Drug Exposure Definition and Healthy User Bias Impacts on the Evaluation of Oral Anti Hyperglycemic Therapies

by

Maxim Eskin

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in
Health Policy Research

School of Public Health

University of Alberta

ABSTRACT

Accurate estimation of medication use is an essential component of any pharmacoepidemiological research as exposure misclassification will threaten study validity and lead to spurious associations. Many pharmacoepidemiological studies use simple definitions, such as the categorical "any versus no use" to classify oral antihyperglycemic medications exposure, which has potentially serious drawbacks. This approach has led to numerous highly publicized observational studies of metformin effect on health outcomes reporting exaggerated relationships that were later contradicted by randomized controlled trials. Although selection bias, unmeasured confounding, and many other factors contribute to the discrepancies, one critical element, which is often overlooked, is the method used to define exposure.

Another factor that may provide additional explanation for the discrepancy between randomized controlled trials and observational study results of the association between metformin use and various health outcomes is the healthy user effect. Aspects of a healthy lifestyle such as following a balanced diet, exercising regularly, avoiding tobacco use, etc., are not recorded in typical administrative databases and failure to account for these factors in observational studies may introduce bias and produce spurious associations. The influence of a healthy user bias has not been fully examined when evaluating the effect of oral antihyperglycemic therapies in observational studies.

It is possible that some, if not all, of the benefit, observed with metformin in observational studies may be related to analytic design and exposure definitions. Thus, our first objective is to explore the potential impact of exposure definition on estimates of the association

between metformin and all-cause mortality risk, using a large administrative health database, similar to databases used in previous studies. The variety of exposure definitions tested in our analysis produced a wide range of associations between metformin and mortality risk, therefore, we recommend that pharmacoepidemiological studies should include at least two exposure definitions and sensitivity analyses of different exposure definitions. Moreover, our second objective is to explore the healthy user effect in metformin users versus non-users on various health outcomes that should not be associated with metformin use. Results of this study suggest that metformin users are more likely to initiate preventive therapies and engage in other healthy behaviors, therefore we recommend that the influence of these behaviors should be accounted for pharmacoepidemiological studies evaluating the effect of oral antihyperglycemic therapies.

AKNOWLEDGMENTS

I would like to express my deepest gratitude to the Alliance for Canadian Health Outcomes Research in Diabetes (ACHORD) and especially to my supervisors Prof. Dean T Eurich and Prof. Scot H Simpson, and I would like to extend my thanks to those who offered collegial guidance and support during the program, especially Ben T. Dubois.

This study was funded through an operating grant provided by the Canadian Diabetes Association and supported by an Alberta Diabetes Institute Studentship for Diabetes Research.

Additionally, I am thankful to Drug Safety and Effectiveness Training Program (DSECT), McMaster University for provision of the valuable knowledge, education and tools, that were useful in my research work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACE - Angiotensin-converting Enzyme

BMI - Body Mass Index

CDA - Canadian Diabetes Association

CI – Confidence Interval

CV – Cardio Vascular

FOB – Fecal occult Blood Test

F/U – Follow-up

HR – Hazard Ratio

ICD - International Classification of Diseases

IRR - Incidence Rate Ratio

NCC – Nested Case Control

OD – Odds Ratio

PIN - Pharmaceutical Information Network

PSA - Prostate-specific Antigen

RCT – Randomized Controlled Study

Rx - Prescription

SD – Standard Deviation

TVA – Time-varying Analysis

CHAPTER 1:

INTRODUCTION

1.1 Statement of the Problem

Few head-to-head randomized trials comparing oral antihyperglycemic medications on hard clinical outcomes exist. As a result, a large amount of evidence comparing the effectiveness of antihyperglycemic medication in diabetes comes from observational studies, which have become increasingly popular and make an important contribution to medical knowledge.(1) However, a major limitation of observational studies is their subjectivity to a number of potential biases that may cause erroneous results and produce spurious associations.

Although selection bias, unmeasured confounding, and many other factors contribute to the bias in pharmacoepidemiological research, one critical element, which is often overlooked, is the method used to define drug exposure.(2) Many pharmacoepidemiological studies use simple definitions, such as the categorical "any versus no use" to classify exposure, which has potentially serious drawbacks.(3, 4) This approach has led to numerous highly publicized observational studies of the effect of diabetes medications on health outcomes reporting exaggerated relationships that were later contradicted by randomized controlled trials (RCT).(5)

An additional potential source of bias is the 'Healthy User Effect' or 'Healthy User Bias', which is a tendency of healthier patients to initiate preventive therapies and engage in behaviors consistent with healthy lifestyles, which in turn can affect their outcomes. Aspects of a healthy

lifestyle include following a balanced diet, exercising regularly, moderation of alcohol consumption and avoidance of tobacco use, etc. Many of these healthy lifestyle behaviors are not recorded in typical administrative databases and failure to account for these factors in observational studies of preventive therapies could introduce bias and produce spurious associations.(6, 7) The healthy user effect is still not fully appreciated within pharmacoepidemiology research and has scarcely been considered in the evaluation of oral antihyperglycemic therapies. Therefore, healthy user bias may provide an alternative explanation for the discrepancy in results of metformin use and various health outcomes observed in RCTs and observational studies.

Although, exposure misclassification and healthy user effects are not the only possible explanations for the discrepancy between RCTs and observational study results,(8, 9) it is possible that some, if not all, of the benefit, observed with metformin in observational studies may be related to analytic design and exposure definition.(10-12)

1.2 Drug Exposure Definition in Pharmacoepidemiology

1.2.1 Exposure Definition Importance

In pharmacoepidemiology, the use of drugs is the determinant of interest when studying exposure outcome associations. The increased availability of administrative databases has greatly facilitated analyses of drug effects on a population-based scale. The early days of pharmacoepidemiology research relied heavily on the use of a simple ‘ever’ versus ‘never’ exposure definition for drug use. It is now clear that this definition has potentially serious

drawbacks, including the introduction of misclassification bias into research results.(3, 4) For example, for the evaluation of many drug/event associations, the dose, the duration, and timing of use (which would not be captured in a simple 'ever' versus 'never' definition) are very important in understanding the drug's potential effect.(3, 4) Moreover, switching between different therapies and types of administration are also important and not captured in the traditional exposure definition.(1, 4) Therefore, accurate exposure definition is an important requirement in every pharmacoepidemiological study, since inaccurately assessed drug exposure can introduce misclassification, bias study results, and produce spurious associations, that consequently have a negative impact on patient safety and treatment effectiveness assessment.(1, 4, 13, 14)

1.2.2 Methods to Define Drug Exposure

Based on a comprehensive search of the literature, a wide range of exposure definitions, and their variants, have been used in pharmacoepidemiological research of oral antihyperglycemic medications in diabetes (Table 1-1). Three general approaches to an exposure definition include time-fixed approaches, time-varying approaches, and a nested case control (NCC) approach. Each method has both strengths and weakness associated with the definition and importantly, each method provides a slightly different interpretation in the evaluation of drug effect.

Table 1-1. Exposure Methods Reference Sources

Exposure Method		Chaiteerakij, 2016	Tseng CH, 2015	Lee et al, 2011	Mamtani, 2014	Margel, 2013	Bowker, 2010	Simpson, 2016	Abdelmoneim, 2016	Eurich, 2013	Calip, 2015	Failie, 2014	MacDonald, 2010	Bensimon, 2014	Azoulai, 2016	Lipscombe, 2007	Shih, 2015
Time-fixed analysis	'Ever' versus 'never' users	✓	✓														
	Filling at least 2 prescriptions			✓	✓												
	Cumulative exposure in years					✓											
Time-varying analysis	Legacy effect	✓								✓	✓						
	Discontinue the follow up if no evidence of ongoing exposure									✓		✓					
	Any prescription fill within a time window						✓										
	Any use within a time window based on days of supply							✓	✓	✓							
	Cary over effect of 10%									✓							
Nested Case Control	Any prescription fill prior to the event date												✓	✓			
	Any use prior to the event date														✓		
	Current use, Past use or Never															✓	✓

Definitions using a time-fixed approach establish medication exposure at a single point based on a portion or all of the prescription records in the study observation period. This exposure definition does not change during the follow-up period and is entered into multivariable models as a dichotomous variable to describe exposure status. Examples of this approach include the 'ever' versus 'never' method,(11, 15-17) and the requirement of at least 2 prescription records within a defined interval, such as the entire study period(18) or within 180 days.(19, 20) A variation of these definitions is to use either the interval between first and last prescription record or the cumulative days of supply information to define exposure as a continuous variable.(11, 21-23)

Time-varying approaches examine the patient's prescription records at multiple points during the follow-up period to establish exposure status. Examples of this approach include: a) legacy effect, where subjects are considered "unexposed" until the first prescription record, then considered "exposed" until the end of the follow-up, regardless of subsequent prescription information;(15, 24, 25) b) a variation of this definition is to discontinue follow-up (i.e., censor patients) if there is no evidence of ongoing medication use;(17, 25, 26) and, c) dividing the follow-up period into set intervals or 'windows', determining exposure status within each window, and using this information as a time dependent variable in the analytical model.(27) To determine exposure status within a window, several different methods exist, including a single prescription record within the window(13); any use within a window based on expected availability from the prescription date and days of supply information(25, 28, 29); and any use within a window based on expected availability from the prescription date, days of supply information, plus a 'carry-over effect' of 10% to allow for poor adherence.(25) Other variations of these also exist but by analyzing exposure to drugs as a time-varying variable in a Cox regression model, cohort studies with complete coverage of all filled prescriptions can provide us with a better understanding of the potential benefits or risks associated with the drug.

Exposure definitions in nested case control studies use either a prescription record or evidence of medication use (based on prescription date and days of supply information) within a set time period (usually 30, 90, 180 or 365 days) prior to case event date.(30-33) In addition, some NCC studies categorized exposure status as current (prescription date plus days of supply overlap the case event date), past (prescription date pus days of supply end before the case event date) or never (no prescription records prior to the case event date).(34, 35) The NCC approach is commonly used in pharmacoepidemiology because of the time varying nature of drug use and

the long duration of follow-up, moreover, this approach provides an efficient and flexible analysis of the association between immediate exposure and outcomes.(36) However, this approach was initially developed to overcome limitations in computing power which is not required in today's digital world.

1.2.3 Exposure Misclassification

How differences in exposure definitions introduce bias into pharmacoepidemiological studies stem from misclassification bias. Exposure misclassification can be either differential or non-differential. Differential bias occurs when the error rate or probability of being misclassified differs across groups of study subjects. This bias has historically been perceived as shifting the results away from the null, however, the direction of bias is completely driven by the amount and direction of misclassification within the study groups. Non-differential misclassification occurs when the frequency of errors is approximately the same in the groups being compared and tends to bias the results towards the null.(37, 38)

Another type of bias which relates to how drug exposure is defined in studies but is not a drug exposure misclassification per se is the concept of immortal time bias. Immortal time bias refers to a period of follow-up during which, by design, the study outcome cannot occur. In pharmacoepidemiological studies, immortal time typically arises when the determination of an individual's treatment exposure involves a delay or wait period during which follow-up time is accrued. Therefore, the timing of follow-up initiation after reaching the exposure definition is of a special importance.(1, 14, 39) Time-related biases in pharmacoepidemiological research, such as immortal time bias and others have been previously described in studies of diabetes

treatment,(40, 41) and the major concern with the analysis of such studies is that patients, who reach the outcome during immortal time period is misclassified as unexposed, leading to the exaggeration of the medication effect.(1, 18, 40, 41)

Exposure misclassification and immortal time bias can be introduced into a pharmacoepidemiological study by the method used to define drug exposure. Table 1-2 outlines the various biases which can be introduced for the most common exposure definitions used in the evaluation of oral antihyperglycemic medications in diabetes and the theoretical impact on the observed estimates.

Table 1-2: Potential Biases Associated with Exposure

Exposure Definition Retrospective Cohort	Immortal Time	Exposure Misclassification	Impact on the estimate ₸
Retrospective Cohort Analysis			
Ever vs. Never metformin use			
Follow up starting from index prescription	Medium	Non Differential	Overestimation
Follow up starting from first metformin prescription	Short	Non Differential	Neutral
Filling at least 2 prescriptions			
all study period (starting from index prescription)	Long	Differential	Overestimation
all study period (starting from 1st prescription)	Medium	---	Neutral
all study period (starting from 2nd prescription)	Short	Differential	Overestimation
180 days (starting from index prescription)	Long	Differential	Overestimation
180 days (starting from 1st prescription)	Medium	---	Neutral
180 days (starting from 2nd prescription)	Short	Differential	Overestimation
Metformin use in years as a continuous variable			
Metformin use in years	Short	Non Differential/Differential	Overestimation

Time Varying Exposure Definition Analyses			
Legacy effect			
30 days window size	---	Differential	Underestimation
90 days window size	---	Differential	Underestimation
180 days window size	---	Differential	Underestimation
365 days window size	---	Differential	Underestimation
Any metformin prescription fill regardless of days supplied			
30 days	---	Differential	Overestimation
90 days	---	Differential	Overestimation
180 days	---	Differential	Overestimation
365 days	---	Differential	Overestimation
Any metformin use based on days' supply			
30 days	---	---	Neutral
90 days	---	---	Neutral
180 days	---	---	Neutral
365 days	---	---	Neutral
Discontinue the follow up if there is no evidence of ongoing exposure			
30 days	---	Differential	Overestimation
90 days	---	Differential	Overestimation
180 days	---	Differential	Overestimation
365 days	---	Differential	Overestimation
Carry Over effect of 10% of the medication			
30 days' time windows	---	Differential	Underestimation
90 days' time windows	---	Differential	Underestimation
180 days' time windows	---	Differential	Underestimation
365 days' time windows	---	Differential	Underestimation
Nested Case Control Exposure Definition Analyses			
Any Metformin prescription claim within time windows prior to the event			
30 days	---	---	Neutral
90 days	---	---	Neutral
180 days	---	---	Neutral
365 days	---	---	Neutral
Any Metformin use within time windows prior to the event			
30 days	---	---	Neutral
90 days	---	---	Neutral
180 days	---	---	Neutral
365 days	---	---	Neutral
Current and Past use versus never use			
Current Use	---	Non Differential	Underestimation
Past Use	---	Non Differential	Underestimation

1.3 Healthy User Bias

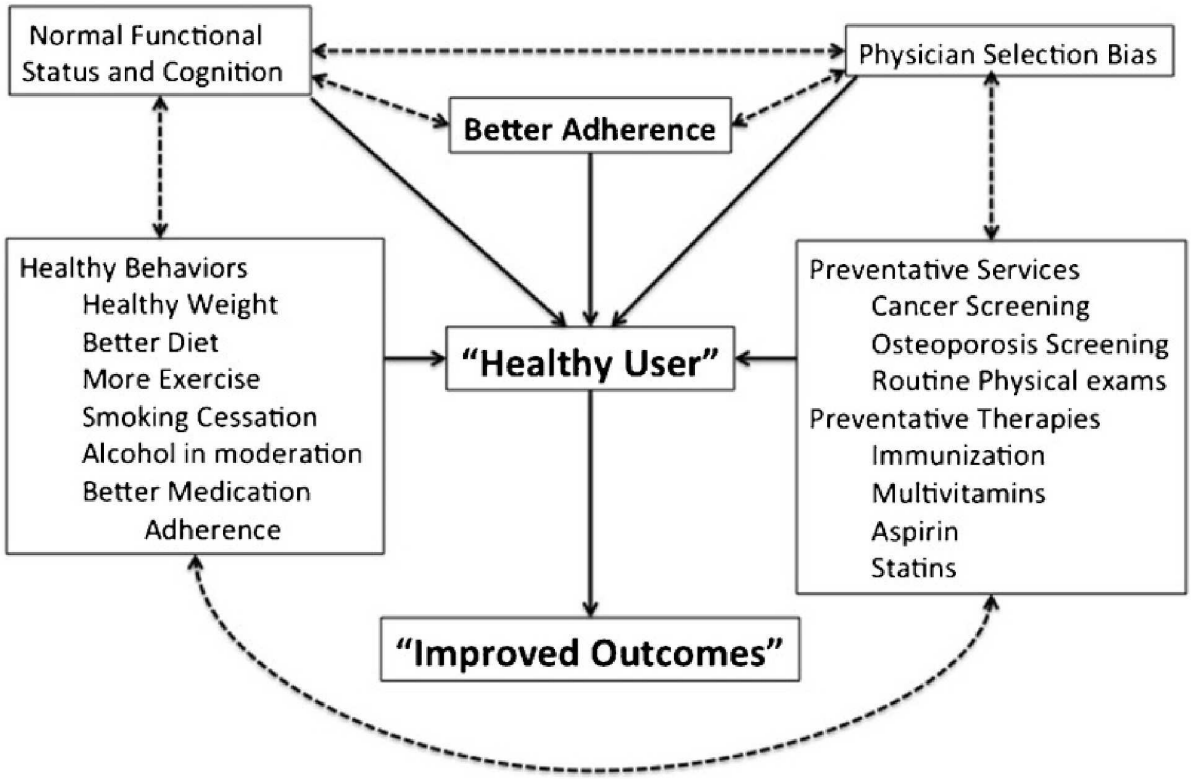
1.3.1 Healthy User Effect Definition

Pharmacoepidemiological studies are conducted to evaluate the safety and effectiveness of drugs used in actual clinical settings using large administrative databases. These studies are of a high importance, since RCTs often exclude populations most likely to use a drug, do not account for the complexity of pharmacotherapies, or are simply not feasible in specific settings. One of the main limitations of such studies is bias, i.e., systematic differences in prognosis between a group of patients using a medication and the comparator group of patients.(6)

Healthy User Effect or Healthy User Bias is a tendency of healthier patients to initiate preventive therapies and engage in behaviors consistent with healthy lifestyles, which in turn can affect their health outcomes.(6, 7, 42) This could occur through either selective prescribing of preventive medications to patients in better health and/or through more health conscious patients seeking out prescriptions for such medications.(7)

Aspects of a healthy lifestyle include following a balanced diet, exercising regularly, moderation of alcohol consumption, avoidance of tobacco use, avoidance of risky behaviors, participation in vaccination programs, using preventive therapies and health services, overall higher functional status and cognition, and interaction with the health system (e.g., annual visits). These characteristics are usually unmeasured in the administrative databases, and, therefore, may be associated with health outcomes in observational studies of vaccinations, diets, screening procedures and preventative therapies (Figure 1).(7, 43, 44)

Figure 1: Conceptual model explaining the healthy-user effect[§]:



[§] Copyright © Society of General Internal Medicine 2011

1.3.2 Healthy User Bias Impact on the Drug Exposure and Health Outcome Association

Healthy lifestyle behaviors are not recorded in typical administrative databases and failure to account for these factors in observational studies of preventive therapies could introduce bias and produce spurious associations.(6, 7) The healthy user effect has gained increasing attention as a potential source of bias in observational studies of preventive therapies and treatment of asymptomatic conditions.(43) Indeed, numerous studies have demonstrated the

healthy user effect can be a plausible explanation for the discrepancy between RCTs and observational studies.(6, 7, 45)

The effect of healthy user bias on the observed association can often be anticipated. When the effect is inversely associated with the outcome, study results will tend to be smaller than the true underlying effect. A failure to account for the healthy user effect in observational studies will make null effects appear protective, exaggerate protective effects, and attenuate harmful effects or even make them appear protective.(7)

1.3.3 Healthy User Effect in the Previous Studies

Healthy user effect has been most notably explored in users of 'statin' therapies.(42) Indeed, statins (5-hydroxy-3-methylglutaryl-coenzyme A reductase inhibitors) therapy is associated with increased use of a number of other preventative therapies and health services, and a reduction in clinical events which are not intuitively associated with statin use, including reduced risk of cancer, sepsis, Alzheimer disease and hip fractures.(6, 7, 46) Furthermore, the healthy user effect is thought to fully explain the large benefits observed in observational studies of hormone replacement therapies and the subsequent RCTs which showed significant harms.(45) Additionally, it has also been extensively discussed as a potential source of bias in observational studies showing large reductions in all-cause mortality associated with influenza vaccination in elderly.(44, 47)

1.3.4 Methods to Control for the Healthy User Effect

There are several analytical strategies proposed in the literature to control for healthy user effect. The optimal method to control for the bias is randomization,(9) however, it is not always feasible and the majority of the research presently performed in observational studies requires other strategies to address the bias. Therefore, other analytical strategies were proposed, such as new user design,(46, 48) using an active comparator,(9, 49, 50) controlling for past medical history by adjustments for prior use of preventive therapies and diagnostic tests, using an instrumental variable,(43, 48) using a propensity score,(51) and refining the exposure definition by using a time-dependent method after using an 'ever' versus 'never' definition.(49)

The variety of these analytical strategies highlight the fundamental difficulties in evaluating preventive therapies within observational studies and emphasize the fact that several approaches are required in order to control for healthy user effect.(6)

1.4 Biases in Oral Antihyperglycemic Agents Research

1.4.1 Literature Review

Numerous observational studies in patients with diabetes have consistently shown metformin to be associated with ~30% reduction in all cause or cardiovascular specific mortality compared to other oral antihyperglycemic agents.(52, 53) Yet, only one small, sub-study within the UKPDS randomized controlled trial has suggested a similar benefit in obese patients with type 2 diabetes.(54) Indeed, meta-analyses of RCT's have failed to replicate the strong association between metformin use and all-cause or CV mortality reported in observational

studies.(2, 55) Importantly, the vast majority of these RCT studies have not been designed to look at mortality associated with metformin *per se*. As a result, a large body of observational studies have further explored the issue and have consistently shown clinically important benefits, including substantial reductions in all-cause and CV-related mortality, as well as beneficial effects in patients with diabetes and heart failure.(55-57) However, not all observational studies of antihyperglycemic therapies have found benefits with metformin. The reason for the discrepancy between RCT's of metformin and within observational studies of metformin are not fully known. One plausible explanation is that these differences may be driven, at least in part, by the underlying definitions used to define metformin exposure in observational studies.

Although less studied, healthy user bias may also explain the discrepancies in results between observational studies and RCTs of metformin use in patients with type 2 diabetes. Although healthy-user bias is most commonly associated with statin and HRT use, there have been suggestions in observational studies of metformin that health user bias may also be at play. Indeed, metformin has been shown to reduce a wide range of outcomes in patients with diabetes, including cancer-related outcomes which one would not normally expect.(12, 40, 41) Therefore, it is possible some, if not all, of the benefit, observed with metformin in observational studies may be related to the healthy user effect and exposure definitions.(10, 11) and it may also explain the discrepancies in results between observational studies and RCTs of metformin use in patients with type 2 diabetes.

1.5 Summary

This research is directly relevant to diabetes, specifically, and health outcomes research in general.(55) In order to accurately estimate drug exposure and health outcomes association, it is crucial to be able precisely define drug exposure and address patient characteristics, like healthy behaviors in a pharmacoepidemiological study. Either during clinical trial or post-marketing drug safety and effectiveness evaluation, proper methodology would contribute to the validity of study results and consequently to safer health care and more effective patient treatment. The availability of computerized information about drug use on an individual basis (e.g., administrative databases) facilitated analyses of oral antihyperglycemic medications effect on a population-based scale. Furthermore, reliable and accurate drug exposure definition along with accounting for other cofounding factors related to patient behavior are of paramount importance in pharmacoepidemiology, since any misclassification will potentially bias study results and produce spurious associations. Moreover, this unique research will shift the current paradigm on how observational studies define exposure of diabetes medications, account for patient characteristics, and may also reconcile the differences between observational studies and RCTs on the potential benefits of metformin on mortality in patients with diabetes.

1.6 Objectives

The two objectives of this program of research were:

1) To evaluate the influence of exposure definitions used in pharmacoepidemiology on estimates, using the association between metformin and all-cause mortality as a proto-typical model

2) To explore the healthy user effect in metformin users versus non-users on various health outcomes that should not be associated with metformin use

The first objective was achieved using a large population-based cohort of 64,293 new oral antihyperglycemic users (≥66 years) between 1998 and 2010 in Alberta, Canada. We examined oral antihyperglycemic medications use and all-cause mortality as a prototypical model to evaluate the effect of different exposure definitions on the estimated association using three major groups for exposure definition (time fixed, time-varying and nested case control) (more comprehensively elaborated in chapter 2). We hypothesized that the association between metformin and mortality will be significantly influenced by the method used to define 'exposure'.

The second objective was addressed with a large population-based cohort of 135,301 new users of oral antihyperglycemic agents in Alberta (≥30 years), Canada based on administrative data augmented with clinical laboratory data between 2008 and 2015 (more comprehensively elaborated in chapter 3). We hypothesized that metformin use, in part, is a marker of "healthy

users" and will be associated with beneficial effects on different health outcomes that could not plausibly be related to metformin use *per se*.

1.7 References

1. Suissa S. Immortal Time Bias in Pharmacoepidemiology. *Am J Epidemiol*. 2008;167(4):492-99.
2. Boussageon R, Supper I, Bejan-Angoulvant T, et al. Reappraisal of Metformin Efficacy in the Treatment of Type 2 Diabetes: a Meta-Analysis of Randomized Controlled Trials. *PLoS Medicine*. 2012;9(4):e1001204.
3. Walker JJ, Johnson JA, Wild SH. Diabetes treatments and cancer risk: the importance of considering aspects of drug exposure. *Lancet Diabetes Endocrinol*. 2013;1(2):132-9.
4. Stricker B, Stijnen T. Analysis of individual drug use as a time-varying determinant of exposure in prospective population-based cohort studies. *Eur J Epidemiol*. 2010;25:245-51.
5. Bannister CA, Holden SE, Jenkins-Jones S, Morgan CL, Halcox JP, Schernthaner GM, J Currie, CJ. Can People With Type 2 Diabetes Live Longer Than Those Without? A Comparison of Mortality in People Initiated With Metformin or Sulphonylurea Monotherapy and Matched Non-diabetic Controls *Diabetes Obes Metab*. 2014;16(11):1165-73.
6. Dormuth C, Patrick A, Shrank W, et al. Statin Adherence and Risk of Accidents A Cautionary Tale. *Circulation*. 2009;119(15):2051-7.
7. Brookhart M, Patrick A, Dormuth C, et al. Adherence to Lipid-lowering Therapy and the Use of Preventive Health Services: An Investigation of the Healthy User Effect. *American Journal of Epidemiology* 2006;166(3):348-54.
8. Vandenbroucke J. When are observational studies as credible as randomised trials? *Lancet*. 2004;363:1728–31.
9. Schneeweiss S, Patrick A, Stürmer T, et al. Increasing Levels of Restriction in Pharmacoepidemiologic Database Studies of Elderly and Comparison With Randomized Trial Results. *Med Care*. 2007;45(10):131-42.
10. Brunelli S. Use of prescription drug claims data to identify lipid-lowering medication exposure in pharmacoepidemiology studies: potential pitfalls. *Pharmacoepidemiol Drug Saf*. 2016;25(7):844-6.
11. Tournier M, Bégaud B, Cougnard A, et al. Influence of the drug exposure definition on the assessment of the antipsychotic metabolic impact in patients initially treated with mood-stabilizers. *Br J Clin Pharmacol*. 2012;74(1):189-96.
12. Zanders M, van Herk-Sukel M, Vissers P, Herings R, Haak H, van de Poll-Franse L. Are metformin, statin and aspirin use still associated with overall mortality among colorectal cancer patients with diabetes if adjusted for one another? *BJC*. 2015;113:403–10.
13. Bowker S, Yasui Y, Veugelers P, Johnson J. Glucose-lowering agents and cancer mortality rates in type 2 diabetes: assessing effects of time-varying exposure. *Diabetologia*. 2010;53(8):1631–37.
14. Suissa S. Inhaled steroids and mortality in COPD: bias from unaccounted immortal time. *Eur Respir J*. 2004;23:391–5.
15. Chaiteerakij R, Petersen G, Bamlet W, et al. Metformin Use and Survival of Patients With Pancreatic Cancer: A Cautionary Lesson. *J Clin Oncol*. 2016;34(16):1898-904.
16. Tseng C. Metformin and endometrial cancer risk in Chinese women with type 2 diabetes mellitus in Taiwan. *Gynecol Oncol*. 2015;138(1):147-53.
17. Faillie J, Azoulay L, Patenaude V, Hillaire-Buys D, Suissa S. Incretin Based Drugs and Risk of Accute Pancreatitis in Patients with Type 2 Diabetes: cohort Study. *BMJ*. 2014;348(g2780).
18. Lee M, Hsu C, Wahlqvist M, Tsai H, Chang Y, Huang Y. Type 2 diabetes increases and metformin reduces total, colorectal, liver and pancreatic cancer incidences in Taiwanese: a representative population prospective cohort study of 800,000 individuals. *BMC Cancer*. 2011;11(20).
19. Ferrara A, Lewis J, Quesenberry C, et al. Cohort Study of Pioglitazone and Cancer Incidence in Patients with Diabetes. *Diabetes Care* 2011;34(4):923–9.
20. Mamtani R, Pfanzelter N, Haynes K, et al. Incidence of bladder cancer in patients with type 2 diabetes treated with metformin or sulfonylureas *Diabetes Care*. 2014;37(7):1910-7.

21. Ruiter R, Visser L, Van Herk-Sukel M, et al. Lower Risk of Cancer in Patients on Metformin in Comparison With Those on Sulfonylurea Derivatives. *Diabetes Care*. 2012;35(1):119–24.
22. Margel D, Urbach D, Lipscombe L, Bell C, Kulkarni G, Austin P. Metformin use and all-cause and prostate cancer-specific mortality among men with diabetes *J Clin Oncol* 2013;31(25):3069-75.
23. Nayana M, Macdonaldb E, Juurlink D, et al. Medication use and survival in diabetic patients with kidney cancer: A population-based cohort study. *Pharmacol Res*. 2016;113(Pt A):468-74.
24. Calip G, Yu O, Hoskins K, Boudreau D. Associations between diabetes medication use and risk of second breast cancer events and mortality. *Cancer Causes Control*. 2015;26(8):1065–77.
25. Eurich D, Simpson S, Senthilselvan A, Asche C, Sandhu-Minhas J, McAlister F. Comparative Safety and Effectiveness of Sitagliptin in Patients with type 2 diabetes: retrospective population based cohort study. *BMJ*. 2013;346(f2267).
26. Garg R, Chen W, Pendergrass M. Acute Pancreatitis in Type 2 Diabetes Treated With Exenatide or Sitagliptin. *Diabetes Care*. 2010;33(11):2349-54.
27. Vittinghoff E, Glidden D, Shiboski S, McCulloch C. Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models. New York, NY:Springer 2012:405-6.
28. Simpson S, Lin M, Eurich D. Medication Adherence Affects Risk of New Diabetes Complications: A Cohort Study. *Annals of Pharmacotherapy*. 2016;5(9):741–6.
29. Abdelmoneim A, Eurich D, Senthilselvan A, Qiu W, Simpson S. Dose-response relationship between sulfonylureas and major adverse cardiovascular events in elderly patients with type 2 diabetes. *Pharmacoepidemiol Drug Saf*. 2016;25(10):1186-95.
30. Bensimon L, Yin H, Suissa S, Pollak M, Azoulay L. The Use of Metformin in Patients with Prostate Cancer and the Risk of Death. *Cancer Epidemiol Biomarkers Prev*. 2014;23(10):2111-8.
31. MacDonald M, Eurich D, Majumdar S, et al. Treatment of type 2 diabetes and outcomes in patients with heart failure: a nested case-control study from the U.K. General Practice Research Database. *Diabetes Care*. 2010;33(6):1213-8.
32. Azoulay L, Filion K, Platt R, et al. Association Between Incretin-Based Drugs and the Risk of Acute Pancreatitis. *JAMA Intern Med*. 2016;176(10):1464-73.
33. Eurich D, Weir D, Simpson S, Senthilselvan S, McAl ister F. Risk of new-onset heart failure in patients using sitagliptin: a population-based cohort study. *Diabet Med*. 2016;33(5):621–30.
34. Lipscombe L, Gomes T, Lévesque L, Hux J, Juurlink D, Alter D. Thiazolidinediones and cardiovascular outcomes in older patients with diabetes. *JAMA*. 2007;298(22):2634-43.
35. Shih C, Wu Y, Chao P, et al. Association between Use of Oral Anti-Diabetic Drugs and the Risk of Sepsis: A Nested Case-Control Study. *Scientific Reports*. 2015;5.
36. Azoulay L, Yin H, Filion K, et al. The use of pioglitazone and the risk of bladder cancer in people with type 2 diabetes: nested case-control study. *BMJ*. 2012;344(e3645).
37. Porta M. A Dictionary of Epidemiology (Fifth ed.). New York: Oxford University Press; 2008.
38. Flegal K, Keyl P, Nieto J. Differential Misclassification Arising from Non differential Errors in Exposure Measurement *Am J Epidemiol*. 1991;134(10):1233-46.
39. Lévesque L, Hanley J, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ*. 2010;340.
40. Suissa S, Azoulay L. Metformin and the Risk of Cancer Time-related biases in observational studies *Diabetes Care*. 2012;36(6):e88.
41. Suissa S, Azoulay L. Metformin and Cancer: Mounting Evidence Against an Association. *Diabetes Care*. 2014;37(7):1786-8.
42. Eurich D, Majumdar S. Statins and sepsis - scientifically interesting but clinically inconsequential. *J Gen Intern Med*. 2012;27(3):268-9.
43. Patrick A, Shrank W, Glynn R, et al. The Association between Statin use and Outcomes Potentially Attributable to an Unhealthy Lifestyle in Older Adults. *Value Health*. 2011;14(4):513-20.
44. Jackson L, Jackson M, Nelson J, Neuzil K, Weiss N. Evidence of bias in estimates of influenza vaccine effectiveness in seniors. International Journal of Epidemiology. 2006;35:337-44.

45. Ray W. Observational studies of drugs and mortality. *N Engl J Med*. 2005;353(22):2319-21.
46. Brassard P, Wu J, Ernst P, Dell'Aniello S, Smiechowski B, Suissa S. The effect of statins on influenza-like illness morbidity and mortality. *Pharmacoepidemiology and Drug Safety* 2016.
47. Eurich D, Marrie T, Johnstone J, Majumdar S. Mortality reduction with influenza vaccine in patients with pneumonia outside "flu" season: pleiotropic benefits or residual confounding? *Am J Respir Crit Care Med*. 2008;178(5):527-33.
48. Brookhart M, Wang P, Solomon D, Schneeweiss S. Evaluating Short-Term Drug Effects Using a PhysicianSpecific Prescribing Preference as an Instrumental Variable. *Epidemiology*. 2006;17(3):268-75.
49. Bykov K, Yoshida K, Weisskopf M, Gagne J. Confounding of the association between statins and Parkinson disease: systematic review and meta-analysis. Pharmacoepidemiology and Drug Safety 2016.
50. Margel D, Urbach D, Lipscombe L, et al. Metformin Use and All-Cause and Prostate Cancer–Specific Mortality Among Men With Diabetes. *J Clin Oncol*. 2013;31:3069-75.
51. Gleason C, Dowling N, Friedman E, Wharton W, Asthana S. Using predictors of hormone therapy use to model the healthy user bias: how does healthy user status influence cognitive effects of hormone therapy? *Menopause*. 2012;19(5):524-33.
52. Eurich D, McAlister F, Blackburn D, et al. Benefits and harms of antidiabetic agents in patients with diabetes and heart failure: systematic review. *BMJ*. 2007; 335(7618):497.
53. Johnson J, Majumdar S, Simpson S, Toth E. Decreased Mortality Associated with the Use of Metformin Compared with Sulfonilurea Monotherapy in Type 2 Diabetes. *Diabetes Care* 2002;25(12):2244-8.
54. Group UPDSU. Effect of intensive blood glucose control with metformin on complications in overweight patients with type 2 diabetes *Lancet*. 1998;352(9131):854–65.
55. UKPDS. Effect of intensive blood glucose control with metformin on complications in overweight patients with type 2 diabetes *The Lancet*. 1998;352(9131):854–65.
56. Eurich D, McAlister F, Blackburn D, et al. Benefits and harms of antidiabetic agents in patients with diabetes and heart failure: systematic review. *BMJ*. 2007;335(7618):497.
57. Jasim S, Smith S. Review: Metformin is linked to reduced mortality in type 2 diabetes with comorbid CKD and CHF. *Ann Intern Med*. 2017;166(8).

CHAPTER 2:

Impact of Drug Exposure Definitions on an Observed Association Pharmacoepidemiology Research

2.1 Introduction

Few head-to-head randomized trials comparing oral antihyperglycemic medications on hard clinical outcomes (e.g. mortality, cardiovascular outcomes) exist. As a result, a large amount of evidence comparing antihyperglycemic medications effectiveness in diabetes comes from pharmacoepidemiological studies. Accurate estimation of medication use is an essential component of any pharmacoepidemiological research as exposure misclassification will threaten study validity and lead to spurious associations.(10)

Many pharmacoepidemiological studies, however, use crude definitions, such as the categorical "any versus no use" to classify exposure, which has potentially serious drawbacks.(3) This approach has led to numerous highly publicized observational studies of the effect of diabetes medications on health outcomes reporting exaggerated relationships(5, 55) that were later contradicted by randomized controlled trials.(54) Although selection bias, unmeasured confounding, and many other factors contribute to the discrepancies, one critical element, which is often overlooked, is the method used to define exposure.(11, 55)

The association between metformin and all-cause mortality is a good proto-typical model to examine the influence of exposure definitions because of the differences observed in pharmacoepidemiological studies and randomized controlled trials. Numerous observational

studies in patients with diabetes have consistently shown metformin to be associated with ~30% reduction in all cause or cardiovascular specific mortality compared to other oral antihyperglycemic agents.(51, 52) Yet, only one small, sub-study within the UKPDS randomized controlled trial has suggested a similar benefit in obese patients with type 2 diabetes.(53) Indeed, a recent meta-analysis of 13 RCTs found no benefit of metformin *per se* relative to other treatments.(54)

It is possible some, if not all, of the benefit, observed with metformin in observational studies may be related to analytic design and exposure definitions.(10, 11) Thus, using a large administrative health database, similar to databases used to evaluate outcomes associated with metformin therapy in previous studies, we explored the potential impact of exposure definition on estimates of the association between metformin and all-cause mortality risk. Although we are using diabetes medication as our prototypical example, our results would be expected to apply to almost all pharmacoepidemiological studies of drug safety and effectiveness.

2.2 Methods

Between January 1, 1998 and December 31, 2010, all new users of an oral antihyperglycemic agent aged 66 years and older were identified using the administrative health databases from Alberta, Canada. We used a standard approach by defining new users as those with no prescription record for any oral antihyperglycemic or insulin for one year before their index date.(52, 59) The restriction to patients 66 years or older was required as in Alberta only patients 65 years of age and older are eligible for universal drug coverage, allowing for 1 year to

establish our new user cohort. Socio demographic information was extracted from Alberta Registry database and mortality was ascertained from Vital Statistics Registry data.

Exposure Definitions

Based on a comprehensive search of the literature, we identified the most common exposure definitions, and their variants, used in pharmacoepidemiological research of oral antihyperglycemic agents in diabetes (Table 1-1). Three general approaches to exposure definitions were identified: 1) time-fixed approaches (Figure 1-A), 2) time-varying approaches (Figure 1-B), and 3) nested case control approach (Figure 1-C).

Definitions using a time-fixed approach establish medication exposure at a single point based on a portion or all of the prescription records in the study observation period. This exposure definition does not change during the follow-up period. Examples of this approach include 'any versus no' prescription record (i.e., 'ever versus never' users);(11, 14-16) and filling at least 2 prescriptions within a defined interval, such as the entire study period(17) or within 180 days.(18, 19) These exposure definitions are entered into multivariable models as a dichotomous variable to describe exposure status. A variation of these definitions is to use either the interval between first and last prescription record or the cumulative days of supply information to define exposure as a continuous variable.(11, 20-22)

Definitions using a time-varying approach examine a patient's prescription records at multiple points during the follow-up period to establish exposure status. The simplest method to define exposure using this approach is the legacy effect, where subjects are considered "unexposed" until the first prescription record, then considered "exposed" until the end of the

follow-up, regardless of subsequent prescription information.(14, 23, 24) A variation of this definition is to discontinue follow-up (i.e., censor patients) if there is no evidence of ongoing medication use among exposed and unexposed patients.(16, 24, 25)

Other time-varying definitions divide the follow-up period into set intervals or 'windows', determine exposure within each window, and use this information as a time dependent variable in the analytical model.(26) Numerous examples of cohort studies using Cox proportional hazards models were identified in the literature, with the time windows ranging from 1 day to 1 year, or according to actual prescription records.(12, 24, 27, 28, 60) To limit the total number of analyses completed, we elected to focus on the more commonly used windows of 30, 90, 180, 365 days or actual prescription records. We also followed the most common procedure by defining time zero as the start of the first oral antihyperglycemic agent use,(35, 59, 61) then splitting the follow-up time into consecutive windows. To establish exposure status within the window, we found several different methods, including a single prescription record within the window(12); any use within a window based on expected availability from the prescription date and days of supply information(24, 27, 28); and any use within a window based on expected availability from the prescription date, days of supply information, plus a 'carry-over effect' of 10% to allow for poor adherence.(24)

Exposure definitions in nested case control studies used either a prescription record or evidence of medication use (based on prescription date and days of supply information) within a set time period (usually 30, 90, 180 or 365 days) prior to case event date.(29-32) In addition, some nested case control studies categorized exposure status as current (prescription date plus

days of supply overlap the case event date), past (prescription date pus days of supply end before the case event date) or never (no prescription records prior to the case event date).(33, 34)

Analytic Design

In all of our models, we adjusted for age, sex, and a comorbidity score.(62) We used a variation of the Elixhauser comorbidity score, which uses ICD-9/10 codes to identify a defined list of diseases and generate a single ordinal variable that is an independent predictor of mortality risk.(63) All comorbidities were identified based on hospital discharge records and emergency department visit records within one year of staring the first oral antihyperglycemic agent.

Our reference group for all models was subjects who did not meet the definition for metformin exposure under study. Depending on the study design and exposure definition, this unexposed group would be patients who did not receive any metformin prescriptions during the entire follow-up time, during an individual window, or during the follow-up time prior to the first metformin prescription record. In all the models, numerous variants of the exposure definitions (e.g., 2 prescriptions within 180 days, or within the follow-up period; different cut-points based on time, etc.) were explored to determine if these subtle changes in the definitions would substantially affect the estimate. All variants were consistent with the more standard definitions identified in the literature; therefore, these additional analyses are not presented but are available on request from the authors (DTE).

In the cohort studies, individuals were followed from the index date (first prescription record of an oral antihyperglycemic agent) until death or censoring. Individuals were censored at the earliest of the end of the study period (December 31, 2010), or departure from the provincial

database. We used Cox Proportional hazard regression models for the cohort study designs to calculate adjusted hazard ratios (aHR) and 95% confidence intervals (CI).

In the nested case control studies, we followed conventional risk-set sampling methods by defining cases as patients with our event of interest (all-cause mortality) and selecting five controls from among those who have not experienced the event after the same duration of follow-up as the case. Adjusted odds ratios (aOR) and 95% CI were calculated using conditional logistic regression models.

We fully acknowledge that our models are prone to additional confounding factors. However, our goal is not to establish whether metformin is associated with mortality but to explore the effects of different exposure definitions. As we are using the same source of information and the same set of variables for all analyses, we would expect that all models would have the same relative degree of confounding. The only change among models is the method used to define metformin exposure. Thus, any changes in estimates between models would be expected to be driven, in large part, by the underlying biases associate with the definition used to classify metformin exposure.

2.3 Results and Discussion

Our cohort consisted of 64,293 new oral antihyperglycemic agent users. The average age in the cohort was 69 years, and 52% were male. Overall, 86% of the cohort (55,525 patients) filled at least one metformin prescription (Table 1-3). Compared to non-metformin users, metformin users tended to be slightly younger (69 vs. 71 years) and have a lower level of comorbidity (5 vs. 7). After an average follow-up of 6 (SD±4) years, 39% of patients died from

any cause. Fewer metformin users died (19,636; 35.4%) relative to those not using metformin (5,109; 58.3%) at any point in the follow-up.

Time-fixed Definitions

The various definitions using a time-fixed approach to establish exposure status produced consistent estimates of metformin effect on all-cause mortality, relative to those not using metformin (Figure 1-D). When a single prescription record is used to define exposure, regardless of timing in the follow-up period, any metformin use was associated with a substantial reduction in the risk of death (aHR 0.64, 0.62-0.66). However, ignoring the interval between first oral antihyperglycemic agent use and first metformin prescription may introduce unintended bias, in particular with regard to survival, and lead to an overestimation of the effect. Starting follow-up for the exposed group from first metformin prescription date, as depicted in Figure 1-A, will aid in eliminating some survival bias and improve estimation of the effect.(1, 64)

A major limitation of using a single prescription record to define exposure is that all exposed patients are considered similar, regardless of the number of prescription records or duration of use.(3) Using two or more prescription records to define exposure may mitigate this problem; however, patients with only one prescription record will be classified as non-exposed, when in fact they did have some exposure. Furthermore, this method is still dependent on when follow-up for the exposed group is started and provides similar adjusted hazard ratios, regardless of the number of prescription records used to define exposure. For example, as seen in Figure 1-A, starting the follow-up at the first oral antihyperglycemic agent prescription produced an aHR 0.64 (95% CI 0.62-0.66) when a single metformin prescription record is used to define exposure, while the aHR was 0.64 (95% CI 0.59-0.68) when using two metformin prescription records

within 180 days, and 0.62 (95% CI 0.60-0.63) when using two metformin prescription records within the entire study period.

Time-fixed analyses are particularly prone to exposure misclassification and to immortal time bias to varying degrees.(1) For example, an analysis that defines exposure based on 2 prescription records inherently assumes that the person had to be alive long enough to fill a second prescription, thus the period between the subject's cohort entry date and the 2nd prescription would be considered immortal time. The amount of immortal time bias depends on when follow-up is initiated (index oral antihyperglycemic agent prescription, and 1st or 2nd metformin prescription).(65) Therefore, the combination of using two prescription records within a defined period of time and starting follow-up on the date of the second prescription provides a further refinement to the exposure definition. However, selection bias can also be introduced when periods of immortal time are differentially excluded in a time fixed analysis. This can occur when the start of the follow-up is defined as first metformin prescription fill for exposed group and first oral antihyperglycemic agent prescription record for comparator.(38)

Defining the exposure as a continuous variable produced a relatively large risk reduction estimate associated with metformin use: 0.84, 95%CI (0.83-0.85), i.e., each year of metformin use is associated with a 16% reduction in all-cause mortality risk compared to non-users. This exposure definition may circumvent potential survival bias in observational studies and strengthen causality of the association.(21, 66) Furthermore, it provides more precise exposure definition and a more robust variable for statistical analysis, unlike commonly used dichotomous 'any versus no' exposure definitions.(3, 67) However, a continuous exposure definition does not account for gaps between refills, when the supply from one prescription record is finished well

before a subsequent prescription record. For example, the interval between first and last prescription record may be several months or years, but if these are the only two prescription records there would be a substantial unaccounted gap in exposure. Thus a time-fixed definition does not allow for variation in exposure during the follow-up period.

Overall, the estimates in the time-fixed approach were relatively consistent with previous observational studies, showing 30-45% lower risk, with the exception of the continuous variable analysis. Several solutions have been proposed to prevent immortal time bias, a well-known limitation of the time-fixed approach to defining exposure, including using time-varying approaches and nested case control analysis.(22, 38)

Time-varying Definitions

Further refinement of exposure using time-varying Cox analysis produced substantial variation in the estimates (Figure 1-E). The legacy effect analysis, whereby once a person is exposed they are considered always exposed, shifted the estimates to 0.87-0.92. This method for defining exposure is similar to the time-fixed 'ever versus never' analysis because exposure starts with the first metformin prescription. However, there was a substantial difference between the observed associations (0.73 versus 0.87-0.92). One possible explanation for this difference is the treatment of time between index oral antihyperglycemic agent and first metformin prescription. In the time-fixed 'ever versus never' definition, this interval is ignored. The advantage of a legacy-based exposure definition is that all observation time can be used in the analysis. The interval between the patient's index oral antihyperglycemic agent date and first metformin prescription contributes to the "unexposed" group in a legacy effect analysis.(64)

However, the major limitation of the legacy approach is the inability to account for future treatment discontinuation, where the patient is still erroneously characterized as exposed.(10)

An approach to address potential misclassification when treatment is discontinued is to censor patients with no evidence of ongoing therapy, i.e., stopping the follow-up in the exposed and unexposed groups after 1 window without medication use (Figure 1-B)(16, 25). This approach resulted in a substantial decrease in time at risk for cohort participants and low risk estimates ranging from 0.40 to 0.53 (Figure 1-E). Censoring based on absence of a prescription record within a defined period of time is highly dependent on when the patient obtains a refill. If there is a delay because of poor adherence or intermittent medication use, the prescription record could appear after the end of the window used to identify discontinuation of therapy. This can be especially problematic if the windows are shorter than the usual refill interval, leading to erroneous censoring. Moreover, this method for defining exposure violates the Cox proportional hazards model key assumption of non-informative censoring (i.e., participants who drop out of the study do so for reasons unrelated to the study; thus censored patients are considered to have survival probabilities similar to the participants who continued to be followed). Treatment may have been discontinued because of advancing disease (for example, switching from metformin to insulin), thus censored patients may be at a different level of risk than those who continue to be followed in the cohort.

In the time varying approach, notable differences were also observed dependent the method used to determine exposure within the window. For example, defining exposure based on any metformin prescription record within a window resulted in extremely low estimates aHR ranging from 0.23 to 0.43. As discussed with the censoring approach above, this approach is

highly dependent on the timing of a refill and can lead to misclassification, especially if the windows are shorter than the usual refill interval (Figure 1-B). Many provincial health jurisdictions in Canada allow a 100-day supply for each metformin refill. Thus, a time-varying exposure definition with short windows (e.g., 30 days) would create a majority of windows with no prescription fill *per se* and thus misclassify windows as unexposed. In contrast, long windows may introduce misclassification because only a single prescription record is required to define exposure within a window. For example, a patient may obtain only a single prescription for 30 days' supply of drug within a 365-day window; yet, they will be considered exposed for the entire 365-day window, introducing a significant amount of misclassification (335 days are truly unexposed). Windows that more appropriately reflect the utilization of the drug in the real world in terms of prescription refill frequency and days of supply may provide a better estimate of exposure over multiple windows in time-varying exposure definition.

Accounting for days of medication supply can improve exposure accuracy, since a single prescription days of supply can cover several short windows (Figure 1-B). The estimates of any metformin use based on this exposure definition produced estimates ranging from 0.64 to 0.70. An advantage over previous definitions is the ability to describe periods in the follow-up where there are gaps in medication supply and account for intermittent drug use, which is commonly seen in chronic diseases. One limitation, however, is the inability to account for poor adherence, which is quite common in chronic disease management.(68, 69) Poor adherence could extend the duration of exposure beyond the interval defined by the days of supply information in the prescription data resulting in exposed periods which may be inadvertently considered as unexposed.

In order to account for poor adherence, an additional 10% 'carry over' of medication supply has been used in the literature (i.e., a 100-day supply is assumed to cover 110 days from the prescription date). This exposure definition produced lower risk estimates relative to the definition based on days of supply alone, with estimates ranging from 0.39 to 0.62. Although this definition would account for poor adherence, it creates a differential introduction of an additional 10% of time which the patients are considered exposed compared to the unexposed group. This differential introduction would favor the exposed group and lead to larger observed protective effects.

Overall, the time-varying approach provided consistent estimates of a 30%-40% lower risk, when accounting for days of medication supply and follow-up time. These estimates did not materially change when different durations for the windows were used within the same exposure definition. However, censoring patients in the absence of prescription records within a defined period yielded highly biased results and are not advocated. The advantage of the time-varying analysis approach is the ability to obtain precise risk estimates by accounting for intermittent drug use,(22) and minimizing the influence of survival bias, immortal time bias, and confounding by duration.(12, 65)

Nested Case Control

Although the nested case control (NCC) approach has rapidly been taken up in pharmacoepidemiology research, further considerations must be accounted for in this model. In the NCC study design we followed conventional risk-set sampling methods by identifying 24,743 cases as patients who had died and matching up to 5 controls from among those with the same duration of follow-up as each case patient, but who had not died. Analyses of any

metformin prescription record prior to the event date resulted in adjusted odds ratios (aORs) ranging from 0.68 to 0.93 (Figure 1-F). As depicted in Figure 1-C, including days of supply information in the exposure definition may improve accuracy. However, similar issues as noted previously with time-varying analysis will also hold true in the NCC and resulted in aORs ranging from 0.69 to 0.96.

The NCC design provides an efficient and flexible analysis of the association between immediate exposure and outcomes; however, it does not account for long term treatment effects, which may be important in patients who use a medication for many years in the management of a chronic disease. Indeed, the NCC only uses the window immediately before the event to determine exposure status. Moreover, the NCC does not easily account for changes in patient's characteristics over time, changes in pertinent risk factors, disease severity and all of the initial treatment period is excluded from the analysis. For example, for critically ill patients, metformin treatment is often switched to insulin and therefore prior to the event the patient may be misclassified as unexposed despite previous metformin use. Patients who discontinue medication use long before the outcome may have a different risk profile than patients who continue using the medication closer to the event, which may be highly correlated with the outcome of interest, introducing a potential selection bias in the estimates.

One refinement to the NCC design that has frequently been used is to categorize exposure into current, past, or never.(33, 34) This approach attempts to capture exposure which may have occurred immediately prior to the event. Using this approach, a substantially large benefit for current use 0.48, 95% CI (0.46-0.50), and lesser benefit for past use 0.77, 95%CI (0.75-0.79) compared to never users was observed.(29, 31)

Overall, nested case control analyses provided a wide spread of estimates ranging from 0.48 to 0.96, which would produce a wide range of clinical interpretations. Our analyses indicate that the observed association between metformin and mortality risk could change materially not only between the different exposure approaches used in NCC analyses, but also between windows within the same exposure definition. Although the NCC is often preferred due to its computational efficiency over cohort studies, in today's era of 'inexpensive' computing power this relative advantage declines and time-varying approaches may be better suited to capture the full exposure profile of drugs used in the management of chronic diseases.(70)

Despite several strengths of our study, including the large population based sample, the replication of numerous exposure definitions and advanced statistical techniques used in the literature, several limitations are inherent to our work. Firstly, and most importantly, we fully acknowledge that additional unmeasured confounding can be present in our study. Our intent was not to establish whether metformin is or is not associated with mortality, but to fully explore the impact of different exposure definitions used in pharmacoepidemiology. Indeed, as all models used the same data, adjusted for the same covariates, and had the same outcomes, the degree of unmeasured confounding is expected to be similar in all the analyses. The differences in observed estimates are solely driven by exposure definitions and bias introduced by those definitions. Secondly, the administrative databases only indicate the drug was dispensed and do not indicate whether the drugs were taken as prescribed. Our assumption that metformin was used if there was a prescription record may lead to an overestimation with any exposure definition. This limitation is inherent to most observational studies using administrative drug data. Third, although in some instances the degree and deflection of bias is readily identifiable

(e.g., immortal time bias in fixed-time analyses), in many cases the substantial change in estimates is unknown.

2.4 Conclusions

In this prototypical model, the observed estimates ranged from 8% to 77% lower risk of all-cause mortality risk associated with metformin use. The differences in observed estimates were completely driven by the exposure definitions and related biases. It is interesting that all estimates based on a time-fixed approach to defining exposure were relatively consistent, showing 30-45% lower risk, while time-varying approaches and NCC definitions had considerable variation in the observed estimates.

Although unmeasured confounding and biases in pharmacoepidemiology research are well recognized and discussed, less attention has been directed towards the potential impact of exposure definitions. Most agree that time-varying or NCC analyses are preferred to time-fixed analyses, but little guidance has been given on how best to implement these approaches. Our study indicates that observed results are highly influenced by the approach used to define exposure. Given the range of observed estimates, we recommend using at least two different exposure definitions with complementary risk of bias, along with sensitivity analyses of exposure definitions to provide more robust and potentially valid study estimates.

2.5 References

1. Brunelli S. Use of prescription drug claims data to identify lipid-lowering medication exposure in pharmacoepidemiology studies: potential pitfalls. *Pharmacoepidemiol Drug Saf*. 2016;25(7):844-6.
2. Walker JJ, Johnson JA, Wild SH. Diabetes treatments and cancer risk: the importance of considering aspects of drug exposure. The lancet Diabetes & endocrinology. 2013;1(2):132-9.
3. Bannister CA, Holden SE, Jenkins-Jones S, et al. Can People With Type 2 Diabetes Live Longer Than Those Without? A Comparison of Mortality in People Initiated With Metformin or Sulphonylurea Monotherapy and Matched Non-diabetic Controls *Diabetes Obes Metab*. 2014;16(11):1165-73.
4. Gandini S, Puntoni M, Heckman-Stoddard B, et al. Metformin and cancer risk and mortality: A systematic review and meta-analysis taking into account biases and confounders. *Cancer Prev Res (Phila)*. 2014;7(9):867-85.
5. Boussageon R, Supper I, Bejan-Angoulvant T, et al. Reappraisal of Metformin Efficacy in the Treatment of Type 2 Diabetes: a Meta-Analysis of Randomized Controlled Trials. *PLoS Med*. 2012;9(4).
6. Tournier M, Bégaud B, Cougnard A, et al. Influence of the drug exposure definition on the assessment of the antipsychotic metabolic impact in patients initially treated with mood-stabilizers. *Br J Clin Pharmacol*. 2012;74(1):189-96.
7. Eurich D, McAlister F, Blackburn D, et al. Benefits and harms of antidiabetic agents in patients with diabetes and heart failure: systematic review. *BMJ*. 2007; 335(7618):497.
8. Johnson J, Majumdar S, Simpson S, et al. Decreased Mortality Associated with the Use of Metformin Compared with Sulfonilurea Monotherapy in Type 2 Diabetes. *Diabetes Care* 2002;25(12):2244-8.
9. Group UPDSU. Effect of intensive blood glucose control with metformin on complications in overweight patients with type 2 diabetes *Lancet*. 1998;352(9131):854–65.
10. Eurich D, Majumdar S, McAlister F, et al. Improved clinical outcomes associated with metformin in patients with diabetes and heart failure. *Diabetes Care*. 2005;28(10):2345-51.
11. Chaiteerakij R, Petersen G, Bamlet W, et al. Metformin Use and Survival of Patients With Pancreatic Cancer: A Cautionary Lesson. *J Clin Oncol*. 2016;34(16):1898-904.
12. Tseng C. Metformin and endometrial cancer risk in Chinese women with type 2 diabetes mellitus in Taiwan. *Gynecol Oncol*. 2015;138(1):147-53.
13. Faillie J, Azoulay L, Patenaude V, et al. Incretin Based Drugs and Risk of Accute Pancreatitis in Patients with Type 2 Diabetes: cohort Study. *BMJ*. 2014;348(g2780).
14. Lee M, Hsu C, Wahlqvist M, et al. Type 2 diabetes increases and metformin reduces total, colorectal, liver and pancreatic cancer incidences in Taiwanese: a representative population prospective cohort study of 800,000 individuals. *BMC Cancer*. 2011;11(20).
15. Ferrara A, Lewis J, Quesenberry C, et al. Cohort Study of Pioglitazone and Cancer Incidence in Patients with Diabetes. *Diabetes Care* 2011;34(4):923–9.
16. Mamtani R, Pfanzelter N, Haynes K, et al. Incidence of bladder cancer in patients with type 2 diabetes treated with metformin or sulfonylureas *Diabetes Care*. 2014;37(7):1910-7.
17. Ruiter R, Visser L, Van Herk-Sukel M, et al. Lower Risk of Cancer in Patients on Metformin in Comparison With Those on Sulfonylurea Derivatives. *Diabetes Care*. 2012;35(1):119–24.
18. Margel D, Urbach D, Lipscombe L, et al. Metformin use and all-cause and prostate cancer-specific mortality among men with diabetes *J Clin Oncol* 2013;31(25):3069-75.

19. Nayana M, Macdonaldb E, Juurlink D, et al. Medication use and survival in diabetic patients with kidney cancer: A population-based cohort study. *Pharmacol Res*. 2016;113(Pt A):468-74.
20. Calip G, Yu O, Hoskins K, et al. Associations between diabetes medication use and risk of second breast cancer events and mortality. *Cancer Causes Control*. 2015;26(8):1065–77.
21. Eurich D, Simpson S, Senthilselvan A, et al. Comparative Safety and Effectiveness of Sitagliptin in Patients with type 2 diabetes: retrospective population based cohort study. *BMJ*. 2013;346(f2267).
22. Garg R, Chen W, Pendergrass M. Acute Pancreatitis in Type 2 Diabetes Treated With Exenatide or Sitagliptin. *Diabetes Care*. 2010;33(11):2349-54.
23. Vittinghoff E, Glidden D, Shiboski S, et al. Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models. New York, NY:Springer 2012:405-6.
24. Abdelmoneim A, Eurich D, Senthilselvan A, et al. Dose-response relationship between sulfonylureas and major adverse cardiovascular events in elderly patients with type 2 diabetes. *Pharmacoepidemiol Drug Saf*. 2016;25(10):1186-95.
25. Suissa S, Dell'aniello S, Vahey S, et al. Time-window bias in case-control studies: statins and lung cancer. *Epidemiology*. 2011;22:22:228–31.
26. Simpson S, Lin M, Eurich D. Medication Adherence Affects Risk of New Diabetes Complications: A Cohort Study. *Annals of Pharmacotherapy*. 2016;5(9):741–6.
27. Bowker S, Yasui Y, Veugelers P, et al. Glucose-lowering agents and cancer mortality rates in type 2 diabetes: assessing effects of time-varying exposure. *Diabetologia*. 2010;53(8):1631–37.
28. Azoulay L, Yin H, Filion K, et al. The use of pioglitazone and the risk of bladder cancer in people with type 2 diabetes: nested case-control study. *BMJ*. 2012;344(e3645).
29. Johnson J, Simpson S, Toth E, et al. Reduced cardiovascular morbidity and mortality associated with metformin use in subjects with type 2 diabetes. *Diabet Med*. 2005;22(4):497-502.
30. Bensimon L, Yin H, Suissa S, et al. The Use of Metformin in Patients with Prostate Cancer and the Risk of Death. *Cancer Epidemiol Biomarkers Prev*. 2014;23(10):2111-8.
31. MacDonald M, Eurich D, Majumdar S, et al. Treatment of type 2 diabetes and outcomes in patients with heart failure: a nested case-control study from the U.K. General Practice Research Database. *Diabetes Care*. 2010;33(6):1213-8.
32. Azoulay L, Filion K, Platt R, et al. Association Between Incretin-Based Drugs and the Risk of Acute Pancreatitis. *JAMA Intern Med*. 2016;176(10):1464-73.
33. Eurich D, Weir D, Simpson S, et al. Risk of new-onset heart failure in patients using sitagliptin: a population-based cohort study. *Diabet Med*. 2016;33(5):621–30.
34. Lipscombe L, Gomes T, Lévesque L, et al. Thiazolidinediones and cardiovascular outcomes in older patients with diabetes. *JAMA*. 2007;298(22):2634-43.
35. Shih C, Wu Y, Chao P, et al. Association between Use of Oral Anti-Diabetic Drugs and the Risk of Sepsis: A Nested Case-Control Study. *Scientific Reports*. 2015;5.
36. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care. 2005;43(11):1130-9.
37. Van Walraven C, Austin P, Jennings A, et al. A Modification of the Elixhauser Comorbidity Into a Point System for Hospital Death Using Administrative Data. *Med Care*. 2009;47 (6):626-33.

38. Suissa S. Immortal Time Bias in Pharmacoepidemiology. *Am J Epidemiol*. 2008;167(4):492-99.
39. Daniel S, Koren G, Lunenfeld E, et al. Immortal time bias in drug safety cohort studies: spontaneous abortion following nonsteroidal antiinflammatory drug exposure *Am J Obstet Gynecol*. 2015;212(3):307.e1-6.
40. Matok I, Azoulay L, Yin H, et al. Immortal time bias in observational studies of drug effects in pregnancy. *Birth Defects Res A Clin Mol Teratol*. 2014;100(9):658-62.
41. Lévesque L, Hanley J, Kezouh A, et al. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ*. 2010;340.
42. Hill A. The environment and disease: Association or causation? . *Proc R Soc Med*. 1965;58:295-300.
43. Royston P, Altman D, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med*. 2006;25(1):127–41.
44. Cramer J. A Systematic Review of Adherence With Medications for Diabetes. *Diabetes Care* 2004;27(5):1218-24.
45. Krass I, Schieback P, Dhippayom T. Adherence to diabetes medication: a systematic review. *Diabet Med*. 2015;32(6):725–37.
46. Austin P, Anderson G, Cigsar C, et al. Comparing the cohort design and the nested case–control design in the presence of both time-invariant and time-dependent treatment and competing risks: bias and precision. *Pharmacoepidemiol Drug Saf*. 2012;21(7):714-24.

Table 1-3: Patient Characteristics Incident Oral Hypoglycemic Agents Users Cohort, Alberta, Canada, 1998-2010

	>1 Metformin Prescription (n=55,525)	No Metformin Prescriptions (n=8,768)	p-value ₸
Age, mean (SD), yr	69 (±4)	71 (±6)	< 0.001
Men, (%)	28,547 (51.4%)	4,584 (52.3%)	< 0.131
Elixhauser Comorbidity Score, mean (SD)	5 (±7)	7 (±8)	< 0.001
Follow-up, mean (SD), yr	6 (±4)	5 (±4)	< 0.001
Death, no. (%)	19,636 (35.4%)	5,109 (58.3%)	< 0.001

SD = standard deviation, yr - year

₸ p-value is for difference in characteristics between groups, when metformin users and non-users within a treatment cohort are compared by analysis of variance or by chi^2 test.

Figure 1-A: Time-fixed Exposure Definitions Schematic Representation

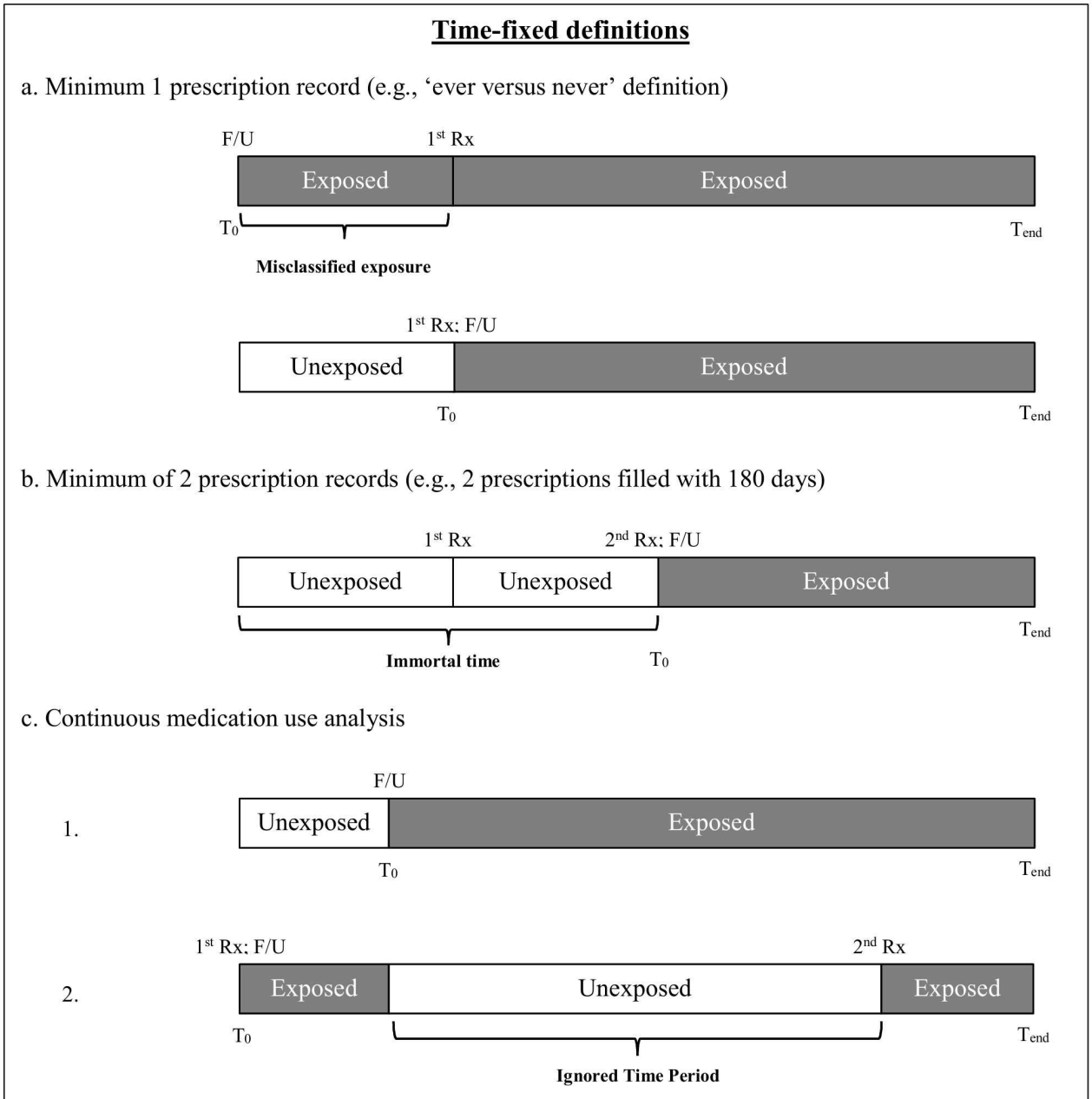


Fig.1 – A: Schema depicting analysis methods and exposure definitions: **'a'** depicts 'ever versus never' exposure definition in time-fixed analysis, when the follow-up starts from the index day (i.e., first ever prescribed oral hypoglycemic agent (OHA)). However, such analysis includes misclassification of unexposed time between first OHA and first metformin prescription, which should be attributed to an unexposed group; **'b'** exposure is defined by 2 consecutive prescription fills within 180 days allows avoidance of immortal time bias where the patient to be exposed has to survive a certain time period at the beginning; **'c'** continuous exposure analysis accounts for a cumulative medication exposure during all follow-up time (e.g., an ideal case **'c1'**) , however it does address possible gaps between prescription fills, non-adherence and a large follow-up period is ignored (e.g., case **'c2'**)

T_0 - cohort start; F/U - follow-up start; Rx – prescription; T_{end} – patient follow-up ends, censored, event occurs.

Figure 1-B: Time-varying Exposure Definitions Schematic Representation

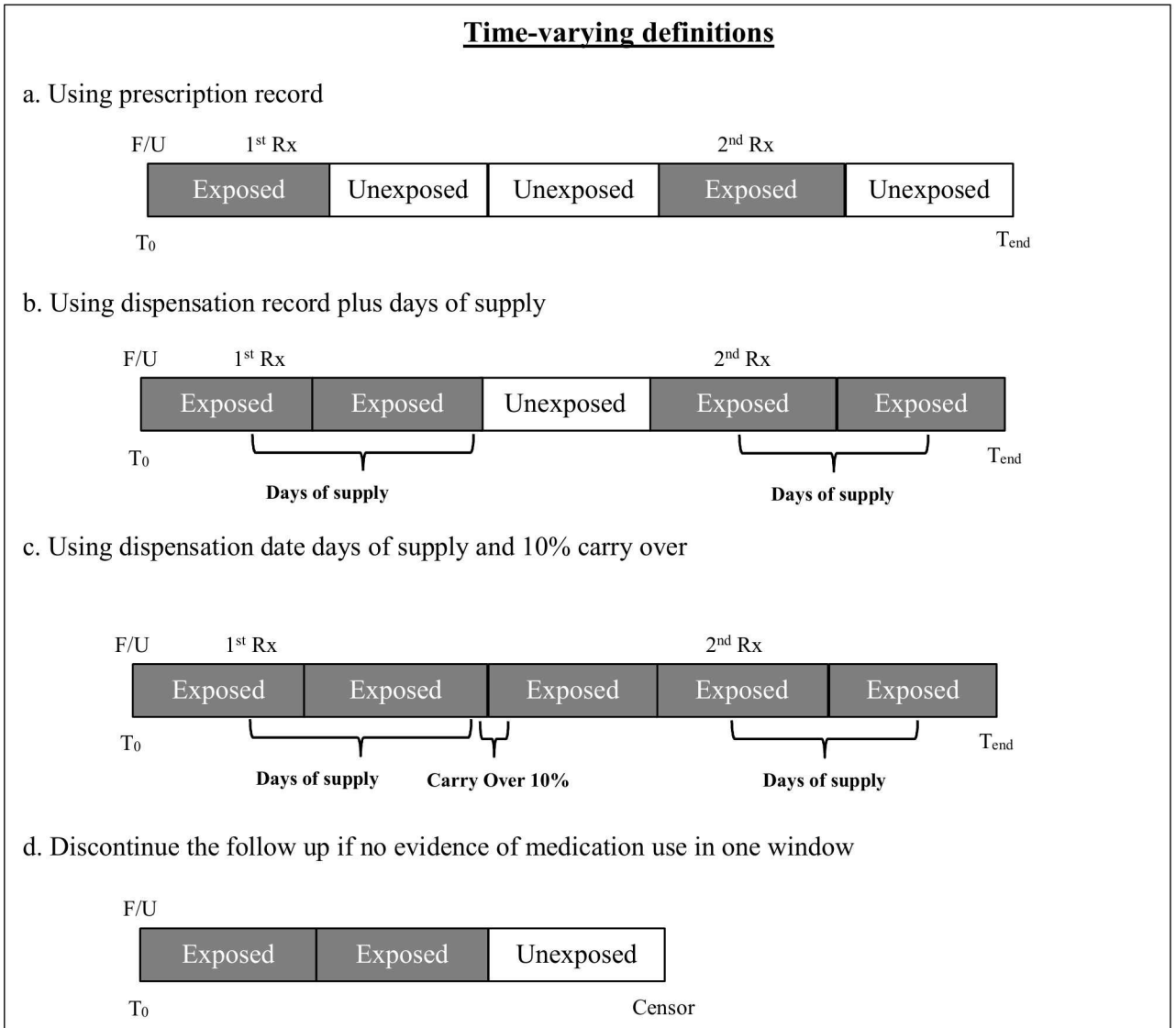


Fig.1- B: Schema depicting analysis methods and exposure definitions in time-varying analysis, which represents follow-up time divided to regular windows with dichotomous exposure defined: **'a'** depicts exposure based on prescription record within windows; **'b'** adding days of supply to exposure definition improves exposure accuracy, when one prescription fill supplies the medication for several windows; while in **'c'** adding 10% of medication supply carry over to the consecutive time window based on previous window estimated supply to correct for non-adherence; **'d'** discontinuation of follow-up if no evidence of medication use in one time window does not account for non-adherence and gaps in medication supply and may selectively censor the patients from the cohort despite possible future treatment continuation

T_0 - cohort start; F/U - follow-up start; Rx – prescription; T_{end} – patient follow-up ends, censored, event occurs.

Figure 1-C: Nested Case Control Exposure Definitions Schematic Representation

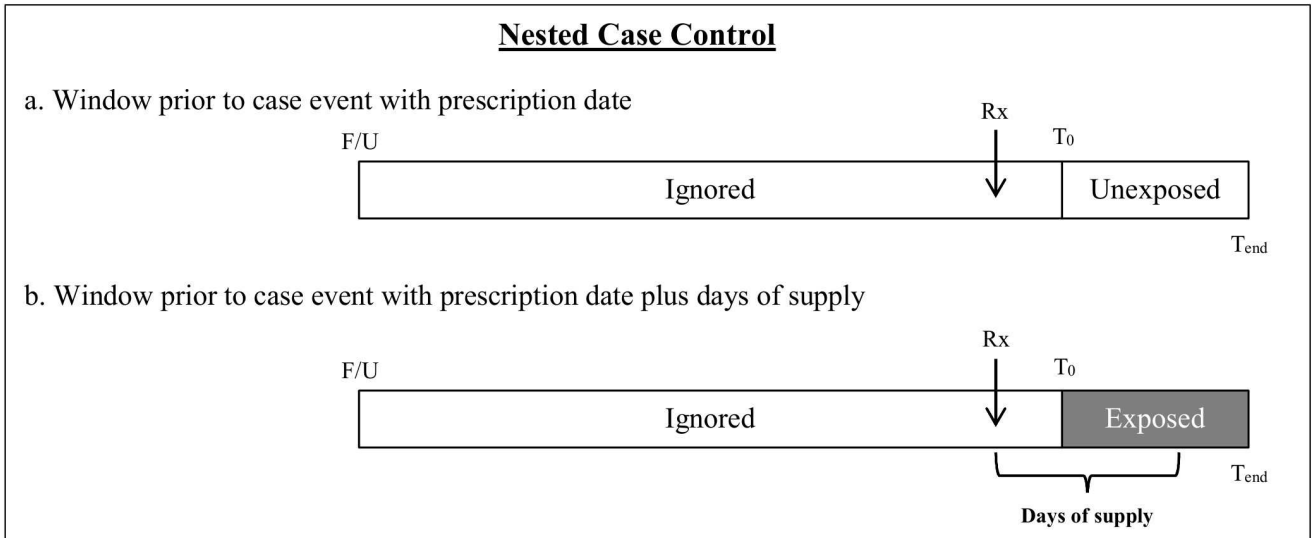


Fig.1-C: Schema depicting exposure definition in nested case control based on prescription fill dates in time window prior to event date; **'a'** depicts exposure status definition using prescription date before time window may mistakenly define a patient as unexposed; while **'b'** is accounting for medications supply and correct for this bias. Furthermore, all time period between follow-up start and beginning of time window prior the event is ignored.

T_0 - cohort start; F/U - follow-up start; Rx – prescription; T_{end} – patient follow-up ends, censored, event occurs.

Figure 1-D: Association of Metformin on All-Cause Mortality using Time-Fixed Exposure Definitions

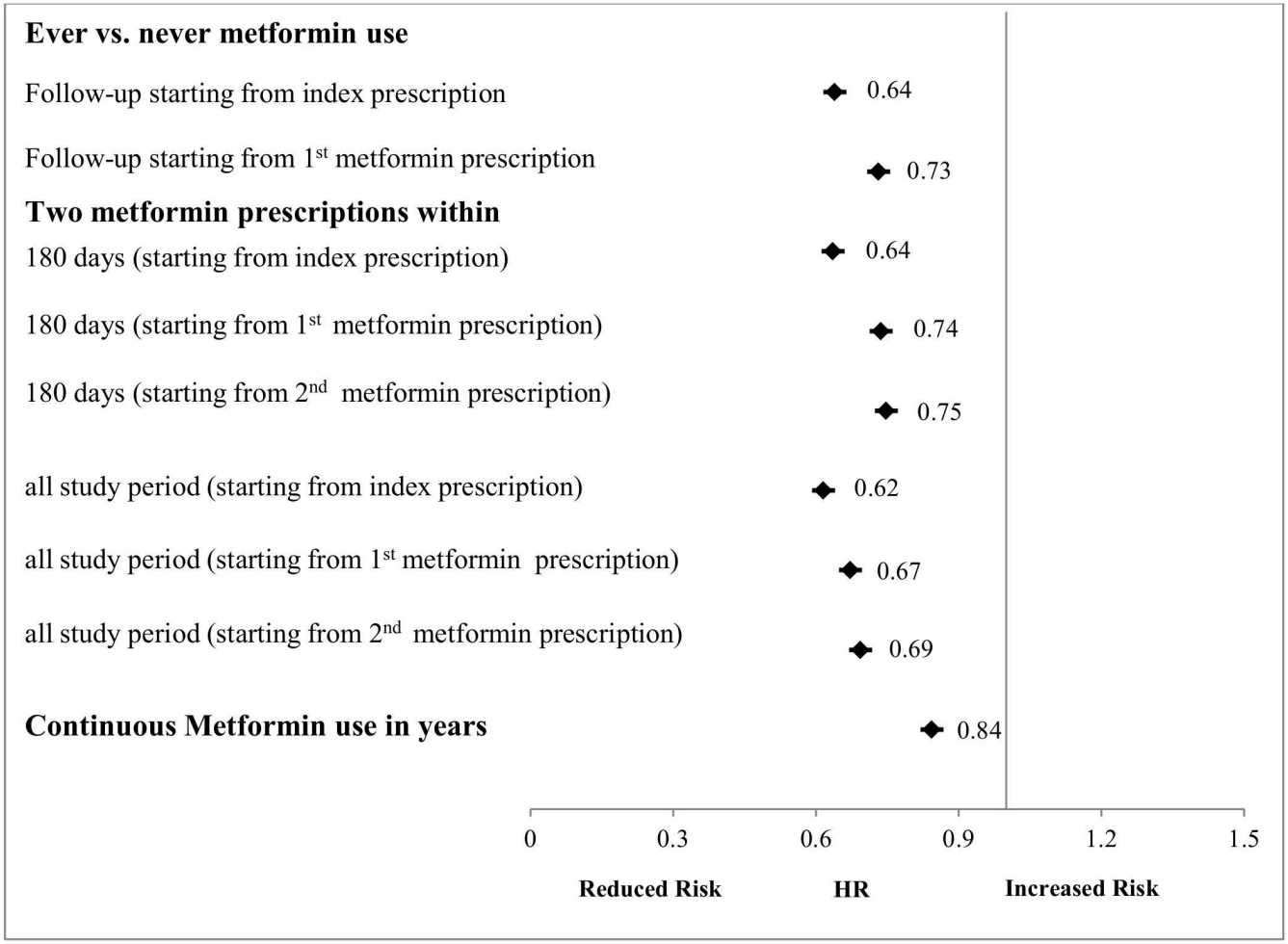


Figure 1-E: Association of Metformin on All-Cause Mortality using Time-Varying Exposure Definitions

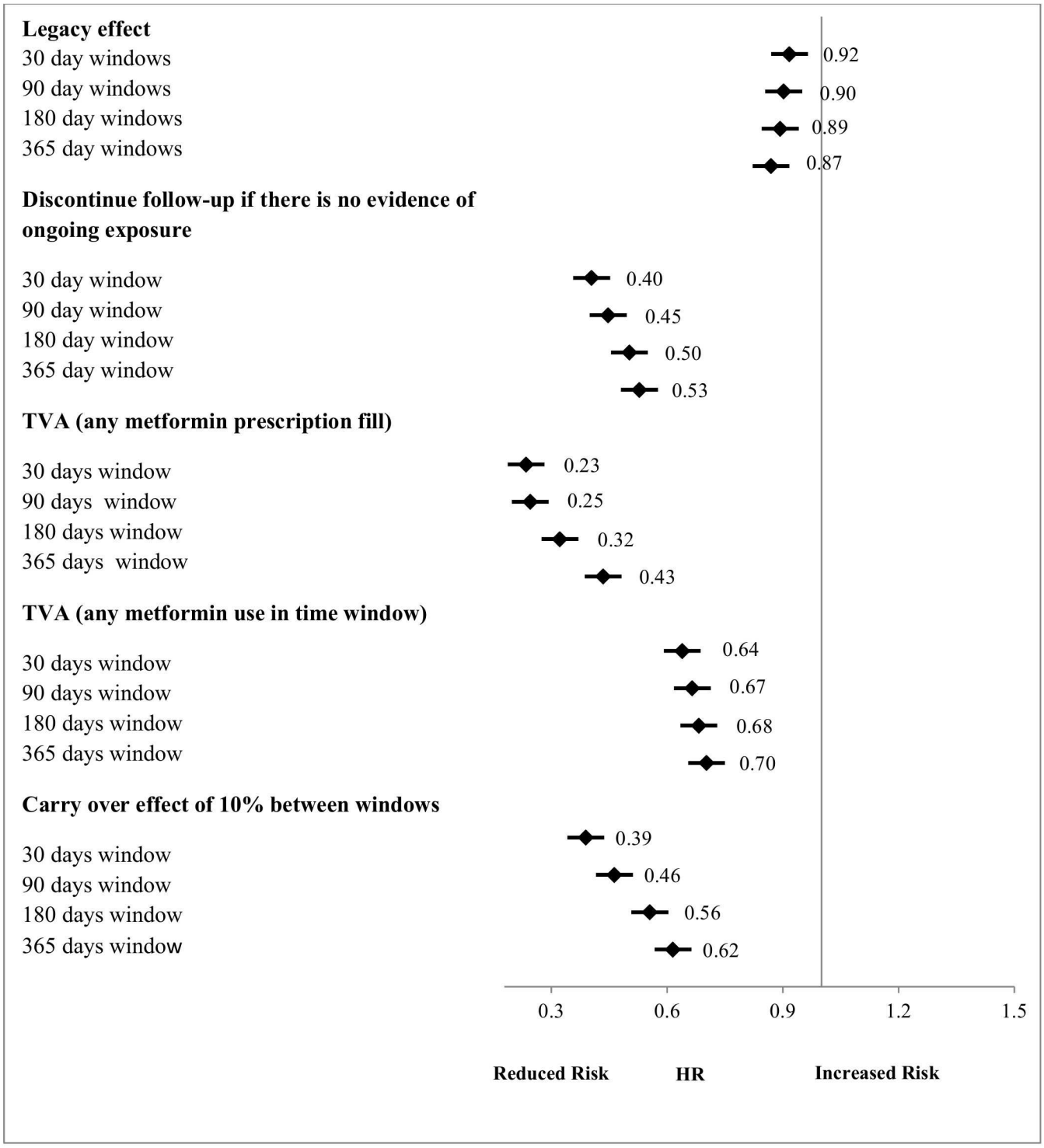
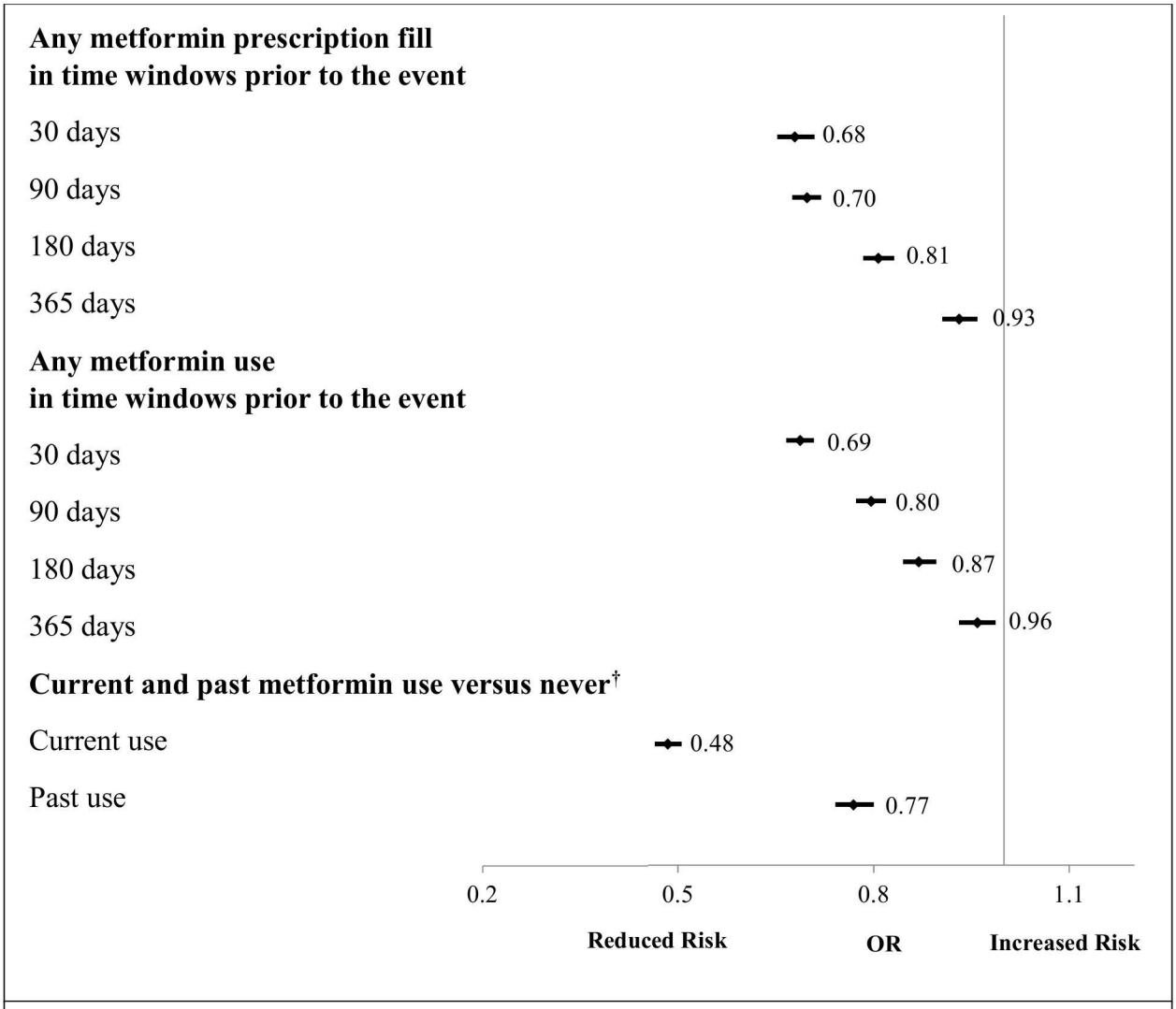


Figure 1-F: Association of Metformin on All-Cause Mortality using Nested Case Control Exposure Definitions



† Current (prescription date plus days of supply overlap the case event date), past (prescription date pus days of supply end before the case event date), or never (no evidence of metformin use 2 years prior to the outcome

CHAPTER 3:

Evaluating Healthy User Effect in Type 2 Diabetes Patients Using Oral Antihyperglycemic Agents

3.1 Introduction

Observational studies are commonly used in modern pharmacoepidemiology to evaluate drug safety and effectiveness. They have become increasingly popular and make an important contribution to medical knowledge, especially when 'gold standard' randomized controlled trials (RCT) are not always feasible. However, a major limitation of observational studies is their subjectivity to a number of potential biases that may cause erroneous results.

One source of bias has been termed the 'Healthy User Effect' or 'Healthy User Bias', which is a tendency of healthier patients initiate preventive therapies and engage in behaviors consistent with healthy lifestyles, which in turn can affect their outcomes.(6, 71) Aspects of a healthy lifestyle include balanced diet, regular exercise, moderation of alcohol consumption and tobacco use, avoidance of risky behaviors, participation in vaccination programs, and using preventive therapies and health services.(71) Many of these healthy lifestyle behaviors are not recorded in typical administrative databases and failure to account for these factors in observational studies of preventive therapies could introduce bias and produce spurious associations.(6, 7)

The healthy user effect has gained increasing attention as a potential source of bias in observational studies of preventive therapies and treatment of asymptomatic conditions. (42)

Indeed, numerous studies have demonstrated the healthy user effect can be a plausible explanation for the discrepancy between RCTs and observational studies.(6, 7, 44) For example, hormone replacement therapies(72) and vitamin E studies among others.(73) Healthy user bias has been most notably explored in users of 'statin' therapies.(41) Indeed, statin therapy is associated with increased use of a number of other preventative therapies and health services, and a reduction in clinical events which are not intuitively associated with statin use, including reduced risk of cancer, sepsis, Alzheimer disease and hip fractures.(6, 7, 45) Further, healthy user bias is thought to fully explain the large benefits observed in observational studies of hormone replacement therapies and the subsequent RCTs which showed significant harms.(44) It has also been extensively discussed as a potential source of bias in observational studies showing large reductions in all-cause mortality associated with influenza vaccination in elderly.(43, 46)

Although less studied, healthy user bias may also explain the discrepancies in results between observational studies and RCTs of metformin use in patients with type 2 diabetes. Meta-analyses of RCT's have failed to replicate the strong association between metformin use and all-cause or CV mortality reported in observational studies.(2, 56) Importantly, the vast majority of these studies have not been designed to look at mortality associated with metformin per se. As a result, a large body of observational studies have further explored the issue and have consistently shown clinically important benefits, including substantial reductions in all-cause and CV-related mortality, as well as beneficial effects in patients with diabetes and heart failure.(56-58)

The underlying reasons for the substantial discrepancy between experimental and observational studies in diabetes is unknown, however, we surmise that unaccounted healthy user bias may explain part, or fully, the difference between the study designs.

Therefore, our primary objective in this study was to gather evidence of healthy user bias among metformin users in a large unselected population of patients with type 2 diabetes. We hypothesized that type 2 diabetes patients who use metformin, as compared to those not using metformin, would be at lower risk of events which should not be associated with metformin (e.g., accident events, certain diseases, etc.) and conversely would be more likely to participate in healthier behaviors (e.g., using preventative services, vaccines, and diagnostic tests).

3.2 Methods

We conducted our analysis in a large population-based cohort of new users of oral antihyperglycemic agents in Alberta, Canada based on administrative data augmented with clinical laboratory data between 2008 and 2015 provided by Alberta Health. Alberta Health collects all medical information as part of the universal health care provided to patients within the province (~3.5 million). The database included laboratory tests results, Pharmaceutical Information Network (PIN), Practitioner Payments (coded with the International Classification of Diseases [ICD9] for medical conditions and with Alberta Medical Association services billing codes for medical services), Ambulatory Care and Inpatient information (coded with the ICD10 for medical conditions). Demographics and mortality data were collected from Population Registry and Vital Statistic databases. The study was approved by the ethics review board of the University of Alberta, Edmonton, Alberta, Canada.

Cohort Selection

Patients were eligible for inclusion if they were at least 30 years old on the index date (date of first dispensing record for an oral antihyperglycemic medication).(74) We identified new oral antihyperglycemic medication users between April 1, 2009 and March 31, 2015. A 1-year washout period was used to define new users by ensuring that there were no prior dispensing records for antihyperglycemic medications, including insulin, before the index date. Women using metformin as a single oral antihyperglycemic agent for poly cystic ovary syndrome were excluded from the analysis based on ICD coding.

Outcomes

We evaluated a variety of clinical events and health behaviors that have been postulated in the literature to be associated with healthy user bias and that could be identified in the administrative databases (Appendix).(6, 7, 41, 42) Clinical events were identified using ICD9 and ICD10 codes in hospitalization records, emergency department visit records, and physician visit records. Receipt of screening tests and vaccines were identified from provincial laboratory records and procedure codes in physician visit records.

Clinical events and preventive services were selected a priori using information from previous studies and Canadian Diabetes Association (CDA) recommendations,(75) and were grouped into 5 broad categories: accident events, screening events, vaccination, other events not expected to be associated with metformin exposure, and other events for which a possible association with metformin exposure could be expected.(6, 7, 42) Additionally, the occurrence of annual laboratory screening for hemoglobin A1C, lipids and kidney function were evaluated.(76, 77)

Exposure

Patients were considered as metformin users if they had at least 1 dispensation record for metformin during the follow-up period. Metformin exposure was modeled as time-dependent variable by considering the patient unexposed until the date of the first metformin dispensation. After this date, the patient was coded as a metformin user until the end of the follow-up.(14) metformin was modeled in this manner as we were not necessarily interested in the true underlying effect of metformin on the outcomes per se, but were interested in metformin as a marker of healthy user bias.

Covariates

We obtained baseline demographics, physician visits, health services utilization, hospitalizations and medication management during the year prior to the initiation of first oral antihyperglycemic agent use. The covariates were defined at the index date and included age, sex, and other antihyperglycemic medications use. As most observational studies include comorbidity scores to control for confounding, we also included the Elixhauser comorbidity score based on chronic conditions identified in the year prior to the first oral antihyperglycemic agent use.(78)

Statistical Analyses

For the majority of outcomes, the associations between metformin exposure and outcomes were examined using a multivariable Cox Proportional Hazard model. For these analyses, the time to each outcome was determined, with subjects censored by death (for non-mortality endpoints), or departure from the provincial benefits program (moved out of the

province) or reaching end of our observation period, March 31, 2015. Crude and adjusted Hazard Ratios (HR), along with the 95% confidence interval (95% CI) were calculated for each outcome. In addition, we also completed Poisson regression in the evaluation of hemoglobin A1C, lipid, and kidney function tests, which may be expected to be repeated several times a year. In these analyses, the adjusted incidence rate ratio (aIRR – number of events observed divided by the time at risk of event during the observation period for metformin compared to non-metformin users) was evaluated using Poisson regression. All analyses were conducted using STATA, version 13.1, statistical software (StataCorp LLC, College Station, Texas, USA).

3.3 Results

Our cohort consisted of 135,301 new oral antihyperglycemic agent users. The average age was 55 years, 75,949 (56%) were male, 97% of the cohort (130,725 patients) filled at least one metformin prescription and after an average follow-up of 3.4 (SD±2) years, 6,677 (5%) cohort patients died. Compared to non-users, metformin users tended to be younger (55 vs. 60 years), and had a lower comorbidity index (0.6 vs. 1.4) (Table 2-1). During the year prior to the index date metformin users had fewer physician visits than non-users (16 vs. 24), had fewer ambulatory care service visits (5 vs.11), but had a similar number of hospitalizations (1.5 and. 1.7).

In time-dependent Cox proportional multivariable models we found statistically significant lower risks for accident events aHR 0.84 (95% CI 0.79-0.89), and other clinical outcomes without an expected association with metformin aHR 0.90 (95% CI 0.84-0.97), including all-cause mortality aHR 0.54 (95% CI 0.50-0.59) for metformin users compared to

non-users (Table 2-2). Furthermore, metformin users were more likely to have diagnostic screening tests and procedures applicable for both sexes aHR 1.17 (95% CI 1.12-1.22) as well as those specific for women aHR 1.13 (95% CI 1.08-1.19) and for men aHR 1.64 (95% CI 1.53-1.75). However, clinical events possibly associated with metformin use showed no association aHR 1.01 (95% CI 0.96-1.06) and the association between metformin use and *vaccination* rates was not statistically significant aHR 1.10 (95% CI 0.98-1.23). The likelihood of initiating preventive therapies recommended by the CDA (e.g., angiotensin-converting-enzyme (ACE) inhibitors and statins)(76) after diabetes diagnosis did not differ between the groups aHR 0.98 (95% CI 0.94-1.02). However, in Poisson models hemoglobin A1C and lipids blood tests were significantly more frequent in metformin users compared to non-users aIRR 1.10 (95% CI 1.09-1.11) and 1.09 (95% CI 1.07-1.10), respectively.

Moreover, among the diagnostic tests mammography and prostate-specific antigen (PSA) test were the most commonly used by the patients with 74.46 and 17.48 cases per 100 person years, respectively, while sigmoidoscopy and fecal occult blood (FOB) test were the least common with only 0.55 and 0.80 cases per 100 person years, respectively (Table 2-2).

The trend in the analyses tended towards decreased risk for clinical events possibly associated with unhealthy lifestyle among metformin users. The exception was drug dependency related events with aHR 1.18 (95% CI 1.07-1.31) and myocardial infarctions aHR 1.11 (95% CI 0.93-1.34) conversely showing increased risk for metformin users. Among 36 clinical events, diagnostic services and tests that we analyzed, 15 showed a statistically significantly lower risk for metformin users, 14 showed non-significant trend towards lower risk, 4 estimates were not

associated with the events and 3 showed an opposite association with general risk reduction trend.

3.4 Discussion

In our cohort of new oral antihyperglycemic agent users, metformin users were less likely to experience accident events or suffer from clinical outcomes that are not expected to be associated with metformin exposure, and were more likely to use preventative services and screening, and diagnostic tests. Collectively, these data suggest metformin may be a strong marker of a 'healthy user'. These findings have important implications for observational studies involving oral antihyperglycemic agent exposure where metformin often constitutes a high proportion of use and is often a user group of interest in the study.

Numerous observational studies in diabetes have consistently shown metformin to be associated with ~30% reduction in all cause or cardiovascular specific mortality compared to other oral antihyperglycemic agents.(57, 74) Yet, only one small, sub-study within the UKPDS randomized controlled trial has suggested a similar benefit in obese patients with type 2 diabetes.(56) Indeed, a recent meta-analysis of 13 RCTs found no benefit of metformin per se relative to other treatments.(2)

In our study, although metformin use was associated with 45% lower risk of all-cause mortality compared to non-users and consistent with previously done observational studies(57), the collective data would strongly suggest this estimate is heavily influenced due to healthy user bias associated with metformin use. Moreover, our data also adds credence to the hypothesis that

major discrepancies observed between observational studies of metformin and meta-analyses of RCTs may in fact be related to difficult to control health user bias.

Although the observed association between metformin and clinical events is likely attributed to unmeasured confounding by health seeking behaviors, other explanations are also possible for these findings. For example, the association between metformin and reduced risk of fracture could be due to differences in actual health status and particularly due to body mass index (BMI), where metformin tends to be used in more obese patients who are at lower risk of osteoporotic fractures but is rarely included in observational studies; falls, an outcome that was previously used in healthy user bias studies with statins(7) can be associated with other medications side effects (syncope from blood pressure medications) or hypoglycemic events with non-metformin therapies; and malignant melanoma is subject to screening bias, such that patients who are more engaged in their own health are more likely to have their lesions detected.(42) Therefore, although these outcome events might be less appropriate markers for healthy behaviors and not attributable to healthy user bias per se, they are nevertheless major potential source of bias which maybe differential between metformin and non-metformin users.

Future observational studies of oral antihyperglycemic agent need to better control for healthy user bias, particularly where metformin is concerned. Beyond traditional methods new user designs(45, 47), using an active comparator,(9, 48, 49), and time-dependent exposures(48), researchers should consider controlling for prior use of preventive tests and utilizing, propensity score approaches that include markers of the healthy user intro the prediction of drug exposure.(50) Additional methods such as instrumental variable utilization(42, 47) have also

been put forth, however, the identification of an appropriate instrumental variable for many observational studies is difficult.

Despite several strengths of our study, including the large population based sample, and rich clinical data, several limitations are inherent to our work. First, and most importantly, we fully acknowledge that additional unmeasured confounding can be present in our study, e.g., Body Mass Index, tobacco use, alcohol consumption and socio-economic status were not available in our database. Second, our data did not allow us to evaluate the patients who were diagnosed with diabetes, but were prescribed with lifestyle modifications by the family physician or patients, who were prescribed with an oral antihyperglycemic agent, but were non-compliant and never filled the prescription at the pharmacy. Last, we acknowledge that ICD coding may not be sensitive or specific for all outcomes evaluated within the administrative data; however, there is no reason to believe the coding would be differential based on metformin exposure status and therefore the results should not be biased.

This study suggests that metformin users take better care of themselves by engaging in various healthy behaviors and initiating preventive therapies. Therefore, a failure to account for these behaviors can introduce bias to observational studies evaluating the effect of oral anti hyperglycemic therapies on health outcomes. Moreover, interpretation of observational studies that attribute surprisingly protective effects to antihyperglycemic therapies requires cautious cautions. Further work is required to gain a better understanding of the heathy user effect in diabetic patients and to develop methods for guarding against this effect in observational studies.

Table 2-1: Oral Anti Hyperglycemic Agents Incident Users Cohort Characteristics, Alberta, Canada, 2008-2015

	>=1 Metformin Prescription	**No Metformin Prescriptions**	**p-value**[Ŧ]
No.	130,725	4,576	
Age, yr, mean (SD)	55 (±13)	60 (±16)	< 0.001
Men, (%)	73,680 (56%)	2,269 (50%)	< 0.001
Elixhauser Comorbidity Score, mean (SD)	0.6 (±2.5)	1.4 (±5)	< 0.001
Death, no. (%)	6,025 (4.6%)	652 (14.3%)	< 0.001
Follow-up[¶], yr, mean (SD)	3.4 (±2)	3.1 (±2.1)	< 0.022
No. of physician visits in prior year, mean (SD)	16 (±24)	24 (±40)	< 0.001
No. of hospitalizations in prior year, mean (SD)	1.4 (±0.9)	1.7(±1.2)	< 0.001
No. of ambulatory care services in prior year, mean (SD)	5(±11)	12(±31)	< 0.001

Abbreviations: SD = standard deviation, No. – number, yr – years.

¶ Follow-up time from first Oral Anti Hyperglycemic use till death, censoring or study end.

Ŧ p-value is for difference in characteristics between groups, when metformin users and non-users within a treatment cohort are compared by analysis of variance or by chi^2 test.

Table 2-2. Association Between Metformin Therapy and Risk of Health-Related Events, Incident Oral Anti Hyperglycemic Agents Users, Alberta, Canada, 2008-2015

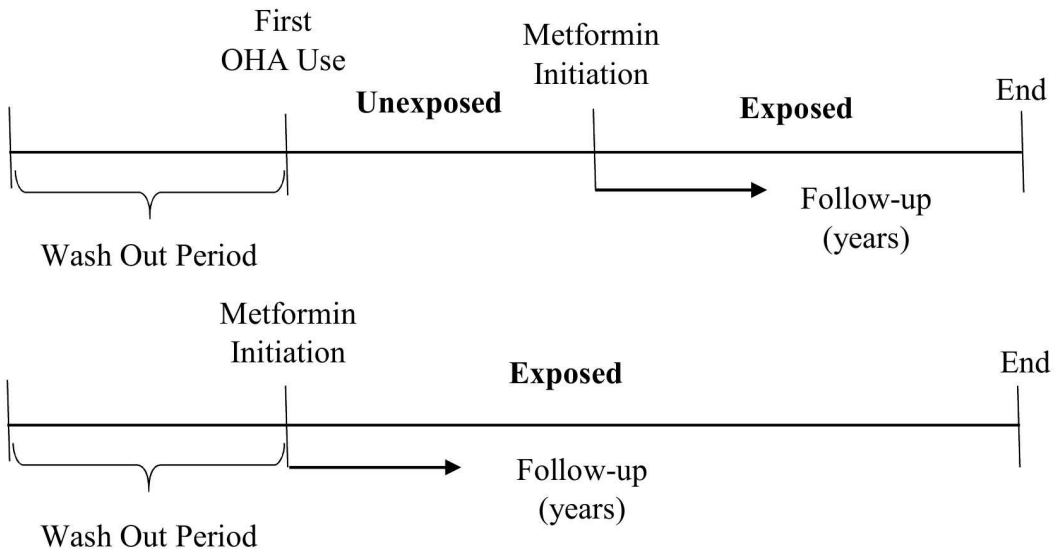
Metformin Use and Health Outcomes Association Based on ICD 10 Codes from Hospitalization Data

Outcome	Event Rate, /100 Person-Years	Time-varying Exposure Definition Model	
		Unadjusted	Adjusted
Accident Events			
Burns	0.37	0.82 (0.66-1.02)	0.90 (0.73-1.12)
Falls	0.04	0.56 (0.33-0.96)	0.77 (0.44-1.35)
Fractures	0.73	0.49 (0.44-0.55)	**0.78 (0.69-0.89)**
Motor vehicle accidents	0.18	0.83 (0.63-1.09)	0.77 (0.58-1.03)
Open wound	2.26	0.80 (0.74-0.88)	**0.88 (0.80-0.96)**
Poisoning	0.73	0.81 (0.69-0.93)	0.90 (0.78-1.05)
All (first occurrence)	*5.82*	*0.96 (0.91-1.02)*	***0.84 (0.79-0.89)***
Screening Events			
Both Sexes (n=135,301)			
Eye examinations	10.95	0.92 (0.88-0.96)	**1.14 (1.09-1.19)**
Fecal occult blood test	0.80	1.29 (1.09-1.53)	**1.60 (1.65-1.91)**
Sigmoidoscopy	0.55	0.77 (0.65-0.91)	0.87 (0.73-1.03)
All (first occurrence)	*12.4*	*0.94 (0.90-0.98)*	***1.17 (1.12-1.22)***
Women (n=59,352)			
Papanicolaou test	13.14	1.38 (1.29-1.48)	**1.11 (1.03-1.19)**
Mammography	74.46	1.24 (1.17-1.31)	**1.21 (1.14-1.29)**
Bone mineral density test	9.49	1.01 (0.95-1.09)	**1.17 (1.09-1.26)**
All (first occurrence)	*33.22*	*1.26 (1.20-1.32)*	*1.13 (1.08-1.19)*
Men (n=75,949)			
Prostate-specific antigen test	17.48	1.87 (1.75-1.99)	**1.64 (1.53-1.75)**
Laboratory Tests			
Hemoglobin A1C	39.63	0.98 (0.96-0.99)	**1.10 (1.09-1.12)**
Creatinine	39.83	0.74 (0.73-0.75)	0.95 (0.94-0.96)
Lipids	43.76	0.97 (0.81-1.16)	1.02 (0.97-1.06)
All	*33.89*	*0.80 (0.80-0.81)*	*0.97 (0.96-0.98)*
Preventative Therapies			

Vaccination			
Pneumonia	0.11	1.50 (0.98-2.28)	1.22 (0.79-1.88)
Influenza	2.04	1.35 (1.20-1.51)	1.10 (0.98-1.24)
All (first occurrence)	*2.14*	*1.32 (1.18-1.48)*	*1.10 (0.98-1.23)*
Preventative pharmacotherapies			
Statin use	9.76	0.75 (0.71-0.78)	1.00 (0.95-1.04)
ACE inhibitors use	6.07	0.71 (0.67-0.75)	0.89 (0.84-0.94)
All (first occurrence)	*13.85*	*0.73 (0.71-0.76)*	*0.98 (0.94-1.02)*
Other Events, Possible association expected			
Ambulatory Services Use	42.41	0.89 (0.85-0.94)	0.96 (0.92 -1.01)
Emergency Department admissions	14.38	0.69 (0.65-0.72)	**0.92 (0.87-0.97)**
Lung cancer	0.33	0.75 (0.61-0.92)	1.01 (0.81-1.25)
Mortality (all-cause)	1.49	0.30 (0.28-0.32)	**0.54 (0.50-0.59)**
Myocardial infarction	0.44	0.69 (0.58-0.82)	1.11 (0.93-1.34)
All (first occurrence)	*42.56*	*0.89 (0.85-0.93)*	*1.01 (0.96-1.06)*
Other Events, No association expected			
Asthma/COPD	5.38	0.82 (0.78-0.87)	**0.91 (0.86-0.97)**
Bacterial Infection	0.62	0.51 (0.45-0.58)	**0.85 (0.74-0.98)**
DVT and PE	0.84	0.69 (0.61-0.79)	0.88 (0.76-1.00)
Dental problems	2.1	0.91 (0.83-0.99)	0.95 (0.86-1.04)
Diverticulitis	0.48	0.76 (0.63-0.91)	0.86 (0.72-1.04)
Drug dependency	2.56	1.36 (1.23-1.49)	1.18 (1.07-1.31)
Food-borne bacterial infection	1.68	0.78 (0.71-0.86)	0.94 (0.86-1.04)
Gallstones	0.88	0.87 (0.76-0.99)	0.92 (0.80-1.06)
Gastrointestinal bleeding	0.40	0.48 (0.41-0.56)	**0.81 (0.69-0.96)**
Gout	0.18	0.72 (0.65-0.78)	**0.72 (0.65-0.79)**
Kidney stones	0.73	1.05 (0.90-1.22)	1.01 (0.86-1.19)
Migraine	0.91	0.96 (0.83-1.10)	0.87 (0.76-1.00)
Skin infection	5.11	0.82 (0.77-0.87)	**0.91 (0.85-0.96)**
Sexually transmitted diseases	351	0.92 (0.84-1.01)	0.99 (0.90-1.09)
All (first occurrence)	*25.14*	*0.95 (0.88-1.02)*	***0.90 (0.84-0.97)***

Abbreviations**:** COPD – Chronic Obstructive Pulmonary Disease, DVT – Deep Vein Thrombosis, PE – Pulmonary Embolism

Figure 2-A: Time-Dependent Exposure Definition



Schema depicting the methods used to define time-dependent exposure: The index date is assigned to the date of first metformin or OHA prescription dispensation after 1 year wash out period. The patient is considered exposed only after a receipt of first metformin prescription and remains exposed till the follow-up end, i.e., diagnostic test date, hospital admission with desired outcome diagnosis, death date, censoring from the study or reaching the study end point March 31, 2015.
Abbreviation: OHA – oral anti hyper glycemic agent

3.5 References

1. Eurich D, Majumdar S. Statins and Sepsis – Scientifically Interesting but Clinically Inconsequential *J Gen Intern Med* 2011;27(3):268-9.
2. Dormuth C, Patrick A, Shrank W, et al. Statin Adherence and Risk of Accidents A Cautionary Tale. *Circulation*. 2009;119(15):2051-7.
3. Brookhart M, Patrick A, Dormuth C, et al. Adherence to Lipid-lowering Therapy and the Use of Preventive Health Services: An Investigation of the Healthy User Effect. *American Journal of Epidemiology* 2006;166(3):348-54.
4. Patrick A, Shrank W, Glynn R, et al. The Association between Statin use and Outcomes Potentially Attributable to an Unhealthy Lifestyle in Older Adults. *Value Health*. 2011;14(4):513-20.
5. Ray W. Observational studies of drugs and mortality. *N Engl J Med*. 2005;353(22):2319-21.
6. Garbe E, Suissa S. Hormone replacement therapy and acute coronary outcomes: methodological issues between randomized and observational studies. *Hum Reprod*. 2004;19(1):8-13.
7. Lee I, Cook N, Gaziano J, et al. Vitamin E in the primary prevention of cardiovascular disease and cancer: the Women's Health Study: a randomized controlled trial. *JAMA*. 2005;294(1):56-65.
8. Eurich D, Majumdar S. Statins and sepsis - scientifically interesting but clinically inconsequential. *J Gen Intern Med*. 2012;27(3):268-9.
9. Brassard P, Wu J, Ernst P, Dell'Aniello S, Smiechowski B, Suissa S. The effect of statins on influenza-like illness morbidity and mortality. *Pharmacoepidemiology and Drug Safety* 2016.
10. Jackson L, Jackson M, Nelson J, Neuzil K, Weiss N. Evidence of bias in estimates of influenza vaccine effectiveness in seniors. International Journal of Epidemiology. 2006;35:337-44.
11. Eurich D, Marrie T, Johnstone J, Majumdar S. Mortality reduction with influenza vaccine in patients with pneumonia outside "flu" season: pleiotropic benefits or residual confounding? *Am J Respir Crit Care Med*. 2008;178(5):527-33.
12. Boussageon R, Supper I, Bejan-Angoulvant T, et al. Reappraisal of Metformin Efficacy in the Treatment of Type 2 Diabetes: a Meta-Analysis of Randomized Controlled Trials. *PLoS Medicine*. 2012;9(4):e1001204.
13. UKPDS. Effect of intensive blood glucose control with metformin on complications in overweight patients with type 2 diabetes *The Lancet*. 1998;352(9131):854–65.
14. Eurich D, McAlister F, Blackburn D, et al. Benefits and harms of antidiabetic agents in patients with diabetes and heart failure: systematic review. *BMJ*. 2007;335(7618):497.
15. Jasim S, Smith S. Review: Metformin is linked to reduced mortality in type 2 diabetes with comorbid CKD and CHF. *Ann Intern Med*. 2017;166(8).
16. Johnson J, Majumdar S, Simpson S, Toth E. Decreased mortality associated with the use of metformin compared with sulfonylurea monotherapy in type 2 diabetes. *Diabetes Care*. 2002;25(12):2244-8.
17. Canadian Diabetes Association Clinical Practice Guidelines Committees. *Can J Diabetes*. 2013;37.

18. Canadian Diabetes Association Clinical Practice Guidelines Expert Committee, Stone J, Fitchett D, Grover S, Lewanczuk R, Lin P. Vascular protection in people with diabetes. *Can J Diabetes*. 2013;37:S100-4.
19. Canadian Diabetes Association Clinical Practice Guidelines Expert Committee, McFarlane P, Gilbert R, MacCallum L, Senior P. Chronic kidney disease in diabetes. *Can J Diabetes*. 2013;37:S129-36.
20. Chaiteerakij R, Petersen G, Bamlet W, et al. Metformin Use and Survival of Patients With Pancreatic Cancer: A Cautionary Lesson. *J Clin Oncol*. 2016;34(16):1898-904.
21. Quan H, Sundararajan V, Halfon P, et al. Coding Algorithms for Defining Comorbidities in ICD-9-CM and ICD-10 Administrative Data. *Med Care*. 2005;43(11):1130-9.
22. Brookhart M, Wang P, Solomon D, Schneeweiss S. Evaluating Short-Term Drug Effects Using a PhysicianSpecific Prescribing Preference as an Instrumental Variable. *Epidemiology*. 2006;17(3):268-75.
23. Schneeweiss S, Patrick A, Stürmer T, et al. Increasing Levels of Restriction in Pharmacoepidemiologic Database Studies of Elderly and Comparison With Randomized Trial Results. *Med Care*. 2007;45(10):131-42.
24. Bykov K, Yoshida K, Weisskopf M, Gagne J. Confounding of the association between statins and Parkinson disease: systematic review and meta-analysis. Pharmacoepidemiology and Drug Safety 2016.
25. Margel D, Urbach D, Lipscombe L, et al. Metformin Use and All-Cause and Prostate Cancer–Specific Mortality Among Men With Diabetes. *J Clin Oncol*. 2013;31:3069-75.
26. Gleason C, Dowling N, Friedman E, Wharton W, Asthana S. Using predictors of hormone therapy use to model the healthy user bias: how does healthy user status influence cognitive effects of hormone therapy? *Menopause*. 2012;19(5):524-33.

CHAPTER 4:
SUMMARY

4.1 Overview of the Research

A large amount of evidence comparing the safety and effectiveness of antihyperglycemic medications in diabetes comes from pharmacoepidemiological studies. These study designs are commonly used to evaluate drug safety and effectiveness, and make an important contribution to medical knowledge, especially when 'gold standard' randomized controlled trials (RCT) are not always feasible.(1-3) However, diabetes research literature includes numerous highly publicized observational studies reporting exaggerated relationships between diabetes medications and health outcomes(4, 5) that were later contradicted by RCTs.(6) These discrepancies are mainly due to the subjectivity of observational studies to a number of potential biases that may cause erroneous results and produce spurious associations. Although, selection bias, unmeasured confounding, and many other factors contribute to the discrepancies, one critical element, which is often overlooked, is the method used to define exposure.(5, 7) Therefore, an accurate estimation of medication use is an essential component of any pharmacoepidemiological research as exposure misclassification will threaten study validity and lead to spurious associations,(8) this may be particularly the case with metformin. Although numerous observational studies have suggested dramatic reductions in mortality and/or cardiovascular events, meta-analyses of RCTs have failed to replicated the strong association between metformin use and all-cause or CV mortality reported in observational studies.(6, 9) It is possible some, if not all, of the benefit, observed with metformin in observational studies may be related to analytic design and exposure definitions.(7, 8)

Another potential source of bias is the 'Healthy User Effect' or 'Healthy User Bias', which is a tendency of healthier patients to initiate preventive therapies and engage in behaviours consistent with healthy lifestyles, which in turn can affect their outcomes.(1-3, 10) The healthy user effect has gained increasing attention as a potential source of bias in observational studies of preventive therapies (e.g., statins and hormone replacement therapy) and treatment of asymptomatic conditions.(1-3) Indeed, numerous studies have demonstrated that the healthy user effect can be a plausible explanation for the discrepancy between many RCTs and observational studies.(1-3, 11) Although less studied, healthy user bias may also explain the discrepancies in results between observational studies and RCTs of metformin use in patients with type 2 diabetes.

Therefore, given the strong possibility that the effects of metformin in observational studies may be tied to these two often overlooked biases, a research program was undertaken to fully explore their effects. The evaluation of exposure definition effect on the observed association and an assessment of the healthy user effect in oral antihyperglycemic medications users and health outcomes provided information regarding the discrepancy between observational studies and RCTs in diabetes and help improve the validity of study results.

4.2 Objectives

Our objectives were accomplished using two large retrospective population-based cohorts of new oral antihyperglycemic medications users in Albert, Canada, based on administrative data provided by Alberta Health.

In our first objective we hypothesized that the association between metformin and mortality will be significantly influenced by the method used to define 'exposure'. However, our hypothesis has never been formally evaluated in diabetes research, therefore, we explored the influence of exposure definitions on estimates, using the association between metformin and all-cause mortality as a proto-typical model.

In our second objective we hypothesized that metformin use, in part, is a marker of "healthy users" and would be associated with beneficial effects on different health outcomes that could not plausibly be related to metformin per se. Therefore, exploring the healthy user effect in metformin users versus non-users we evaluated the risk of events which should not be associated with metformin as well the use of preventative services or diagnostic procedures and occurrence of various health outcomes.

4.3 Summary of the Findings

In our first part of the research we used the prototypical model assessing metformin exposure and all-cause mortality. The observed estimates widely ranged from 8% to 77% lower risk of all-cause mortality risk associated with metformin use, and since the only difference between the models is the method that we use to define the exposure, the variation in the observed estimates was completely driven by the exposure definitions and their related biases. The estimates based on a time-fixed approach to define drug exposure were relatively consistent, showing 30-45% lower risk, while time-varying approaches and NCC definitions had considerable variation in the observed estimates. Risk decrease in the time-fixed approach was consistent with previous observational studies results.(12-14) These results support our initial

hypothesis and highlights the importance of exposure definition as a potential source of bias to fully, or partly explain the discrepancies observed in observational studies and RCTs of metformin. The fact that the estimates were so sensitive to the way metformin use was defined is concerning as few pharmacoepidemiological studies fully explore alternative exposure definitions when evaluating drug effects. Our research suggest that no single approach is sufficient.

In the second part of the research, we observed that metformin users were less likely to experience accident events or suffer from clinical outcomes that were not expected to be associated with metformin exposure. Metformin users were also more likely to initiate most preventive therapies and engage in other healthy behaviors. The exceptions to this pattern were was drug dependency related events and myocardial infarction, which conversely showed increased risk for metformin users. Overall, among 36 clinical events, diagnostic services and tests that we analyzed, 15 showed a statistically significantly lower risk for metformin users, 14 showed a non-significant lower risk, 4 estimates were not associated with the events and 3 showed an opposite association with an increased risk. Collectively the data suggests that the estimates are heavily influenced by healthy user bias and that metformin could be considered a strong marker of a 'healthy user'. Furthermore, these results support our initial hypothesis and presents evidence for a healthy user effect existing in diabetes research of oral antihyperglycemic medications therapies. This is an increasing and novel finding as, to our knowledge, no research has suggested a strong healthy user bias associated with metformin use pre se. These findings will have important implications of future pharmacological research related to oral antihyperglycemic medications use.

Taking observations from both parts of this research, our data adds credence to the hypothesis that major discrepancies observed between observational studies of metformin and meta-analyses of RCTs may in fact be related to the difficulty of controlling for the health user effect and the method used to define drug exposure of oral antihyperglycemic medications therapies. How much each bias may be contributing to this discrepancy is unclear, but our data strongly suggest that the biases observed from these factors likely play a significant role in the overestimation of metformin's effects in observational studies.

4.4 Implications for Future Research

Our current research work is directly related to evaluation of drug safety and effectiveness in pharmacoepidemiological studies. In order to accurately estimate the effect size of a potential adverse drug reaction, it is crucial to be able to define exposure to the medication. For post-marketing drug safety and effectiveness evaluation to be meaningful, proper methodology in exposure definition is essential to ensure the validity of study results and consequently to safer health care and more effective patient treatment. The ability to reliably and accurately identify drug exposure is of paramount importance in pharmacoepidemiology as misclassification of exposure will lead to biased effect estimates. Given the range of observed estimates in our prototypical model of metformin and all-cause mortality in the first part of the research (Chapter 2), we recommend using at least two different exposure definitions with complementary risk of bias in pharmacoepidemiological studies, along with sensitivity analyses of exposure definitions to provide more robust and potentially valid study estimates.

Careful evaluation of the association between drug exposure and health outcomes should also consider important differences in patient characteristics between those who are exposed and those who are not exposed. Our research also suggests that in the real-world metformin users may be fundamentally different than users of other oral antihyperglycemic medications. Metformin users tend to take better care of themselves by engaging in various healthy behaviors and initiating more preventive therapies and procedures. Therefore, failure to account for these behaviours can introduce major bias to pharmacoepidemiological studies evaluating the effect of oral antihyperglycemic medications therapies on health outcomes. Like statins and hormone replacement therapies(1-3, 15, 16) metformin should also be considered as a marker of healthy users and could be used in studies not directly evaluating metformin's effects *per se* to account for healthy user bias in observational studies.

Collectively, these findings have important implications for pharmacoepidemiological studies involving oral antihyperglycemic medications exposure, where metformin often constitutes a high proportion of use and is often a user group of interest in the study. Implementation of our research findings will help to decrease the discrepancy between observational studies and RCTs, contribute to the validity of study results, and reduce the bias associated with exposure misclassification and the healthy user effect. Moreover, observational studies that attribute surprisingly protective effects to antihyperglycemic therapies require cautious interpretation. Further work is required to gain a better understanding of the exposure misclassification and the heathy user effect in diabetic patients and to develop methods for guarding against these biases in observational studies.

4.5 Implications for Clinical Practice

Our aim was to gather information on the biases and confounding that may be introduced by the design of pharmacoepidemiological studies, which in turn could negatively impact clinical decision-making.(7, 17-19) Since exposure definition misclassification and the healthy user effect are not fully examined in diabetes research, are often overlooked as a potential source of bias and possibly contribute to the discrepancy between RCTs and observational studies, we decided to focus on these two problems.(5, 7)

First, due to a wide range of estimates observed in the prototypical model of metformin use and all-cause mortality, utilization of results from pharmacoepidemiological studies with a single exposure definition (e.g., 'ever' versus 'never')(17) should be avoided. We recommend that clinical decision-making should be guided by studies that have conducted sensitivity analysis by examining the influence of different approaches to defining drug exposures, especially those with complementary risks of bias. Accurate exposure definition is important for clinicians to make informed decisions based on pharmacoepidemiological studies results. Second, the healthy user effect has not been fully examined in oral antihyperglycemic medications therapies and not considered as a potential source of bias in pharmacoepidemiological studies. However, based on our study results, metformin users are more likely to initiate preventive therapies and engage in other healthy behaviors. Thus, studies including methodological strategies accounting for this effect (e.g., new user designs(20, 21), time-dependent exposures,(22) propensity score approaches that include markers of the healthy user into the prediction of drug exposure,(16) etc.) will minimize the healthy user effect on the estimates and provide higher results validity. Third, studies using multivariable models including co-variates that are suspected as confounders in the association of interest (e.g., age, sex, and a comorbidity

score),(23) are an important part of any pharmacological research that contributes to confounding reduction and an increase in the validity of study results; however, consideration should be given to routinely using drugs such as statins, hormone replacement therapy, and now metformin in models to further account for healthy user bias.

Our research shows possible pitfalls in modern pharmacoepidemiological studies of oral antihyperglycemic medications therapies and evaluates the discrepancy between 'gold standard' RCTs and observational studies, indicating several potential biases that should be accounted for when appraising research papers and integrating their results into clinical practice. Moreover, an interpretation of observational studies that attribute surprisingly protective effects to antihyperglycemic therapies should be regarded with high skepticism. However, further research is required to evaluate the healthy user effect and the methods to guard against its negative impact on oral antihyperglycemic medications therapies evaluation.

4.6 Conclusion

In the first part of our research, a variety of exposure definitions were tested and produced a wide range of associations between metformin and mortality risk, therefore, pharmacoepidemiological studies should include at least two different exposure definitions with complementary risk of bias and implement sensitivity analyses of exposure definitions to provide more robust and potentially valid study estimates.

Furthermore, our research suggests that metformin users take better care of themselves by engaging in various healthy behaviors and initiating preventive therapies, and therefore, in the studies evaluating the effect of oral antihyperglycemic medications therapies a failure to account for

behaviours consistent with healthy lifestyle may introduce healthy user bias. However, further work is required to gain a better understanding of the heathy user effect in diabetic patients and to develop methods for guarding against this effect.

The reason for the discrepancies between observational studies and RCTs on oral antihyperglycemic medications therapies are unknown, however, our research indicated a few potential additional sources for the problem, such as unmeasured confounding, exposure misclassification and the difficulty to control patient characteristics. Moreover, our research has highlighted several methodologic strategies to be considered in the design of pharmacoepidemiological studies evaluating oral antihyperglycemic medications therapies that can improve research methods and enhance clinician decision-making on the management of diabetic patients, which we hope in turn will improve health outcomes and therapy safety in these patients.

4.7 References

1. Brookhart M, Patrick A, Dormuth C, et al. Adherence to Lipid-lowering Therapy and the Use of Preventive Health Services: An Investigation of the Healthy User Effect. *American Journal of Epidemiology* 2006;166(3):348-54.
2. Dormuth C, Patrick A, Shrank W, et al. Statin Adherence and Risk of Accidents A Cautionary Tale. *Circulation*. 2009;119(15):2051-7.
3. Patrick A, Shrank W, Glynn R, et al. The Association between Statin use and Outcomes Potentially Attributable to an Unhealthy Lifestyle in Older Adults. *Value Health*. 2011;14(4):513-20.
4. Bannister CA, Holden SE, Jenkins-Jones S, Morgan CL, Halcox JP, Schernthaner GM, J Currie, CJ. Can People With Type 2 Diabetes Live Longer Than Those Without? A Comparison of Mortality in People Initiated With Metformin or Sulphonylurea Monotherapy and Matched Non-diabetic Controls *Diabetes Obes Metab*. 2014;16(11):1165-73.
5. Gandini S, Puntoni M, Heckman-Stoddard B, et al. Metformin and cancer risk and mortality: A systematic review and meta-analysis taking into account biases and confounders. *Cancer Prev Res (Phila)*. 2014;7(9):867-85.
6. Boussageon R, Supper I, Bejan-Angoulvant T, et al. Reappraisal of Metformin Efficacy in the Treatment of Type 2 Diabetes: a Meta-Analysis of Randomized Controlled Trials. *PLoS Med*. 2012;9(4).
7. Tournier M, Bégaud B, Cougnard A, et al. Influence of the drug exposure definition on the assessment of the antipsychotic metabolic impact in patients initially treated with mood-stabilizers. *Br J Clin Pharmacol*. 2012;74(1):189-96.
8. Brunelli S. Use of prescription drug claims data to identify lipid-lowering medication exposure in pharmacoepidemiology studies: potential pitfalls. *Pharmacoepidemiol Drug Saf*. 2016;25(7):844-6.
9. UKPDS. Effect of intensive blood glucose control with metformin on complications in overweight patients with type 2 diabetes *The Lancet*. 1998;352(9131):854–65.
10. Eurich D, Majumdar S. Statins and Sepsis – Scientifically Interesting but Clinically Inconsequential *J Gen Intern Med* 2011;27(3):268-9.
11. Ray W. Observational studies of drugs and mortality. *N Engl J Med*. 2005;353(22):2319-21.
12. Johnson J, Simpson S, Toth E, Majumdar S. Reduced cardiovascular morbidity and mortality associated with metformin use in subjects with type 2 diabetes. *Diabet Med*. 2005;22(4):497-502.
13. Johnson J, Majumdar S, Simpson S, Toth E. Decreased Mortality Associated with the Use of Metformin Compared with Sulfonilurea Monotherapy in Type 2 Diabetes. *Diabetes Care* 2002;25(12):2244-8.
14. Eurich D, Majumdar S, McAlister F, Tsuyuki R, Johnson J. Improved clinical outcomes associated with metformin in patients with diabetes and heart failure. *Diabetes Care*. 2005;28(10):2345-51.
15. Garbe E, Suissa S. Hormone replacement therapy and acute coronary outcomes: methodological issues between randomized and observational studies. *Hum Reprod*. 2004;19(1):8-13.
16. Gleason C, Dowling N, Friedman E, Wharton W, Asthana S. Using predictors of hormone therapy use to model the healthy user bias: how does healthy user status influence cognitive effects of hormone therapy? *Menopause*. 2012;19(5):524-33.
17. Walker JJ, Johnson JA, Wild SH. Diabetes treatments and cancer risk: the importance of considering aspects of drug exposure. *Lancet Diabetes Endocrinol*. 2013;1(2):132-9.
18. Matok I, Azoulay L, Yin H, Suissa S. Immortal time bias in observational studies of drug effects in pregnancy. *Birth Defects Res A Clin Mol Teratol*. 2014;100(9):658-62.
19. Eurich D, Majumdar S. Statins and sepsis - scientifically interesting but clinically inconsequential. *J Gen Intern Med*. 2012;27(3):268-9.
20. Brookhart M, Wang P, Solomon D, Schneeweiss S. Evaluating Short-Term Drug Effects Using a PhysicianSpecific Prescribing Preference as an Instrumental Variable. *Epidemiology*. 2006;17(3):268-75.

21. Brassard P, Wu J, Ernst P, Dell'Aniello S, Smiechowski B, Suissa S. The effect of statins on influenza-like illness morbidity and mortality. *Pharmacoepidemiology and Drug Safety* 2016.
22. Bykov K, Yoshida K, Weisskopf M, Gagne J. Confounding of the association between statins and Parkinson disease: systematic review and meta-analysis. Pharmacoepidemiology and Drug Safety 2016.
23. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care. 2005;43(11):1130-9.