs relies on the exchangeability assumption that is not empirically testable. Some information on potential prognostic factors associated both with censoring and the study outcome is not available in the data (e.g. smoking, alcohol consumption, income, education). However, we built our censoring models with the best available variables such as chronic obstructive pulmonary disease, diseases related to alcohol abuse, hypertension, atherosclerosis, and the use of cholesterol-lowering drugs (statins). Next, we matched individuals on warfarin to those on DOACs, thus modifying the population of interest from all individuals with atrial fibrillation to only those for whom we found a match. This results in lacking external validity for all AF patients. Indeed, before matching, individuals on warfarin had more comorbidities, including renal failure, coronary atherosclerosis, and heart failure, while individuals on DOACs were younger and had fewer comorbidities. It must be noted that not all warfarin initiators may receive treatment with DOACs, as there are some absolute contraindications to DOAC therapy, such as hemodialysis or mechanical heart valves. Therefore, the parameter of interest of this study is more precisely described as "the safety and effectiveness of continuous use of warfarin compared to DOACs among warfarin users eligible to receive a DOAC". This restriction of the external validity is, by definition, applicable to all comparative effectiveness studies of warfarin vs DOACs. Finally, in our study, we did not have exact

measurements of kidney function, instead, we used ICD-9 and ICD-10 codes for diagnoses of chronic kidney disease and procedural codes for initiation of hemodialysis. This may restrict the identification of patients with renal failure in whom DOACs were contraindicated, i.e., dabigatran is contraindicated when the estimated glomerular filtration rate (eGFR) is <30 mL/min, rivaroxaban when it <30 mL/min (at time of the study) or 15 mL/min (at present), and apixaban when it is <15 mL/min. Therefore, the results of our study may be subject to residual confounding. This may bias the study estimate in favor of the DOAC group.

Importantly, our study does not contain a simulation to show that using exposure-stratified weights reduces bias. Such a simulation model was provided elsewhere by our group for a related example and showed that ignoring the distinct associations of prognostic variables with censoring according to exposure groups may lead to bias[161]. In this study, we sought to illustrate the distinct associations of prognostic variables with censoring in DOACs and warfarin users obtained from a real-life dataset.

Our study has notable strengths. First, our case study is of clinical importance as DOACs represent the most significant change in atrial fibrillation management of the last decade. In the US Medicare beneficiaries, utilization of DOACs increased from 13 % of all anticoagulants in 2012 to 53 % in 2017[296]. However, the results from studies comparing the safety and effectiveness of sustained treatment with DOACs vs. warfarin may be biased because of the failure to adjust for different censoring dynamics. In clinical practice, physicians are well-aware of the specific contra-indications to DOACs, yet those are not always correctly considered in epidemiological studies. A recent study reported that compared to DOACs, continuous use of warfarin was associated with a

40% higher risk of the composite of ischemic/hemorrhagic stroke, gastrointestinal bleeding, and all-cause mortality (HR:1.40, 95%CI:1.33, 1.45)[297]. This estimate, which is unadjusted for informative censoring, is close to our HR estimates obtained with unadjusted and adjusted with unstratified censoring weights analyses. Another strength of our study is the use of a large pharmaceutical dispensation claims database that represents about 40 % of the Québec population and is recognized for its high accuracy, thus minimizing the measurement error. The information on exposure to medications is almost complete because, in Québec, all prescription drugs are covered by the public drug insurance plan. Further, because of universal entitlement, disenrollment from the RAMQ drug plan is uncommon[97]. An exception is those aged under 65 years who may switch to private insurance in case of changes in their employment status. However, the influence of such cases on the study results is likely to be negligible as 75% of our study population was aged over 70 years (Table 7.2). For this reason, we did not consider censoring due to disenrollment as informative. In other settings, particularly where healthcare insurance depends on employment status, disenrollment should be considered informative censoring.

In this study, we showed that contextualizing and understanding the specifics of clinical management according to treatment groups is important for the correct specification of a censoring weight model. We propose stratifying probabilities of censoring by exposure as a simple and effective method to identify the distinct associations of prognostic variables with censoring in treatment groups, which, if present, must be accounted for to effectively correct selection bias.

6.6 Tables and figures

Table 6.1. ICD codes used to identify the cohort of individuals treated with warfarin vs DOACs in the Régie de l'assurance-maladie du Québec (RAMQ) databases between 2010 and 2016.

	ICD-9	**ICD-10**
Atrial fibrillation and flutter	427.3	I48.x
Venous thromboembolism (VTE), including deep vein thrombosis (DVT) and pulmonary embolism (PE)	415.1, 451.1, 451.2, 451.81, 451.83, 451.89, 451.9, 453.2, 453.8, 453.9	I80.2, I80.3, I80.1, I80.8, I80.9, I82.8, I82.9, O22.3, O22.30, O22.9, O22.90, O87.1, O87.10, I26.9, I26.0
Valvular heart disease	391.x, 394.x, 395.x, 396.x, 397.x, 398.x, 424.0, 424.1, 424.2, 424.3, 746.x	I01.x, I05.x, I06.x, I07.x, I08.x, I09.x, I43.x, I35.x, I36.x, I37.x, I38, I39.x, Q24.x

Table 6.2. Characteristics of the base and matched cohorts of individuals treated with warfarin vs DOACs. Québec, 2010-2016.

Characteristic	Base cohort					Matched cohort				
	DOACs		Warfarin		SD	DOACs		Warfarin		SD
	No	%	No	%		No	%	No	%	
Number of patients	40,101		30,743			23,927		23,927		
Age in years, median (IQR) [a]	76 (69-83)		78 (71-84)		0.183	78 (71-84)		78 (71-84)		0.005
Male sex [b]	20,802	51.9	15,703	51.1	0.016	12,172	50.9	12,172	50.9	0.000
Comorbidities [b]										
Hypertension	37,460	93.4	29,442	95.8	0.104	22,832	95.4	22,858	95.5	0.005
Congestive heart failure	6,283	15.7	7,964	25.9	0.254	4,637	19.4	5,181	21.7	0.056
Coronary atherosclerosis	12,909	32.2	14,692	47.8	0.323	9,486	39.6	9,939	41.5	0.039
Diabetes	10,182	25.4	10,234	33.3	0.174	6,928	29.0	7,297	30.5	0.034
Peripheral vascular disease	3,864	9.6	5,234	17.0	0.219	2,987	12.5	3,290	13.8	0.038
Chronic renal failure	2,908	7.3	5,790	18.8	0.349	2,590	10.8	2,966	12.4	0.049
Cancer	6,357	15.9	4,869	15.8	0.000	3,607	15.1	3,878	16.2	0.031
Acute kidney injury	2,369	5.9	3,773	12.3	0.223	1,809	7.6	2,300	9.6	0.073
COPD	6,149	15.3	6,369	20.7	0.140	4,148	17.3	4,544	19.0	0.043
Liver disease	738	1.8	780	2.5	0.048	472	2.0	565	2.4	0.027
Dementia	2,885	7.2	3,005	9.8	0.093	2,028	8.5	2,205	9.2	0.026
Diseases related to alcohol abuse	1,387	3.5	1,188	3.9	0.022	799	3.3	909	3.8	0.025
Bleeding	3,304	8.2	3,245	10.6	0.079	2,141	8.9	2,374	9.9	0.033
Stroke	4,114	10.3	4,817	15.7	0.162	3,074	12.8	3,244	13.6	0.021
Medical procedures [b]										
Pacemaker implantation/ Catheter ablation	2,664	6.6	2,410	7.8	0.046	1,708	7.1	1,713	7.2	0.001
PCI/CAB	1,011	2.5	2,731	8.9	0.277	866	3.6	1,106	4.6	0.050

Table 6.2. Characteristics of the base and matched cohorts of individuals treated with warfarin vs DOACs. Québec, 2010-2016.

Characteristic	Base cohort					Matched cohort				
	DOACs		Warfarin		SD	DOACs		Warfarin		SD
	No	%	No	%		No	%	No	%	
Medications [b]										
Antiplatelets	20,803	51.9	18,978	61.7	0.200	14,112	59.0	14,001	58.5	0.009
Antipsychotics	2,103	5.2	1,697	5.5	0.012	1,197	5.0	1,297	5.4	0.019
NSAIDs	2,010	5.0	1,234	4.0	0.048	1,000	4.2	1,013	4.2	0.003
Statins	21,335	53.2	18,458	60.0	0.138	13,659	57.1	13,932	58.2	0.023
Benzodiazepines	9,078	22.6	8,697	28.3	0.130	6,172	25.8	6,268	26.2	0.009
Systemic corticosteroids	1,878	4.7	1,721	5.6	0.041	1,193	5.0	1,309	5.5	0.022
H-2 receptor blockers	389	1.0	333	1.1	0.011	233	1.0	257	1.1	0.010
Proton pump inhibitors	16,977	42.3	15,239	49.6	0.146	10,618	44.4	11,227	46.9	0.051
Opioids	3,603	9.0	2,532	8.2	0.027	1,801	7.5	1,947	8.1	0.023
SSRI	3,139	7.8	2,487	8.1	0.010	1,832	7.7	1,951	8.2	0.018

Abbreviations: CAB, coronary artery bypass surgery; COPD, chronic obstructive pulmonary disease; IQR, interquartile range; NSAIDs, non-steroidal anti-inflammatory drugs; No, number; PCI, percutaneous coronary intervention; SD, absolute standardized differences; SSRIs, selective serotonin reuptake inhibitors.

[a] continuous variable, starting from the year 2010

[b] dichotomous variable

Table 6.3. Odds ratio with 95%CI of being censored associated with patient characteristics using unstratified and exposure-stratified logistic regression, n=47,854, Québec, 2010-2016.

Characteristic	Unstratified		Exposure-stratified			
			DOACs		Warfarin	
	OR	95% CI	OR	95% CI	OR	95% CI
DOACs vs warfarin	0.39	0.37, 0.40	-	-	-	-
Month of follow-up [a]	0.97	0.97, 0.97	0.96	0.96, 0.96	0.97	0.97, 0.97
Year of cohort entry [a]	1.00	0.99, 1.01	0.88	0.87, 0.90	1.05	1.04, 1.06
Age [a]	0.99	0.99, 0.99	1.00	1.00, 1.00	0.98	0.98, 0.98
Male sex () [b]	1.06	1.03, 1.08	1.14	1.09, 1.20	1.01	0.97, 1.04
Comorbidities (yes vs no) [b]						
Hypertension	0.76	0.71, 0.81	0.61	0.56, 0.67	0.89	0.82, 0.97
Congestive heart failure	1.04	1.01, 1.07	1.14	1.08, 1.20	0.99	0.96, 1.03
Coronary atherosclerosis	1.01	0.98, 1.04	1.01	0.96, 1.06	1.01	0.98, 1.05
Diabetes	0.93	0.91, 0.96	1.01	0.96, 1.06	0.89	0.86, 0.92
Peripheral vascular disease	1.03	0.99, 1.07	1.02	0.96, 1.09	1.03	0.98, 1.07
Chronic renal failure	1.01	0.97, 1.06	1.34	1.25, 1.43	0.86	0.82, 0.91
Cancer	1.20	1.16, 1.24	1.21	1.15, 1.27	1.19	1.14, 1.23
Acute kidney injury	1.01	0.96, 1.05	1.17	1.08, 1.26	0.96	0.90, 1.01
COPD	1.01	0.98, 1.04	0.99	0.93, 1.04	1.02	0.98, 1.06
Liver disease	1.04	0.96, 1.12	1.12	0.99, 1.27	0.98	0.89, 1.07
Dementia	1.25	1.19, 1.30	1.52	1.42, 1.63	1.12	1.07, 1.18
Diseases related to alcohol abuse	1.12	1.05, 1.18	1.12	1.02, 1.24	1.11	1.03, 1.20
Medical procedures (yes vs no)[b]						
PCI/CAB	1.48	1.39, 1.57	1.45	1.30, 1.61	1.47	1.37, 1.59
Pacemaker implantation/ Catheter ablation	1.07	1.03, 1.12	1.11	1.04, 1.18	1.06	1.01, 1.12
Medications (yes vs no)[b]						
Antiplatelets	1.44	1.40, 1.48	1.87	1.77, 1.97	1.25	1.21, 1.30
Antipsychotics	1.02	0.97, 1.08	1.08	0.98, 1.18	0.97	0.91, 1.04
NSAIDs	1.26	1.17, 1.35	1.33	1.20, 1.47	1.20	1.09, 1.32
Statins	0.80	0.77, 0.82	0.64	0.61, 0.67	0.89	0.86, 0.93
Benzodiazepines	0.99	0.96, 1.02	1.00	0.95, 1.05	0.99	0.95, 1.03
Systemic corticosteroids	1.24	1.17, 1.30	1.22	1.12, 1.34	1.24	1.16, 1.33
H-2 receptor blockers	1.04	0.92, 1.17	0.92	0.74, 1.14	1.06	0.91, 1.23
Proton pump inhibitors	0.93	0.90, 0.95	0.89	0.85, 0.93	0.96	0.92, 0.99
Opioids	1.53	1.46, 1.60	2.16	2.02, 2.31	1.21	1.14, 1.28
SSRI	0.93	0.89, 0.97	0.98	0.90, 1.06	0.91	0.86, 0.96

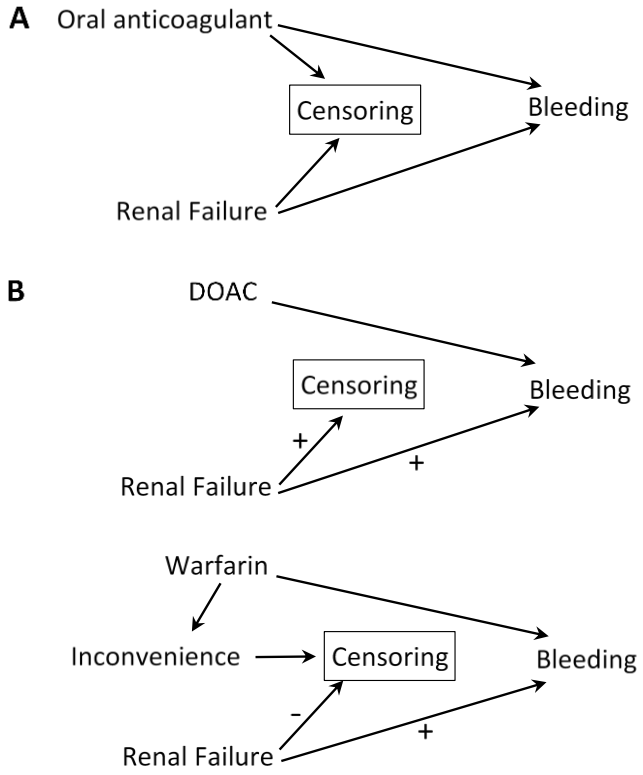
Abbreviations: CAB, coronary artery bypass surgery; CI, confidence interval; COPD, chronic obstructive pulmonary disease; NSAIDs, non-steroidal anti-inflammatory drugs; OR, odds ratio; PCI, percutaneous coronary intervention; SSRIs, selective serotonin reuptake inhibitors
[a] continuous variable, [b] dichotomous variable.

Table 6.4. Hazard ratios of the composite of stroke, major bleeding, MI, and death in individuals treated with warfarin vs DOACs (n=47,854) using different definitions of exposure and censoring adjustment. Québec, 2010-2016.

Estimate	HR	95% CI	SE
Unweighted	1.35	1.28, 1.42	0.0252
Weighed with unstratified weights	1.41	1.34, 1.48	0.0261
Weighted with exposure-stratified weights	1.26	1.20, 1.33	0.0264

Abbreviations: DOACs, direct oral anticoagulants; CI, confidence interval; HR, hazard ratio; MI, myocardial infarction; SE, standard error.

Figure 6.1 . Structure of bias due to informative censoring in a study of the risk of bleeding in individuals treated with oral anticoagulants.



A. Treatment with oral anticoagulants (DOACs or warfarin) increases the risk of bleeding.
Renal failure also increases the risk of bleeding, and it is a contraindication for DOACs. When an individual on a DOAC develops renal failure, the DOAC is discontinued and the follow-up of this individual is censored. Censoring is a collider on the path DOAC→Censoring←Renal Failure→Bleeding. Restriction of the analysis to those uncensored (a box around censoring) will result in biased estimates.
B. The effect of renal failure on censoring is different in two exposure groups. In the DOACs group, those with renal failure would be censored. As such, individuals without renal failure who have a low risk for bleeding would remain in the study. The cumulative rate of bleeding in the DOAC group would be underestimated.
In the warfarin group, individuals may be censored when they switch to a DOAC for convenience purposes and in the absence of renal failure. Therefore, those with renal failure and a higher risk of bleeding would remain in the study. The cumulative rate of bleeding in the warfarin group would be overestimated.
Adapted from Howe et al.[291]

Figure 6.2. Kaplan-Meier plot of the composite outcome-free survivorship by treatment, n=47,854, Québec, 2010-2016.

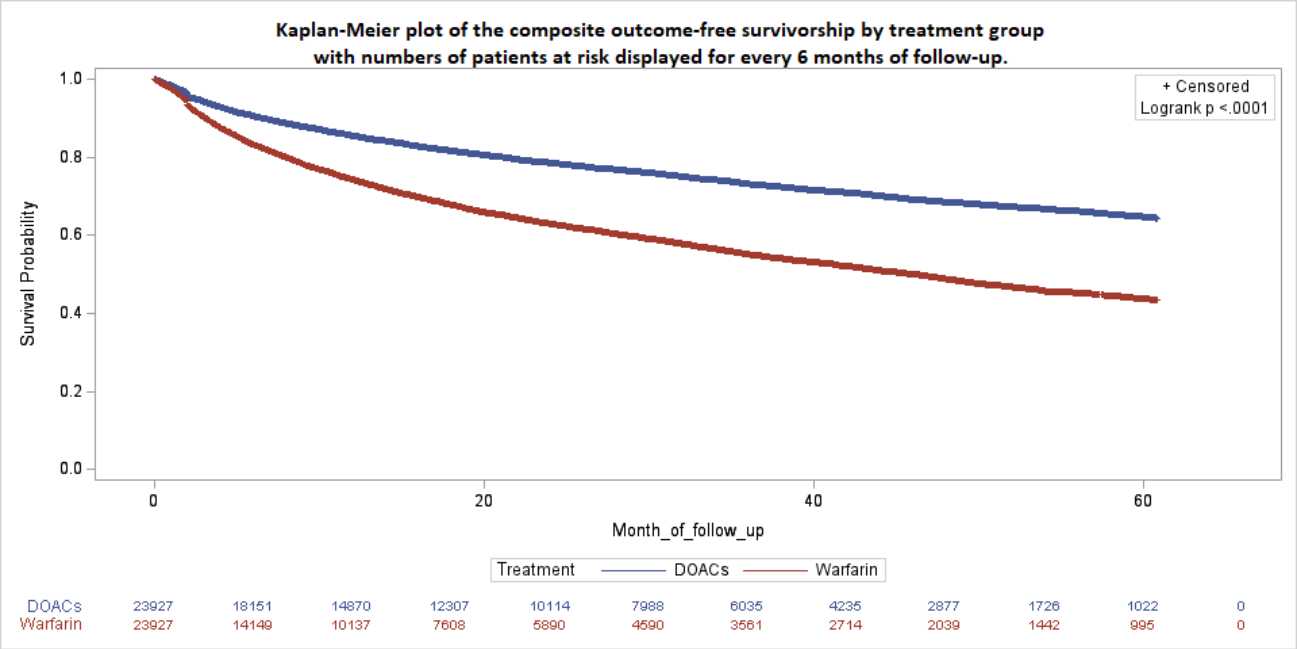
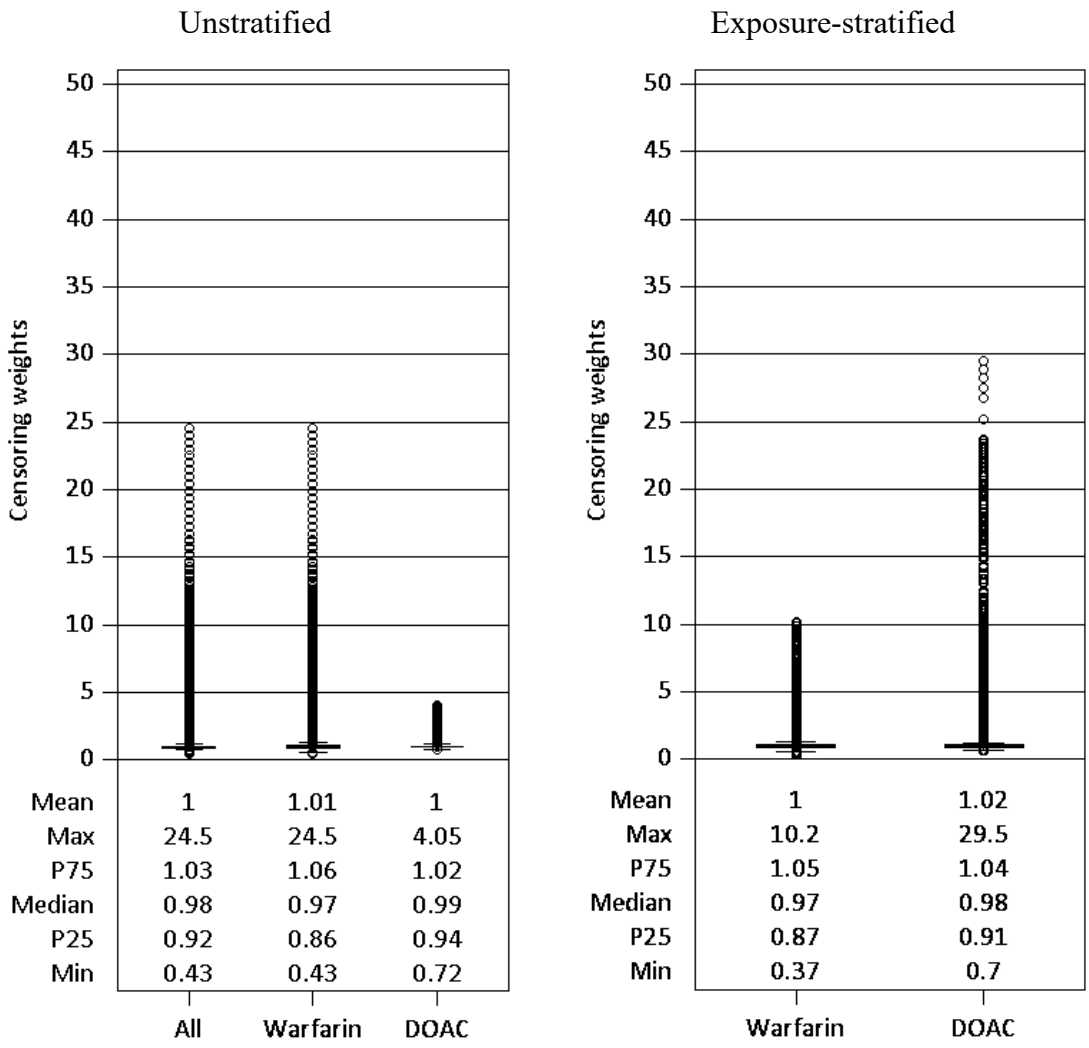


Figure 6.3. Distribution of IPCWs for individuals treated with warfarin and DOACs, estimated with unstratified and exposure-stratified censoring models, Québec, 2010-2016.



Abbreviations: DOACs, direct oral anticoagulants; P25, 25th percentile; P75, 75th percentile.

CHAPTER 7: Discussion

7.1 Summary of research

The overarching goal of this work was to advance methodological practices of pharmacoepidemiological studies on the comparative effectiveness and safety of oral anticoagulants in AF. Using the Québec administrative databases, we addressed two biases: misclassification of exposure to a dose-varying warfarin therapy when using pharmacy claims data, and selection bias due to censoring of follow-up at the time of treatment discontinuation in individuals exposed to warfarin and those exposed to DOACs.

The first two studies explored the potential for misclassification of exposure when using pharmacy claims, and its effect on the measure of association comparing the safety of DOACs and warfarin. The objectives of the first study were to characterize dispensation patterns and to measure gaps between expected and observed refill dates for warfarin, rivaroxaban, dabigatran, and apixaban. We found that for all oral anticoagulants, the most common duration of dispensation was seven days, suggesting the use of pharmacist-prepared weekly pillboxes. After matching individuals treated with warfarin and each DOAC, the average refill gap was greater for warfarin compared to DOACs. The variance of refill gaps was also greater among warfarin users than among DOAC users. Based on expert knowledge and clinical experience, we concluded that these patterns of refill gaps in individuals treated with warfarin may reflect the adjustments of warfarin posology by health professionals based on INR results. Therefore, defining the duration of warfarin dispensation based on the days supplied value may lead to misclassification of exposure to warfarin.

From the literature, we identified that a data-driven approach (longitudinal coverage approximation, COV) to define the duration of dispensation may account for the past medication use patterns and accurately approximate the duration of dispensation for medications with large dosage variability, such as vitamin K antagonist phenprocoumon. Thus, the objectives of my second study were to evaluate if the data-driven approach better captures the variability in the duration of warfarin dispensations compared to the days supplied approach; and to investigate if the choice of the approach affects the strength of association between major bleeding and treatment with dabigatran or rivaroxaban versus warfarin. The study found that the data-driven approach captured more variability for warfarin compared to the days supplied approach. However, the choice of the approach had no impact on the HR estimates. These findings may be explained by the fact that misclassification of exposure, which was more important for the patients exposed to warfarin than those exposed to DOACs, was nonetheless non-differential, i.e., unrelated to the study outcome. Another explanation is that in our study, a relatively small number of individuals had long gaps between dispensations. Indeed, most of the dispensations were for 7 days, suggesting the use of pharmacist-prepared weekly pillboxes aimed to improve adherence. Finally, the lack of impact of the chosen approach on the HR estimates may be explained by the low sensitivity of the data-driven definition to capture the true duration of warfarin dispensation, and more research is needed to assess its validity.

The third study was designed to address the selection bias due to censoring at the time of treatment discontinuation in individuals exposed to warfarin and those exposed to DOACs. The first objective was to describe the associations between the prognostic variables and censoring in both treatment

groups. The second objective was to compare two IPCW strategies to estimate the measure of association between the continuous treatment with warfarin versus DOACs and the composite of thromboembolism, major bleeding, myocardial infarction, and all-cause mortality. We compared three estimates of the hazard ratio obtained by using unweighted, weighted with exposure-unstratified IPCWs, and weighted with exposure-stratified IPCWs analyses. We found that prognostically meaningful covariates, such as renal and heart failure had the opposite direction of association with the probability of censoring in the DOACs and the warfarin groups. Censoring weights not accounting for this differential attrition failed to correct for selection bias and resulted in the estimate of the comparative treatment effectiveness and safety that favored DOACs. Using exposure-stratified censoring weights changed the estimate of comparative effectiveness and safety by 15% compared to unstratified censoring weights. With this study, we were able to demonstrate that not accounting for the different associations between the predictors and informative censoring across exposure groups may lead to misspecification of censoring weights and biased estimate of comparative effectiveness and safety.

7.2 Contribution to literature

There is a growing number of publications focusing on specific limitations inherent in studies based on the analysis of large administrative healthcare datasets[97,125]. Our research adds to this body of literature by highlighting how the specific clinical considerations for the use of warfarin and DOACs must be accounted for when designing an observational study to ensure the validity of study conclusions.

To our knowledge, no previous studies assessed the potential for misclassification of exposure to warfarin when using pharmacy claims. However, some recognized that the true duration of warfarin dispensation may last longer than the recorded duration, and considered this issue in their design. For example, Go et al.[48] developed and validated an algorithm that combined data from pharmacy claims with data on INR measurement from laboratory files. Azoulay et al.[176,177] adapted the algorithm developed by Go et al.[48] in their studies using the General Practice Research Database (CPRD). Webster-Clark et al.[178,179] in their studies based on an analysis of the US Medicare data, used a definition of exposure time that was tailored to the specific classes of oral anticoagulants. Thus, the length of the grace period was different for dabigatran and warfarin. Additionally, the Medicare *current procedural terminology* (CPT) codes for anticoagulation management were used to "refresh" the warfarin supply for another 30 days[178]. However, information on the validation of this method is not available.

The results of this thesis extend beyond identifying the potential for misclassification of exposure to warfarin using the days supplied variable (First study) by examining the possible impact of this misclassification on the estimates of the comparative safety of DOACs and warfarin, and by implementing the data-driven approach that better captured the variability in the duration of warfarin dispensations compared to the days supplied (Second study). We recommend the data-driven approach of defining exposure to warfarin for settings where no laboratory data on INR measurement is available, or for settings where laboratory data contains substantial missing information on INR measurement.

The results of the third study improved the understanding of the source and direction of selection bias due to informative censoring that may differentially affect the warfarin and the DOACs treatment groups. Previously, Webster-Clark et al.[179] in their study comparing dabigatran and warfarin-treated individuals, recognized this bias and corrected it in the analysis by using exposure-stratified IPCWs. However, to our knowledge, no previous literature described the conceptual framework for this bias and demonstrated the differential effect of prognostic factors on treatment discontinuation in warfarin and DOACs-exposed individuals. Our study is the first to formally illustrate its structure for each treatment group using directed acyclic graphs (DAG) and provide insights into the different dynamics of censoring in two treatment groups (Figure 6.1, Chapter 6:). In general, DAGs are intended to elucidate and visualize the possible source of confounding or bias in a study with the ultimate goal of improving its validity and transparency[298]. Thus, our study adds to the literature encouraging the usage of DAGs in different areas of medical research, for example, in cardiovascular[299], pulmonary medicine[300], or pediatric research[301]. Next, to address the selection bias due to censoring, we focused on the correct specification of the statistical model for the estimation of censoring weights, the exposure-stratified model. In the settings of anticoagulation in NVAF, Webster-Clark et al.[179] used exposure-stratified censoring weights "to account for differential discontinuation and switching in the two groups" when comparing dabigatran vs warfarin initiators. However, in our study, we showed that censoring weights that were not stratified by treatment group produced an estimate of comparative effectiveness and safety that was biased in favor of DOACs.

7.3 Informing methodological best practices

Previous observational studies on the comparative effectiveness and safety of warfarin and DOACs have acknowledged the potential for information bias when using the days supplied value from pharmacy claims, and the selection bias due to the differential effect of prognostic factors on treatment discontinuation across exposure groups.

The important contribution of this work to the literature is the recommendation of the data-driven approach to define the duration of dispensation as an alternative to the approach based on the days supplied value. We concluded that the data-driven approach that relies on the individual dispensing history may better account for between- and within-individual variability of warfarin dosage. Furthermore, when using administrative databases without supporting information on INR measurements, the data-driven approach may reduce the misclassification of exposure to warfarin. Future studies based on the analysis of administrative data may use the data-driven approach to define exposure to warfarin for their main analysis or sensitivity analyses.

Currently, there are limited recommendations on the correct specification of the statistical model to estimate censoring weights. These recommendations mainly concern the choice of predictors of censoring to include in the model and their functional form. In our study the functional form was stratification by exposure; alternatively, interaction terms of predictors of censoring with exposure can be used. To select the correct functional form of the predictor, one needs to understand the mechanism of censoring in each exposure group. In our study, we recommend using DAGs and examining the association of important prognostic variables with the probability of censoring in each exposure group separately. This recommendation should be applied in all observational

studies comparing newer with older therapies. In such settings, the newer therapy may have specific target groups and contraindications, whereas, the older therapy may be applied to broader groups of patients and be regarded as a mainstay. Furthermore, our recommendation may be pertinent not only to observational but also to experimental studies that perform an analysis with censored data.

7.4 Strengths and limitations

The strengths and limitations of each study are described in the relevant chapters. However, several global strengths and limitations of this thesis are restated in this section.

7.4.1 Misclassification of exposure

The refill gaps were used to capture the deviations from the posology from the days supplied recorded in pharmacy claims. We hypothesized that these deviations would be common in those on warfarin due to frequent dosage adjustments. The underlying assumptions of the research were, first, similar adherence to warfarin and DOACs in the study population, and second, accurate and complete information on the days supplied variable recorded in pharmacy claims. Similar adherence in the study population assumes that the difference in the variation of refill gaps is explained by inadequate capture of the true medication regimen. Thus, hypothetically, different levels of adherence to warfarin and DOACs would influence the conclusions of the study. Indeed, some previous studies based on the analysis of pharmacy claims data reported higher levels of adherence in those taking DOACs than those taking warfarin[189,302]. However, in these studies, the adherence rates were assessed using the refill dates and the days supplied variable that may not reflect the true treatment patterns of patients treated with warfarin. Nevertheless, a study using medical records of patients taking oral anticoagulants reported similar adherence between DOACs and warfarin-treated patients[276].

In the first study, the median follow-up for apixaban was the shortest (0.86 years) among all oral anticoagulants (2.1 years for dabigatran, 1.4 years for rivaroxaban, and 1.5 years for warfarin). This can be explained by the fact that in Quebec, apixaban became available only in 2014, whereas

dabigatran in 2010 and rivaroxaban in 2012. In the first study, because of data availability, the end of study date was March 31, 2015. Therefore, for patients on apixaban, we captured only the first months of their treatment, during which the medication adherence may be higher than during the later months[283]. However, in the matched individuals, among all DOACs, the mean refill gap was only slightly shorter for apixaban (3.08 days) compared to dabigatran (3.7 days) and rivaroxaban (3.15 days).

In the first study, the refill gaps were investigated not to measure adherence but to capture deviations from prescribed posology. Under the assumption that adherence to warfarin and DOACs is similar at a population level, any significant increase in refill gaps for a given medication would indicate that days supplied data is an imprecise measure of the actual number of days of availability of that drug. The first study found that the cluster-adjusted average refill gap was more than twice as long for warfarin (9.28 days) than for apixaban (3.08 days), dabigatran (3.7 days), and rivaroxaban (3.15 days), and the within- and between-patient variation of refill gaps was more than twice as higher in warfarin compared to DOAC patients. These findings suggest that such difference in refill gaps for warfarin may be due to inadequate capture in the dispensation data of the periods covered by the number of dispensed pills.

We believe that the assumption of similar adherence was held in our population given the strong recommendations from the healthcare providers for users of DOACs, and routine monitoring visits to healthcare facilities of those treated with warfarin. Moreover, in the first study, we found that before matching, at least half of patients taking any oral anticoagulant, had a duration of the dispensation of 7 days and a refill gap of 0 days, which gives a medication possession ratio of

100%. This high adherence to DOACs and warfarin strengthens our conclusions; however, our study was not designed to assess the population adherence rates, as we had no data that measured the true adherence.

Next, we assumed that information in pharmacy claims accurately represents the information from physician prescriptions. We did not perform a prior validation study, and to our knowledge, there is no published research on the validity of information on oral anticoagulants in the Québec pharmacy claims. The lack of validation studies may be explained by the strict data protection regulations in Québec, which make such research operationally difficult and time-consuming. However, we believe that information in the Québec pharmacy claims is accurate since is it used for billing purposes and RAMQ performs routine internal validation of claims data, including inspection visits and pharmacy prescription requests[202]. Moreover, an external validation study of all dispensations for a random sample of 65,349 elderly Québec residents found that information in the pharmacy claims is accurate and comprehensive[131]. The missing information was less than 0.5% for commonly used variables, including an individual identifier, drug name, dispensed quantity, date of dispensation, and intended duration of dispensation. Furthermore, the same study verified data from pharmacy claims and medical files for 306 individuals. The study found that the dispensed quantity and the duration of dispensation were accurate in 69.1% and 72.1% of records, respectively. The likely explanation for some inaccurate information on the dispensed quantity and duration is that the pharmacist may split a prescription with an extended duration on several dispensations. This is done for various reasons: to encourage treatment adherence by increasing the number of communications of pharmacists with the patient; to restrict the available drug supply

for patient's safety purposes; to avoid the financial penalties for dispensations with more than 30-day supply[131].

The next limitation of this research is that in the first and second studies, the interpretation of the results was founded on expert opinion and clinical experience on utilization patterns of DOACs and warfarin. There is no "gold standard" to conclude that refill gaps may be associated with the prolongation of warfarin supply due to changes in warfarin dosage based on INR results. Ideally, the performance of different approaches for estimation of the duration of dispensation should be validated against the duration of dispensation estimated with the actual warfarin doses derived from medical records.

We defined the users of weekly pillboxes as any patients with four successive dispensations of seven-day duration. However, in practice, some patients get the drugs in pillboxes, while others get their drugs in a plastic vial (this is the case when the dose of the drug can be changed) even if their other drugs are in the pillbox. Therefore, adherence of those who receive their 7-day supply of warfarin in a plastic vial can be compromised if the patient does not manage well with the plastic vial.

The findings of the first and second studies are specific to the RAMQ databases and may not be generalizable to other administrative databases. Thus, in our database, for all oral anticoagulants, the most common duration of dispensation was 7 days, for most dispensations, there are no refill gaps. Again, this may be explained by the usage of pharmacy-prepared weekly pillboxes or plastic vials at community pharmacies in Québec, and are aimed to improve clinical outcomes in the population[277]. It is possible, that in settings with different community pharmacy interventions and

levels of adherence, the length of refill gaps might be longer. Hence, the approach used to define the duration of dispensations may have a different impact on the comparative safety estimates.

In this research, we assumed that patients followed prescribed treatment based on information from dispensation claims, but the true exposure status, as well as TTR for patients on warfarin, was unknown. Thus, the medication dispensed does not necessarily indicate exposure to this medication, as a drug that is dispensed may not be ingested by the patients[118]. The same may apply to users of weekly pill boxes for whom prescriptions may be automatically renewed at pharmacies even if the previous supply is unused. Furthermore, in patients on warfarin, who are compliant with the prescribed therapy, dietary fluctuations, food supplements, or over-the-counter agents may lead to alteration of the optimal therapeutic INR level. The information on patient diet, use of food supplements or over-the-counter agents, as well as INR measurements are not available in RAMQ data.

7.4.2 Selection bias due to censoring

In the third study, the IPCWs were used to correct the selection bias introduced by censoring at the time of discontinuation of treatment dispensed at baseline. In general, the IPCWs might fail to eliminate the selection bias in case of violation of the exchangeability assumption. The exchangeability assumption implies that those who are censored and those who stay in the study are exchangeable, i.e., if these groups were swapped, those who were censored would experience the same outcome as those who stayed in the study. This assumption holds under the condition that all predictors of censoring and the outcome event are measured and included in the statistical model estimating censoring weights. As such, this assumption is not empirically testable and depends on

the recorded information in the dataset and expert knowledge. As described in the third study, some potential prognostic factors associated both with censoring and the study outcome are not available in the data (e.g., smoking, diet, alcohol consumption, income, education, creatinine measures). However, we built our censoring models with the best available variables such as diagnoses of chronic obstructive pulmonary disease, chronic renal failure, acute kidney injury, diseases related to alcohol abuse, hypertension, atherosclerosis, or use of cholesterol-lowering drugs statins. Thus, we assume that the most important prognostic factors associated with censoring or their proxies were identified using expert knowledge and were included in the model for the estimation of censoring weights.

This study aimed to improve the methodology of observational studies comparing DOACs and warfarin. The composite effectiveness and safety outcome was used to make our argument more compelling in making a stronger case for the application of exposure-stratified censoirng weights. Having more outcome events improved the precision of the HR estimate. In this study, we hypothesized that because of differential selection bias across the treatment groups, each component of the composite outcome may be biased in favor of DOACs. However, there are some disadvantages when using the composite outcome. In general, in studies on comparative effectiveness and safety, the study conclusions may be misleading, and demonstrated treatment benefits may be assumed to relate to all components when, in fact, only a few events of an individual component may contribute to the composite outcome. This may be more starkly if the components include the most important clinical events (e.g., myocardial infarction, a serious bleeding event requiring hospitalization, or death) and less serious components (e.g., bleeding not

requiring hospitalization). To communicate the study findings in a publication more clearly, it is recommended to present data for all individual components[263]. In our research, each component of the composite outcome was assessed in a previous study conducted by our research group (Durand, *CMAJ*, 2020)[264]. The study found no difference in the incidence of ischemic stroke or systemic embolization in patients with NVAF (HR: 0.98, 95%CI: 0.77-1.25), however, the use of DOACs was associated with fewer major bleeding events (HR: 0.71, 95%CI: 0.64-0.80)[264].

7.4.3 Other strengths and limitations

In the second and the third study we presented estimates on the comparative effectiveness and safety of DOACs and warfarin. We found that at baseline, compared to patients from the DOACs groups, those from the warfarin group were frailer and older, and a higher proportion had a history of previous stroke or bleeding. To reduce the confounding, we used PS methods. We estimated the PSs from variables that may be associated with the treatment choice (DOACs or warfarin) and the probability of the outcome event. These variables included co-morbidities, medical procedures, and co-medications. In the second study, the hazard ratio of the association between treatment with DOACs versus warfarin and major bleedings was adjusted for the deciles of the PSs and year of cohort entry. In the third study, the HRs were estimated using the cohort of individuals treated with DOACs and warfarin that were matched on age, sex, year of cohort entry, and PSs. Yet, the HR estimates may be subject to residual confounding. For example, the RAMQ databases lack the exact measurements of kidney function. Instead, we used proxy variables, such as diagnoses of chronic kidney disease or acute renal failure. However, these proxy variables may not precisely represent the exact measurements of kidney function to determine patients that have

contraindications to DOACs. Hence, the HR estimates in the second and third studies may favor DOACs, because DOACs are contraindicated in patients with end-stage or severe chronic renal failure. Furthermore, because of the low overall sensitivity of ICD codes, some comorbidities may be missed in our studies, including hypertension, diabetes, chronic renal disease, or dementia. We made an effort to enhance the ascertainment of comorbidities by capturing information from physician claims, hospital discharge data, as well as the use of prescribed medications as proxies for diseases (e.g., use of medications prescribed for treatment of hypertension, or antidiabetic medication).

Unmeasured confounding may arise due to the omission of some potential confounders, such as socioeconomic status, smoking, or body mass index. This may bias the study estimates in an unpredictable direction. For example, the RAMQ public prescription drug insurance imposes monthly deductibles of $22.25 and a patient portion of the co-insurance corresponding to 35% of the cost of covered drugs[303]. Thus, socioeconomically disadvantaged patients may have lower access to DOACs. Furthermore, individuals with low socioeconomic status (SES) are likely to have more chronic conditions, worse cardiovascular outcomes, and lower life expectancy[304]. Therefore, SES is an important confounding that may influence the choice of oral anticoagulant and the risk of thromboembolic or bleeding events. The determinants of socioeconomic status (e.g., education, occupation, or income) are not recorded in administrative data, including RAMQ data. In Quebec, the deprivation index was created to use as a proxy for lacking socioeconomic data in administrative databases[305]. The deprivation index is based on geographical units (enumeration areas or dissemination areas) with a relatively homogeneous population in terms of

socioeconomic status. The geographical units can be matched with the three-digit postal code variable recorded in the RAMQ database. In our research, we did not use the deprivation index because we measured most of the comorbidities included in the CHA2DS2-VASc and HAS-BLED scores, thus, we directly measured characteristics that may be affected by SES. Overall, we believe that all strong confounders were identified either from the previously published literature or based on the clinical experience of our research team. Thus, even if some unmeasured confounding exists, its effect on the HR estimates is likely negligible.

The next limitation is that the coding of diagnoses AF in administrative healthcare data may be inaccurate. A systematic review of previous validation studies reported that the PPV of the ICD-9 code for AF (427.3) ranged from 70% to 96%, with a median of 89%, and the sensitivity ranged from 57% to 95%, with a median of 79%[205]. Thus, some errors may be produced when identifying patients with AF. This may be particularly challenging in studies of incident cases of the disease[306]. In our research, we used a combination of information from pharmacy claims, physician claims, and hospital discharge data. Likely, this improved precision of our case definition[307]. Furthermore, we excluded those with other health conditions requiring oral anticoagulation (i.e., venous thromboembolism, valvular disease, or prior cardiac valve surgery). However, these measures could not improve the sensitivity of the case definition, and likely some patients initiating oral anticoagulation for atrial fibrillation were not included in our study. We assume that this misclassification was non-differential between users of warfarin and DOACs. Next, the accuracy of ICD codes for valvular disease is also limited. A study found that in general, the specificity of ICD-9 codes for valvular heart disease (394.x-397.x, 398.4.x) was high (97%), but sensitivity was

low (41%)[308,309]. Therefore, some patients with valvular disease remained in our study cohort, and likely most of these patients were in the warfarin group. Because patients with the valvular disease have a higher risk of stroke or bleeding, this misclassification may bias the study results in favor of DOACs.

We identified the outcome events of major bleeding, thromboembolism, or myocardial infarction only from hospital discharge data using ICD-10 codes for the primary or principal diagnoses, or the cause of in-hospital death. We did not extract the outcome events from the physician claims to prevent detection bias, since regular outpatient visits of individuals treated with warfarin may increase the likelihood of less important outcomes (especially bleeding) being recorded in the physician claims database.

Joos et al. assessed the accuracy of ICD-10 codes for acute anticoagulation therapy-related bleeding in patients admitted to the University of Utah hospital using patients' medical records as a reference standard. ICD-10 codes were extracted for diagnoses at any position[256]. In total, the study identified 71 bleeding events from the medical records. The reported sensitivity, specificity, PPV, and NPV for intracranial hemorrhage were 100% (95%CI: 78%–100%), 98% (95%CI: 97%–99%), 58% (95%CI: 37%–77%), and 100% (95%CI: 99.%–100%), respectively. For gastrointestinal bleeding, reported sensitivity, specificity, PPV, and NPV were 96% (95%CI: 78%–100%), 97% (95%CI: 96%–98%), 55% (95%CI: 38%–79%), and 100% (95%CI: 99.%–100%), respectively. Hall et al. provided a validation study of ICD-10 codes for any stroke and TIA for 5,270 events identified from Ontario hospital discharge data against the information on these patients in the Ontario Stroke Registry (a part of the province-wide system of stroke care management)[310]. The reported overall

sensitivity on ICD-10 code was 82.2% (95%CI:81.0-83.3) and a PPV was 68.8% (95% CI:67.5-70.0). Arnason et al. investigated the accuracy of coding (based on ICD-9 codes) of possible warfarin complications in the chart abstractions submitted by a university-associated teaching hospital in Ottawa for billing purposes[257]. Thus, the abstracts for 1,964 hospitalizations were validated against hospital medical records as a reference standard. In this study, the possible warfarin complications were major bleeding or thromboembolism. Major bleeding was defined as an event of intracranial bleeds, GI bleeds, hematuria, vaginal bleeding, epistaxis, hemoptysis, hemorrhage not otherwise specified, hemarthrosis, or hemopericardium. Thromboembolism was defined as an event of stroke/ TIA, peripheral vascular disease, or venous thrombosis. The study reported that for major bleeding, the sensitivity, specificity, PPV, and NPV were 94% (95%CI: 91%–96%), 83% (95%CI: 78%–87%), 87% (95%CI: 83%–90%), and 92% (95%CI: 88%–95%), respectively. For thromboembolism, the reported sensitivity, specificity, PPV, and NPV were 97% (95%CI: 94%–99%), 75% (95%CI: 70%–79%), 63% (95%CI: 57%–68%), and 98% (95%CI: 96%–99%), respectively. Based on the findings from these previous studies, we may have some misclassification of the outcome. For example, the reported PPV for major bleeding was 87% (95%CI: 83%–90%), this may suggest that up to 20% of our major bleeding events may be falsely positive. On the other hand, the reported sensitivity for major bleeding was 94% (95%CI: 91%–96%) indicating that up to 10% of patients with major bleeding were not identified with our outcome definition. However, we assumed that this outcome misclassification in non-differential across the treatment groups.

The second and third studies did not contain a simulation to show that using the data-driven approach to define the duration of dispensation or the exposure-stratified weights reduces bias. However, some simulations and their results are provided by previously published studies. Thus, the study by Meid et al. investigated the agreement of the data-driven approach (longitudinal coverage approximation, COV) with the simulated ("true") duration of dispensation[251]. The input parameters for the simulation were driven from the claims data and the data collected for a large cohort study (ESTHER study[311]). In this study, oral anticoagulants (dabigatran, rivaroxaban, apixaban, warfarin, and a second vitamin K antagonist phenprocoumon) were treated as a group, however, up to 80% of all dispensations of oral anticoagulants were of phenprocoumon. This study found that compared to other approaches of approximation of the duration of dispensation (defined daily dose and one tablet per day), the data-driven approach (COV) more closely approximated the simulated duration (based on the estimated mean duration, mean relative bias, and mean relative absolute error)[251]. Next, a simulation model on the distinct effect of a prognostic variable (renal failure) on the probability of censoring in the DOACs and warfarin groups was provided elsewhere by our group[161]. Using the generated data, this study aimed to estimate the odds ratio of major bleeding in those treated with warfarin versus those treated with DOACs. It was shown that not accounting for the different effects of renal failure on censoring in the DOACs and warfarin groups, the produced odds ratio estimate was biased in favor of DOACs[161].

As defined by the US National Stroke Association (NSA), medication compliance is the "act of taking medication on schedule or taking medication as prescribed", whereas, medication adherence is the "act of filling new prescriptions or refilling prescriptions on time"[312]. Poor

compliance with warfarin therapy and suboptimal anticoagulation control at the population level may partially explain the differences in the magnitude of comparative safety and effectiveness estimates that we found in our population and those from pivotal RCTs - generally, our estimates are more in favor of DOACs. Other explanation – higher proportion of patients with reduced dose. Thus, in our second study, the hazards of major bleeding were 30 to 40% significantly lower in dabigatran vs. warfarin-treated individuals (e.g., HR: 0.67, 95% CI: 0.56-0.80 with a 15-day grace period and the data-driven approach for the definition of duration of dispensation). The hazards were 9 % to 20% lower in rivaroxaban vs. warfarin-treated individuals with borderline statistical significance (e.g., HR: 0.91, 95% CI: 0.77-1.00 with a 15-day grace period and the data-driven approach for the definition of duration of dispensation). Whereas, in the RE-LY trial, the reported HR estimates were 0.80, 95% CI: 0.69-0.93 and HR: 0.93, 95% CI: 0.81-1.07 for dabigatran 110 mg and dabigatran 150 mg, respectively[76]. ROCKET-AF trial reported the HR for major bleeding in patients treated with rivaroxaban versus those treated with DOACs was 1.04 (95% CI: 0.90-1.20)[293]. Furthermore, in our third study, the composite of stroke, systemic embolism, major bleeding, MI, and any cause of death in DOACs vs warfarin was HR: 0.74 (95% CI: 0.67-0.80). The ENGAGE AF-TIMI trial reported that compared to warfarin, the composite of systemic embolism, major bleeding, and any cause of death for apixaban was HR: 0.85 (95%CI: 0.78–0.92)[79]. RE-LY trial reported that compared to warfarin, the HR for the composite of a major vascular event, major bleeding, and any cause of death was 0.91 (95%CI: 0.82–1.00) for dabigatran 150 mg, and HR: 0.92 (95%CI: 0.84–1.02) for dabigatran 110 mg[76]. Thus, likely due to poor compliance in real-world clinical settings, the results from clinical trials cannot be mapped directly to real-world patients, and the estimates from real-world studies

measure not the biological effect of DOACs versus warfarin, but the effect of engagement in clinical care. Furthermore, in clinical practice, patients on warfarin with poor anticoagulation control and a high risk of thromboembolism would likely switch to a DOAC. This may bias the comparative safety and effectiveness estimates in studies using a time-varying exposure definition.

In our studies, we used warfarin as the comparator for three DOACs. Another coumarin derivative that Health Canada approved for stroke prevention in AF is acenocoumarol. However, before the wide use of DOACs, warfarin was prescribed more frequently than acenocoumarol due to its longer half-life, and more stable anticoagulation[313], and it was preferable for use in AF patients requiring a prolonged duration of oral anticoagulant therapy[314]. Indeed, we found that in our data, only 173 patients with AF initiated acenocoumarol between January 1, 2010, and December 31, 2016. Furthermore, acenocoumarol was not included in the pivotal randomized controlled trials comparing the efficacy and safety of DOACs to vitamin K antagonist warfarin. In line with the randomized trials, we did not include patients on acenocoumarol because the research interest was to see if our findings would support finding from the pivotal RCT.

The significant strength of this research is using claims data from a large Canadian province with over 2.5 million dispensations for oral anticoagulants (first study), and about 50,000 individuals exposed to oral anticoagulants (second and third studies) identified. This large sample gave us the statistical power to find the difference in estimates and provide more accurate results for our methodological studies. Because of the linkage capacities between the databases, we characterized our patients in terms of comorbidity profile, surgical treatment, and current use of medications.

Moreover, the RAMQ collects comprehensive follow-up data permitting not only control for the baseline characteristics but also the longitudinal changes of prognostic factors. The high accuracy of the Québec prescription claims data, including dates of dispensation and the dispensation duration, allows for modeling the periods of continuous use to approximate the exposure status.

7.5 Conclusion

In this thesis, I explored important methodological issues that may arise when conducting a study on the comparative effectiveness and safety of DOACs and a vitamin K antagonist warfarin based on the analysis of administrative data. Specifically, I studied the potential exposure misclassification when using pharmacy dispensation data, as well as selection bias induced by differential attrition across the study groups.

My findings suggest that, based on the examination of the refill patterns, the days supplied value may not adequately reflect the true duration of warfarin dispensations. Further validation studies are needed for quantitative measurements of validity of the days supplied value when defining the duration of the dispensation of the dose-varying therapies such as warfarin. For example, such studies may be based on the review of outpatient records where the updated daily dose of warfarin is recorded by physicians. Next, a greater exploration of the predictors of refill gaps and their variation in patients treated with warfarin is warranted. For example, the use of cytochrome P-450 inhibitors (e.g., SSRIs, clopidogrel), loop diuretics, and aspirin may change the INR readings thereby leading to modification of warfarin dosage.

In our first study, we found that patients classified as users of pharmacist-prepared weekly pillboxes had markedly shorter refill gaps and their variation compared to non-users. Thus, for patients on DOACs, the weighted average of refill gaps was one day in the users of weekly pillboxes versus four to five days in non-users. For warfarin, the weighted averages were two and 13 days, respectively. Implementation of pharmacy-prepared weekly pillboxes is an intervention to improve patient adherence to the treatment[210]. Previous studies found that poor adherence to oral anticoagulants in patients with AF may be associated with a higher risk of thromboembolism and bleeding[234,315]. In the second study, we found that the incidence rates of major bleeding were among the lowest of those reported in previous observational studies of individuals with NVAF. Better adherence to the therapy in our population may partially explain these findings. Future research should evaluate the effectiveness and cost-effectiveness of this intervention in improving patients' adherence and ultimately improving patients' outcomes. Moreover, more research is needed to validate the definition for the user of pharmacy-prepared weekly pillboxes when using administrative data. In studies using RAMQ data, for any medication, exposure definition based on dispensation data may be more reliable for the users of pharmacy-prepared weekly pillboxes.

To add value and sophistication to the research based on RAMQ data, it would be desirable to increase the scope of the available information. For example, a linkage with individual medical records and laboratory files would allow the collection of additional data on warfarin dosage modifications, referrals to INR tests, and their results. Presently, in Quebec, pharmacists get more access to clinical data such as INR measurement for warfarin, or renal function for direct oral anticoagulants. This allows for better monitoring of the dosage of warfarin or a DOAC according

to the characteristics of the person. Furthermore, these practices may improve the accuracy of the information in pharmacy claims, including the duration of supply, thereby minimizing misclassification of exposure. This "period effect" should be accounted for in the design of analysis phases of future studies.

Currently, point-of-care INR testing is available for patients taking warfarin[316]. This testing can be done at home (self-testing), at the nearest physician's office, or anticoagulation clinic. Based, on the test results, the warfarin dose may be adjusted by the patient, his/her clinician, or a pharmacist. Presently, these patient-physician or patient-pharmacist interactions are not captured in the RAMQ administrative database. However, this information would allow improving the definition of exposure to warfarin by incorporating intervening INR measurements to bridge the gaps between warfarin dispensations[48,317].

While oral anticoagulants were the focus of this research, the results of this thesis have significant practical implications for other areas of pharmacoepidemiologic research. For example, the refill gaps and their variation can be explored for other medications with high between-individual dose variability and within-individual dose adjustments, therapies with time-varying patterns, or symptomatic treatments. Examples of such drugs are inhaled corticosteroids, beta-agonists, insulin, NSAIDs, benzodiazepines, or psychotropics. Future studies may explore other factors that influence drug adherence or utilization patterns, such as sex, age, or type of prescriber. Moreover, refill gaps may differ for each administrative claims data depending on existing insurance plan policies, standards of practice at pharmacies, characteristics of the study population, etc. Measuring

the refill gaps, and understanding their determinants and dispensation patterns can help in making decisions on the definition of exposure, including the choice of an appropriate grace period.

It will be worth exploring refill gaps in various populations of patients taking other long-term treatments (e.g. beta-blockers, statins, antidiabetic drugs) in future studies on the validity of pharmacy claims data or studies evaluating new methods for defining the periods of drug exposure.

Future studies may further investigate the determinants of refill gaps by including predictors in the multilevel model. The predictors may be included at both the micro-level and macro-level. For example, to account for the within-patient variation of refill gaps, the possible predictors are diet, use of drug supplements, concomitant treatment with cytochrome P-450 inhibitors

. To account for between-patient variation, the predictors may include sex, age, or the presence of comorbidities, area of residence (urban versus rural), types of prescribers, or types of prescription drug insurance plans[318].

Moreover, patient characteristics, laboratory results, or blood pressure measurements can be added to a multilevel model to create a predictive model providing the estimation of the average daily dose. These predictive models may be used to estimate the duration of dispensation. Furthermore, these models may be used in day-to-day clinical practice to determine the most robust possible starting dose of a drug that should be safe and effective in a patient with a given set of characteristics.

The results of my thesis showed that the dynamic of censoring is different in the DOACs and the warfarin treatment groups. Failure to recognize this and correctly specify the model for estimation of censoring weights may result in false study conclusions. In our study, we did not

account for the dose of DOACs (standard or reduced). Nonetheless, compared to patients treated with standard doses of DOACs, those treated with reduced doses are more likely to be older, have renal disease and lower eGFR[65,319], and hence, have a higher baseline risk of thromboembolism and bleeding. Patients with reduced doses may also have a higher probability of discontinuation of a DOAC than those with standard dosing, for example, due to progressive chronic renal disease or a new condition that increases the risk of bleeding. These may have an impact on the probability of censoring. Thus, additional research is encouraged to investigate and compare the predictors of censoring in patients treated with standard doses of DOACs and those treated with reduced doses.

In the third study, we discussed censoring that may introduce a selection bias that may be differential between treatment groups. In particular, we discussed a scenario when the prognoses of thromboembolism or/and bleeding differ in patients who switch from a DOAC to warfarin and those who switch from warfarin to a DOAC. However, there are other situations when treatment discontinuation may result in selection bias. For example, patients with a low risk of thromboembolism as those with transient AF require a few months of anticoagulation. Another group is patients developing a new health condition that increases their risk of bleeding when no form of anticoagulation is allowed. In general, contextualizing and understanding the specifics of clinical management is essential in the study, including the development of the censoring weight model.

It is reasonable to assume that the differential censoring mechanism is not unique to DOACs and warfarin. To ensure the validity of study conclusions, a greater exploration of censoring

mechanisms across other therapeutic areas is warranted. Furthermore, the prior step of all studies with an active comparator should be a thoughtful investigation of the mechanisms of censoring in each treatment group separately. Areas of particular interest may be newer versus older anticancer medications, angiotensin-converting enzyme (ACE) inhibitors versus diuretics in the treatment of hypertension, and others. Furthermore, the availability of additional data from laboratory files or individual clinical records, including free-text physician notes, would improve the identification of the predictors of censoring. To increase the validity of the study design the implementation of more robust approaches for measuring prognostic characteristics is warranted. For example, to identify severe chronic kidney disease, validated algorithms may be used that were shown to have sensitivity from 82.5% to 89.0%, specificity from 97.1% to 98.9%, PPV from 94.5% to 97.7%, and NPV from 91.4% to 94.2%. These algorithms use information extracted by linkage of three administrative databases, physician claims, hospital discharge data, and pharmacy claims[320]. Finally, the development and implementation of more flexible machine learning approaches may further improve the specification of the statistical model for the estimation of censoring weights. These may enhance the identification of unknown predictors of censoring or their proxies, as well as the determination of the best fitting functional forms for the data. The results of the third study have direct implications for research utilizing administrative claims data and for randomized clinical trials undertaking the per protocol analysis.

Références bibliographiques

1. Wolf PA, Mitchell JB, Baker CS, Kannel WB, D'Agostino RB. Impact of atrial fibrillation on mortality, stroke, and medical costs. *Archives of internal medicine* 1998; **158**(3): 229-34.

2. Iwasaki Y-k, Nishida K, Kato T, Nattel S. Atrial Fibrillation Pathophysiology. *Circulation* 2011; **124**(20): 2264-74.

3. Staerk L, Sherer JA, Ko D, Benjamin EJ, Helm RH. Atrial fibrillation: epidemiology, pathophysiology, and clinical outcomes. *Circulation research* 2017; **120**(9): 1501-17.

4. Newell A, Haynes J, Smith R. Evaluation of asymptomatic atrial fibrillation. *American family physician* 2012; **86**(6): Online.

5. January CT, Wann LS, Alpert JS, et al. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on practice guidelines and the Heart Rhythm Society. *Circulation* 2014; **130**(23): 2071-104.

6. Heidenreich PA, Estes NAM, Fonarow GC, et al. 2020 Update to the 2016 ACC/AHA Clinical Performance and Quality Measures for Adults With Atrial Fibrillation or Atrial Flutter: A Report of the American College of Cardiology/American Heart Association Task Force on Performance Measures. *Circulation: Cardiovascular quality and outcomes* 2021; **14**(1): e000100.

7. Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2017 (GBD 2017) Results. Seattle, United States: Institute for Health Metrics and Evaluation (IHME), 2018. Available from http://ghdx.healthdata.org/gbd-results-tool.

8. Chugh SS, Havmoeller R, Narayanan K, et al. Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. *Circulation* 2014; **129**(8): 837-47.

9. Colilla S, Crow A, Petkun W, Singer DE, Simon T, Liu X. Estimates of current and future incidence and prevalence of atrial fibrillation in the U.S. adult population. *The American journal of cardiology* 2013; **112**(8): 1142-7.

10. Renoux C, Patenaude V, Suissa S. Incidence, mortality, and sex differences of non-valvular atrial fibrillation: a population-based study. *Journal of the American Heart Association* 2014; **3**(6): e001402.

11. Schnabel RB, Yin X, Gona P, et al. 50 year trends in atrial fibrillation prevalence, incidence, risk factors, and mortality in the Framingham Heart Study: a cohort study. *Lancet (London, England)* 2015; **386**(9989): 154-62.

12. Murphy NF, Simpson CR, Jhund PS, et al. A national survey of the prevalence, incidence, primary care burden and treatment of atrial fibrillation in Scotland. *Heart (British Cardiac Society)* 2007; **93**(5): 606-12.

13. Naderi S, Wang Y, Miller AL, et al. The impact of age on the epidemiology of atrial fibrillation hospitalizations. *American Journal of Medicine* 2014; **127**(2): 158.e1-7.

14. Marcus GM, Alonso A, Peralta CA, et al. European ancestry as a risk factor for atrial fibrillation in African Americans. *Circulation* 2010; **122**(20): 2009-15.

15. Dewland TA, Olgin JE, Vittinghoff E, Marcus GM. Incident atrial fibrillation among Asians, Hispanics, blacks, and whites. *Circulation* 2013; **128**(23): 2470-7.

16. Fox CS, Parise H, D'Agostino RB, Sr., et al. Parental atrial fibrillation as a risk factor for atrial fibrillation in offspring. *Journal of the American Medical Association (JAMA)* 2004; **291**(23): 2851-5.

17. Fuster V, Ryden LE, Cannom DS, et al. 2011 ACCF/AHA/HRS focused updates incorporated into the ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation: a report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. *Circulation* 2011; **123**(10): e269-367.

18. Benjamin EJ, Levy D, Vaziri SM, D'Agostino RB, Belanger AJ, Wolf PA. Independent risk factors for atrial fibrillation in a population-based cohort. The Framingham Heart Study. *Journal of the American Medical Association (JAMA)* 1994; **271**(11): 840-4.

19. Dublin S, Glazer NL, Smith NL, et al. Diabetes mellitus, glycemic control, and risk of atrial fibrillation. *Journal of general internal medicine* 2010; **25**(8): 853-8.

20. Kannel WB, Wolf PA, Benjamin EJ, Levy D. Prevalence, incidence, prognosis, and predisposing conditions for atrial fibrillation: population-based estimates. *The American journal of cardiology* 1998; **82**(8a): 2n-9n.

21. Chao TF, Leu HB, Huang CC, et al. Thiazolidinediones can prevent new onset atrial fibrillation in patients with non-insulin dependent diabetes. *International journal of cardiology* 2012; **156**(2): 199-202.

22. Wachtell K, Lehto M, Gerdts E, et al. Angiotensin II receptor blockade reduces new-onset atrial fibrillation and subsequent stroke compared to atenolol: the Losartan Intervention For End Point Reduction in Hypertension (LIFE) study. *Journal of the American College of Cardiology* 2005; **45**(5): 712-9.

23. Schmieder RE, Kjeldsen SE, Julius S, McInnes GT, Zanchetti A, Hua TA. Reduced incidence of new-onset atrial fibrillation with angiotensin II receptor blockade: the VALUE trial. *Journal of hypertension* 2008; **26**(3): 403-11.

24. Pan G, Zhou X, Zhao J. Effect of telmisartan on atrial fibrillation recurrences in patients with hypertension: a systematic review and meta-analysis. *Cardiovascular Therapeutics* 2014; **32**(4): 184-8.

25. Gami AS, Hodge DO, Herges RM, et al. Obstructive sleep apnea, obesity, and the risk of incident atrial fibrillation. *Journal of the American College of Cardiology* 2007; **49**(5): 565-71.

26. Fein AS, Shvilkin A, Shah D, et al. Treatment of obstructive sleep apnea reduces the risk of atrial fibrillation recurrence after catheter ablation. *Journal of the American College of Cardiology* 2013; **62**(4): 300-5.

27. Wanahita N, Messerli FH, Bangalore S, Gami AS, Somers VK, Steinberg JS. Atrial fibrillation and obesity--results of a meta-analysis. *American heart journal* 2008; **155**(2): 310-5.

28. Pathak RK, Middeldorp ME, Meredith M, et al. Long-term effect of goal-directed weight management in an atrial fibrillation cohort: a long-term follow-up study (LEGACY). *Journal of the American College of Cardiology* 2015; **65**(20): 2159-69.

29. Chamberlain AM, Agarwal SK, Folsom AR, et al. Smoking and incidence of atrial fibrillation: results from the Atherosclerosis Risk in Communities (ARIC) study. *Heart rhythm* 2011; **8**(8): 1160-6.

30. O'Keefe JH, Bhatti SK, Bajwa A, DiNicolantonio JJ, Lavie CJ. Alcohol and cardiovascular health: the dose makes the poison...or the remedy. *Mayo Clinic proceedings* 2014; **89**(3): 382-93.

31. Bekwelem W, Connolly SJ, Halperin JL, et al. Extracranial systemic embolic events in patients with nonvalvular atrial fibrillation: incidence, risk factors, and outcomes. *Circulation* 2015; **132**(9): 796-803.

32. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 1991; **22**(8): 983-8.

33. Kirchhof P, Benussi S, Kotecha D, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *European Heart Journal* 2016; **37**(38): 2893-962.

34. Macle L, Cairns J, Leblanc K, et al. 2016 focused update of the Canadian Cardiovascular Society Guidelines for the management of atrial fibrillation. *Canadian Journal of Cardiology* 2016; **32**(10): 1170-85.

35. Martin RIR, Bates MGD. Management of atrial fibrillation and concomitant coronary artery disease. *Continuing Cardiology Education* 2017; **3**(2): 47-55.

36. Andrade JG, Macle L, Nattel S, Verma A, Cairns J. Contemporary atrial fibrillation management: a comparison of the current AHA/ACC/HRS, CCS, and ESC guidelines. *Canadian Journal of Cardiology* 2017; **33**(8): 965-76.

37. Blustin JM, McBane RD, Ketha SS, Wysokinski WE. Distribution of thromboembolism in valvular versus non-valvular atrial fibrillation. *Expert Review Cardiovascular Therapy* 2014; **12**(10): 1129-32.

38. Steinberg BA, Hellkamp AS, Lokhnygina Y, et al. Higher risk of death and stroke in patients with persistent vs. paroxysmal atrial fibrillation: results from the ROCKET-AF Trial. *European Heart Journal* 2015; **36**(5): 288-96.

39. Olesen JB, Lip GY, Kamper AL, et al. Stroke and bleeding in atrial fibrillation with chronic kidney disease. *The New England journal of medicine* 2012; **367**(7): 625-35.

40. Aursulesei V, Costache, II. Anticoagulation in chronic kidney disease: from guidelines to clinical practice. *Clinical Cardiology* 2019; **42**(8): 774-82.

41. Kwon Y, Norby FL, Jensen PN, et al. Association of smoking, alcohol, and obesity with cardiovascular death and ischemic stroke in atrial fibrillation: the Atherosclerosis Risk in Communities (ARIC) study and Cardiovascular Health Study (CHS). *The public library of science (PLOS) One* 2016; **11**(1): e0147065.

42. Kirchhof P. Integrated care of patients with atrial fibrillation: the 2016 ESC atrial fibrillation guidelines. *Heart (British Cardiac Society)* 2017; **103**(10): 729-31.

43. Makaryus JN, Halperin JL, Lau JF. Oral anticoagulants in the management of venous thromboembolism. *Nature Reviews Cardiology* 2013; **10**(7): 397-409.

44. Dempfle CE. Direct oral anticoagulants--pharmacology, drug interactions, and side effects. *Seminars in hematology* 2014; **51**(2): 89-97.

45. Government of Canada. Health Canada. Drug Product Database online query. https://translate.google.ca/?hl=en&tab=rT&sl=en&tl=ru&text=Government%20of%20Canada.%20Health%20Canada.%20Drug%20Product%20Database%20online%20query.&op=translate (accessed June 10 2022).

46. Hart RG, Benavente O, McBride R, Pearce LA. Antithrombotic therapy to prevent stroke in patients with atrial fibrillation: a meta-analysis. *Annals of internal medicine* 1999; **131**(7): 492-501.

47. Segal JB, McNamara RL, Miller MR, et al. Prevention of thromboembolism in atrial fibrillation. A meta-analysis of trials of anticoagulants and antiplatelet drugs. *Journal of general internal medicine* 2000; **15**(1): 56-67.

48. Go AS, Hylek EM, Chang Y, et al. Anticoagulation therapy for stroke prevention in atrial fibrillation: how well do randomized trials translate into clinical practice? *Journal of the American Medical Association (JAMA)* 2003; **290**(20): 2685-92.

49. Dumont Z, Mordasiewicz M, Kosar L, Schuster B. Warfarin: its highs and lows. *Canadian Family Physician* 2013; **59**(8): 856-60.

50. Hylek EM, Singer DE. Risk factors for intracranial hemorrhage in outpatients taking warfarin. *Annals of internal medicine* 1994; **120**(11): 897-902.

51. Hylek EM, Skates SJ, Sheehan MA, Singer DE. An analysis of the lowest effective intensity of prophylactic anticoagulation for patients with nonrheumatic atrial fibrillation. *The New England journal of medicine* 1996; **335**(8): 540-6.

52. Hylek EM, Go AS, Chang Y, et al. Effect of intensity of oral anticoagulation on stroke severity and mortality in atrial fibrillation. *The New England journal of medicine* 2003; **349**(11): 1019-26.

53. Schulman S, Beyth RJ. Risk of bleeding with long-term antithrombotic therapy in atrial fibrillation. *European Heart Journal Supplements* 2005; **7 (Supplement C)**: C34–C40.

54. Rosendaal FR, Cannegieter SC, van der Meer FJM, Briët E. A method to determine the optimal intensity of oral anticoagulant therapy. *Thrombosis and Haemostasis* 1993; **69**(03): 236-9.

55. White HD, Gruber M, Feyzi J, et al. Comparison of outcomes among patients randomized to warfarin therapy according to anticoagulant control: results from SPORTIF III and V. *Archives of internal medicine* 2007; **167**(3): 239-45.

56. Connolly SJ, Pogue J, Eikelboom J, et al. Benefit of oral anticoagulant over antiplatelet therapy in atrial fibrillation depends on the quality of international normalized ratio control achieved by centers and countries as measured by time in therapeutic range. *Circulation* 2008; **118**(20): 2029-37.

57. Rosenstein R, Parra D. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *The New England journal of medicine* 2011; **365**(24): 2334; author reply 5.

58. Meyer B. Time in therapeutic range for warfarin—a European success story. *Journal Watch cardiology* 2011; **19**.

59. van Walraven C, Jennings A, Oake N, Fergusson D, Forster AJ. Effect of study setting on anticoagulation control: a systematic review and metaregression. *Chest* 2006; **129**(5): 1155-66.

60. Palareti G, Leali N, Coccheri S, et al. Bleeding complications of oral anticoagulant treatment: an inception-cohort, prospective collaborative study (ISCOAT). Italian Study on Complications of Oral Anticoagulant Therapy. *Lancet (London, England)* 1996; **348**(9025): 423-8.

61. Hutten BA, Prins MH, Gent M, Ginsberg J, Tijssen JG, Buller HR. Incidence of recurrent thromboembolic and bleeding complications among patients with venous thromboembolism in relation to both malignancy and achieved international normalized ratio: a retrospective analysis. *Journal of Clinical Oncology* 2000; **18**(17): 3078-83.

62. Levine MN, Raskob G, Beyth RJ, Kearon C, Schulman S. Hemorrhagic complications of anticoagulant treatment: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. *Chest* 2004; **126**(3 Suppl): 287s-310s.

63. White RH, McKittrick T, Takakuwa J, Callahan C, McDonell M, Fihn S. Management and prognosis of life-threatening bleeding during warfarin therapy. National Consortium of Anticoagulation Clinics. *Archives of internal medicine* 1996; **156**(11): 1197-201.

64. Crader MF, Johns T, Arnold JK. Warfarin drug interactions. StatPearls. Treasure Island (FL): StatPearls Publishing. Copyright © 2020, StatPearls Publishing LLC.; 2020.

65. Andrade JG, Aguilar M, Atzema C, et al. The 2020 Canadian Cardiovascular Society/Canadian Heart Rhythm Society comprehensive guidelines for the management of atrial fibrillation. *Canadian Journal of Cardiology* 2020; **36**(12): 1847-948.

66. St-Pierre F. Usage des anticoagulants oraux directs et de la warfarine dans le contexte de la fibrillation auriculaire et de la thromboembolie veineuse. Rapport en soutien aux guides d'usage optimal / rédigé par Frédéric St-Pierre. Québec, Québec; 2019.

67. Mar PL, Gopinathannair R, Gengler BE, et al. Drug interactions affecting oral anticoagulant use. *Circulation: arrhythmia and electrophysiology* 2022; **15**(6): e007956.

68. Mekaj YH, Mekaj AY, Duci SB, Miftari EI. New oral anticoagulants: their advantages and disadvantages compared with vitamin K antagonists in the prevention and treatment of patients with thromboembolic events. *Ther Clinical Risk Management* 2015; **11**: 967-77.

69. Antithrombotic agents for the prevention of stroke and systemic embolism in patients with atrial fibrillation [Internet]. Ottawa (ON): Canadian Agency for Drugs and Technologies in Health; 2013 Mar. (CADTH Therapeutic Review, No. 1.1B.) Available from: https://www.ncbi.nlm.nih.gov/books/NBK168988/.

70. Noviyani R, Youngkong S, Nathisuwan S, et al. Economic evaluation of direct oral anticoagulants (DOACs) versus vitamin K antagonists (VKAs) for stroke prevention in patients with atrial fibrillation: a systematic review and meta-analysis. *BMJ Evidence-based medicine* 2021.

71. Mannarino MG. Cost of warfarin versus direct oral anticoagulants: Canadian family physician. 2020 Jan;66(1):9.

72. Ortiz-Cartagena I, Gotay AO, Acevedo J. Cost effectiveness of oral anticoagulation therapy for non-valvular atrial fibrillation patients: warfarin versus the new oral anticoagulants rivaroxaban, dabigatran and apixaban. *Journal of the American College of Cardiology* 2018; **71**(11_Supplement): A490-A.

73. Tangedal K, Semchuk W, Bolt J. How should Canadian pharmacy departments utilize idarucizumab? *Canadian Journal of Hospital Pharmacy* 2016; **69**(5): 429-30.

74. Siegal DM, Curnutte JT, Connolly SJ, et al. Andexanet Alfa for the Reversal of Factor Xa Inhibitor Activity. *New England Journal of Medicine* 2015; **373**(25): 2413-24.

75. Ansell J, Laulicht BE, Bakhru SH, et al. Ciraparantag, an anticoagulant reversal drug: mechanism of action, pharmacokinetics, and reversal of anticoagulants. *Blood* 2021; **137**(1): 115-25.

76. Connolly SJ, Ezekowitz MD, Yusuf S, et al. Dabigatran versus warfarin in patients with atrial fibrillation. *New England Journal of Medicine* 2009; **361**(12): 1139-51.

77. Patel MR, Mahaffey KW, Garg J, et al. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *New England Journal of Medicine* 2011; **365**(10): 883-91.

78. Granger CB, Alexander JH, McMurray JJV, et al. Apixaban versus warfarin in patients with atrial fibrillation. *New England Journal of Medicine* 2011; **365**(11): 981-92.

79. Giugliano RP, Ruff CT, Braunwald E, et al. Edoxaban versus warfarin in patients with atrial fibrillation. *New England Journal of Medicine* 2013; **369**(22): 2093-104.

80. Carnicelli AP, Hong H, Connolly SJ, et al. Direct oral anticoagulants versus warfarin in patients with atrial fibrillation: patient-level network meta-analyses of randomized clinical trials with interaction testing by age and sex. *Circulation* 2022; **145**(4): 242-55.

81. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. *Lancet (London, England)* 2017; **390**(10092): 415-23.

82. Avorn J. The promise of pharmacoepidemiology in helping clinicians assess drug risk. *Circulation* 2013; **128**(7): 745-8.

83. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". *Lancet (London, England)* 2005; **365**(9453): 82-93.

84. Hagg L, Johansson C, Jansson JH, Johansson L. External validity of the ARISTOTLE trial in real-life atrial fibrillation patients. *Cardiovascular therapeutics* 2014; **32**(5): 214-8.

85. Desmaele S, Steurbaut S, Cornu P, Brouns R, Dupont AG. Clinical trials with direct oral anticoagulants for stroke prevention in atrial fibrillation: how representative are they for real life patients? *European Journal of Clinical Pharmacology* 2016; **72**(9): 1125-34.

86. Marietta M. Direct oral anticoagulants in atrial fibrillation: can data from randomized clinical trials be safely transferred to the general population? No. *Internal and Emergency Medicine* 2015; **10**(6): 647-50.

87. Courtright K. POINT: do randomized controlled trials ignore needed patient populations? Yes. *Chest* 2016; **149**(5): 1128-30.

88. Sjögren V, Grzymala-Lubanski B, Renlund H, et al. Safety and efficacy of well managed warfarin. *Thrombosis and Haemostasis* 2015; **113**(06): 1370-7.

89. Abraham NS, Noseworthy PA, Inselman J, et al. Risk of gastrointestinal bleeding increases with combinations of antithrombotic agents and patient age. *Clinical gastroenterology and hepatology* 2020; **18**(2): 337-46 e19.

90. Graham DJ, Reichman ME, Wernecke M, et al. Cardiovascular, bleeding, and mortality risks in elderly Medicare patients treated with dabigatran or warfarin for nonvalvular atrial fibrillation. *Circulation* 2015; **131**(2): 157-64.

91. Xavier D, Hanna M, Wallentin L, et al. Abstract 18003: Patients with good adherence to study medication have better outcomes: insights from the ARISTOTLE trial. *Circulation* 2017; **136**(suppl_1).

92. Pham PN, Brown JD. Real-world adherence for direct oral anticoagulants in a newly diagnosed atrial fibrillation cohort: does the dosing interval matter? *BMC Cardiovascular disorders* 2019; **19**(1): 64.

93. Strom BL, Kimmel SE, Hennessy S. Pharmacoepidemiology. Chapter 1. What is pharmacoepidemiology? 5th ed. Chichester, West Sussex, UK: Wiley-Blackwell,; 2012.

94. Booth CM, Tannock IF. Randomised controlled trials and population-based observational research: partners in the evolution of medical evidence. *British Journal of Cancer* 2014; **110**(3): 551-5.

95. Chan EW, Liu KQ, Chui CS, Sing CW, Wong LY, Wong IC. Adverse drug reactions - examples of detection of rare events using databases. *British journal of clinical pharmacology* 2015; **80**(4): 855-61.

96. Strom BL. Overview of electronic databases in pharmacoepidemiology. Pharmacoepidemiology; 2019: 203-10.

97. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *Journal of clinical epidemiology* 2005; **58**(4): 323-37.

98. Cohen A, Goto S, Schreiber K, Torp-Pedersen C. Why do we need observational studies of everyday patients in the real-life setting?: Table 1. *European Heart Journal Supplements* 2015; **17**.

99. Southworth MR, Reichman ME, Unger EF. Dabigatran and postmarketing reports of bleeding. *The New England journal of medicine* 2013; **368**(14): 1272-4.

100. Martinez BK, Sood NA, Bunz TJ, Coleman CI. Effectiveness and safety of apixaban, dabigatran, and rivaroxaban versus warfarin in frail patients with nonvalvular atrial fibrillation. *Journal of the American Heart Association* 2018; **7**(8): e008643.

101. Escobar C, Marti-Almor J, Perez Cabeza A, Martinez-Zapata MJ. Direct oral anticoagulants versus vitamin K antagonists in real-life patients with atrial fibrillation. A systematic review and meta-analysis. *Revista española de cardiología (Engl Ed)* 2019; **72**(4): 305-16.

102. Bai Y, Deng H, Shantsila A, Lip GY. Rivaroxaban versus dabigatran or warfarin in real-world studies of stroke prevention in atrial fibrillation: systematic review and meta-analysis. *Stroke* 2017; **48**(4): 970-6.

103. Hirschl M, Kundi M. Safety and efficacy of direct acting oral anticoagulants and vitamin K antagonists in nonvalvular atrial fibrillation – a network meta-analysis of real-world data. *Vasa* 2019; **48**(2): 134-47.

104. Lowenstern A, Al-Khatib SM, Sharan L, et al. Interventions for Preventing Thromboembolic Events in Patients With Atrial Fibrillation: A Systematic Review. *Annals of internal medicine* 2018; **169**(11): 774-87.

105. Deitelzweig S, Farmer C, Luo X, et al. Risk of major bleeding in patients with non-valvular atrial fibrillation treated with oral anticoagulants: a systematic review of real-world observational studies. *Current Medical Research and Opinion* 2017; **33**(9): 1583-94.

106. Cochran WG. The combination of estimates from different experiments. *Biometrics* 1954; **10**(1): 101-29.

107. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ (Clinical research ed)* 2003; **327**(7414): 557-60.

108. Sedgwick P. Meta-analyses: what is heterogeneity? *BMJ : British Medical Journal* 2015; **350**: h1435.

109. Dekkers OM, von Elm E, Algra A, Romijn JA, Vandenbroucke JP. How to assess the external validity of therapeutic trials: a conceptual approach. *International journal of epidemiology* 2010; **39**(1): 89-94.

110. Porta M, ed. A Dictionary of epidemiology. 6th Edition. Oxford: Oxford University Press, 2014.

111. Lund JL, Richardson DB, Stürmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. *Current Epidemiology Reports* 2015; **2**(4): 221-8.

112. Petri H, Urquhart J. Channeling bias in the interpretation of drug effects. *Statistics in medicine* 1991; **10**(4): 577-81.

113. Larsen TB, Rasmussen LH, Skjoth F, et al. Efficacy and safety of dabigatran etexilate and warfarin in "real-world" patients with atrial fibrillation: a prospective nationwide cohort study. *Journal of the American College of Cardiology* 2013; **61**(22): 2264-73.

114. Abraham NS, Singh S, Alexander GC, et al. Comparative risk of gastrointestinal bleeding with dabigatran, rivaroxaban, and warfarin: population based cohort study. *BMJ : British Medical Journal* 2015; **350**: h1857.

115. Lauffenburger JC, Farley JF, Gehi AK, Rhoney DH, Brookhart MA, Fang G. Effectiveness and safety of dabigatran and warfarin in real-world US patients with non-valvular atrial fibrillation: a retrospective cohort study. *Journal of the American Heart Association* 2015; **4**(4).

116. Brais C, Larochelle J, Turgeon MH, et al. Predictors of direct oral anticoagulants utilization for thromboembolism prevention in atrial fibrillation. *Journal of Pharmaceutical Sciences* 2017; **20**: 8-14.

117. Ankarfeldt MZ, Thorsted BL, Groenwold RHH, Adalsteinsson E, Ali MS, Klungel OH. Assessment of channeling bias among initiators of glucose-lowering drugs: A UK cohort study. *Journal of clinical epidemiology* 2017; **9**: 19-30.

118. Funk MJ, Landi SN. Misclassification in administrative claims data: quantifying the impact on treatment effect estimates. *Current Epidemiology Reports* 2014; **1**(4): 175-85.

119. Arfè A, Corrao G. The lag-time approach improved drug-outcome association estimates in presence of protopathic bias. *Journal of clinical epidemiology* 2016; **78**: 101-7.

120. Renoux C, Dell'Aniello S, Brenner B, Suissa S. Bias from depletion of susceptibles: the example of hormone replacement therapy and the risk of venous thromboembolism. *Pharmacoepidemiol Drug Saf* 2017; **26**(5): 554-60.

121. Suissa S, Dell'Aniello S. Time-related biases in pharmacoepidemiology. *Pharmacoepidemiol Drug Saf* 2020; **29**(9): 1101-10.

122. Suissa S. Immortal time bias in pharmacoepidemiology. *Am J Epidemiol* 2007; **167**(4): 492-9.

123. Suissa S, Azoulay L. Metformin and the risk of cancer: time-related biases in observational studies. *Diabetes care* 2012; **35**(12): 2665-73.

124. Suissa S, Dell'Aniello S, Vahey S, Renoux C. Time-window bias in case-control studies: statins and lung cancer. *Epidemiology* 2011; **22**(2): 228-31.

125. Prada-Ramallal G, Takkouche B, Figueiras A. Bias in pharmacoepidemiologic studies using secondary health care databases: a scoping review. *BMC Medical Research Methodology* 2019; **19**(1): 53.

126. Rothman KJ. Epidemiology : an introduction. New York, NY: Oxford University Press; 2012.

127. West SL, Savitz DA, Koch G, Strom BL, Guess HA, Hartzema A. Recall accuracy for prescription medications: self-report compared with database information. *Am J Epidemiol* 1995; **142**(10): 1103-12.

128. Warholak TL, McCulloch M, Baumgart A, Smith M, Fink W, Fritz W. An exploratory comparison of medication lists at hospital admission with administrative database records. *The Journal of Managed Care & Specialty Pharmacy* 2009; **15**(9): 751-8.

129. Lau HS, Florax C, Porsius AJ, De Boer A. The completeness of medication histories in hospital medical records of patients admitted to general internal medicine wards. *British journal of clinical pharmacology* 2000; **49**(6): 597-603.

130. Koponen M, Ilomäki J, Aarnio E, Taipale H. Methodological considerations in pharmacoepidemiology and pharmacovigilance studies. In: Z-U-D B, ed. Encyclopedia of Pharmacy Practice and Clinical Pharmacy: Elsevier - Mosby; 2019.

131. Tamblyn R, Lavoie G, Petrella L, Monette J. The use of prescription claims databases in pharmacoepidemiological research: the accuracy and comprehensiveness of the prescription claims database in Québec. *Journal of clinical epidemiology* 1995; **48**(8): 999-1009.

132. Gardarsdottir H, Souverein PC, Egberts TC, Heerdink ER. Construction of drug treatment episodes from drug-dispensing histories is influenced by the gap length. *Journal of clinical epidemiology* 2010; **63**(4): 422-7.

133. Nielsen LH, Keiding N. Validation of methods for identifying discontinuation of treatment from prescription data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 2010; **59**(4): 707-22.

134. Ray WA, Thapa PB, Gideon P. Misclassification of current benzodiazepine exposure by use of a single baseline measurement and its effects upon studies of injuries. *Pharmacoepidemiol Drug Saf* 2002; **11**(8): 663-9.

135. Støvring H, Pottegård A, Hallas J. Determining prescription durations based on the parametric waiting time distribution. *Pharmacoepidemiol Drug Saf* 2016; **25**(12): 1451-9.

136. Blais L, Vilain A, Kettani FZ, et al. Accuracy of the days' supply and the number of refills allowed recorded in Québec prescription claims databases for inhaled corticosteroids. *BMJ Open* 2014; **4**(11): e005903.

137. Kirkman MS, Rowan-Martin MT, Levin R, et al. Determinants of adherence to diabetes medications: findings from a large pharmacy claims database. *Diabetes care* 2015; **38**(4): 604-9.

138. Lip GYH, Banerjee A, Boriani G, et al. Antithrombotic therapy for atrial fibrillation: CHEST guideline and expert panel report. *Chest* 2018; **154**(5): 1121-201.

139. Delaney JAC, Seeger JD. Sensitivity analysis. In: Velentgas P, Dreyer NA, Nourjah P, et al., editors. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. Rockville (MD): Agency for Healthcare Research and Quality (US); 2013 Jan. Chapter 11. Available from: https://www.ncbi.nlm.nih.gov/books/NBK126178/.

140. Tang W, Chang H-Y, Zhou M, Singh S. Risk of gastrointestinal bleeding among dabigatran users – a self controlled case series analysis. *Scientific Reports* 2017; **7**(1): 40120.

141. Go AS, Singer DE, Toh S, et al. Outcomes of dabigatran and warfarin for atrial fibrillation in contemporary practice: a retrospective cohort study. *Annals of internal medicine* 2017; **167**(12): 845-54.

142. Fox MP. Creating a demand for bias analysis in epidemiological research. *Journal of epidemiology and community health* 2009; **63**(2): 91.

143. Sen PK. Censoring in theory and practice: statistical perspectives and controversies. Analysis of censored data IMS Lecture Notes Monogr Ser. Hayward, CA: Institute of Mathematical Statistics; 1995.

144. Cox DR, Oakes D. Analysis of survival data: Taylor & Francis; 1984.

145. Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. 2nd ed: John Wiley & Sons.; 2002.

146. Klein JP, Moeschberger ML. Survival analysis. Techniques for censored and truncated data: Springer-Verlag New York; 2003.

147. Kleinbaum DG, Klein M. Survival analysis, a self-learning text. Third ed: Springer-Verlag New York; 2012.

148. Joffe MM. Administrative and artificial censoring in censored regression models. *Statistics in medicine* 2001; **20**(15): 2287-304.

149. Leung KM, Elashoff RM, Afifi AA. Censoring issues in survival analysis. *Annual review of public health* 1997; **18**: 83-104.

150. Butler AM, Todd JV, Sahrmann JM, Lesko CR, Brookhart MA. Informative censoring by health plan disenrollment among commercially insured adults. *Pharmacoepidemiol Drug Saf* 2019; **28**(5): 640-8.

151. Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics* 2000; **56**(3): 779-88.

152. Morden JP, Lambert PC, Latimer N, Abrams KR, Wailoo AJ. Assessing methods for dealing with treatment switching in randomised controlled trials: a simulation study. *BMC Medical Research Methodology* 2011; **11**: 4.

153. Khosrow-Khavar F, Filion KB, Bouganim N, Suissa S, Azoulay L. Aromatase inhibitors and the risk of cardiovascular outcomes in women with breast cancer. *Circulation* 2020; **141**(7): 549-59.

154. Filion KB, Lix LM, Yu OH, et al. Sodium glucose cotransporter 2 inhibitors and risk of major adverse cardiovascular events: multi-database retrospective cohort study. *BMJ : British Medical Journal* 2020; **370**: m3342.

155. Hernan MA, Hernandez-Diaz S. Beyond the intention-to-treat in comparative effectiveness research. *Clinical Trials* 2012; **9**(1): 48-55.

156. Austin PC, Lee DS, Fine JP. Introduction to the analysis of survival data in the presence of competing risks. *Circulation* 2016; **133**(6): 601-9.

157. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004; **15**(5): 615-25.

158. International conference on harmonisation; guidance on statistical principles for clinical trials; availability--FDA. Notice. *Federal register* 1998; **63**(179): 49583-98.

159. Hernan MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology* 2008; **19**(6): 766-79.

160. Mansournia MA, Higgins JPT, Sterne JAC, Hernán MA. Biases in Randomized Trials: A Conversation Between Trialists and Epidemiologists. *Epidemiology (Cambridge, Mass)* 2017; **28**(1): 54-9.

161. Schnitzer ME, Platt RW, Durand M. A tutorial on dealing with time-varying eligibility for treatment: Comparing the risk of major bleeding with direct-acting oral anticoagulants vs warfarin. *Statistics in medicine* 2020.

162. Hernán MA, Brumback BA, Robins JM. Estimating the causal effect of zidovudine on CD4 count with a marginal structural model for repeated measures. *Statistics in medicine* 2002; **21**(12): 1689-709.

163. Rotnitzky A, Robins JM. Semiparametric regression estimation in the presence of dependent censoring. *Biometrika* 1995; **82**(4): 805-20.

164. Wooldridge JM. Econometric analysis of cross section and panel data. Chapter 21 Estimating average treatment effects.: The MIT Press; 2010.

165. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol* 2008; **168**(6): 656-64.

166. Cook NR, Cole SR, Buring JE. Aspirin in the primary prevention of cardiovascular disease in the Women's Health Study: effect of noncompliance. *European journal of epidemiology* 2012; **27**(6): 431-8.

167. Hernan MA, Robins JM. Estimating causal effects from epidemiological data. *Journal of epidemiology and community health* 2006; **60**(7): 578-86.

168. Weuve J, Tchetgen Tchetgen EJ, Glymour MM, et al. Accounting for bias due to selective attrition: the example of smoking and cognitive decline. *Epidemiology* 2012; **23**(1): 119-28.

169. Danaei G, Rodriguez LA, Cantero OF, Logan R, Hernan MA. Observational data for comparative effectiveness research: an emulation of randomised trials of statins and primary prevention of coronary heart disease. *Statistical Methods in Medical Research* 2013; **22**(1): 70-96.

170. Hernan MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. *Basic and Clinical Pharmacology and Toxicology* 2006; **98**(3): 237-42.

171. Bourgeois S, Jorgensen A, Zhang EJ, et al. A multi-factorial analysis of response to warfarin in a UK prospective cohort. *Genome Medicine* 2016; **8**(1): 2.

172. Linder MW, Looney S, Adams JE, et al. Warfarin dose adjustments based on CYP2C9 genetic polymorphisms. *Journal of Thrombosis and Thrombolysis* 2002; **14**(3): 227-32.

173. Salmasi S, Loewen PS, Tandun R, Andrade JG, De Vera MA. Adherence to oral anticoagulants among patients with atrial fibrillation: a systematic review and meta-analysis of observational studies. *BMJ Open* 2020; **10**(4): e034778.

174. Ansell J, Hirsh J, Hylek E, Jacobson A, Crowther M, Palareti G. Pharmacology and management of the vitamin K antagonists: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th Edition). *Chest* 2008; **133**(6 Suppl): 160s-98s.

175. Casciano JP, Dotiwala ZJ, Martin BC, Kwong WJ. The costs of warfarin underuse and nonadherence in patients with atrial fibrillation: a commercial insurer perspective. *Journal of Managed Care Pharmacy* 2013; **19**(4): 302-16.

176. Azoulay L, Dell'Aniello S, Simon TA, Langleben D, Renoux C, Suissa S. A net clinical benefit analysis of warfarin and aspirin on stroke in patients with atrial fibrillation: a nested case-control study. *BMC Cardiovascular Disorders* 2012; **12**: 49-.

177. Azoulay L, Dell'Aniello S, Simon TA, Renoux C, Suissa S. Initiation of warfarin in patients with atrial fibrillation: early effects on ischaemic strokes. *European heart journal* 2013; **35**(28): 1881-7.

178. Webster-Clark M, Lund JL, Stürmer T, Poole C, Simpson RJ, Edwards JK. Reweighting oranges to apples: transported RE-LY trial versus nonexperimental effect estimates of anticoagulation in atrial fibrillation. *Epidemiology* 2020; **31**(5): 605-13.

179. Webster-Clark M, Stürmer T, Edwards JK, Poole C, Simpson RJ, Jr., Lund JL. Real-world on-treatment and initial treatment absolute risk differences for dabigatran vs warfarin in older US adults. *Pharmacoepidemiol Drug Saf* 2020; **29**(8): 832-41.

180. Ruiz-Nodar JM, Marin F, Manzano-Fernandez S, et al. An evaluation of the CHADS(2) stroke risk score in patients with atrial fibrillation who undergo percutaneous coronary revascularization. *Chest* 2011; **139**(6): 1402-9.

181. Rollins A, Hanigan S, Pogue K, Renner E, Barnes G, Dorsch M. Identifying clinical predictors of switching from direct oral anticoagulants to warfarin. *Clinical Medicine Insights: Therapeutics* 2019; **11**: 1179559X19831287.

182. Manzoor BS, Walton SM, Sharp LK, Galanter WL, Lee TA, Nutescu EA. High number of newly initiated direct oral anticoagulant users switch to alternate anticoagulant therapy. *Journal of Thrombosis and Thrombolysis* 2017; **44**(4): 435-41.

183. Gouvernement du Québec. Régie de l'assurance-maladie du Québec; données et statistiques. 2022. http://www.ramq.gouv.qc.ca/fr/donnees-statistiques/Pages/donnees-statistiques.aspx (accessed 21st July 2022.

184. Régie de l'assurance maladie du Québec. Statistiques annuelles. 2022. https://www.prod.ramq.gouv.qc.ca/IST/CD/CDF_DifsnInfoStats/CDF1_CnsulInfoStatsCNC_iut/DifsnInfoStats.aspx?LANGUE=fr-CA (accessed 21 July 2022.

185. Régie de l'assurance-maladie du Québec. Connaître les conditions d'admissibilité au régime public. 2022. https://www.ramq.gouv.qc.ca/fr/citoyens/assurance-medicaments/connaitre-conditions-admissibilite-regime-public21 July 2022).

186. Renoux C, Coulombe J, Suissa S. Long-term vitamin K antagonists treatment patterns of Non-Valvular Atrial Fibrillation (NVAF): a population-based cohort study. *BMC Cardiovascular Disorders* 2016; **16**: 84.

187. Jun M, Lix LM, Durand M, et al. Comparative safety of direct oral anticoagulants and warfarin in venous thromboembolism: multicentre, population based, observational study. *BMJ (Clinical research ed)* 2017; **359**: j4323.

188. Douros A, Renoux C, Yin H, Filion KB, Suissa S, Azoulay L. Concomitant use of direct oral anticoagulants with antiplatelet agents and the risk of major bleeding in patients with nonvalvular atrial fibrillation. *American Journal of Medicine* 2019; **132**(2): 191-9.e12.

189. Perreault S, de Denus S, White-Guay B, et al. Oral anticoagulant prescription trends, profile use, and determinants of adherence in patients with atrial fibrillation. *Pharmacotherapy* 2020; **40**(1): 40-54.

190. Bérard A, Lacasse A. Validity of perinatal pharmacoepidemiologic studies using data from the RAMQ administrative database. *The Canadian Journal of Clinical Pharmacology* 2009; **16**(2): e360-9.

191. Blais C, Rochette L, Hamel D, Poirier P. Prevalence, incidence, awareness and control of hypertension in the province of Quebec: perspective from administrative and survey data. *Can J Public Health* 2014; **105**(1): e79-e85.

192. Wilchesky M, Tamblyn RM, Huang A. Validation of diagnostic codes within medical services claims. *Journal of clinical epidemiology* 2004; **57**(2): 131-41.

193. Cadieux G, Tamblyn R. Accuracy of physician billing claims for identifying acute respiratory infections in primary care. *Health Services Research* 2008; **43**(6): 2223-38.

194. Oskoui M, Ng P, Dorais M, et al. Accuracy of administrative claims data for cerebral palsy diagnosis: a retrospective cohort study. *Canadian Medical Association Journal* 2017; **5**(3): E570-E5.

195. Lambert L, Blais C, Hamel D, et al. Evaluation of care and surveillance of cardiovascular disease: can we trust medico-administrative hospital data? *Canadian Journal of Cardiology* 2012; **28**(2): 162-8.

196. Andrade SE, Harrold LR, Tjia J, et al. A systematic review of validated methods for identifying cerebrovascular accident or transient ischemic attack using administrative data. *Pharmacoepidemiol Drug Saf* 2012; **21 Suppl 1**(Suppl 1): 100-28.

197. Mayo NE, Danys I, Carlton J, Scott SC. Accuracy of hospital discharge coding for stroke. *Canadian Journal of Cardiology* 1993;9(supplement D): 121D-124D.

198. Canadian Institute for Health Information. Physicians in Canada, 2016: Summary Report. Ottawa, ON: CIHI; 2017.

199. Thacker EL, Soliman EZ, Pulley L, Safford MM, Howard G, Howard VJ. Investigation of selection bias in the association of race with prevalent atrial fibrillation in a national cohort study: REasons for Geographic And Racial Differences in Stroke (REGARDS). *Annals of Epidemiology* 2016; **26**(8): 534-9.

200. Toh S, Andrade SE, Raebel MA, et al. Examples of existing automated databases. Textbook of Pharmacoepidemiology; 2013: 123-77.

201. Commission d'accès à l'information du Québec. Access to Information and Privacy Process. . https://www.cai.gouv.qc.ca/ (accessed July, 24 2022.

202. Wilchesky M, Suissa S. The Régie de l'assurance maladie du Québec (RAMQ) databases. In: Sturkenboom M, Schink T, eds. Databases for pharmacoepidemiological research: Springer series on epidemiology and public health; 2021.

203. Institut de la statistique du Québec. Research data access servoces. . https://statistique.quebec.ca/research/#/accueil.

204. Johnson ES, Bartman BA, Briesacher BA, et al. The incident user design in comparative effectiveness research. *Pharmacoepidemiol Drug Saf* 2013; **22**(1): 1-6.

205. Jensen PN, Johnson K, Floyd J, Heckbert SR, Carnahan R, Dublin S. A systematic review of validated methods for identifying atrial fibrillation using administrative data. *Pharmacoepidemiol Drug Saf* 2012; **21 Suppl 1**(0 1): 141-7.

206. Otto CM, Nishimura RA, Bonow RO, et al. 2020 ACC/AHA Guideline for the Management of Patients With Valvular Heart Disease: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation* 2021; **143**(5): e72-e227.

207. Smilowitz NR, Mega JL, Berger JS. Duration of anticoagulation for venous thromboembolic events. *Circulation* 2014; **130**(25): 2343-8.

208. Levin A, Hemmelgarn B, Culleton B, et al. Guidelines for the management of chronic kidney disease. *Canadian Medical Association Journal* 2008; **179**(11): 1154-62.

209. Rassen JA, Shelat AA, Franklin JM, Glynn RJ, Solomon DH, Schneeweiss S. Matching by propensity score in cohort studies with three treatment groups. *Epidemiology* 2013; **24**(3): 401-9.

210. Demers H, Giguere T. 15 médicaments + 5 moments différents = 30 prises par jour, un pilulier svp. . *Le Médecin du Québec* 2008; **43**(12): 65-71.

211. Elseviers M, Wettermark B, Almarsdóttir A, et al. Chapter 36: Assessment of adherence to drug treatment in database research. Drug utilization research : methods and applications. Chichester, West Sussex Hoboken, NJ: John Wiley & Sons Inc.,; 2016.

212. Rosenbaum PR. Observational studies. 2nd ed. New York, NY: Springer-Verlag; 2002.

213. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical statistics* 2011; **10**(2): 150-61.

214. Snijders TAB, Bosker RJ. Multilevel analysis : an introduction to basic and advanced multilevel modeling. London: Sage Publications; 1999.

215. Merlo J, Chaix B, Yang M, Lynch J, Rastam L. A brief conceptual tutorial of multilevel analysis in social epidemiology: linking the statistical concept of clustering to the idea of contextual phenomenon. *Journal of epidemiology and community health* 2005; **59**(6): 443-9.

216. Strom, Brian L.; Kimmel, Stephen E; Hennessy, Sean. (2011). Pharmacoepidemiology. Chapter 37. Use of pharmacoepidemiology to study benefitial drug effects. Hoboken, NJ: Wiley-Blackwell.

217. Kahlert J, Gribsholt SB, Gammelager H, Dekkers OM, Luta G. Control of confounding in the analysis phase - an overview for clinicians. *Clinical Epidemiology* 2017; **9**: 195-204.

218. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983; **70**(1): 41-55.

219. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research* 2011; **46**(3): 399-424.

220. Elze MC, Gregson J, Baber U, et al. Comparison of Propensity Score Methods and Covariate Adjustment: Evaluation in 4 Cardiovascular Studies. *Journal of the American College of Cardiology* 2017; **69**(3): 345-57.

221. Rubin DB. Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services and Outcomes Research Methodology* 2001; **2**(3): 169-88.

222. Rubin DB. Estimating causal effects from large data sets using propensity scores. *Annals of internal medicine* 1997; **127**(8 Pt 2): 757-63.

223. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Sturmer T. Variable selection for propensity score models. *Am J Epidemiol* 2006; **163**(12): 1149-56.

224. Austin PC, Grootendorst P, Anderson GM. A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: a Monte Carlo study. *Statistics in medicine* 2007; **26**(4): 734-53.

225. Flores B, Trivedi HD, Robson SC, Bonder A. Hemostasis, bleeding and thrombosis in liver disease. *Journal of Translational Science* 2017; **3**(3).

226. Zaorsky NG, Zhang Y, Tchelebi LT, Mackley HB, Chinchilli VM, Zacharia BE. Stroke among cancer patients. *Nature Communications* 2019; **10**(1): 5172.

227. Al-Samkari H, Connors JM. Managing the competing risks of thrombosis, bleeding, and anticoagulation in patients with malignancy. *Blood advances* 2019; **3**(22): 3770-9.

228. Gaudino M, Angiolillo DJ, Franco AD, et al. Stroke after coronary artery bypass grafting and percutaneous coronary intervention: incidence, pathogenesis, and outcomes. *Journal of the American Heart Association* 2019; **8**(13): e013032.

229. Ryad R, Saad-Omer SM, Khan F, Limbana T, Jahan N. Does catheter ablation lower the long-term risk of stroke and mortality in patients with atrial fibrillation? A concise review of the current state of knowledge. *Cureus* 2020; **12**(8): e9701.

230. Healey JS, Connolly SJ, Gold MR, et al. Subclinical atrial fibrillation and the risk of stroke. *New England Journal of Medicine* 2012; **366**(2): 120-9.

231. Merck Manal Professional Version. Merck & Co., Inc. Retrieved February 18, 2021, from https://www.merckmanuals.com/en-ca/professional.

232. Fernando E, Fraser M, Hendriksen J, Kim CH, Muir-Hunter SW. Risk factors associated with falls in older adults with dementia: a systematic review. *Physiotherapy Canada* 2017; **69**(2): 161-70.

233. El-Saifi N, Moyle W, Jones C, Tuffaha H. Medication adherence in older patients with dementia: a systematic literature review. *Journal of Pharmacy Practice* 2018; **31**(3): 322-34.

234. Yao X, Abraham NS, Alexander GC, et al. Effect of adherence to oral anticoagulants on risk of stroke and major bleeding among patients with atrial fibrillation. *Journal of the American Heart Association* 2016; **5**(2): e003074.

235. Hénon H, Pasquier F, Durieu M, et al. Preexisting dementia in stroke patients. *Stroke* 1997; **28**(12): 2429-36.

236. Preen DB, Holman CD, Spilsbury K, Semmens JB, Brameld KJ. Length of comorbidity lookback period affected regression model performance of administrative health data. *Journal of clinical epidemiology* 2006; **59**(9): 940-6.

237. Fortin Y, Crispo JAG, Cohen D, McNair DS, Mattison DR, Krewski D. Optimal look back period and summary method for Elixhauser comorbidity measures in a US population-based electronic health record database. Open Access Medical Statistics. 2017;7:1-13

https://doi.org/10.2147/OAMS.S120426.

238. Yurkovich M, Avina-Zubieta JA, Thomas J, Gorenchtein M, Lacaille D. A systematic review identifies valid comorbidity indices derived from administrative health data. *Journal of clinical epidemiology* 2015; **68**(1): 3-14.

239. Wooldridge, JM. Econometric analysis of cross section and panel data. MIT Press; Cambridge, MA: 2001.

240. McDonald JH. Handbook of biological statistics. 3rd ed. Baltimore, Maryland: Sparky house publishing; 2014.

241. West BT. Linear mixed models : a practical guide using statistical software. In: Welch KB, Gałecki AT, Gillespie BW, editors. Second edition. ed.

242. Cochran, W., 1977. Sampling Techniques, third ed. Wiley, New York.

243. Zięba A. Effective number of observations and unbiased estimators of variance for autocorrelated data - an overview. *Metrology and Measurement Systems* 2010; **XVII**.

244. SAS Institute Inc. SAS/STAT® 14.1 User's Guide: Cary, NC: SAS Institute Inc.; 2015.

245. Lumley T, Diehr P, Emerson S, Chen L. The importance of the normality assumption in large public health data sets. *Annual review of public health* 2002; **23**: 151-69.

246. Schielzeth H, Dingemanse NJ, Nakagawa S, et al. Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution* 2020; **11**(9): 1141-52.

247. Singer JD, Willett JB. Applied longitudinal data analysis: modeling change and event occurrence. Chapter 4. Doing data analysis with the multilevel model for change: Oxford University Press; 2003.

248. Hernandez I, Baik SH, Piñera A, Zhang Y. Risk of bleeding with dabigatran in atrial fibrillation. *The Journal of the American Medical Association Internal Medicine* 2015; **175**(1): 18-24.

249. Noseworthy PA, Yao X, Abraham NS, Sangaralingham LR, McBane RD, Shah ND. Direct comparison of dabigatran, rivaroxaban, and apixaban for effectiveness and safety in nonvalvular atrial fibrillation. *Chest* 2016; **150**(6): 1302-12.

250. Schjerning Olsen AM, Gislason GH, McGettigan P, et al. Association of NSAID use with risk of bleeding and cardiovascular events in patients receiving antithrombotic therapy after myocardial infarction. *The Journal of the American Medical Association* 2015; **313**(8): 805-14.

251. Meid AD, Heider D, Adler JB, et al. Comparative evaluation of methods approximating drug prescription durations in claims data: modeling, simulation, and application to real data. *Pharmacoepidemiol Drug Saf* 2016; **25**(12): 1434-42.

252. Skanes AC, Healey JS, Cairns JA, et al. Focused 2012 update of the Canadian Cardiovascular Society atrial fibrillation guidelines: recommendations for stroke prevention and rate/rhythm control. *Canadian Journal of Cardiology* 2012; **28**(2): 125-36.

253. Avgil-Tsadok M, Jackevicius CA, Essebag V, et al. Dabigatran use in elderly patients with atrial fibrillation. *Thrombosis and Haemostasis* 2016; **115**(1): 152-60.

254. Wigle P, Hein B, Bloomfield HE, Tubb M, Doherty M. Updated guidelines on outpatient anticoagulation. *American family physician* 2013; **87**(8): 556-66.

255. Hernandez I, Zhang Y, Saba S. Comparison of the effectiveness and safety of apixaban, dabigatran, rivaroxaban, and warfarin in newly diagnosed atrial fibrillation. *The American journal of cardiology* 2017; **120**(10): 1813-9.

256. Joos C, Lawrence K, Jones AE, Johnson SA, Witt DM. Accuracy of ICD-10 codes for identifying hospitalizations for acute anticoagulation therapy-related bleeding events. *Thrombosis Research* 2019; **181**: 71-6.

257. Arnason T, Wells PS, van Walraven C, Forster AJ. Accuracy of coding for possible warfarin complications in hospital discharge abstracts. *Thrombosis Research* 2006; **118**(2): 253-62.

258. Quan H, Parsons GA, Ghali WA. Validity of procedure codes in International Classification of Diseases, 9th revision, Clinical Modification Administrative Data. *Medical Care* 2004; **42**(8): 801-9.

259. Kirkwood BR, Sterne JAC, Kirkwood BR. Essential medical statistics. 2003.

260. Toh S, Hernandez-Diaz S, Logan R, Robins JM, Hernan MA. Estimating absolute risks in the presence of nonadherence: an application to a follow-up study with baseline randomization. *Epidemiology* 2010; **21**(4): 528-39.

261. January CT, Wann LS, Alpert JS, et al. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on practice guidelines and the Heart Rhythm Society. *Circulation* 2014; **130**(23): e199-267.

262. Kirchhof P, Lip GY, Van Gelder IC, et al. Comprehensive risk reduction in patients with atrial fibrillation: Emerging diagnostic and therapeutic options. Executive summary of the report from the 3rd AFNET/EHRA consensus conference. *Thrombosis and Haemostasis* 2011; **106**(6): 1012-9.

263. Cordoba G, Schwartz L, Woloshin S, Bae H, Gøtzsche PC. Definition, reporting, and interpretation of composite outcomes in clinical trials: systematic review. *BMJ (Clinical research ed)* 2010; **341**: c3920.

264. Durand M, Schnitzer ME, Pang M, et al. Comparative effectiveness and safety of direct oral anticoagulants versus vitamin K antagonists in nonvalvular atrial fibrillation: a Canadian multicentre observational cohort study. *CMAJ Open* 2020; **8**(4): E877-e86.

265. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol* 2016; **183**(8): 758-64.

266. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in medicine* 2009; **28**(25): 3083-107.

267. Hernán MA, Robins JM (2020). Causal Inference: What If. Boca Raton: Chapman & Hall/CRC.

268. Thompson WA, Jr. On the treatment of grouped observations in life studies. *Biometrics* 1977; **33**(3): 463-70.

269. Cole SR, Hernán MA, Anastos K, Jamieson BD, Robins JM. Determining the effect of highly active antiretroviral therapy on changes in human immunodeficiency virus type 1 RNA viral load using a marginal structural left-censored mean model. *Am J Epidemiol* 2007; **166**(2): 219-27.

270. Sinyavskaya L, Matteau A, Johnson S, Durand M. Methodological challenges in assessment of current use of warfarin among patients with atrial fibrillation using dispensation data from administrative health care databases. *Pharmacoepidemiol Drug Saf* 2018; **27**(9): 979-86.

271. Holbrook AM, Pereira JA, Labiris R, et al. Systematic overview of warfarin and its drug and food interactions. *Archives of internal medicine* 2005; **165**(10): 1095-106.

272. Villines TC, Peacock WF. Safety of direct oral anticoagulants: insights from postmarketing studies. *American Journal of Medicine* 2016; **129**(11S): S41-S6.

273. Vaughan Sarrazin MS, Jones M, Mazur A, Chrischilles E, Cram P. Bleeding rates in Veterans Affairs patients with atrial fibrillation who switch from warfarin to dabigatran. *American Journal of Medicine* 2014; **127**(12): 1179-85.

274. Yao X, Abraham NS, Sangaralingham LR, et al. Effectiveness and safety of dabigatran, rivaroxaban, and apixaban versus warfarin in nonvalvular atrial fibrillation. *Journal of the American Heart Association* 2016; **5**(6).

275. LeLorier J. Pharmacology: the neglected half of pharmacoepidemiology. The Canadian network for observational drug effect studies (CNODES), 2015.

276. Patel SI, Cherington C, Scherber R, et al. Assessment of patient adherence to direct oral anticoagulant vs warfarin therapy. *The Journal of the American Osteopathic Association* 2017; **117**(1): 7-15.

277. Ryan R, Santesso N, Lowe D, et al. Interventions to improve safe and effective medicines use by consumers: an overview of systematic reviews. *Cochrane Database of Systematic Reviews* 2014; (4): Cd007768.

278. Sorensen R, Jamie Nielsen B, Langtved Pallisgaard J, Ji-Young Lee C, Torp-Pedersen C. Adherence with oral anticoagulation in non-valvular atrial fibrillation: a comparison of vitamin K antagonists and non-vitamin K antagonists. *European Heart Journal - Cardiovascular Pharmacotherapy* 2017; **3**(3): 151-6.

279. Schulman S. Advantages and limitations of the new anticoagulants. *Journal of internal medicine* 2014; **275**(1): 1-11.

280. Zaheer-Ud-Din B. Encyclopedia of pharmacy practice and clinical pharmacy. 1st ed: Huddersfield, ELSEVIER Academic Press; 2019.

281. Furu K, Wettermark B, Andersen M, Martikainen JE, Almarsdottir AB, Sørensen HT. The Nordic countries as a cohort for pharmacoepidemiological research. *Basic & clinical pharmacology & toxicology* 2010; **106**(2): 86-94.

282. Lum KJ, Newcomb CW, Roy JA, et al. Evaluation of methods to estimate missing days' supply within pharmacy data of the Clinical Practice Research Datalink (CPRD) and The Health Improvement Network (THIN). *European Journal of Clinical Pharmacology* 2017; **73**(1): 115-23.

283. Brown JD, Shewale AR, Talbert JC. Adherence to rivaroxaban, dabigatran, and apixaban for stroke prevention for newly diagnosed and treatment-naive atrial fibrillation patients: an update using 2013-2014 data. *The Journal of Managed Care & Specialty Pharmacy* 2017; **23**(9): 958-67.

284. Sinyavskaya L, Schnitzer M, Renoux C, Guertin JR, Talbot D, Durand M. Evidence of the different associations of prognostic factors with censoring across treatment groups and impact on censoring weight model specification: the example of anticoagulation in atrial fibrillation. *The American journal of cardiology* 2021.

285. Templeton AJ, Amir E, Tannock IF. Informative censoring — a neglected cause of bias in oncology trials. *Nature Reviews Clinical Oncology* 2020; **17**(6): 327-8.

286. Hernán MA, Hernández-Díaz S, Robins JM. Randomized trials analyzed as observational studies. *Annals of internal medicine* 2013; **159**(8): 560-2.

287. Chung H, Deshpande G, Zolotarjova J, et al. Health plan enrollment and disenrollment of individuals with and without established chronic disease in a U.S. commercially insured and

medicare advantage population. *The Journal of Managed Care & Specialty Pharmacy* 2019; **25**(5): 612-20.

288. Butler AM, Todd JV, Sahrmann JM, Lesko CR, Brookhart MA. Informative censoring by health plan disenrollment among commercially insured adults. *Pharmacoepidemiology and drug safety* 2019; **28**(5): 640-8.

289. Willems S, Schat A, van Noorden MS, Fiocco M. Correcting for dependent censoring in routine outcome monitoring data by applying the inverse probability censoring weighted estimator. *Statistical Methods in Medical Research* 2018; **27**(2): 323-35.

290. American Heart Association. Atrial Fibrillation. https://www.heart.org/en/health-topics/atrial-fibrillation (accessed October 26, 2020.

291. Howe CJ, Cole SR, Lau B, Napravnik S, Eron JJ, Jr. Selection bias due to loss to follow up in cohort studies. *Epidemiology* 2016; **27**(1): 91-7.

292. Stuart EA. Matching methods for causal inference: A review and a look forward. *Statistical Science* 2010; **25**(1): 1-21.

293. Patel MR, Mahaffey KW, Garg J, et al. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *New England Journal of Medicine* 2011; **365**(10): 883-91.

294. Connolly SJ, Ezekowitz MD, Yusuf S, et al. Dabigatran versus warfarin in patients with atrial fibrillation. *New England Journal of Medicine* 2009; **361**(12): 1139-51.

295. Granger CB, Alexander JH, McMurray JJ, et al. Apixaban versus warfarin in patients with atrial fibrillation. *New England Journal of Medicine* 2011; **365**(11): 981-92.

296. Colacci M, Tseng EK, Sacks CA, Fralick M. Oral anticoagulant utilization in the United States and United Kingdom. *Journal of general internal medicine* 2020; **35**(8): 2505-7.

297. Lee SR, Choi EK, Park CS, et al. Direct oral anticoagulants in patients with nonvalvular atrial fibrillation and low body weight. *Journal of the American College of Cardiology* 2019; **73**(8): 919-31.

298. Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Medical Research Methodology* 2008; **8**(1): 70.

299. Thornley S, Marshall RJ, Wells S, Jackson R. Using directed acyclic graphs for investigating causal paths for cardiovascular disease. *Journal of Biometrics and Biostatistics* 2013; **4**(182).

300. Etminan M, Collins GS, Mansournia MA. Using causal diagrams to improve the design and interpretation of medical research. *Chest* 2020; **158**(1s): S21-s8.

301. Williams TC, Bach CC, Matthiesen NB, Henriksen TB, Gagliardi L. Directed acyclic graphs: a tool for causal studies in paediatrics. *Pediatric Research* 2018; **84**(4): 487-93.

302. McHorney CA, Ashton V, Laliberté F, et al. Adherence to rivaroxaban compared with other oral anticoagulant agents among patients with nonvalvular atrial fibrillation. *The Journal of Managed Care & Specialty Pharmacy* 2017; **23**(9): 980-8.

303. Régie de l'assurance maladie. Amount to pay for prescription drugs. https://www.ramq.gouv.qc.ca/en/citizens/prescription-drug-insurance/amount-pay-prescription-drugs (accessed July 23 2022.

304. Ravvaz K, Weissert JA, Jahangir A, Ruff CT. Evaluating the effects of socioeconomic status on stroke and bleeding risk scores and clinical events in patients on oral anticoagulant for new onset atrial fibrillation. *The public library of science (PLOS) One* 2021; **16**(3): e0248134.

305. Blaser, P., 2020. Material and social deprivation index: A summary - OVERVIEW OF THE METHODOLOGY, Institut national de santé publique du Québec. Retrieved from https://policycommons.net/artifacts/2064276/material-and-social-deprivation-index/2817367/ on 24 Jul 2022. CID: 20.500.12592/p93fmh.

306. Bourgon Labelle J, Farand P, Vincelette C, Dumont M, Le Blanc M, Rochefort CM. Validation of an algorithm based on administrative data to detect new onset of atrial fibrillation after cardiac surgery. *BMC Medical Research Methodology* 2020; **20**(1): 75.

307. Wang MC, Laud PW, Macias M, Nattinger AB. Strengths and limitations of International Classification of Disease Ninth Revision Clinical Modification codes in defining cervical spine surgery. *Spine (Phila Pa 1976)* 2011; **36**(1): E38-44.

308. Birman-Deych E, Waterman AD, Yan Y, Nilasena DS, Radford MJ, Gage BF. Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors. *Medical Care* 2005; **43**(5): 480-5.

309. Elena B-D, Amy DW, Yan Y, David SN, Martha JR, Brian FG. Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors. *Medical Care* 2005; **43**(5): 480-5.

310. Hall R, Mondor L, Porter J, Fang J, Kapral MK. Accuracy of administrative data for the coding of acute stroke and TIAs. *Canadian Journal of Neurological Sciences* 2016; **43**(6): 765-73.

311. Schöttker B, Haug U, Schomburg L, et al. Strong associations of 25-hydroxyvitamin D concentrations with all-cause, cardiovascular, cancer, and respiratory disease mortality in a large cohort study. *The American journal of clinical nutrition* 2013; **97**(4): 782-93.

312. Kosmas CE, Silverio D, Ovalle J, Montan PD, Guzman E. Patient adherence, compliance, and perspectives on evolocumab for the management of resistant hypercholesterolemia. *Patient Preference and Adherence* 2018; **12**: 2263-6.

313. Barcellona D, Vannini ML, Fenu L, Balestrieri C, Marongiu F. Warfarin or acenocoumarol: which is better in the management of oral anticoagulants? *Thrombosis and Haemostasis* 1998; **80**(6): 899-902.

314. Pattacini C, Manotti C, Pini M, Quintavalla R, Dettori AG. A comparative study on the quality of oral anticoagulant therapy (warfarin versus acenocoumarol). *Thrombosis and Haemostasis* 1994; **71**(2): 188-91.

315. Yang S-Y, Kang D-W, Nam JH, et al. Adherence is an optimal factor for maximizing the effective and safe use of oral anticoagulants in patients with atrial fibrillation. *Scientific Reports* 2022; **12**(1): 3413.

316. Canadian Agency for Drugs and Technologies in Health (CADTH). Point-of-care testing of the International Normalized Ratio (INR) for patients taking warfarin or other vitamin K antagonists. Health technology review. https://www.cadth.ca/point-care-testing-international-normalized-ratio-inr-patients-taking-warfarin-or-other-vitamin-k (accessed 21 July 2022.

317. Lin KJ, Schneeweiss S, Pawar A, Singer DE, Liu J, Gagne JJ. Using a simple prescription gap to determine warfarin discontinuation can lead to substantial misclassification. *Thrombosis and Haemostasis* 2022; **122**(3): 386-93.

318. Kvarnström K, Westerholm A, Airaksinen M, Liira H. Factors contributing to medication adherence in patients with a chronic condition: a scoping review of qualitative research. *Pharmaceutics* 2021; **13**(7).

319. Briasoulis A, Gao Y, Inampudi C, et al. Characteristics and outcomes in patients with atrial fibrillation receiving direct oral anticoagulants in off-label doses. *BMC Cardiovascular Disorders* 2020; **20**(1): 42.

320. Roy L, Zappitelli M, White-Guay B, Lafrance JP, Dorais M, Perreault S. Agreement between administrative database and medical chart review for the prediction of chronic kidney disease G category. *Canadian Journal of Kidney Health and Disease* 2020; **7**: 2054358120959908.

321. Laird NM, Mosteller F. Some statistical methods for combining experimental results. *International Journal of Technology Assessment in Health Care* 1990; **6**(1): 5-30.

Annexes

Annex 1. Methods and the list of references used for the creation of forest plots displaying hazard ratios of stroke and systemic embolism, major bleeding, and GI bleeding reported in observational studies comparing DOACs and warfarin.

In the scope of these theses, I extracted some findings from the systematic review and meta-analyses of observational studies comparing the effectiveness and safety of COACs versus warfarin conducted by Escobar et al.[101], Bai et al.[102], Hirschl et al.[103], Deitelzweig et al.[105], and Lowenstern[104]. The objective of this analysis was to investigate the consistency in the direction and magnitude of the reported effect estimates across the individual observational studies comparing dabigatran, rivaroxaban, or apixaban with warfarin. The secondary objective was to combine the estimates to produce summary statistics of the effect of exposure to each DOAC vs warfarin on the risk of selected effectiveness and safety outcomes. The test for heterogeneity was done with the Cochran Q test (X^2 test) with a significance level of 0.10. This test examines the null hypothesis that the effect estimates are the same across the participating studies[106]. Additionally, I^2 test measured the percentage of the total variation in the effect estimates that was not due to chance[107]. The summary statistics were estimated using two methods: the fixed-effect method (Mantel-Haenszel), and the random effect method (DerSimonian-Laird) [321]. The fixed-effect method is based on the assumption that the true effect is the same in all participating studies, whereas, the random effect method has an assumption that the participating studies are a random sample from the hypothetical population of studies, and the magnitude of the true estimate may vary across the studies[321].

The reference list of individual studies used to create the forest plots.

Abraham NS, Singh S, Alexander GC, Heien H, Haas LR, Crown W, Shah ND. Comparative risk of gastrointestinal bleeding with dabigatran, rivaroxaban, and warfarin: population based cohort study. *BMJ* 2015; 350: h1857.

Adeboyeje G, Sylwestrzak G, Barron JJ, White J, Rosenberg A, Abarca J, Crawford G, Redberg R. Major bleeding risk during anticoagulation with warfarin, dabigatran, apixaban, or rivaroxaban in patients with nonvalvular atrial fibrillation. *The Journal of Managed Care & Specialty Pharmacy* 2017; 23: 968-78.

Amin A, Keshishian A, Trocio J, Dina O, Le H, Rosenblatt L, Liu X, Mardekian J, Zhang Q, Baser O, Vo L. Risk of stroke/systemic embolism, major bleeding and associated costs in non-valvular atrial fibrillation patients who initiated apixaban, dabigatran or rivaroxaban compared with warfarin in the United States Medicare population. *Current medical research and opinion* 2017; 33: 1595-604.

Bouillon K, Bertrand M, Maura G, Blotière PO, Ricordeau P, Zureik M. Risk of bleeding and arterial thromboembolism in patients with non-valvular atrial fibrillation either maintained on a vitamin K antagonist or switched to a non-vitamin K-antagonist oral anticoagulant: a retrospective, matched-cohort study. *Lancet Haematol*ogy 2015; 2: e150-9.

Chan YH, Kuo CT, Yeh YH, Chang SH, Wu LS, Lee HF, Tu HT, See LC. Thromboembolic, bleeding, and mortality risks of rivaroxaban and dabigatran in Asians with nonvalvular atrial fibrillation. *Journal of the American College of Cardiology* 2016; 68: 1389-401.

Chang HY, Zhou M, Tang W, Alexander GC, Singh S. Risk of gastrointestinal bleeding associated with oral anticoagulants: population based retrospective cohort study. *BMJ* 2015; 350: h1585.

Coleman CI, Peacock WF, Bunz TJ, Alberts MJ. Effectiveness and Safety of Apixaban, Dabigatran, and Rivaroxaban Versus Warfarin in Patients With Nonvalvular Atrial Fibrillation and Previous Stroke or Transient Ischemic Attack. *Stroke* 2017; 48: 2142-49.

Deitelzweig S, Luo X, Gupta K, Trocio J, Mardekian J, Curtice T, Lingohr-Smith M, Menges B, Lin J. Comparison of effectiveness and safety of treatment with apixaban vs. other oral anticoagulants among elderly nonvalvular atrial fibrillation patients. *Current medical research and opinion* 2017; 33: 1745-54.

Ellis MH, Neuman T, Bitterman H, Dotan SG, Hammerman A, Battat E, Eikelboom JW, Ginsberg JS, Hirsh J. Bleeding in patients with atrial fibrillation treated with dabigatran, rivaroxaban or warfarin: A retrospective population-based cohort study. *European Journal of Internal Medicine* 2016: 33: 55-9.

Forslund TB Wettermark MA, Hjemdahl P. Stroke and bleeding with non-vitamin K antagonist oral anticoagulant or warfarin treatment in patients with non-valvular atrial fibrillation: a population-based cohort study. *Europace* 2018; 20: 420-28.

Graham DJ, Reichman ME, Wernecke M, Zhang R, Southworth MR, Levenson M, Sheu TC, Mott K, Goulding MR, Houstoun M, MaCurdy TE, Worrall C, Kelman JA. Cardiovascular, bleeding, and mortality risks in elderly Medicare patients treated with dabigatran or warfarin for nonvalvular atrial fibrillation. *Circulation* 2015; 131: 157-64.

Halvorsen S, Ghanima W, Tvete IF, Hoxmark C, Falck P, Solli O, Jonasson C. A nationwide registry study to compare bleeding rates in patients with atrial fibrillation being prescribed oral anticoagulants. *European Heart Journal - Cardiovascular Pharmacotherapy* 2017; 3: 28-36.

Hernandez I, Baik SH, Piñera A, Zhang Y. Risk of bleeding with dabigatran in atrial fibrillation. *The Journal of the American Medical Association Internal Medicine* 2015; 175: 18-24.

Hernandez I, Zhang Y, Saba S. Comparison of the effectiveness and safety of apixaban, dabigatran, rivaroxaban, and warfarin in newly diagnosed atrial fibrillation. *The American Journal of Cardiology* 2017; 120: 1813-19.

Ho CW, Ho MH, Chan PH, Hai JJ, Cheung E, Yeung CY, Lau KK, Chan KH, Lau CP, Lip GY, Leung GK, Tse HF, Siu CW. Ischemic stroke and intracranial hemorrhage with aspirin, dabigatran, and warfarin: impact of quality of anticoagulation control. *Stroke* 2015; 46: 23-30.

Kohsaka S, Murata T, Izumi N, Katada J, Wang F, Terayama Y. Bleeding risk of apixaban, dabigatran, and low-dose rivaroxaban compared with warfarin in Japanese patients with non-valvular atrial fibrillation: a propensity matched analysis of administrative claims data. *Current Medical Research and Opinion* 2017; 33: 1955-63.

Laliberté F, Cloutier M, Nelson WW, Coleman CI, Pilon D, Olson WH, Damaraju CV, Schein JR, Lefebvre P. Real-world comparative effectiveness and safety of rivaroxaban and warfarin in nonvalvular atrial fibrillation patients. *Current Medical Research and Opinion* 2014: 30: 1317-25.

Larsen TB, Skjøth F, Nielsen PB, Kjældgaard JN, Lip GY. Comparative effectiveness and safety of non-vitamin K antagonist oral anticoagulants and warfarin in patients with atrial fibrillation: propensity weighted nationwide cohort study. *BMJ* 2016; 353: i3189.

Lauffenburger JC, Farley JF, Gehi AK, Rhoney DH, Brookhart MA, Fang G. Effectiveness and safety of dabigatran and warfarin in real-world US patients with non-valvular atrial fibrillation: a retrospective cohort study. *Journal of the American Heart Association* 2015; 4:e001798 doi: 10.1161.

Li XS, Deitelzweig S, Keshishian A, Hamilton M, Horblyuk R, Gupta K, Luo X, Mardekian J, Friend K, Nadkarni A, Pan X, Lip GY. Effectiveness and safety of apixaban versus warfarin in non-valvular atrial fibrillation patients in "real-world" clinical practice. A propensity-matched analysis of 76,940 patients. *Thrombosis and Haemostasis* 2017; 117: 1072-82.

Lip GY, Skjøth F, Nielsen PB, Kjældgaard JN, Larsen TB. Effectiveness and Safety of Standard-Dose Nonvitamin K Antagonist Oral Anticoagulants and Warfarin Among Patients With Atrial Fibrillation With a Single Stroke Risk Factor: A Nationwide Cohort Study. *JAMA Cardiology* 2017; 2: 872-81.

Lip GY, Pan X, Kamble S, Kawabata H, Mardekian J, Masseria C, Bruno A, Phatak H. Major bleeding risk among non-valvular atrial fibrillation patients initiated on apixaban, dabigatran, rivaroxaban or warfarin: a "real-world" observational study in the United States. *International Journal of Clinical Practice* 2016; 70: 752-63.

Maura G, Blotière PO, Bouillon K, Billionnet C, Ricordeau P, Alla F, Zureik M. Comparison of the short-term risk of bleeding and arterial thromboembolic events in nonvalvular atrial fibrillation patients newly treated with dabigatran or rivaroxaban versus vitamin K antagonists: a French nationwide propensity-matched cohort study. *Circulation* 2015; 132: 1252-60.

Naganuma M, Shiga T, Nagao T, Suzuki A, Murasaki K, Hagiwara N. Effectiveness and safety of dabigatran versus warfarin in "real-world" Japanese patients with atrial fibrillation: A single-center observational study. *Journal of Arrhythmia* 2017; 33: 107-10.

Nishtala PS, Gnjidic D, Jamieson HA, Hanger HC, Kaluarachchi C, Hilmer SN. 'Real-world' haemorrhagic rates for warfarin and dabigatran using population-level data in New Zealand. International Journal of Cardiology 2016; 203: 746-52.

Seeger JD, Bykov K, Bartels DB, Huybrechts K, Zint K, Schneeweiss S. Safety and effectiveness of dabigatran and warfarin in routine care of patients with atrial fibrillation. *Thrombosis and Haemostasis* 2015; 114: 1277-89.

Staerk L, Fosbøl EL, Lip GY, Lamberts M, Bonde AN, Torp-Pedersen C, Ozenne B, Gerds TA, Gislason GH, Olesen JB. Ischaemic and haemorrhagic stroke associated with non-vitamin K antagonist oral anticoagulants and warfarin use in patients with atrial fibrillation: a nationwide cohort study. *European Heart Journal* 2017; 38: 907-15.

Yao X, Abraham NS, Sangaralingham LR, Bellolio MF, McBane RD, Shah ND, Noseworthy PA. Effectiveness and safety of dabigatran, rivaroxaban, and apixaban versus warfarin in nonvalvular atrial fibrillation. *Journal of the American Heart Association* 2016; 5:e003725 doi: 10.1161.

Yap SH, Ng YP, Roslan A, Kolanthaivelu J, Koh KW, P'Ng SH, Boo YL, Hoo FK, Yap LB. A comparison of dabigatran and warfarin for stroke prevention in elderly Asian population with nonvalvular atrial fibrillation: An audit of current practice in Malaysia. *Medical Journal of Malaysia* 2017; 72: 360-64.

Annex 2. List of Drug Identification Numbers (DIN) used to define exposure to anticoagulants in all studies of this thesis.

DIN	Ingredient	Brand name	Strength, mg
Warfarin			
10308	Warfarin sodium	Warfilone	5
1918311	Warfarin sodium	Coumadin	1
1918338	Warfarin sodium	Coumadin	2
1918346	Warfarin sodium	Coumadin	2.5
1918354	Warfarin sodium	Coumadin	5
1918362	Warfarin sodium	Coumadin	10
2007959	Warfarin sodium	Coumadin	4
2229741	Warfarin sodium	Coumadin	5
2240205	Warfarin sodium	Coumadin	3
2240206	Warfarin sodium	Coumadin	6
2242680	Warfarin sodium	Taro-warfarin	1
2242681	Warfarin sodium	Taro-warfarin	2
2242682	Warfarin sodium	Taro-warfarin	2.5
2242683	Warfarin sodium	Taro-warfarin	3
2242684	Warfarin sodium	Taro-warfarin	4
2242685	Warfarin sodium	Taro-warfarin	5
2242686	Warfarin sodium	Taro-warfarin	6
2242687	Warfarin sodium	Taro-warfarin	10
2242697	Warfarin sodium	Taro-warfarin	7.5
2242924	Warfarin sodium	Apo-warfarin	1
2242925	Warfarin sodium	Apo-warfarin	2
2242926	Warfarin sodium	Apo-warfarin	2.5
2242927	Warfarin sodium	Apo-warfarin	4
2242928	Warfarin sodium	Apo-warfarin	5
2242929	Warfarin sodium	Apo-warfarin	10
2244462	Warfarin sodium	Mylan-warfarin	1
2244463	Warfarin sodium	Mylan-warfarin	2
2244464	Warfarin sodium	Mylan-warfarin	2.5
2244465	Warfarin sodium	Mylan-warfarin	4
2244466	Warfarin sodium	Mylan-warfarin	5
2244467	Warfarin sodium	Mylan-warfarin	10
2245618	Warfarin sodium	Apo-warfarin	3
2265273	Warfarin sodium	Novo-warfarin	1
2265281	Warfarin sodium	Novo-warfarin	2
2265303	Warfarin sodium	Novo-warfarin	2.5
2265311	Warfarin sodium	Novo-warfarin	3
2265338	Warfarin sodium	Novo-warfarin	4

2265346	Warfarin sodium	Novo-warfarin	5
2265354	Warfarin sodium	Novo-warfarin	6
2265362	Warfarin sodium	Novo-warfarin	7.5
2265370	Warfarin sodium	Novo-warfarin	10
2287498	Warfarin sodium	Mylan-warfarin	3
2287501	Warfarin sodium	Mylan-warfarin	6
2287528	Warfarin sodium	Mylan-warfarin	7.5
2335611	Warfarin sodium	Nu-warfarin	1
2335638	Warfarin sodium	Nu-warfarin	2
2335646	Warfarin sodium	Nu-warfarin	2.5
2335654	Warfarin sodium	Nu-warfarin	3
2335662	Warfarin sodium	Nu-warfarin	4
2335670	Warfarin sodium	Nu-warfarin	5
2335689	Warfarin sodium	Nu-warfarin	10
2344025	Warfarin sodium	Warfarin	1
2344033	Warfarin sodium	Warfarin	2
2344041	Warfarin sodium	Warfarin	2.5
2344068	Warfarin sodium	Warfarin	3
2344076	Warfarin sodium	Warfarin	4
2344084	Warfarin sodium	Warfarin	5
2344092	Warfarin sodium	Warfarin	6
2344106	Warfarin sodium	Warfarin	7.5
2344114	Warfarin sodium	Warfarin	10
9296	Warfarin	Coumadin	2
9318	Warfarin	Coumadin	2.5
9342	Warfarin	Coumadin	10
585629	Warfarin	Coumadin	5
585637	Warfarin	Coumadin	10
585645	Warfarin	Coumadin	2.5
585653	Warfarin	Coumadin	2
861634	Warfarin	Coumadin	1
2152460	Warfarin sodium	Endo warfarin	1
2152479	Warfarin sodium	Endo warfarin	2
2152487	Warfarin sodium	Endo warfarin	2.5
2152495	Warfarin sodium	Endo warfarin	4
2152509	Warfarin sodium	Endo warfarin	5
2152517	Warfarin sodium	Endo warfarin	10
2242881	Warfarin sodium	Coumadin	1
2242882	Warfarin sodium	Coumadin	2
2242883	Warfarin sodium	Coumadin	2.5
2242884	Warfarin sodium	Coumadin	3
2242885	Warfarin sodium	Coumadin	4
2242886	Warfarin sodium	Coumadin	5

2242887	Warfarin sodium	Coumadin	6
2242888	Warfarin sodium	Coumadin	7.5
2242889	Warfarin sodium	Coumadin	10
2245817	Warfarin sodium	Lin warfarin	3
2245818	Warfarin sodium	Lin warfarin	6
2311070	Warfarin sodium	Coumadin	1
2311089	Warfarin sodium	Coumadin	2
2311097	Warfarin sodium	Coumadin	2.5
2311100	Warfarin sodium	Coumadin	3
2311119	Warfarin sodium	Coumadin	4
2311127	Warfarin sodium	Coumadin	5
2311135	Warfarin sodium	Coumadin	10
Apixaban			
2377233	Apixaban	Eliquis	2.5
2397714	Apixaban	Eliquis	5
Dabigatran			
2312433	Dabigatran etexilate	Pradaxa	75
2312441	Dabigatran etexilate	Pradaxa	110
2358808	Dabigatran etexilate	Pradaxa	150
Rivaroxaban			
2316986	Rivaroxaban	Xarelto	10
2378604	Rivaroxaban	Xarelto	15
2378612	Rivaroxaban	Xarelto	20
2441535	Rivaroxaban	Xarelto	15

Annex 3. List of the International Classification of Diseases (ICD) codes used in all studies of the thesis.

Condition	ICD-9	ICD-10
Venous thromboembolism (includes deep vein thrombosis and pulmonary embolism)	415.1, 451.1, 451.2, 451.81, 451.83, 451.8, 451.9, 453.2, 453.8, 453.9	I80.2, I80.3, I80.1, I80.8, I80.9, I82.8, I82.9, O22.3, O22.9, O87.1, I26.9, I26.0
Valvular disease	391.x, 394.x, 395.x, 396.x, 397.x, 398.x (rheumatic), 424.0, 424.1, 424.2, 424.3 (valve disease), 746.x (congenital diseases)	I01.x, I05.x, I06.x, I07.x, I08.x, I09.x (rheumatic) I43.x, I35.x, I36.x, I37.x, I38, I39.x (valve disease) Q23.x (congenital diseases)
Ischemic Stroke	434.x (ischemic)	I63.x(ischemic), I64.x (unspecified)
Systemic embolization	444.x	I74.x
Intracranial bleeding	430, 431, 432.x	I60.x, I61.x, I62.x
Gastrointestinal bleeding	456.0, 531.0, 531.2, 531.4, 531.6, 532.0, 532.2, 532.4, 532.6, 533.0, 533.2, 533.4, 533.6, 534.0, 534.2, 534.4, 534.6, 569.3, 578.x	I85.0, I98.3, K25.0, K25.2, K25.4, K25.6, K26.0, K26.2, K26.4, K26.6, K27.0, K27.2, K27.4, K27.6, K28.0, K28.2, K28.4, K28.6, K29.0, K55.2, K62.5, K63.8, K92.0, K92.1, K92.2
Ocular bleeding	362.8, 363.6, 376.3, 379.2, 377.4	H31.3, H35.6, H43.1, H45.0
Other bleeding causing hospitalization	459.0, 596.7, 599.7, 627.1, 719.1,729.9, 784.7, 784.8, 786.3	D68.3, K66.1, M25.0, N02.x, N93.8, N93.9, N95.0, R04.x, R31.x, R58
Myocardial infarction	410.x	I21.x
Congestive heart failure	428.x	I50.x
Hypertension	401.x, 402.x, 403.x, 404.x	I10.x, I11, I12, I13
Diabetes	250.x	E10.x, E11.x, E13.x, E14.x
Stroke	433.x, 434.x, 430, 431, 432.x, 436, 438.x	I60.x, I61.x, I62.x, I63.x, I64.x, I65.x, I66.x, I67.x, I69.x
Peripheral vascular disease	440.x, 441.x, 443.x, 444.x, 445.x, 437.x, 557.x,	I70.x, I71.x, I73 I74.x K55.1

Transient ischemic attack	435.x	G45.x
Chronic renal failure	403.x, 585.x	E10.2x, E11.2x, E13.2x, E14.2x, I12, I13, N08.x, N18.x, N19
Acute kidney injury	NA	N17.x
Liver disease	570, 571.x, 070.2, 070.3, 070.4x, 070.5x, 070.6. 070.7x, 070.9, 572.2, 572.3, 572.4, 572.8	B16.x, B17.x, B18.x, B19.x, K70.x, K71.x, K72.x, K73.x, K74.x, K76.0, K76.2, K76.6, K76.7
Active cancer	140.x - 172.x, 174.x - 209.x	C00.x - C43.x, C45.x - C97.x
Chronic Obstructive pulmonary disease	490, 491.x, 492.x 496.x	J40, J41.x, J42, J43.x, J44.x
Bleeding	430, 431, 432.0, 432.1, 432.9, 456.0, 531.0, 531.2, 531.4, 531.6, 532.0, 532.2, 532.4, 532.6, 533.0, 533.2, 533.4, 533.6, 534.0, 534.2, 534.4, 534.6, 569.3, 578.0, 578.1, 578.9, 287.8, 287.9, 596.7, 784.8, 786.3, 599.7, 627.1, 459.0, 719.1,	I60, I61, I62.0, I62.1, I62.9, K92.0, K92.1, I85.0, I98.20, I98.3, K22.1, K25.0, K25.2, K25.4, K25.6, K26.0, K26.2, K26.4, K26.6, K27.0, K27.2, K27.4, K27.6, K28.0, K28.2, K28.4, K28.6, K29.0, K63.8, K31.8 K55.2, K62.5, K92.2 N02.0, N02.1, N02.2, N02.3, N02.4, N02.5, N02.6, N02.7, N02.8, N02.9, K66.1, N93.8, N93.9, N95.0, R04.1, R04.2, R04.8, R04.9, R31.0, R31.1, R31.8, R58, D68.3, H35.6, H43.1, H45.0, M25.0
Dementia	290.x, 331.0, 331.1, 331.2, 797	F00.0, F00.1, F00.2, F00.9, F01.0, F01.1, F01.2, F01.3, F01.8, F01.9, F02.0, F02.1, F02.2, F02.3, F02.4, F02.8, F03, F05.1, F06.5, F06.6, F06.8, F06.9, F09, G30.0, G30.1, G30.8, G30.9, G31.0 G31.1, R54
Diseases related to alcohol abuse	305.0, 303.0, 303.1, 303.2, 303.3, 303.9, 291.4, 291.0, 291.1, 291.2, 291.3, 291.5, 291.8, 291.9, 571.2, 571.1, 571.3, 331.7	F10.0, F10.1, F10.2, F10.3, F10.4, F10.5, F10.6, F10.7, F10.8, Z71.4, F19.2, K70.0, K70.1, K70.2, K70.3, K70.4, K70.9, R78.0, G31.2, G62.1, G72.1, I42.6, K29.2, K85.2, K86.0, T51.0, T51.9, Y91.1, Y91.2, Y91.3, Y91.9, Z50.2, Z71.4, Z72.1, Z86.40
Coronary Atherosclerosis	411 412 413 414	I25

Annex 4. List of Anatomical Therapeutic Chemical (ATC) classification codes used in studies of this thesis.

Drug	Anatomical Therapeutic Chemical (ATC) classification codes
Aspirin	B01AC06, N02BA01
Other antiplatelet agents	B01AC (excluding: B01AC06, B01AC09, B01AC11, B01AC13, B01AC16, B01AC17, B01AC19, B01AC21, B01AC25, B01AC27)
NSAIDS	M01A
H2 receptor blockers	A02BA
Proton pump inhibitors	A02BC, M01AE52, A02BD07
Anti-hypertensive drugs	C02AB, C02AC01, C02CA, C02D, C02L, C03A, C03D, C07, C08C, C08D, C08G, C09,
Antidiabetic drugs	A10B, A10X
Insulin	A10A
Statins	C10AA, C10BA, C10BX
Selective serotonin reuptake inhibitors	N06AB
Benzodiazepines	N03AE, N05BA, N05CD
Antipsychotics (typical and atypical)	N05A
Cholinesterase inhibitors (anti-dementia drugs)	N06DA, N06DX
Systemic corticosteroids	H02AB, H02B, A07EA

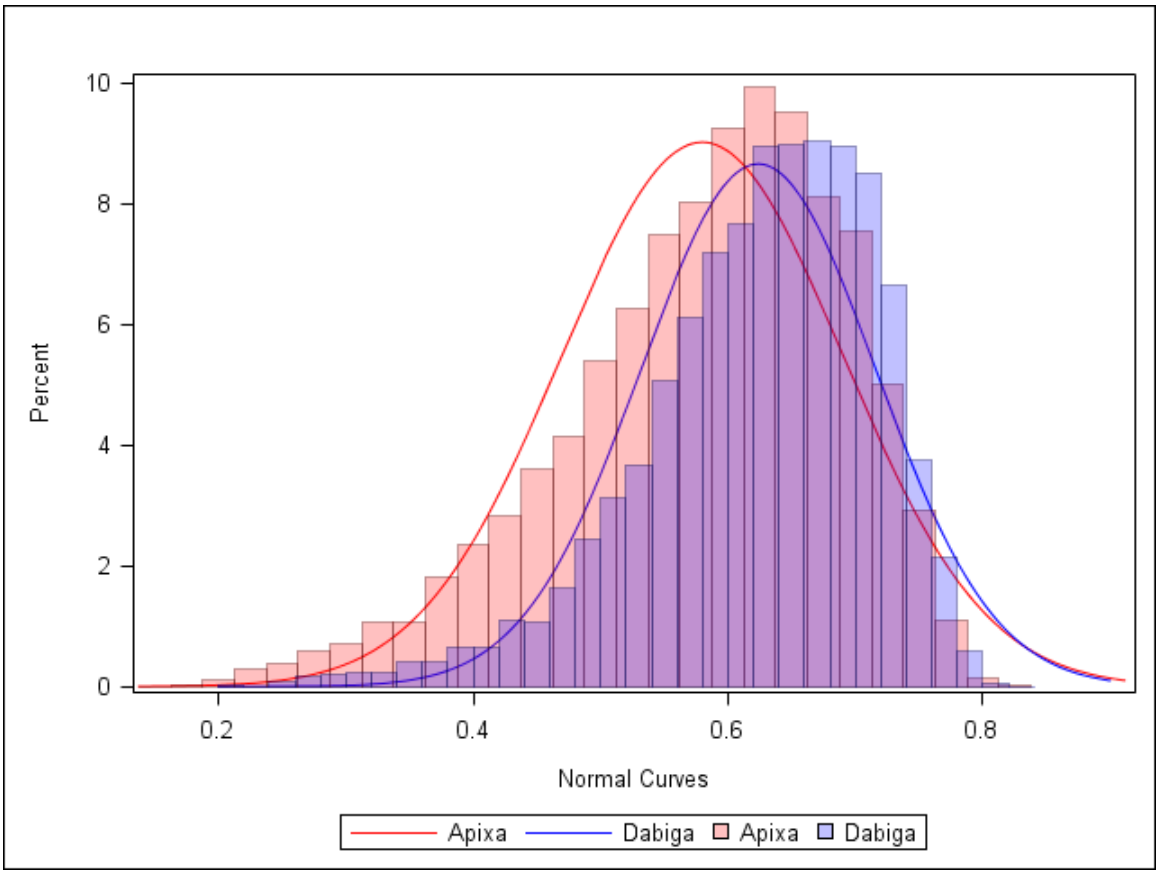
Drug Product Database (DPD) online query managed by Health Canada was used to convert ATC codes to DIN codes. DPD contains information on all drugs approved in Canada (https://health-products.canada.ca/dpd-bdpp/).

Annex 5. Number of individuals who switched from their assigned anticoagulant during the follow-up by their first switch (N=13,472) for the first study. Québec, 2010-2015.

Switched FROM		Switched TO (Number and percentage from all switchers)			
		Apixaban	**Dabigatran**	**Rivaroxaban**	**Warfarin**
	Number at baseline				
Apixaban	5,432		22 (0.4)	149 (2.7)	177 (3.3)
Dabigatran	8,947	634 (7.1)		1217 (13.6)	736 (8.2)
Rivaroxaban	10,071	495 (4.9)	128 (1.3)		407 (4.0)
Warfarin	37,066	1591 (4.3)	4725 (12.7)	3191 (8.6)	

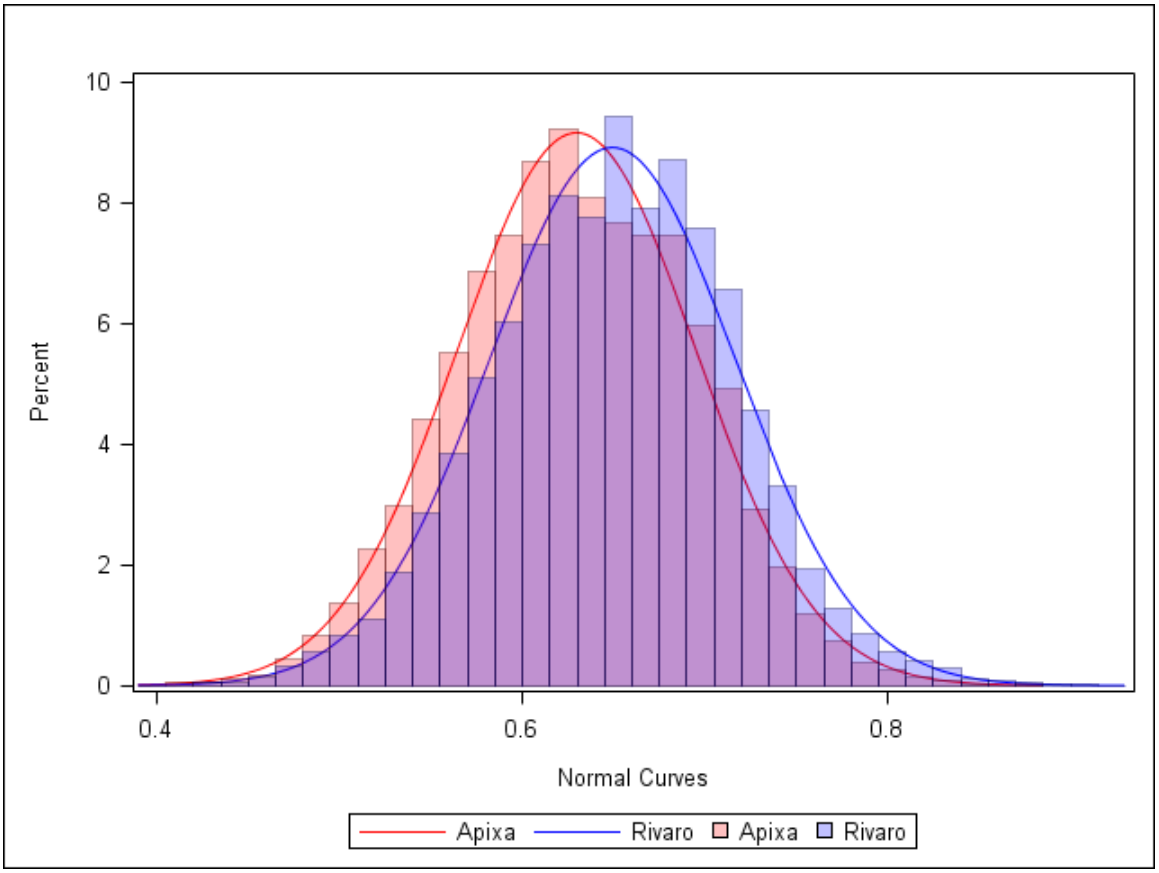
Annex 6. Distribution of propensity scores in treatment groups.

Annex figure A. Distribution of propensity scores in individuals treated with apixaban and those treated with dabigatran before matching (First study). Québec, 2010-2015.



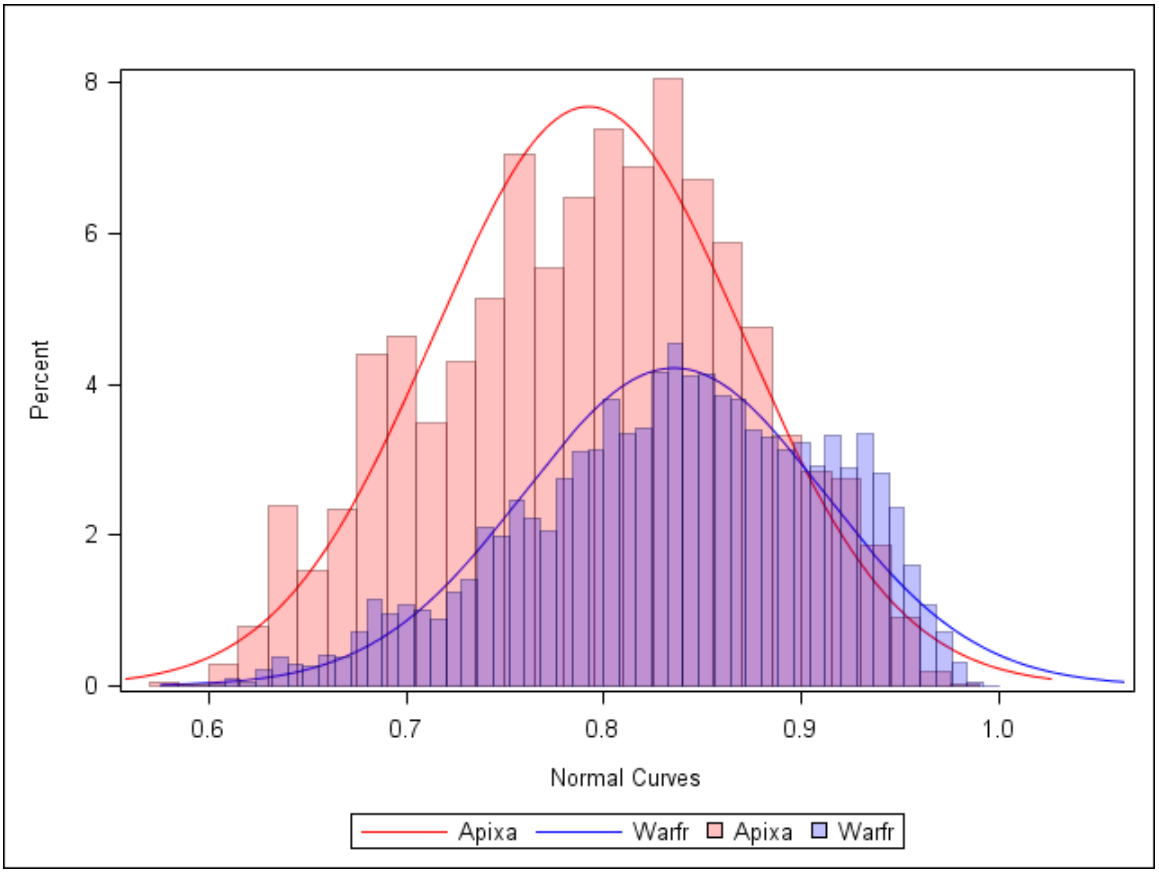
				Percentiles					
	Mean	SD	Min	5th	25th	Median	75th	95th	Max
Apixaban	0.58	0.11	0.14	0.37	0.51	0.59	0.66	0.73	0.83
Dabigatran	0.62	0.09	0.20	0.45	0.57	0.64	0.69	0.75	0.83

Annex figure B. Distribution of propensity scores in individuals treated with apixaban and those treated with rivaroxaban before matching (First study). Québec, 2010-2015.



				Percentiles					
	Mean	SD	Min	5th	25th	Median	75th	95th	Max
Apixaban	0.63	0.06	0.40	0.52	0.58	0.63	0.68	0.73	0.88
Rivaroxaban	0.65	0.07	0.39	0.54	0.60	0.65	0.69	0.76	0.92

Annex figure C. Distribution of propensity scores in individuals treated with apixaban and those treated with warfarin before matching (First study). Québec, 2010-2015.

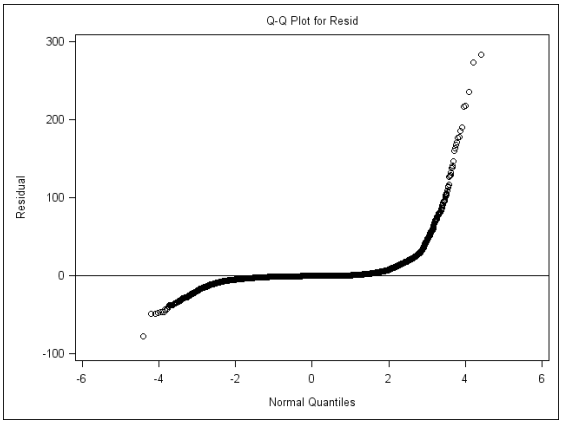


				Percentiles					
	Mean	SD	Min	5th	25th	Median	75th	95th	Max
Apixaban	0.79	0.08	0.57	0.66	0.74	0.80	0.85	0.92	0.98
Warfarin	0.83	0.07	0.58	0.70	0.79	0.84	0.89	0.95	0.99

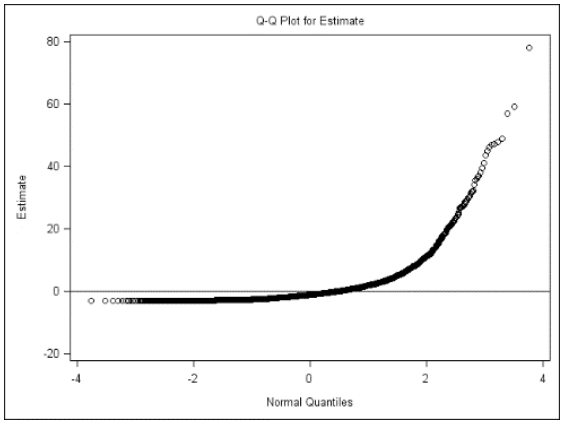
Annex 7. Q-Q (quantiles-quantiles) plots of the distributions of level 1 and level 2 raw residuals (First study) for every study cohort. Québec, 2010-2015.

Apixaban

Level 1 residuals

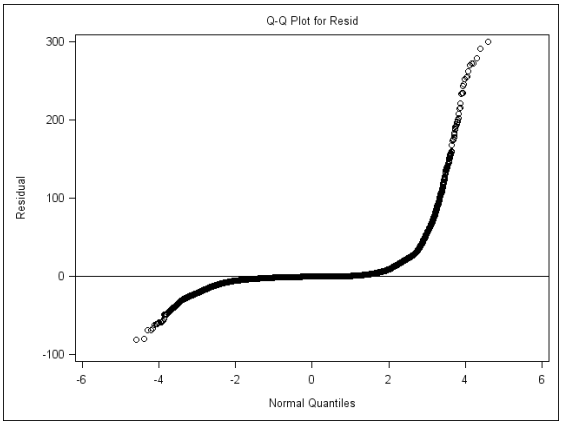


Level 2 residuals

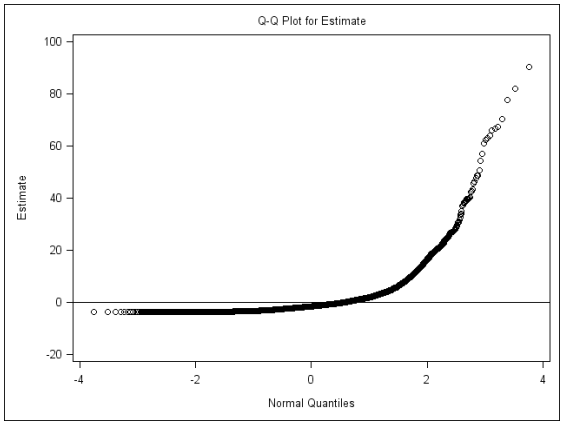


Dabigatran

Level 1 residuals

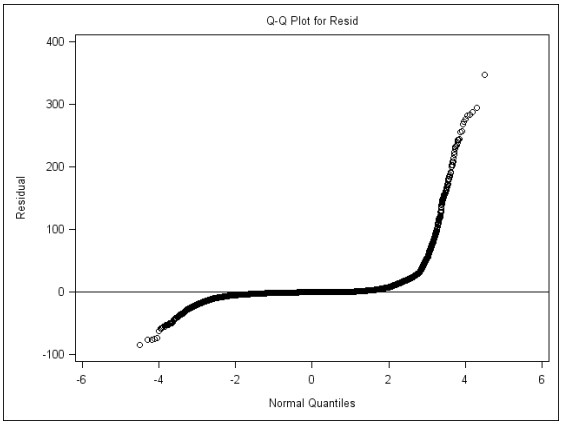


Level 2 residuals

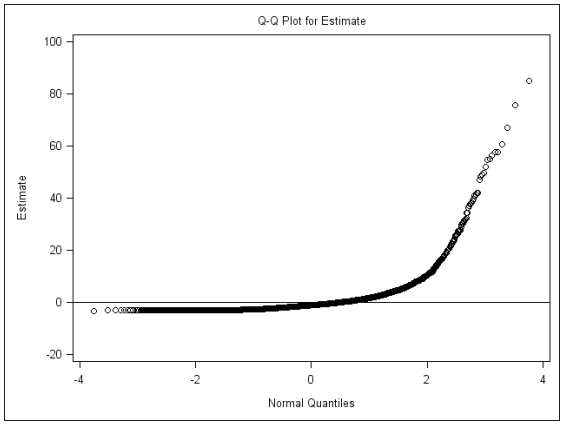


Rivaroxaban

Level 1 residuals

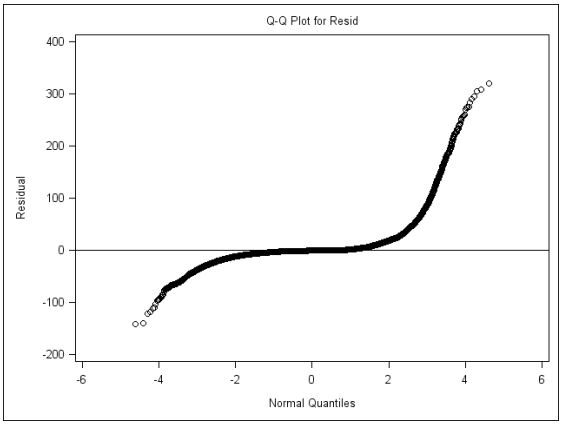


Level 2 residuals

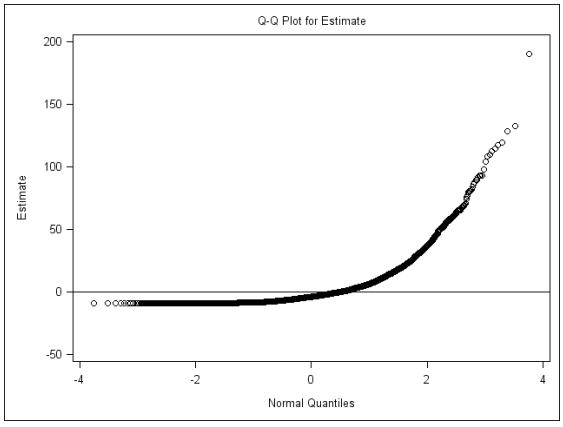


Warfarin

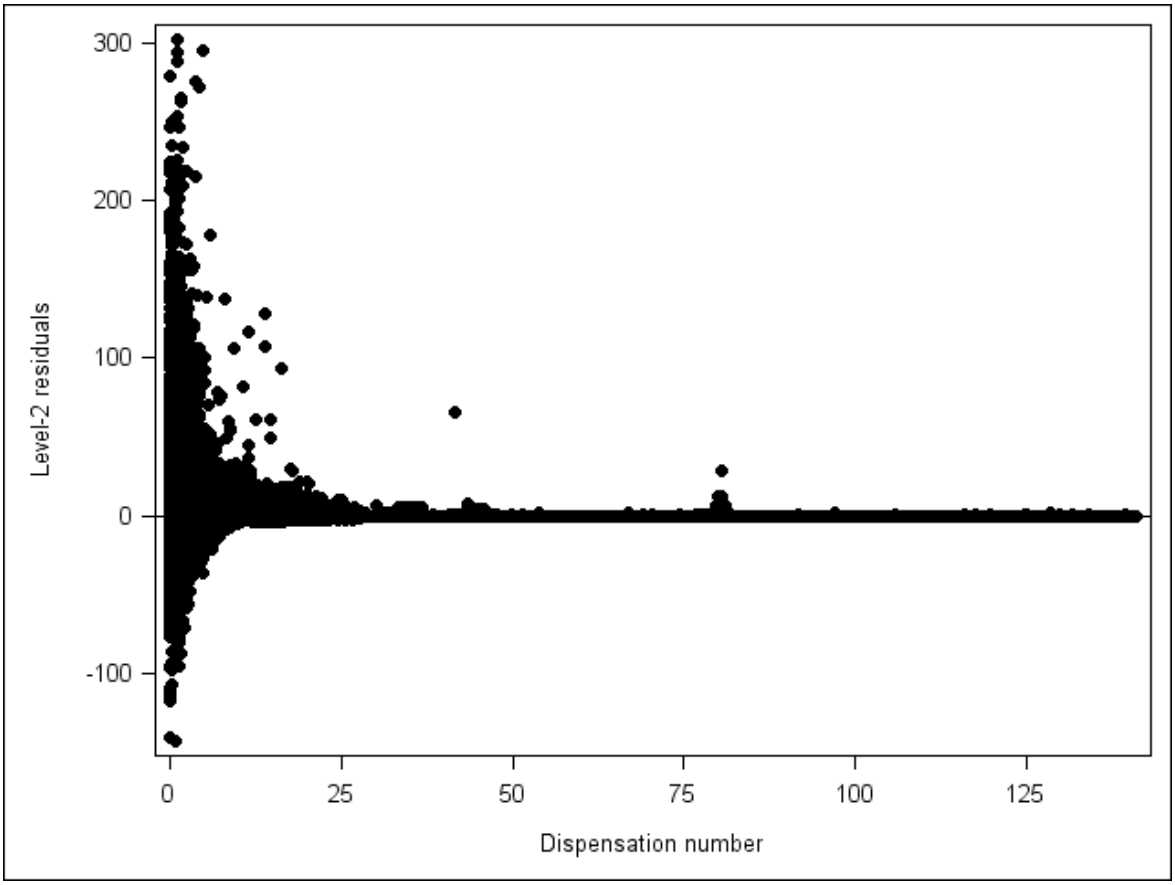
Level 1 residuals



Level 2 residuals



Annex 8. Distribution of level-2 residuals versus the dispensation number in the cohort of warfarin users. Québec, 2010-2015.



Annex 9. The flow of study participants for the second study, Québec, 2010-2016.

	Dabigatran	Rivaroxaban	Warfarin	Total
Total number of new users of oral anticoagulants between 01/01/2010 and 31/12/2016, aged ≥18 years, and covered by the RAMQ public drug insurance plan at least one year before treatment initiation	12,603	58,903	88,250	159,756
Diagnosis of AF or atrial flutter in the year before the date of cohort entry	9,716	18,209	46,173	74,098
Excluded patients with a diagnosis of valvular disease in the 12 months before cohort entry	749	1,509	10,037	12,295
Excluded patients with valvular surgery in the 12 months before cohort entry	4	11	315	330
Excluded patients with VTE in the three months before the date of cohort entry	152	980	4230	5,362
Excluded patients who underwent hemodialysis three months before cohort entry	6	10	818	834
Number of new users of oral anticoagulants in the analysis	8,805	15,699	30,773	55,277

Excluded those with the imputed date of death which is ≤ the date of cohort entry	5	16	30	51	
Number of new users of oral anticoagulants in the analysis	8,800	15,683	30,743	55,226	

Abbreviations: RAMQ, *Régie de l'assurance-maladie du Québec*; AF, atrial fibrillation; VTE, venous thromboembolism

[†]Each treatment switch between oral anticoagulants defines a new treatment episode

Annex 10. Statistics on the fit of the Cox regression model comparing the association between treatment with dabigatran (Annex table 10.1) or rivaroxaban (Annex table 10.2) versus warfarin and the risk of major bleeding (Second study).

Annex table A. Statistics on the fit of the Cox regression model comparing the association between treatment with dabigatran versus warfarin and the risk of major bleeding.

Definition of exposure	Statistical model	Model fit statistics	Testing global null hypothesis: Beta=0, Likelihood Ratio Chi-Square (p-value)	Type-3 test Wald Chi-Square (p-value)
Days supplied				
Grace period 15-day				
	No predictors	-2 LOG L = 19932.42		
	Major bleeding= treatment	-2 LOG L = 19840.114 AIC = 19842.114 SBC = 19847.051	92.3095 (<.0001)	Treatment-78.8366 (<.0001)
	Major bleeding= treatment+ year of cohort entry	-2 LOG L =19827.469 AIC = 19841.469 SBC = 19876.030	104.9547 (<.0001)	Treatment-78.5762 (<.0001) Year-13.6097 (0.0343)
	Major bleeding= treatment + year of cohort entry + deciles of PS	-2 LOG L = 19684.402 AIC = 19716.402 SBC = 19795.399	248.0208 (<.0001)	Treatment -25.9227 (<.0001) Year-8.9362 (0.1772) Dec of PS-130.9498 (<.0001)
Grace period 30-day				
	No predictors	-2 LOG L = 25991.38		
	Major bleeding= treatment	-2 LOG L = 25871.691 AIC = 25873.691 SBC = 25878.880	119.6931 (<.0001)	Treatment -100.5768 (<.0001)

	Major bleeding= treatment+ year of cohort entry	-2 LOG L = 25855.886 AIC = 25869.886 SBC = 25906.216	135.4971 (<.0001)	Treatment-101.9642 (<.0001) Year-17.1813 (0.0086)
	Major bleeding= treatment + year of cohort entry + deciles of PS	-2 LOG L = 25689.524 AIC = 25721.524 SBC = 25804.563	301.8593 (<.0001)	Treatment-40.4658 (<.0001) Year-10.4799 (0.1058) Dec of PS-150.3218 (<.0001)
Grace period 60-day				
	No predictors	-2 LOG L = 30882.09		
	Major bleeding= treatment	-2 LOG L = 30751.173 AIC = 30753.173 SBC = 30758.526	130.9169 (<.0001)	Treatment- 109.2456 (<.0001)
	Major bleeding= treatment+ year of cohort entry	-2 LOG L = 30732.993 AIC = 30746.993 SBC = 30784.465	149.0971 (<.0001)	Treatment-104.2770 (<.0001) Year-19.8788 (0.0029)
	Major bleeding=treatment + year of cohort entry + deciles of PS	-2 LOG L = 30562.600 AIC = 30594.600 SBC = 30680.249	319.4906 (<.0001)	Treatment- 45.5785 (<.0001) Year-13.1653 (0.0405) Dec of PS-154.5807 (<.0001)
Data-driven				
Grace period 15-day				
	No predictors	-2 LOG L = 19376.96		
	Major bleeding= treatment	-2 LOG L = 19297.715 AIC = 19299.715 SBC = 19304.623	79.2453 (<.0001)	Treatment - 68.3395 (<.0001)

	Major bleeding= treatment+ year of cohort entry	-2 LOG L = 19283.144 AIC = 19297.144 SBC = 19331.506	93.8156 (<.0001)	Treatment - 65.7921 (<.0001) Year - 15.3677 (0.0176)
	Major bleeding= treatment + year of cohort entry + deciles of PS	-2 LOG L = 19129.068 AIC = 19161.068 SBC = 19239.608	247.8923 (<.0001)	Treatment - 19.0172 (<.0001) Year -9.9363 (0.1274) Dec of PS - 135.9258 (<.0001)
Grace period 30-day				
	No predictors	-2 LOG L = 27130.04		
	Major bleeding= treatment	-2 LOG L = 27009.811 AIC = 27011.811 SBC = 27017.041	120.2263 (<.0001)	Treatment -100.1565 (<.0001)
	Major bleeding= treatment+ year of cohort entry	-2 LOG L = 26995.122 AIC = 27009.122 SBC = 27045.731	134.9151 (<.0001)	Treatment -99.9886 (<.0001) Year -15.9280 (0.0141)
	Major bleeding= treatment + year of cohort entry + deciles of PS	-2 LOG L = 26828.101 AIC = 26860.101 SBC = 26943.778	301.9367 (<.0001)	Treatment - 41.9920 (<.0001) Year - 9.3946 (0.1526) Dec of PS - 153.7316 (<.0001)
Grace period 60-day				
	No predictors	-2 LOG L = 31954.09		
	Major bleeding= treatment	-2 LOG L = 31827.114 AIC = 31829.114 SBC = 31834.500	126.9762 (<.0001)	Treatment - 106.0942 (<.0001)

Major bleeding= treatment+ year of cohort entry	-2 LOG L = 31808.871 AIC = 31822.871 SBC = 31860.576	145.2189 (<.0001)	Treatment - 101.0998 (<.0001) Year -19.7498 (0.0031)
Major bleeding= treatment + year of cohort entry + deciles of PS	-2 LOG L = 31629.778 AIC = 31661.778 SBC = 31747.961	324.3122 (<.0001)	Treatment - 43.9228 (<.0001) Year - 13.0990 (0.0415) Dec of PS - 161.1970 (<.0001)

Abbreviations: -2 LOG L, log-likelihood ratio; AIC, Akaike Information Criterion; SBC, Schwarz information criterion

Annex table B. Statistics on the fit of the Cox regression model comparing the association between treatment with rivaroxaban versus warfarin and the risk of major bleeding (Second study).

Definition of exposure	**Statistical model**	**Model fit statistics**	**Testing global null hypothesis: Beta=0, Likelihood Ratio Chi-Square (p-value)**	**Type-3 test Wald Chi-Square (p value)**
Days supplied				
Grace period 15-day				
	No predictors	-2 LOG L = 23903.64		
	Major bleeding= treatment	-2 LOG L = 23859.024 AIC = 23861.024 SBC = 23866.122	44.6141 (<.0001)	Treatment - 42.3616 (<.0001)
	Major bleeding= treatment+ year	-2 LOG L = 23845.607 AIC = 23859.607 SBC = 23895.295	58.0313 (<.0001)	Treatment - 35.7312 (<.0001) Year - 13.2346 (0.0395)
	Major bleeding= treatment + year of cohort entry + deciles of PS	-2 LOG L = 23678.278 AIC = 23710.278 SBC = 23791.852	225.3595 (<.0001)	Treatment - 1.2354 (0.2664) Year - 9.9935 (0.1249) Dec of PS - 155.8051 (<.0001)
Grace period 30-day				
	No predictors	-2 LOG L = 30290.5		
	Major bleeding= treatment	-2 LOG L = 30227.310 AIC = 30229.310 SBC = 30234.635	63.1916 (<.0001)	Treatment - 59.1021 (<.0001)
	Major bleeding= treatment+ year	-2 LOG L = 30211.840 AIC = 30225.840 SBC = 30263.116	78.6614 (<.0001)	Treatment - 54.9798 (<.0001)

				Year - 15.5451 (0.0164)
	Major bleeding= treatment + year of cohort entry + deciles of PS	-2 LOG L = 30044.793 AIC = 30076.793 SBC = 30161.996	245.7080 (<.0001)	Treatment - 7.2441 (0.0071) Year - 10.0899 (0.1209) Dec of PS - 158.2763 (<.0001)
Grace period 60-day				
	No predictors	-2 LOG L = 35366.63		
	Major bleeding= treatment	-2 LOG L = 35296.948 AIC = 35298.948 SBC = 35304.421	69.6849 (<.0001)	Treatment - 64.7331 (<.0001)
	Major bleeding= treatment+ year	-2 LOG L = 35280.168 AIC = 35294.168 SBC = 35332.480	86.4647 (<.0001)	Treatment - 60.3831 (<.0001) Year - 17.0684 (0.0090)
	Major bleeding= treatment + year of cohort entry + deciles of PS	-2 LOG L = 35106.102 AIC = 35138.102 SBC = 35225.671	260.5313 (<.0001)	Treatment - 10.3858 (0.0013) Year - 11.2860 (0.0799) Dec of PS - 165.4101 (<.0001)
Data-driven				
Grace period 15-day				
	No predictors	-2 LOG L = 23000.69		
	Major bleeding= treatment	-2 LOG L = 22962.133 AIC = 22964.133	38.5551 (<.0001)	Treatment - 36.6158 (<.0001)

		SBC = 22969.194		
	Major bleeding= treatment+ deciles of PS	-2 LOG L = 22945.458 AIC = 22959.458 SBC = 22934.881	55.2305 (<.0001)	Treatment - 34.3467 (<.0001) Year - 16.4643 (0.0115)
	Major bleeding= treatment + year of cohort entry + deciles of PS	-2 LOG L = 22774.841 AIC = 22806.841 SBC = 22887.809	225.8475 (<.0001)	Treatment - 1.1523 (0.2831) Year - 11.5657 (0.0724) Dec of PS - 158.2877 (<.0001)
Grace period 30-day				
	No predictors	-2 LOG L = 31378.9		
	Major bleeding= treatment	-2 LOG L = 31318.340 AIC = 31320.340 SBC = 31325.699	60.5637 (<.0001)	Treatment - 56.4715 (<.0001)
	Major bleeding= treatment+ deciles of PS	-2 LOG L = 31304.191 AIC = 31318.191 SBC = 31355.702	74.7134 (<.0001)	Treatment - 52.5400 (<.0001) Year - 14.2601 (0.0269)
	Major bleeding= treatment + year of cohort entry + deciles of PS	-2 LOG L = 31130.867 AIC = 31162.867 SBC = 31248.609	248.0367 (<.0001)	Treatment - 7.1637 (0.0074) Year - 8.6177 (0.1962) Dec of PS - 165.0179 (<.0001)
Grace period 60-day				
	No predictors	-2 LOG L = 36397.55		
	Major bleeding= treatment	-2 LOG L = 36330.881 AIC = 36332.881 SBC = 36338.383	66.6698 (<.0001)	Treatment - 61.9377 (<.0001)

Major bleeding= treatment+ deciles of PS	-2 LOG L = 36313.798 AIC = 36327.798 SBC = 36366.310	83.7529 (<.0001)	Treatment - 55.9931 (<.0001) Year - 17.2132 (0.0085)
Major bleeding= treatment + year of cohort entry + deciles of PS	-2 LOG L = 36128.559 AIC = 36160.559 SBC = 36248.586	268.9916 (<.0001)	Treatment - 8.8573 (0.0029) Year - 11.4248 (0.0761) Dec of PS - 175.8959 (<.0001)

Abbreviations: -2 LOG L, log-likelihood ratio; AIC, Akaike Information Criterion; SBC, Schwarz information criterion

Annex 11. Number of events, person-years of follow-up, and hazard ratios of the composite of ischemic stroke or systemic thromboembolism associated with the use of dabigatran versus warfarin and rivaroxaban versus warfarin using two approaches to define the duration of dispensation (Second study).

Grace period	Exposure	Events	Person-years	Incidence rate (95% CI)*	Crude HR	Adjusted HR (95% CI)
15-day Days supplied	Dabigatran	88	14,109	0.62 (0.51-0.77)	0.64	0.93 (0.72-1.22)
	Rivaroxaban	138	19,856	0.69 (0.59-0.82)	0.69	1.17 (0.90-1.52)
	Warfarin	333	31,997	1.04 (0.93-1.16)	Ref	Ref
Data-driven	Dabigatran	86	13,519	0.64 (0.51-0.78)	0.66	0.92 (0.71-1.19)
	Rivaroxaban	136	18,599	0.73 (0.62-0.86)	0.73	1.18 (0.90-1.53)
	Warfarin	325	31,638	1.03 (0.92-1.04)	Ref	Ref
30-day Days supplied	Dabigatran	109	17,786	0.61 (0.51-0.74)	0.65	0.88 (0.70-1.11)
	Rivaroxaban	168	23,736	0.71 (0.61-0.82)	0.73	1.07 (0.84-1.35)
	Warfarin	442	45,112	0.98 (0.89-1.07)	Ref	Ref
Data-driven	Dabigatran	108	17,686	0.61 (0.50-0.74)	0.67	0.88 (0.70-1.10)
	Rivaroxaban	165	23,366	0.71 (0.61-0.82)	0.74	1.10 (0.87-1.39)
	Warfarin	466	49.070	0.95 (0.87-1.04)	Ref	Ref
60-day Days supplied	Dabigatran	120	20,094	0.60 (0.50-0.71)	0.67	0.88 (0.71-1.08)
	Rivaroxaban	182	26,246	0.69 (0.60-0.80)	0.73	1.09 (0.87-1.35)
	Warfarin	533	57,721	0.92 (0.85-1.01)	Ref	Ref
Data-driven	Dabigatran	120	20,118	0.60 (0.50-0.71)	0.67	0.86 (0.69-1.06)
	Rivaroxaban	182	26,174	0.69 (0.60-0.80)	0.73	1.06 (0.85-1.32)
	Warfarin	557	61,031	0.91 (0.84-0.99)	Ref	Ref

*per 100 person-years
Abbreviations: CI, confidence interval; HR, hazard ratio

Annex 12. The flow of study participants for the third study, Québec, 2010-2016.

	All DOACs	Warfarin	Total
Total number of new users of oral anticoagulants between 01/01/2010 and 31/12/2016, aged ≥18 years, and covered by the RAMQ public drug insurance plan at least one year before treatment initiation	95,215	88,250	183,465
Diagnosis of AF or atrial flutter in the year before the date of cohort entry	46,523	46,173	92,696
Excluded patients with a diagnosis of valvular disease in the 12 months before cohort entry	4,622	10,037	14,659
Excluded patients with valvular surgery in the 12 months before cohort entry	39	315	354
Excluded patients with VTE in the three months before the date of cohort entry	1,679	4230	5,909
Excluded patients who underwent hemodialysis three months before cohort entry	33	818	851
Number of new users of oral anticoagulants in the analysis	40,150	30,773	70,923
Excluded those with the imputed date of death which is before or equal to the date of cohort entry	49	30	79
Number of new users of oral anticoagulants in the analysis	40,101	30,743	70,844
Number of patients in the matched cohort	23,927	23,927	47,854

Abbreviations: RAMQ, *Régie de l'assurance-maladie du Québec*; AF, atrial fibrillation; VTE, venous thromboembolism

[†]Each treatment switch between oral anticoagulants defines a new treatment episode

Annex 13. Incidence of some major prognostic factors in the study cohort during the follow-up in the third study, Québec, 2010-2017 (Third study).

	All (No=47,854) Total follow-up = 76,177 years	DOACs (No=23,927) Total follow-up= 44,647 years	Warfarin (No=23,927) Total follow-up= 31,530 years
	No(%)	No(%)	No(%)
Chronic renal failure	2,664 (5.6)	1,364 (5.7)	1,300 (5.4)
Hypertension	939 (2.0)	554 (2.3)	385 (1.6)
Congestive heart failure	4,523 (9.4)	2,418 (10.1)	2,105 (8.8)
Cancer	3,261 (6.8)	1,914 (8.0)	1,347 (5.6)
Dementia	2,817 (5.9)	1,444 (6.0)	1,373 (5.7)
Diseases related to alcohol abuse	657 (1.4)	343 (1.4)	314 (1.3)
Diabetes	1,632 (3.4)	929 (3.9)	703 (2.9)
percutaneous coronary intervention/ coronary artery bypass surgery	417 (0.9)	242 (1.0)	175 (0.7)
Pacemaker implantation/ Catheter ablation	2,351 (4.9)	1,418 (5.9)	933 (3.9)

Abbreviations: No, number

Annex 14. Partial output from SAS PROC LOGISTIC to run a model to estimate probabilities of censoring (Third study).

Annex table A. Output from SAS PROC LOGISTIC to run a model to estimate probabilities of censoring using exposure unstratified model.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	231960.22	218853.85
SC	231971.98	219230.27
-2 Log L	231958.22	218789.85

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	13168.3671	31	<.0001
Score	13290.6222	31	<.0001
Wald	12288.0285	31	<.0001

Characteristic	Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
Intercept	-0.2752	9.1619	0.0009	0.9760
DOACs vs warfarin	-0.9515	0.0150	4041.5936	<.0001
Month of follow-up	-0.0317	0.000596	2819.7792	<.0001
Year of cohort entry	-0.00062	0.00456	0.0183	0.8924
Age	-0.0134	0.000799	280.8092	<.0001
Male sex	0.0542	0.0139	15.2683	<.0001
Hypertension	-0.2739	0.0321	72.6021	<.0001
Congestive heart failure	0.0391	0.0155	6.3879	0.0115
Coronary atherosclerosis	0.00763	0.0143	0.2846	0.5937
Diabetes	-0.0706	0.0147	23.0250	<.0001
Peripheral vascular disease	0.0304	0.0184	2.7258	0.0987
Chronic renal failure	0.0135	0.0214	0.3969	0.5287
Cancer	0.1826	0.0159	131.9455	<.0001
Acute kidney injury	0.00597	0.0237	0.0634	0.8012
COPD	0.00879	0.0162	0.2931	0.5882
Liver disease	0.0373	0.0383	0.9486	0.3301

Characteristic	Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
Dementia	0.2196	0.0212	107.4874	<.0001
Diseases related to alcohol abuse	0.1092	0.0308	12.5968	0.0004
PCI/CAB	0.3899	0.0308	160.6238	<.0001
Pacemaker implantation/Catheter ablation	0.0694	0.0203	11.6835	0.0006
Antiplatelets	0.3634	0.0154	554.9908	<.0001
Antipsychotics	0.0217	0.0280	0.6008	0.4383
NSAIDS	0.2289	0.0354	41.7044	<.0001
Statins	-0.2288	0.0143	254.4226	<.0001
Benzodiazepines	-0.00585	0.0152	0.1480	0.7005
Systemic corticosteroids	0.2119	0.0276	58.9130	<.0001
H-2 receptor blockers	0.0388	0.0625	0.3857	0.5346
Proton pump inhibitors	-0.0740	0.0140	27.7911	<.0001
Opioids	0.4236	0.0224	356.5619	<.0001
SSRI	-0.0730	0.0232	9.8814	0.0017

Association of Predicted Probabilities and Observed Responses

Percent Concordant	69.2	Somers' D	0.411
Percent Discordant	28.0	Gamma	0.423
Percent Tied	2.8	Tau-a	0.021
Pairs	23195025256	c	0.706

Annex table B. Output from SAS PROC LOGISTIC to run a model to estimate probabilities of censoring for patients treated with DOACs using an exposure stratified model (Third study).

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	88509.724	84001.604
SC	88520.951	84349.645
-2 Log L	88507.724	83939.604

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	4568.1196	30	<.0001
Score	4968.3518	30	<.0001
Wald	4623.0671	30	<.0001

Characteristic	Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
Intercept	245.2	16.9853	208.4239	<.0001
Month of follow-up	-0.0384	0.00105	1341.9469	<.0001
Year of cohort entry	-0.1235	0.00844	213.9343	<.0001
Age	0.000023	0.00138	0.0003	0.9870
Male sex	0.1334	0.0236	31.9483	<.0001
Hypertension	-0.4929	0.0484	103.4947	<.0001
Congestive heart failure	0.1298	0.0265	23.9434	<.0001
Coronary atherosclerosis	0.0109	0.0245	0.1970	0.6571
Diabetes	0.0127	0.0251	0.2556	0.6132
Peripheral vascular disease	0.0208	0.0318	0.4304	0.5118
Chronic renal failure	0.2915	0.0350	69.3316	<.0001
Cancer	0.1899	0.0266	50.8477	<.0001
Acute kidney injury	0.1551	0.0398	15.1680	<.0001
COPD	-0.0126	0.0280	0.2020	0.6531
Liver disease	0.1170	0.0641	3.3253	0.0682
Dementia	0.4209	0.0346	147.6004	<.0001
Diseases related to alcohol abuse	0.1176	0.0515	5.2059	0.0225
PCI/CAB	0.3690	0.0540	46.7733	<.0001
Pacemaker implantation/Catheter ablation	0.1018	0.0332	9.3912	0.0022
Antiplatelets	0.6238	0.0266	548.6978	<.0001
Antipsychotics	0.0742	0.0465	2.5404	0.1110
NSAIDS	0.2820	0.0527	28.5941	<.0001
Statins	-0.4425	0.0241	337.6873	<.0001

Characteristic	Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
Benzodiazepines	-0.00306	0.0258	0.0141	0.9056
Systemic corticosteroids	0.2025	0.0466	18.9038	<.0001
H-2 receptor blockers	-0.0816	0.1104	0.5462	0.4599
Proton pump inhibitors	-0.1168	0.0238	24.1762	<.0001
Opioids	0.7704	0.0344	500.9038	<.0001
SSRI	-0.0224	0.0396	0.3201	0.5715

Association of Predicted Probabilities and Observed Responses

Percent Concordant	67.2	Somers' D	0.394
Percent Discordant	27.8	Gamma	0.415
Percent Tied	5.1	Tau-a	0.012
Pairs	4685268582	c	0.697

Annex table C. Output from SAS PROC LOGISTIC to run a model to estimate probabilities of censoring for patients treated with warfarin using an exposure stratified model (Third study).

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	137228.82	133502.11
SC	137239.71	133839.48
-2 Log L	137226.82	133440.11

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	3786.7115	30	<.0001
Score	3679.9176	30	<.0001
Wald	3542.6536	30	<.0001

Characteristic	Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
Intercept	-99.1031	10.7384	85.1710	<.0001
Month of follow-up	-0.0276	0.000721	1459.3424	<.0001
Year of cohort entry	0.0487	0.00534	83.2709	<.0001
Age	-0.0202	0.000978	428.2248	<.0001
Male sex	0.00544	0.0172	0.1001	0.7517
Hypertension	-0.1188	0.0430	7.6291	0.0057
Congestive heart failure	-0.00811	0.0191	0.1813	0.6703
Coronary atherosclerosis	0.0120	0.0177	0.4644	0.4956
Diabetes	-0.1149	0.0182	39.9903	<.0001
Peripheral vascular disease	0.0260	0.0226	1.3232	0.2500
Chronic renal failure	-0.1480	0.0268	30.3787	<.0001
Cancer	0.1713	0.0199	74.4314	<.0001
Acute kidney injury	-0.0448	0.0293	2.3357	0.1264
COPD	0.0184	0.0200	0.8490	0.3568
Liver disease	-0.0214	0.0477	0.2001	0.6546
Dementia	0.1162	0.0268	18.8659	<.0001
Diseases related to alcohol abuse	0.1054	0.0384	7.5299	0.0061
PCI/CAB	0.3873	0.0377	105.7075	<.0001
Pacemaker implantation/Catheter ablation	0.0597	0.0257	5.3887	0.0203

Characteristic	Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square
Antiplatelets	0.2243	0.0189	141.2769	<.0001
Antipsychotics	-0.0292	0.0350	0.6932	0.4051
NSAIDS	0.1839	0.0480	14.6740	0.0001
Statins	-0.1127	0.0179	39.5701	<.0001
Benzodiazepines	-0.00957	0.0188	0.2589	0.6109
Systemic corticosteroids	0.2143	0.0343	39.0494	<.0001
H-2 receptor blockers	0.0581	0.0758	0.5873	0.4435
Proton pump inhibitors	-0.0458	0.0174	6.9079	0.0086
Opioids	0.1915	0.0297	41.6030	<.0001
SSRI	-0.0979	0.0287	11.6452	0.0006

Association of Predicted Probabilities and Observed Responses

Percent Concordant	62.7	Somers' D	0.281
Percent Discordant	34.6	Gamma	0.288
Percent Tied	2.7	Tau-a	0.023
Pairs	6236304361	c	0.640

Annex 15. Odds ratio with 95%CI of being censored associated with patient characteristics using exposure-stratified logistic regression, by type of censoring, n=47,854, Québec, 2010-2016 (Third study).

Annex table A. Odds ratio with 95%CI of being censored associated with patient characteristics using exposure-stratified logistic regression, censoring due to switching between treatment groups, n=47,854, Québec, 2010-2016

	DOACs		Warfarin	
Characteristic	**OR**	**95% CI**	**OR**	**95% CI**
Month of follow-up	0.94	0.94, 0.95	0.99	0.99, 0.99
Year of cohort entry	0.75	0.71, 0.78	1.21	1.19, 1.23
Age	1.03	1.02, 1.04	0.98	0.97, 0.98
Male sex	0.85	0.75, 0.97	0.93	0.88, 0.98
Comorbidities				
Hypertension	1.24	0.80, 1.93	1.11	0.96, 1.30
Congestive heart failure	1.93	1.68, 2.20	0.91	0.85, 0.96
Coronary atherosclerosis	1.09	0.95, 1.26	0.93	0.88, 0.99
Diabetes	1.34	1.17, 1.52	0.90	0.85, 0.96
Peripheral vascular disease	1.05	0.89, 1.24	0.98	0.90, 1.05
Chronic renal failure	2.81	2.40, 3.29	0.57	0.51, 0.62
Cancer	1.03	0.88, 1.20	0.99	0.93, 1.06
Acute kidney injury	1.38	1.16, 1.65	0.88	0.79, 0.97
COPD	1.10	0.95, 1.27	1.07	1.00, 1.14
Liver disease	1.09	0.77, 1.52	0.99	0.85, 1.16
Dementia	0.90	0.73, 1.09	1.01	0.92, 1.10
Diseases related to alcohol abuse	0.88	0.64, 1.20	1.17	1.03, 1.32
Medical procedures				
PCI/CAB	1.74	1.32, 2.29	1.06	0.92, 1.21
Pacemaker implantation/Catheter ablation	1.11	0.93, 1.33	1.00	0.91, 1.09
Medications				
Antiplatelets	1.51	1.31, 1.74	1.14	1.07, 1.22
Antipsychotics	0.90	0.69, 1.18	0.94	0.84, 1.06
NSAIDS	0.99	0.68, 1.42	1.40	1.21, 1.62
Statins	0.97	0.85, 1.11	1.00	0.94, 1.06
Benzodiazepines	1.23	1.07, 1.40	1.07	1.01, 1.14
Systemic corticosteroids	1.52	1.22, 1.90	1.14	1.02, 1.28
H-2 receptor blockers	1.10	0.66, 1.83	0.95	0.73, 1.22
Proton pump inhibitors	1.03	0.90, 1.17	1.03	0.97, 1.09
Opioids	0.97	0.76, 1.22	1.11	1.00, 1.22
SSRI	0.94	0.76, 1.16	0.98	0.89, 1.07

	DOACs		Warfarin	
Characteristic	**OR**	**95% CI**	**OR**	**95% CI**

Abbreviations: CAB, coronary artery bypass surgery; CI, confidence interval; COPD, chronic obstructive pulmonary disease; NSAIDs, non-steroidal anti-inflammatory drugs; OR, odds ratio; PCI, percutaneous coronary intervention; SSRIs, selective serotonin reuptake inhibitors

Annex table B. Odds ratio with 95%CI of being censored associated with patient characteristics using exposure-stratified logistic regression, censoring due to permanent treatment discontinuation n=47,854, Québec, 2010-2016.

	DOACs		Warfarin	
Characteristics	**OR**	**95% CI**	**OR**	**95% CI**
Month of follow-up	0.97	0.96, 0.97	0.96	0.96, 0.97
Year of cohort entry	0.91	0.89, 0.92	0.97	0.96, 0.98
Age	1.00	0.99, 1.00	0.98	0.98, 0.98
Male sex	1.19	1.13, 1.25	1.05	1.00, 1.09
Comorbidities				
Hypertension	0.58	0.53, 0.64	0.79	0.72, 0.88
Congestive heart failure	1.04	0.98, 1.10	1.04	1.00, 1.09
Coronary atherosclerosis	1.00	0.95, 1.05	1.06	1.02, 1.11
Diabetes	0.97	0.92, 1.02	0.89	0.85, 0.93
Peripheral vascular disease	1.02	0.95, 1.09	1.06	1.00, 1.11
Chronic renal failure	1.15	1.06, 1.24	1.05	0.99, 1.12
Cancer	1.24	1.17, 1.31	1.30	1.24, 1.36
Acute kidney injury	1.12	1.02, 1.22	1.00	0.93, 1.07
COPD	0.97	0.92, 1.03	0.99	0.95, 1.04
Liver disease	1.14	1.00, 1.30	0.97	0.87, 1.09
Dementia	1.65	1.54, 1.77	1.19	1.12, 1.27
Diseases related to alcohol abuse	1.16	1.05, 1.29	1.09	0.99, 1.19
Medical procedures				
PCI/CAB	1.41	1.26, 1.58	1.74	1.59, 1.90
Pacemaker implantation/Catheter ablation	1.11	1.03, 1.19	1.10	1.03, 1.17
Medications				
Antiplatelets	1.92	1.82, 2.03	1.31	1.25, 1.37
Antipsychotics	1.10	1.00, 1.21	0.98	0.90, 1.07
NSAIDS	1.37	1.23, 1.52	1.09	0.97, 1.23
Statins	0.61	0.58, 0.64	0.84	0.80, 0.88
Benzodiazepines	0.96	0.91, 1.02	0.95	0.91, 0.99
Systemic corticosteroids	1.18	1.07, 1.31	1.29	1.19, 1.40
H-2 receptor blockers	0.90	0.71, 1.14	1.12	0.94, 1.34
Proton pump inhibitors	0.87	0.83, 0.92	0.92	0.88, 0.96
Opioids	2.39	2.22, 2.56	1.27	1.18, 1.36

	DOACs		Warfarin	
Characteristics	**OR**	**95% CI**	**OR**	**95% CI**
SSRI	0.99	0.91, 1.07	0.87	0.81, 0.93

Abbreviations: CAB, coronary artery bypass surgery; CI, confidence interval; COPD, chronic obstructive pulmonary disease; NSAIDs, non-steroidal anti-inflammatory drugs; OR, odds ratio; PCI, percutaneous coronary intervention; SSRIs, selective serotonin reuptake inhibitors

Annex 16. Sensitivity analysis. Hazard ratios of the composite of stroke, major bleeding, MI, and death in individuals treated with warfarin vs DOACs (n=47,854) using different definitions of exposure and censoring adjustment. Québec, 2010-2016 (Third study).

Estimate	HR	95% CI	SE
Unweighted	1.35	1.28, 1.42	0.0252
Weighed with unstratified weights	1.41	1.34, 1.48	0.0261
Weighted with exposure-stratified weights	1.26	1.20, 1.33	0.0264
Weighted with exposure-stratified weights estimated from two probabilities of censoring (switch between treatment groups and permanent treatment discontinuation)	1.26	1.19, 1.35	0.0266

Abbreviations: DOACs, direct oral anticoagulants; CI, confidence interval; HR, hazard ratio; MI, myocardial infarction; SE, standard error.