# Research methods for conducting pharmacoepidemiological studies using medicines claims data


## M Obeng-Kusi

orcid.org/ 0000-0001-6159-6782


Dissertation submitted in fulfilment of the requirements for the degree Master of Master of Pharmacy in Pharmacy Practice at the North West University

| | |
|---|---|
| Supervisor: | Prof JR Burger |
| Co-supervisor: | Prof MS Lubbe |
| Co-supervisor: | Mrs M Cockeran |

Graduation: May 2019

Student number: 27959716

## ACKNOWLEDGEMENTS

First and most importantly, all glory and honour must be given to the Almighty God who continues to fulfil His promises to me and has given me the grace, wisdom and perseverance to pursue this dream with excellence.

My utmost gratitude goes to Prof JR Burger, my supervisor extraordinaire, for her constant counsel and support throughout the course of this study. Although your expertise and leadership were excellent, I appreciate most your belief in me and how you challenged me. Thank you for your constantly opened doors and unrelenting efforts. You have been an inspiring mentor.

I would also like to express my sincerest thanks to Prof MS Lubbe who not only always played her part as co-supervisor effectively but provided timely encouragement and advice. I thank Prof Lubbe especially for her help in conceptualising the research, securing funding for the completion of this study and her expert analyses of the data.

To Mrs M Cockeran, I say thank you for your role as my co-supervisor, taking time to guide me through the statistical analyses of this work. Your expertise and patience are appreciated.

I thank Ms A Bekker for her help with the data analyses and cross-checking my reference list, and for always having a kind word and a welcoming smile. You helped make the weight of this study significantly lighter.

I appreciate Mrs H Hoffman for her assistance with regards to my dissertation. I am thankful for the thorough proofreading through the dissertation and double-checking of the reference list.

Thank you, Ms E Oosthuizen, for your assistance with the technical editing and administrative support. I cannot thank you enough for the friendship you offered throughout this period. The times of sharing coffee and nice chats were always a refreshing break for me and I am grateful.

I thank the Pharmaceutical Benefit Management Company for the data used in this study and the North-West University and National Research Fund for financial assistance.

Special thanks go to Prof. A Combrink for the meticulous language editing of this dissertation.

Finally, words fail me when I think of all my family and friends who have stood by me throughout this study. Your genuine interest in my project, suggestions, corrections and patience did not go unnoticed. God bless you all.

# PREFACE

This dissertation is presented in article format. The chapters in this dissertation are as follows:

Chapter 1 presents a comprehensive background of the study, outlining the research aims and objectives, as well as the research methods employed.

Chapter 2 is the literature review which focuses on pharmacoepidemiology as a field of study and its various measures; the study designs, sources of data and research methods employed in pharmacoepidemiological studies.

Chapter 3 presents the results and discussions of the empirical study presented in the form of two manuscripts, prepared for submission to the following journals for publication:

*(i) Journal of clinical pharmacy and therapeutics*

*(ii) Journal of pharmacoepidemiology and drug safety*

Chapter 4 contains the conclusions of the study, along with recommendations for future investigations and the limitations.

The references and annexures are presented at the end of the dissertation.

The manuscripts were written and referenced according to the specified journal author guidelines. The Harvard style, the required referencing format by the North-West University, was employed for compiling the complete reference list for this dissertation.

The manuscripts in this dissertation were written under the direction and with the approval of the supervisor and co-supervisors, who also acted as co-authors. The contributions of each author are summarised in Chapter 3.

# ABSTRACT

**Title**: Research methods for conducting pharmacoepidemiological studies using medicines claims data.

This study aimed to evaluate the appropriate use of various research methods in pharmacoepidemiological studies using a South African medicines claims database. A literature review was carried out to conceptualise pharmacoepidemiology and the study designs, data sources as well as methods employed in performing pharmacoepidemiological studies. The empirical investigation utilised a quantitative, cross-sectional approach. Medicines claims data covering the period between January 2006 and December 2015 were supplied by a privately-owned Pharmaceutical Benefit Management company (PBM) for analyses in this study. The objectives of the empirical study were to:

(i) Determine the time to onset of treatment of hypertension and hyperlipidaemia in patients with type 2 diabetes mellitus using survival analysis.

(ii) Compare different adherence measures, by determining adherence to montelukast among asthma patients, using data from a medicines claims database in South Africa.

Survival analysis and methods for the determination of adherence were the methods assessed in this study.

Manuscript 1 presented the results of a survival analysis conducted to determine the time to onset of treatment of hyperlipidaemia and hypertension among diabetes patients using a South African medicines claims database. Patients with ICD-10 codes I10, I11, I12, I13, I15, O10 and O11 for hypertension and E78.5 for hyperlipidaemia, who were receiving medications for these conditions according to the National Pharmaceutical Product Index (NAPPI) codes provided by the Monthly Index of Medical Specialities (MIMS), were selected among patients with ICD-10 code E11 for diabetes in conjunction with the NAPPI codes for antidiabetic medications for this study. Retrospective data of patients continuously enrolled with a Pharmaceutical Benefit Management company in South Africa from 1st January 2008 to 31st December 2016 (N = 494) were analysed and the Kaplan-Meier approach was used to compare the survival experience of subjects who developed hypertension and hyperlipidaemia. The mean age of the population was 53.5 ± 11.4 years, with 34.8% (N = 494) being females. Prevalence of hyperlipidaemia and hypertension among diabetic patients were 35.0% and 45.6%, respectively. The mean time to onset of treatment of hyperlipidaemia was 2684.4 ± 42.2 days compared to 2434.2 ±47.6 days for hypertension. There was no statistically significant difference in age and sex among subjects who

developed either of these conditions during the study. This study showed that survival analysis can successfully be conducted using secondary data, provided data fields are accurately documented.

Manuscript 2 reported the findings of the investigation into the methods for the determination of adherence. Six different measures of adherence were compared to the medication possession ratio (MPR), as the reference adherence measure, by analysing claims made for montelukast. Data of patients continuously enrolled with a South African PBM from 1st January 2006 to 31st December 2015 were obtained and analysed to determine the adherence to montelukast. Patients with ICD-10 code J45 for asthma who made at least two consecutive claims for montelukast based on the NAPPI code 10.4.2 were selected and included in this study. Of the total of 9 141 patients with a median age of 13.3 (5.2 – 49.2) years, 52.8% (n = 4 825) were females. These women were significantly older than their male counterparts ($p < 0.0001$; Cohen's $d = 0.5$). Compared to the MPR, continuous multiple interval measures of oversupply (CMOS) and compliance ratio (CR) were found to be equivalent, each producing an adherence value of 86.0%. Higher adherence values of 96.9%, 117.2% and 129.0% were produced by the modified medication possession ratio (MPRm), refill compliance rate (RCR) and mean continuous single interval measure of medication acquisition (CSA), respectively, whereas the proportion of days covered (PDC) capped at 1 yielded a lower adherence value of 76.0%. This study showed that compared to the measures that used the difference between claims dates as denominator, those that had the entire investigation period as the denominator produced consistent results.

From the study, it can be concluded that medicines claims data are vital in in pharmacoepidemiological studies since they provide large and readily available data over a wide period for research. Research methods, such as measures for determining adherence and survival analysis can be effectively conducted using medicines claims data when all relevant data fields are accurately recorded. In carrying out research using medicines claims data, availability of specific parent and medication information determines the methods to be used and the extent of data analysis and interpretation.

**Keywords**: pharmacoepidemiology, research methods, measures for adherence, time to onset of treatment, survival analysis, medicines claims data, South Africa.

# LIST OF ABBREVIATIONS

| | |
|---|---|
| ADE | Adverse drug event |
| ADR | Adverse drug reaction |
| ANOVA | Analysis of variance |
| CDC | Centers for Disease Control and Prevention |
| CMG | Continuous measure of medication gaps |
| CMOS | Continuous multiple interval measure of oversupply |
| CMS | Council for Medical Schemes |
| CR | Compliance ratio |
| CSA | Continuous single interval measure of medication acquisition |
| EMA | European Medicines Agency and Heads of Medicines Agency |
| ENCePP | European Network of Centres for Pharmacoepidemiology and Pharmacovigilance |
| ICD-10 | International Statistical Classification of Diseases and Related Health Problems 10$^{th}$ Revision |
| MIMS | Monthly Index of Medical Specialities |
| MPR | Medicine possession ratio |
| MPRm | Medication possession ratio, modified |
| MUSA | Medicines Usage in South Africa |
| NAPPI | National Pharmaceutical Product Index |
| NCC MERP | National Coordinating Council for Medication Error Reporting and Prevention |
| NWU | North-West University |
| PBM | Pharmaceutical benefit management |
| PEM | Prescription events monitoring |

| | |
|---|---|
| PDC | Proportion of days covered |
| RCR | Refill compliance rate |
| RCT | Randomised controlled trial |
| SA | South Africa |
| SAS | Statistical Analysis System® |
| WHO | World Health Organization |

# LIST OF DEFINITIONS

| | |
|---|---|
| Asthma | Asthma is a long-term inflammatory condition of the airways, which causes obstruction of the airways and is associated with recurrent wheezing, coughing, breathlessness and chest tightness (Wells *et al.*, 2015:821). |
| Bias | Bias is defined as the absence of internal validity or inaccurate estimation of the relationship between an exposure and an event in a specified population resulting in a variable that is equal to its true value (Delgado-Rodríguez & Llorca, 2004:635). |
| Confounding | Confounding occurs when an extraneous variable affects those being examined resulting in an altered relationship between the variables under study (Pourhoseingholi *et al*., 2012:79). |
| Confounding by indication | Confounding by indication occurs when the indication for which a medication being studied has been prescribed, is also a determinant of the outcome being investigated (Garbe & Suissa, 2014:1905). |
| Evidence-based medicine | Evidence-based medicine is the approach to practice that integrates clinical expertise, values of patients and the best available research information in making decisions that are related to healthcare of individual patients (Masic *et al*., 2008:219). |
| Hyperlipidaemia | Hyperlipidaemia is defined as increased levels of triglycerides (Malloy & Kane, 2012:619). |
| Hypertension | Hypertension is defined as “a persistently raised arterial blood pressure” (Wells *et al*., 2015:87). |
| Odds ratio | The odds ratio measures the relationship between an outcome and an exposure, representing the chance that an event will take place in the presence of a specific exposure compared to the chance that the same event will occur when the exposure is absent (Szumilas, 2010:227). |
| Protopathic bias | Protopathic bias is the bias that occurs when a medication is initiated in response to the initial symptoms of a disease which is undiagnosed at the point the medication is given (Faillie, 2015:779). |

| | |
|---|---|
| Quality-adjusted life years (QALYs) | Quality-adjusted life years refers to a standardised burden-of-illness measure which incorporates survival and health-related quality of life as a single statistic and is used in cost-effectiveness analyses as a guide for decision-making regarding the distribution of healthcare resources among competing programmes for a given population (Howren, 2013:104). |
| Regression to the mean | Regression to the mean is a statistical phenomenon that arises as a result of study subjects being selected for extreme values of characteristics that vary over time (Hughes *et al*., 2015:439). |
| Selection bias | Selection bias is bias arising from the selective recruitment of subjects that do not entirely represent the outcome or exposure in the population into a study (ENCePP, 2017:22). |
| Simpson's paradox | Simpson's paradox is "a phenomenon whereby the association between a pair of variables (X; Y) reverses sign upon conditioning of a third variable, Z, regardless of the value taken by Z" (Pearl, 2013:1). |
| Type 2 diabetes mellitus | Type 2 diabetes mellitus is a condition which presents with resistance of tissues to insulin activity and a relative insulin deficiency (Kennedy, 2012:744). |

# TABLE OF CONTENTS

# LIST OF TABLES

## LIST OF FIGURES

# CHAPTER 1: INTRODUCTION

## 1.1 Background and problem statement

Pharmacoepidemiology consists of two words, “pharmaco”, reflecting an aspect of clinical pharmacology and “epidemiology”, which addresses the study of health states among populations (Thaker *et al*., 2015:53). It is, therefore, a study that comprehensively covers aspects of both epidemiology and clinical pharmacology. Strom (2013a:3) defines pharmacoepidemiology as “*the study of the use and the effects of drugs in a large number of people*”. The World Health Organization (WHO, 2003:8) added that the purpose of pharmacoepidemiology is to support cost-effective and rational drug use in a populace to improve health outcomes. Thus, the field of pharmacoepidemiology is vital to health research and practice.

Pharmacoepidemiological studies are beneficial for regulatory, marketing, legal and clinical purposes (Strom, 2013b:56; Thaker *et al*., 2015:57). In regulatory settings, pharmacoepidemiological studies provide a framework for post-marketing studies before drug approval (Strom, 2013b:57) and also serves as ‘legal prophylaxis’ for future drug liability suits (Thaker *et al*., 2015:56; Strom, 2013b:57). There is little information about a drug in a population at the initial stages of marketing, creating the need to answer certain clinical questions which are addressed by pharmacoepidemiological studies. These studies can determine other uses of a drug which were not identified during randomised controlled trials (Fautrel, 2004:175; Strom, 2013b:57). An outcome, such as adverse effects of a drug, is a major consequence of drug use in a population; the quantifying of which pharmacoepidemiological studies offer valuable contributions to, along with testing clinical hypotheses to improve knowledge in the medical field (Berlin *et al*., 2008:1368; Strom, 2013b:58).

Pharmacoepidemiological studies can be divided into two main fields (Eggen & Straand, 2001:3). Firstly, there is the focus on public health impacts of drug use, variations and patterns in drug use, as well as hypotheses generation in exploring these variations. Secondly, these studies address follow-up studies such as adverse drug events (ADEs), side-effects and post-marketing research that investigate long-term of drug effects in a population setting (Eggen & Straand, 2001:3). Basic epidemiological measures are employed in exploring drug use in a given population and determining measures of clinical frequency such as incidence and prevalence of diseases (Fletcher & Fletcher, 2005:60). In the determination of incidence of drug effects, drug utilisation patterns, adherence profiles, the economic implications of drug use and the association between medication use and outcomes, pharmacoepidemiological studies are of significant importance (Strom, 2013a:9).

Bonita *et al*. (2006:18) define incidence as "*the rate of occurrence of new cases arising in a given period in a specified population*". Suruki and Chan (2008:220) add that this population should be at risk for developing the disease. This measure of disease frequency is relevant in identifying the population at risk of developing a disease and the rate or intensity at which the disease occurs (Suruki & Chan, 2008:222). Prevalence, on the other hand, is viewed as the frequency of occurrence of an existing case at a specified time in a given population (Bonita *et al*., 2006:18). This measure is useful for drug utilisation studies and in making decisions related to resource allocation (Suruki & Chan, 2008:225).

The WHO (2003:9) defined drug utilisation research as "*the marketing, distribution, prescription, and use of drugs in a society, with special emphasis on the resulting medical, social and economic consequences*". Wettermark *et al*. (2008:159) further described drug utilisation research as "*an eclectic collection of descriptive and analytical methods for the quantification, the understanding and the evaluation of the processes of prescribing, dispensing and consumption of medicines and the testing of interventions to enhance the quality of these processes*". The principles of pharmacoepidemiology are applied in drug utilisation studies to give insight into the patterns, quality, determinants and outcomes of drug use (WHO, 2003:9).

Adherence to medication is vital to reaching clinical goals and for this reason, knowledge on the extent of a patient's adherence to a treatment protocol is important to both researchers and clinicians (Lam & Fresco, 2015:1). The adherence project of the WHO in 2003 adopted the definition of adherence as "*the extent to which a person's behaviour - taking medication, following a diet and/or executing lifestyle changes, corresponds with agreed recommendations from a healthcare provider*". Adherence is thought to be an umbrella term of three constructs: initiation or acceptance, compliance and persistence (Urquhart, 2001:473-474). Grégoire and Moisan (2016:369-370) further explain initiation as primary adherence whereby the patient accepts treatment; compliance as the implementation phase within which a patient takes the recommended number of doses and follows the recommended treatment schedule; and persistence or continuation as the capability of a patient to maintain the treatment for the prescribed duration. Together, these three constructs of adherence are important factors that determine the success of any therapy.

The rising cost of medical care is a growing source of concern because differences in the effectiveness of a pharmaceutical product compared to its cost are important in differentiating one product from another just as differences in efficacy and safety are (Schulman *et al*., 2013:280). Pharmacoepidemiology has a significant role in evaluating the cost-effectiveness of medications, as analysis of the total cost of a medical intervention and its benefits depends on pharmacoepidemiological techniques that are related to costs (Tanaka *et al*., 2015:3). Decision-

makers can have access to better information and use limited resources available for public health more effectively when pharmacoepidemiological studies, merged with economic aspects, are applied (Schulman *et al*., 2013:290).

Medication use outcomes comprise the safety with which medications are used and the extent to which the therapeutic goal for the medication is achieved (Osheroff, 2009:11). Examples of these outcomes are medication errors and preventable ADEs. The National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP, 2017) defines a medication error as "*any preventable event that may cause or lead to inappropriate medication use or patient harm while the medication is in the control of the health care professional, patient, or consumer. Such events may be related to professional practice, health care products, procedures, and systems, including prescribing, order communication, product labelling, packaging, and nomenclature, compounding, dispensing, distribution, administration, education, monitoring, and use*". According to Aronson (2009:514) medication errors can occur when choosing a medication, writing a prescription, dispensing a formulation, administering or taking a drug.

An adverse drug reaction (ADR) may be defined as "*an unwanted, undesirable effect of a medication that occurs during usual clinical use*" (Schatz & Weber, 2015:5). Adverse drug reactions may significantly impact practice in both clinical and economic terms because of their high frequency of occurrence and potentially serious consequences (Sultana *et al*., 2013:73). Several pharmacoepidemiological strategies such as direct reporting and root cause analysis are applied for identifying and minimizing the occurrence of medication errors and adverse drug events (Gurtwitz & Field, 2008:655). Sequence symmetry analysis is also employed in identifying ADRs. This technique assesses association by examining prescribing or hospitalisation data for symmetry in sequence of dispensed medications and signals of adverse events within a given time frame (Wahab, Pratt, Wiese, *et al*., 2013:496).

Random clinical trials, an example of experimental studies, are the gold standard for epidemiological studies. However, they are expensive in terms of financing and time and often do not mimic practice settings (Waning & Montagne, 2001:64). Observational studies are more practical for evaluating the effects of an intervention or therapy in a given population. These studies have an advantage of being less expensive, easier to conduct and present less ethical concerns (Silverberg, 2015:722). Observational studies may be descriptive, leading to hypothesis generation, or analytical for hypothesis testing (DiPietro, 2010:976). Cohorts, case-controls, cross-sectional studies, case series and case reports are examples of observational study designs employed in pharmacoepidemiology.

Applications of observational studies have led to a greater appreciation of the methodological complexities in conducting research of drug effects and far-reaching consequences of imperfect methods. There have been discrepancies in results of observational studies which may be due to imperfections in the study design and statistical analyses (Suissa, 2009:4). Observational studies may lead to uncertainty about the characteristics of the drug or outcome of study and require further studies for a more reliable conclusion to be drawn (Strom, 2013a:27).

An important aspect of a research approach is the identification of an appropriate source of data. Pharmacoepidemiological investigations may utilise primary data which are collected forward in time for the purpose of the study or administrative data that were already collected for some other purpose (European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP), 2017:8). Primary data sources such as active surveillance, intensive monitoring schemes, surveys, prescription event monitoring and clinical trials have important roles in pharmacoepidemiology (European Medicines Agency and Heads of Medicines Agency (EMA), 2011:22). Data from primary sources have been used to evaluate disease-drug associations for rare complex conditions that require intensive assessments by clinical experts (ENCePP, 2017:8). The case-control study conducted by Roujeau *et al*. (1995:1600) to measure the associated risks with the use of specific drugs is an example of a study that makes use of primary data to assess such relationships. Primary data sources, however, have limitations. There are ethical concerns regarding the privacy and confidentiality of study subjects. Primary data collection may also be costly and time-consuming (Harpe, 2009:139). Underreporting, inaccuracy of information submitted, biases and difficulty in estimating exposure in a population may also limit the use of primary sources of data (Pan *et al*., 2013:113).

Administrative data for pharmacoepidemiological studies may be obtained from databases. Generally, two types of databases are available for pharmacoepidemiological studies, i.e. databases that contain in-depth medical information and those mainly created for administrative purposes. The former usually includes prescriptions, diagnosis and discharge reports, whereas the latter requires a record-linkage between pharmacy claims and medical claims databases (EMA, 2011:26).

In addition to an ideal database's ability to cover a population large enough to discover infrequent outcomes and have in its formulary, the drug under study prescribed in sufficient quantities, it must also contain records from emergency, outpatient, inpatient and mental care; as well as all radiological and laboratory investigations, prescribed and over-the-counter medicines (Strom, 2013c:118). The International Society for Pharmacoepidemiology-approved Guidelines for Good Database Selection and use in pharmacoepidemiology (Hall *et al*., 2011:1-10) is a well-structured

guideline that is aimed at helping investigators identify and select appropriate databases for studies.

The Council for Medical Schemes (CMS) named hypertension, hyperlipidaemia, type 2 diabetes mellitus and asthma among the most prevalent chronic disease list (CDL) conditions affecting beneficiaries of medical aid schemes in South Africa between 2010 and 2015 (CMS, 2017:6). The global report of the WHO (2016:6) reported a 3.8% rise in global diabetes prevalence from 1980 to 2014 and 1.5 million deaths caused by diabetes mellitus in 2012. The prevalence of type 2 diabetes mellitus rose from 20.29 per 1000 to 31.21 per 1000 among beneficiaries of medical schemes between 2010 and 2015 (CMS, 2017:8). Along with its complications, diabetes mellitus has significant socio-economic effects on patients, care givers, healthcare facilities and society as a whole (Simpson *et al*., 2003:1661). Hypertension and hyperlipidaemia are among the costliest complications of diabetes with regards to patient suffering and healthcare expenditure (Sowers *et al*., 2001:1053). Hypertension is a significant risk factor for other morbidities such as heart failure, stroke, myocardial infarction, and renal failure and a challenge to public health (Ong *et al*., 2007:69). According to Wong *et al*. (2006:204), hypertension and hypercholesterolemia usually co-exist and can cause “dyslipidaemic hypertension”, which significantly increases the cardiovascular risks of patients. An analysis of the relationship between these conditions is thus vital for their management. Masoli *et al*. (2004:470) noted an increase in asthma prevalence in both adult and children populations. This is possibly attributed to a possible increase in atopic sensitisation, the adoption of western lifestyles as well as urbanisation. Uncontrolled or poorly controlled asthma can result in severe limitations on quality of life and can be fatal (Bateman *et al*., 2008:143). Maintaining adherence to therapy contributes significantly to good outcomes in the management of asthma, reducing morbidity and mortality and saving costs as well (Gillisen, 2007:205). With its significant prevalence in South Africa coupled with the value that adherence provides to its therapy, asthma is an essential condition for which adherence must be investigated.

Medical aid schemes or their contracted administrators collect data on the provision of healthcare services to their beneficiaries (Gray *et al*., 2016:37). These data are patient-specific and are among the most accurate data sources used for pharmacoepidemiological studies (Lighter, 2004:98). Pharmacoepidemiology provides insight into patterns, quality, determinants and outcomes of medicines use. This insight informs decision-making regarding the use of medicines, which account for a high proportion of healthcare expenditure (Sjöqvist & Birkett, 2004:77).

Databases are employed in pharmacoepidemiological investigations because they provide a large sample size for research, increased methodological flexibility and are relatively inexpensive to use (Harpe, 2009:139; Hess *et al*., 2006:1281). It provides extensive longitudinal data, regular

data collection and minimal observer bias (Yasmina *et al*., 2014:601). Databases offer protection from privacy and confidentiality issues associated with direct contact with subjects. Database studies have been applied in drug utilisation research, studies of physician prescribing, beneficial drug effects, adverse events and health policy research and potentially increase generalisability by examining drug use under real-life situations (Harpe, 2009:139).

There are, however, concerns with the adequacy of study designs, timeframe available for study from databases, relevance of the population used and specificity of clinical outcome assessment which threaten the validity of research findings from database studies limiting the usefulness of studies and their subsequent adoption into policy and practice (Berger *et al*., 2009:1045). Accuracy of an association in a study depends on the accuracy of the study design used. The presence of random errors, biases and confounders weaken the strength of association in a study and subsequently renders its findings inaccurate or unreliable (Strom, 2013a:18). An appropriate research design and the application of relevant analyses will be of significance in addressing these concerns and will improve understanding and relevance of study findings (Berger *et al*., 2009:1045).

Several studies have evaluated the study designs used in pharmacoepidemiological studies (Grzeskowiak *et al*., 2012; Hallas & Pottegård, 2014; Lu, 2009:691). There are, however, limited publications that illustrate the appropriate application of the different methodologies to pharmacoepidemiological studies using a medicine claims database specific to the South African healthcare environment. Increased appreciation of what needs to be considered in designing, analysing and interpreting data from pharmacoepidemiological studies will make researchers better placed to design and implement such studies. This improved insight will also make researchers better resourced to make recommendations based on findings from these studies (DiPietro, 2010:975). This research therefore sought to illustrate the appropriate use of various methodologies for conducting longitudinal pharmacoepidemiological studies using data from a medicines claims database. The research addressed the following research questions:

- What are the study designs available for pharmacoepidemiological research?
- What are the advantages, disadvantages and limitations of each study design in the context in which it is used in pharmacoepidemiological studies?
- What research methodologies can be applied in pharmacoepidemiology using medicines claims databases?

## 1.2 Research aims and objectives

The following sections address the aims and objectives of the study.

### 1.2.1 Research aims

The study aimed to evaluate the appropriate use of various research methods in pharmacoepidemiological studies using a South African medicines claims database. The study consisted of a literature review and an empirical study.

### 1.2.2 Specific research objectives

The specific objectives of the literature review were to:

(i) Conceptualise pharmacoepidemiology; its uses and relevance.

(ii) Conceptualise various study designs; their advantages and disadvantages as well as the statistical analyses applicable to them.

(iii) Conceptualise databases used in pharmacoepidemiological studies.

(iv) Determine, from literature, the application of various research methods in pharmacoepidemiological studies.

The specific objectives of the empirical study were to:

(i) Determine the time to onset of treatment of hypertension and hyperlipidaemia in patients with type 2 diabetes mellitus using survival analysis.

(ii) Compare different adherence measures, by determining adherence to montelukast among asthma patients, using data from a medicines claims database in South Africa.

## 1.3 Research methodology

The subsequent paragraphs focus on the literature review and empirical investigation that were carried out to address the objectives of the study.

### 1.3.1 Literature review

Literature from books and works published in reliable sources such as GoogleScholar®, PubMed®, Scopus®, ScienceDirect® and EBSCOhost® which addressed the outlined objectives and gave more insight into the research, were reviewed. The search for literature involved the combination of key terms such as ‘pharmacoepidemiology’, ‘medicines claims databases’, ‘study designs’,

'research methods', 'adherence', 'non-adherence', 'partial-adherence', 'survival analysis', 'time-to-event', 'disease risk factors' and 'South Africa'. English was the preferred language for the literature search and findings from 2000-2018 were identified as relevant references.

### 1.3.2 Empirical investigation

The following paragraphs focus on study setting, design, population, variables and methods of data analysis that were used to attain outlined objectives of the empirical investigation.

### 1.3.3 Study design

The research followed a quantitative, cross-sectional study design. According to Maree and Pietersen (2016:162), "*quantitative research is a process that is systematic and objective in its ways of using numerical data from only a selected subgroup of a universe (or population) to generalise the findings to the universe that is being studied*". In quantitative studies, a researcher uses numerical data to assess the link between variables, look for probable cause and effect and to answer research questions (Ivankova *et al*., 2016:307). Waning and Montagne (2001:45) describe observational studies as the approaches that have the advantage of identifying, studying and measuring variables devoid of human interventions. Cross-sectional studies, also known as prevalence studies, involve measuring both exposure and outcomes in a group of people at a specified time (Bhopal, 2002:242; Harpe, 2011:45; Verhamme & Sturkenboom, 2010:69; Waning & Montagne, 2001:51) and focus on simultaneous collection of information on disease or drug-related problems, characteristics of the population as well as the risk factors (Bhopal, 2002:242; Harpe, 2011:45). The use of retrospective data for pharmacoepidemiological studies is of benefit since it provides large sample sizes over long periods of study and is relatively less expensive and faster to carry out (Motheral *et al*., 2003:91).

### 1.3.4 Study setting and data source

Data were acquired from the medicines claims database of one South African Pharmaceutical Benefit Management (PBM) company. Administrative databases do not depend on the interviews or recall of patients to obtain information and thus avoids recall bias (Strom, 2013c:120). Matshidze and Hanmer (2007:92) define claims data as "*clinical information collected through claims submitted by health care providers to medical schemes for access to benefits and reimbursement; claims usually contain clinical, financial and administrative information*". Claims data represent the billable interactions that exist between the insured patients and healthcare delivery systems and fall into four general categories: inpatient, outpatient, enrolment and pharmacy or medicines claims data (University of Washington, 2016).

The PBM Company from which the data for the study were obtained is a large independent company that has been servicing the South African private healthcare industry for more than two decades. The company provides medicine claims processing services to about 1.6 million beneficiaries of 42 medical schemes in SA. Ten years' data from 1st January 2006 to 31st December 2015 were obtained from the database provided by the PBM for the study.

#### 1.3.4.1 Validity and reliability of data

Validity and reliability of data are essential aspects of a good research. Validity is the extent to which a measure estimates what it is intended to, whereas reliability is the extent to which the measure remains stable when it is repeatedly measured under homogenous conditions (Waning & Montagne, 2001:123). The PBM which provided data for this study provides gate-keeping, utilisation management, clinical management and eligibility services as well as pricing management to establish data reliability and validity. The PBM also offers data integrity validation and benefits validity, all of which are targeted at ensuring that standards of claiming are met. Unpaid claims were not included in the data as part of a cleaning-up process. Random data checks were conducted after each cleaning process by verifying datasets. Park and Stergachis (2008:519) proposed that the quality of claims databases must be sufficiently high and allow linking of individual subject data across datasets, as well as the ability to trace patients in datasets for protracted follow-up.

Although data from a database provide good resource for studies, some are not necessarily designed for research purposes (Harpe, 2009:139). There may be limitations in the quality of the data from which they are derived and lack of important confounder information such as occupation, alcohol consumption and cigarette smoking (Hartzema *et al*., 2008:4-6). Motheral *et al*. (2003:90-97) proposed a checklist to assist researchers determine how suitable a database is for a study and the appropriate methodology to be employed.

### 1.3.5 Target and study population

The study involved all patients who were continuously enrolled as determined from the datasets obtained from the PBM Company within the time frame selected for this study. The study population consisted of all patients meeting inclusion criteria per specific research objective (refer to Table 1-1).

#### 1.3.5.1 Inclusion criteria

Table 1-1 depicts the inclusion criteria that were used for the selection of the study population per each research objective.

**Table 1-1: Inclusion criteria**

| Objective | Inclusion criteria |
|---|---|
| To determine the time to onset of treatment of hypertension and hyperlipidaemia in patients with type 2 diabetes mellitus using survival analysis | All patients who had a diagnosis code (International Classification of Diseases, Tenth Revision (ICD-10) code E11) for diabetes mellitus and were receiving antidiabetic medication according to Monthly Index of Medical Specialities (MIMS) classification code 19.1 (Snyman, 2017). All antidiabetic medications were considered for this investigation. Patients among these who had ICD-10 codes for hypertension (I10, I11, I12, I13, I15, O10, and O11) and hyperlipidaemia (E78.5) and were receiving medications classified according to the National Pharmaceutical Product Index (NAPPI) codes provided by the Monthly Index of Medical Specialties (MIMS) were considered. |
| To compare different adherence measures, by determining adherence to montelukast among asthma patients, using data from a medicines claims database in South Africa | All patients who had a diagnosis code (ICD-10 code J45) for asthma in conjunction with at least two consecutive claims for montelukast based on the NAPPI code 10.4.2 according to MIMS classification (Snyman, 2017), during the study period. Patients had to be enrolled continuously with the PBM throughout the study period. |

#### 1.3.5.2 Exclusion criteria

Individuals who have incompletely filled age and sex fields were excluded from the study. The study did not include patients who were not enrolled continuously during the study period.

### 1.3.6 Data analysis

The variables included in the study, as well as the descriptive and inferential statistics employed to meet the specific objectives outlined for the empirical study is discussed in subsequent paragraphs.

#### 1.3.6.1 Study variables

Variables are the characteristics which are measured in a study (Joubert, 2014:132). According to the Centers for Disease Control and Prevention (CDC), (2012:22), an independent variable may be defined as any exposure, risk factor or characteristic that is thought to influence a manifestation or an event. Independent variables are perceived to contribute to, or precede particular outcomes (Brink *et al*., 2012:90) and can be manipulated by the researcher so as to obtain an outcome of interest (Heiman, 2014:24). Dependent variables reflect the effect of, or respond to independent variables (Brink *et al*., 2012:90) and have values that are functions of

other variables (CDC, 2012:22). The use of a variable in the context of a research project is what determines whether it is a dependent or an independent variable (Brink *et al*., 2012:90). Table 1-2 depicts the variables to that were employed in this study.

**Table 1-2: List and description of study variables**

| Variables | Description |
|---|---|
| Age | Age, defined as the time passed since birth (Pugh *et al*., 2000:34) is an important variable because almost every health-related outcome is influenced by it (CDC, 2012:40). In this study, age was determined by the date of birth, using 1 January of the following year as a reference date. Ages were stratified as narrowly as possible to allow detection of any age-related trends that may be present. Age categories from the data available were used to ensure even distribution of patients in the various groups. |
| Sex | The WHO (2017b) defines sex as "*the biological and physiological characteristics that define women and men*". In this study, patients were categorised as male and female. Patients whose sexes are unknown were excluded from the study. |
| Time to onset of treatment of disease | This is the time between the index date, when the first claim for at least one medication from the pharmacological drug class of antidiabetics based on MIMS classification code 19.1 (Snyman, 2017) is made, prior to which no claim has been made for such a drug, and the date on which the first claim is made for at least one drug from the pharmacological drug class of antihypertensives and antihyperlipidaemics based on the MIMS classification codes, in conjunction with ICD-10 codes I10, I11, I12, I13, I15, O10, and O11 for hypertension and E78.5 for hyperlipidaemia. |
| Type 2 diabetes mellitus status | Patients were categorised as having diabetes if they had an ICD-10 code E11 for type 2 diabetes mellitus and were receiving antidiabetic medication according to Monthly Index of Medical Specialities (MIMS) classification code 19.1 (Snyman, 2017). |
| Adherence | For the purpose of this study, seven validated adherence measures based on a study by Karve *et al*. (2009:989) were used to determine the adherence to montelukast for the management of asthma with medication possession ratio (MPR) serving as the reference. These adherence measures are as follows:<br>Medication possession ratio (MPR) = $\frac{number\ of\ days'\ supply\ in\ index\ period}{number\ of\ days\ in\ the\ study\ period}$<br>Proportion of days covered (PDC) = $\frac{number\ of\ days'\ suppy\ in\ index\ period}{number\ of\ days\ in\ study\ period} \times 100\ capped\ at\ 1$<br>Refill compliance rate (RCR) = $\frac{number\ of\ days'\ supply}{last\ claim\ date - index\ date} \times 100$<br>Compliance ratio (CR) = $\frac{number\ of\ days'\ supply\ in\ the\ index\ period - last\ days'\ supply}{last\ claim\ date - index\ date}$<br>Medication possession ratio, modified (MPRm) = $\frac{number\ of\ days'\ supply}{last\ claim\ date - index\ date + last\ day's\ supply} \times 100$<br>Continuous measure of medication gaps (CMG) = $\frac{total\ days\ of\ treatment\ gaps}{total\ days\ to\ next\ fill\ or\ end\ of\ observation\ period}$ |

| Variables | Description |
|---|---|
| | Continuous multiple interval measure of oversupply (CMOS) = $\frac{total\ days\ of\ treatment\ gaps\ (+)\ or\ surplus\ (-)}{total\ days\ to\ next\ fill\ or\ end\ of\ observation\ period}$<br><br>Continuous single interval measure of medication acquisition (CSA) = $\frac{days'\ supply\ obtained\ at\ the\ beginning\ of\ the\ interval}{days\ in\ interval}$ |

#### 1.3.6.2 Statistical analyses

The analysis of data was done with the Statistical Analysis System (SAS) 9.4® software (SAS Institute Inc., 2002-2012) and Statistical Package for the Social Sciences (IBM® SPSS® 25) (IBM Corp., 2017). The sections below explore the methods by which the pharmacoepidemiological data were statistically analysed.

##### 1.3.6.2.1 Descriptive statistics

Descriptive statistics describe the processes of organising data in a manner that facilitates effective communication and describes their significant characteristics (Heiman, 2014:21). The descriptive statistics used in this study included frequencies, percentages, arithmetic means (average) and standard deviations, and confidence intervals (CI) for normally distributed data. Kaplan-Meier graphs were used to describe the survival experience of patients (see application of descriptive statistics for data analysis in Table 1-3).

##### 1.3.6.2.2 Inferential statistics

Inferential statistics are applied to decide on whether the obtained data represents a difference in a particular population that is statistically significant (Heiman, 2014:22). The inferential statistical tests applied in this study included two-sample *t*-tests, Bland-Altman plots, log rank tests and chi-squared test.

- Two-sample *t*-test

According to Heiman (2014:264), the *t*-test determines the statistical significance of the difference between the means of two independent groups.

The *t*-test is calculated using the mathematical formula (Swanepoel *et al*., 2010:262):

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Where:

$\bar{x}_1 - \bar{x}_2$ are the sample means

$s_1^2$ and $s_2^2$ are the sample variances

$n_1$ and $n_2$ are the sample sizes

The two-sample *t*-test was used to compare the demographics among various patient categories.

- Bland-Altman plots

Bland and Altman introduced a method that calculates the mean difference between two methods of measurement and limits of agreement using the standard deviation (Bland & Altman, 1986:307). Giavarina (2015:143) adds that a graphical method is used in which the difference between the paired measurements is plotted against the means of these measures and recommends that 95% of the data points must lie within ±2 standard deviations of the average difference.

Different adherence measures were compared, with the medication possession ratio (MPR) as the reference, using the Bland-Altman plots.

- Log-rank test statistic

This tests the null hypothesis that in the event of an outcome occurring at any point in time, groups being examined are identical. The log-rank test is employed in survival analysis to compare the survival curves of two or more independent groups (Sullivan, 2016). The test is based on the times of the events occurrence and is most likely to identify the difference in populations when the risk of an outcome occurring in one group is higher consistently than that of another (Bland & Altman, 2004:1073).

The log-rank statistic is represented by the formula (Goel *et al*., 2010:276):

$$Log\text{-}rank\ test\ statistic = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2}$$

Where:

$E_1$ and $E_2$ are the expected number of events for each group

$O_1$ and $O_2$ are the total number of observed events for each group

The log-rank test was applied to compare the time to onset of treatment of hypertension and hyperlipidaemia in the study population.

- Chi-squared test

The chi-squared test is an essential non-parametric statistic which is used for testing hypotheses when the research involves nominal variables. The chi-square is based on assumptions that frequencies, and not percentages, are used; the study groups are independent; each study subject contributes data to only one cell and the categories of variables are mutually exclusive (McHugh, 2013:143).

The mathematical formula for computing the chi-square statistic is given as (Hoffman, 2015:185; McHugh, 2013:145):

$$\chi^2 = \frac{(O - E)^2}{E}$$

Where:

$\chi^2$ denotes the chi-square statistic

*O* denotes the observed value

*E* denotes the expected value

#### 1.3.6.2.3 Effect size

Effect size is the extent of existence of a phenomenon (Cohen, 1988:4). Cohen's *d*-value and the Pearson's correlation coefficient were used in this study.

- Cohen's *d*-value

According to Cohen and Lea (2004:60), the *d*-value is the absolute difference between the means of two populations divided by the largest standard deviation of the two means. The *d*-value is used to determine the effect size when the *t*-test is applied in a given study population. (Cohen, 1988:24).

The *d*-value is calculated as follows:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s}$$

Where:

$d$ = *d*-value

$\bar{x}_1 \; and \; \bar{x}_2$= averages of the two populations

$s$ = maximum standard deviation of the two averages

The effect size is categorised according to the following (Cohen, 1988:29):

(i) Small effect size, $|d|$ = 0.2

(ii) Medium effect size, $|d|$ = 0.5

(iii) Large effect size, $|d|$ = 0.8.

Based on recommendations by Steyn (2009:30), a $d$-value ≥ 0.8 will be considered practically significant.

A summary of the data analysis plan is provided in Table 1-3.

**Table 1-3: Data analysis plan**

| Objective | Measurement | Variables | | Statistics | | |
|---|---|---|---|---|---|---|
| | | Independent | Dependent | Descriptive | Inferential | Effect size |
| To determine time to onset of treatment of hypertension and hyperlipidaemia in patients with type 2 diabetes mellitus using survival analysis. | Demographics of study populations | | Number of patients<br>Sex<br>Age<br>Age*sex | Frequency (%)<br>Mean ± SD<br>95%CI | Two sample *t*-test | Cohen's *d*- value |
| | Time to onset of treatment of hypertension and hyperlipidaemia | Type 2 diabetes mellitus status | Time to onset of treatment | Frequency (%)<br>Kaplan-Meier graphs | Cox proportional hazards regression model<br>Log-rank test | |
| Compare different adherence measures, by determining adherence to montelukast among asthma patients, using data from a medicines claims database in South Africa | Demographics of study populations | | Number of patients<br>Sex<br>Age<br>Age*sex | Frequency (%)<br>Mean ± SD<br>95%CI | Two sample *t*-test | Cohen's *d*- value |
| | Adherence | | Medication possession ratio (MPR)<br>Proportion of days covered (PDC)<br>Refill compliance rate (RCR)<br>Compliance ratio (CR) | Mean ± SD | Bland-Altman plots | |

| Objective | Measurement | Variables | | Statistics | | |
|---|---|---|---|---|---|---|
| | | Independent | Dependent | Descriptive | Inferential | Effect size |
| | | | Medication possession ratio, modified (MPRm)<br>Continuous multiple interval measure of oversupply (CMOS)<br>Continuous single interval measure of medication acquisition (CSA) | | | |
| | MPR vs PDC vs RCR vs CR vs MPRm vs CMOG vs CSA | MPR<br>PDC<br>RCR<br>CR<br>MPRm<br>CMOG<br>CSA | Agreement between measures of adherence | Mean ± SD 95% CI | Bland-Altman plots | |

## 1.4 Ethical aspects of the study

The Board of Directors of the PBM, the Scientific Committee of the Research Entity, Medicine Usage in South Africa (MUSA), as well as the Health Research Ethics Committee (HREC) of the North-West University (NWU-00179-14-A1-06; Refer to Annexure A) permitted the conduct of the study. Information on the medical schemes, service providers, patients and prescribers were anonymised by the PBM company prior to their release for this research, ensuring data privacy and confidentiality was preserved at all times. The researchers signed confidentiality agreements to use the database for this study.

The study was of low-risk since the use of retrospective medicines claims data ensured no direct contact with identifiable patients. The benefits of the study outweighed the risks. The research was neither funded by the PBM providing the data nor the private pharmaceutical sector thus potential bias was minimised in the study.

## 1.5 Chapter summary

In this chapter, the use of administrative data for carrying out pharmacoepidemiological studies has been acknowledged. There are, however, limited publications on how various methods are applied for pharmacoepidemiological studies that employ administrative data. The aims and objectives of the study as well as method of research have also been established. The subsequent chapter discusses pharmacoepidemiology, the measures studied in pharmacoepidemiology and methods used to study these measures when secondary data are employed.

# CHAPTER 2: LITERATURE REVIEW

## 2.1 Introduction

This chapter summarises pharmacoepidemiology — its uses and relevance, as well as the measures determined in pharmacoepidemiology; conceptualises the study designs employed in pharmacoepidemiological studies with their advantages and disadvantages; and discusses data sources and methods used in pharmacoepidemiological studies.

## 2.2 Pharmacoepidemiology

Pharmacoepidemiology is considered a bridge science that joins the fields of epidemiology and pharmacology, resulting in important studies of drug effects (Spitzer, 1999:352). According to Porta (2014:184), pharmacoepidemiology is defined as "*the study of distribution and determinants of drug-related events in populations and the application of this study to efficacious treatments*". Kongpatanakul and Strom (2001:27) define pharmacoepidemiology as "*the application of epidemiologic methods, knowledge and reasoning to the subject of clinical pharmacology, whereby focussing on the study of the use of and the effects of drugs in large numbers of people*". Stergachis *et al*. (2008:67) added that pharmacoepidemiology provides essential information about the clinical and economic outcomes of drugs, biologicals and devices, especially after they have been approved for clinical use. Pharmacoepidemiology borrows from epidemiology its methods of inquiry and from pharmacology its focus of inquiry, creating a versatile research field that is rapidly developing (Evans, 2012:973; Varanasi, 2012:11; Wettermark, 2013:43; Wise, 2011:95).

The aim of pharmacoepidemiology is to give insight into and to predict drug therapy use and effects in defined populations and involves studies carried out to assess safety, effectiveness, efficacy and utilisation of drugs (Briggs & Levy, 2006:1080). Pharmacoepidemiology is a field of research that not only evaluates and describes drug use but also identifies associations or relationships in drug use and determines causal relationships between exposure to a drug or an intervention and specific outcomes (West-Strum, 2011:8).

### 2.2.1 History of pharmacoepidemiology

The growing concerns about adverse drug reactions emphasised the need to develop methods to study drug safety (Wettermark, 2013:43). Until the 1950s when chloramphenicol was identified as causing aplastic anaemia, there was little attention paid to side effects of medications (Balcik & Kahraman, 2016:58). In 1960, the United States of America (USA) Food and Drug Administration (FDA) started the collection of reports of ADRs, bringing about the institution of hospital-based drug monitoring programs (Strom, 2013a:5). According to Kongpatanakul and Strom (2001:28), the thalidomide disaster was perhaps the most significant historical event that profoundly impacted the drug regulation process. Thalidomide was marketed as a mild hypnotic and also as an anti-emetic for pregnant women in many countries but the USA. Shortly after its marketing, however, there was a surge in the number of phocomelia cases in those countries. A causal relationship between the once rare birth defect and *in utero* exposure to thalidomide was demonstrated by epidemiologic studies leading to the set-up of the Committee on Safety of Medicines in the United Kingdom in 1968 (Strom, 2013a:6).

Although thalidomide was not marketed in the USA, the impact of its side effects was so immense that the USA passed the Kefauver-Harris Amendment in 1962 and this required toxicological and non-clinical pharmacologic testing before a drug could be used in humans, along with three distinct stages of clinical testing to ensure a drug's safety and effectiveness (Kongpatanakul & Strom, 2001:29; Strom, 2013a:6). There were publications on drug utilisation studies in the mid-1960s that described how doctors used drugs resulting in a series of studies on the determinants and frequency of irrational drug prescription (Balcik & Kahraman, 2016:58). After these developments, it was thought that the 1960s birthed pharmacoepidemiology as a discipline (Balcik & Kahraman, 2016:58; Strom, 2013a:6). Although pharmacoepidemiology originated mainly from a concern about ADR documentation and minimisation, its focus has expanded to cover other health economic aspects and clinical outcomes of medication use (Wettermark, 2013:44).

### 2.2.2 Uses of pharmacoepidemiology

Pharmacoepidemiology provides reliable data that give insight into drug utilisation and outcomes, contributing to evidence-based decision-making (Rodriguez & Gutthann, 1998:421) and can be applied in all stages of drug development, marketing and use, in various industries and at different levels of healthcare (Thaker *et al*., 2015:53).

#### 2.2.2.1 Pre-approval drug development stages

At pre-approval drug development stages, pharmacoepidemiology can be used to better understand target populations and target indications and to identify the major comorbid conditions that may exist as well as any existing relationships between the drug under investigation, the comorbidities and the targeted medical condition (Wise, 2011:96). Manufacturers, at this phase, may also conduct pharmacoepidemiological studies with the hope of obtaining an earlier approval for marketing from the regulatory agency, which may be more comfortable with a sooner release of the drug because of the notion that any serious issues would be promptly and reliably detected (Strom, 2013b:56).

#### 2.2.2.2 Drug marketing

During marketing of drugs, pharmacoepidemiological studies are useful for increasing product name recognition and protecting investments made in investigating and testing the new drug. Results from pharmacoepidemiological studies, when presented publicly and published, attract the attention of prescribers and subsequently increases prescription of the drug, increasing sales. Investments made during drug development and testing can be protected when pharmacoepidemiological studies are applied to answer questions about the drug's toxicity as they may arise (Strom, 2013b:57). Thaker *et al*. (2015:56) added that with the results of pharmacoepidemiological studies, market penetration of a drug can be supported, the drug protected from adverse effects 'accusations'. Drugs already on the market can be repositioned when information on unintended beneficial effects, patient populations that were not investigated as well as different outcomes for a drug are obtained from pharmacoepidemiological studies (Thaker *et al*., 2015:56).

#### 2.2.2.3 Ethical uses

In anticipation of future product litigations, pharmacoepidemiological studies can serve as "insurance". Pharmacoepidemiological studies, by documenting the anticipated beneficial and adverse effects of drugs, can be used as a defence in the event where the product's liability becomes an issue of contention (Strom, 2013b:57; Thaker *et al*., 2015:56).

#### 2.2.2.4 Post-approval medication use

Pharmacoepidemiological research supports rational and cost-effective use of medicines and thus improves health outcomes (Lu, 2015:198). To begin with, pharmacoepidemiology, through a population-based approach, can be used to define extent and significance of a clinical problem, identify the significance of a new drug and to assess patterns of medicines use by physician class or

patient type, leading to programs that improve appropriate drug use (Avorn, 2004:82; Lu, 2015:198; Waller, 2001:166). By post-marketing surveillance, pharmacoepidemiology generates data that assess product risk assessment (Thaker *et al*., 2015:52), identify and quantify adverse events frequency and severity, and provide information to reduce problematic use (Avorn, 2004:83; Balcik & Kahraman, 2016:60; Lu, 2015:199). Such information can inform decision-making, reduce the retention of unsafe medicines or banning of safe and effective drugs (Spitzer, 1999:353). According to Qureshi *et al*. (2011:774), 26 drugs were discontinued in the USA between 1980 and 2009 on account of safety reasons, representing a safety discontinuation rate of 3.5%.

Pharmacoepidemiology aims to provide reliable data that expands the knowledge of drug use and effects in humans and is applied to estimate drug utilisation in populations, identifying areas of under- and over-utilisation (Rodriguez & Gutthann, 1998:421). The field is applied to monitor utilisation of therapeutic groups with particular anticipated problems, to evaluate the appropriate use of drug, to monitor the effectiveness of educational and regulatory activities, and to estimate and plan for drug expenditure (Avorn, 2004:83; Gama, 2008:70; Lee *et al*., 2013:400; Waning & Montagne, 2001:102). With knowledge of drug utilisation, pharmacoepidemiology reviews prescribing patterns, documenting the degree of use of a particular medication in its main indication and the extent of inappropriate prescribing and its associated consequences (Gama, 2008:70; Kildemoes *et al*., 2012:1028; Sachdeva & Patel, 2010:12).

#### 2.2.2.5 Pharmacoepidemiology as a tool for monitoring medication behaviour

Measurement of medication behaviour is an important aspect of healthcare, a role that can be played by pharmacoepidemiology (Lam & Fresco, 2015:2). The tools of pharmacoepidemiology are valuable in understanding the components of adherence which is defined as "*the extent to which a person's behaviour - taking medication, following a diet, and/or executing lifestyle changes, corresponds with agreed recommendations from a health care provider"* (Sabate, 2003:3). These tools can help understand patients' compliance with therapy, as well as reasons that may be fundamental to utilisation or access barriers, and to design and evaluate interventional programs to manage specific parts of these utilisation issues in a population (Avorn, 2004:83).

Pharmacoepidemiology provides information on the economic implications of drug use. Using economic evaluations of medicines and devices, pharmacoepidemiology identifies cost-effective products and enhance patients' access to medications by providing information that re-orients drug budgets of healthcare systems (Avorn, 2004:84). Pharmacoepidemiology is a source of data that is

significant to authorities for the setting of priorities and allocation of resources and is useful in making formulary as well as practice policy decisions (Larson & Bjornson, 1996:283).

### 2.2.3 Challenges of pharmacoepidemiology

An important issue with pharmacoepidemiology is the quality of conducting and reporting studies (Etminan *et al*., 2006:7). Pharmacoepidemiological studies' design and conduct are challenging (McMahon & McDonald, 2000:419; Tzimis, 2001:633) and can contribute to diversities and discordance in study results (Abbing-Karahagopian *et al*., 2014:131). Earlier in 2006, Etminan and colleagues (2006:7) reported a lack of standards for the conduct and report of pharmacoepidemiological findings. This problem has, however, been addressed by the introduction of guidelines for good pharmacoepidemiological practices by the International Society for Pharmacoepidemiology (ISPE) (ISPE, 2016:2).

The ISPE guidelines propose procedures and practices relevant to ensure the quality of pharmacoepidemiological studies, addressing issues related to developing study protocol, conducting, communicating, reporting and archiving adverse events (ISPE, 2016:2). The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) guide on methodological standards in pharmacoepidemiology identifies the assessment of exposures, outcomes and covariates as challenges that are linked to the conduct of pharmacoepidemiological studies (ENCePP, 2017:20). According to this guide, accurate assessment of these measures is crucial to a study's validity. The ENCePP guide, however, acknowledges difficulties in measuring exposures due to uncertainties about intake of dispensed drugs and patients' accountability of their drugs, the need for special skills for identification of outcomes using coding systems and disease areas, as well as the specific relevance of timing in assessment of covariates, all of which are challenges to be addressed in pharmacoepidemiological studies.

Availability and quality of data for studies are also areas of concern in pharmacoepidemiology. The need to make data richer and more robust to enhance evidence-based decisions, coupled with the need for rapid data access and analysis is challenging in pharmacoepidemiological studies (Thaker *et al*., 2015:56). Assembling of cases and controls has proved to be an expensive, time-consuming and difficult source of data in pharmacoepidemiological studies (Bergman, 2001:31; Thaker *et al*., 2015:55) and this has increased the appreciation of the use of databases (Andrews *et al*., 2014:199; Hallas, 2001:10; Wettermark, 2013:45). These electronic data sources, although advantageous, are not devoid of challenges. Data from databases are not specifically generated for the purposes of pharmacoepidemiological studies and may not contain relevant information such as potential risk

factors, utilisation of over-the-counter medicines and actual adherence to medicines and thus, need to be validated before they are used (Andrews *et al*., 2014:201; Evans, 2012: 976; Hallas, 2001:10; Wettermark, 2013:45).

In the estimation of the effects of a drug when prescribed, three types of errors may be seen in pharmacoepidemiological studies, namely random errors (related to reliability and precision), bias and confounding (Strom, 2013d:19). These errors challenge accurate estimation of the association between exposure and health status since drug exposure may be associated with factors that influence the event of interest. Biases pose a threat to pharmacoepidemiological studies (Ali, 2013:163; Andrews *et al*., 2014:201; Hallas, 2001:9; Lu, 2015:213; Thaker *et al*., 2015:56; Wettermark, 2013:45) since they interfere with the overall goals of pharmacoepidemiology, which are accuracy and precision, in measuring the value of a significant parameter (ENCePP, 2017:22). According to Ali (2013:162), biases distort the relationship between outcomes and exposures. Observational studies are subject to bias, identified as selection bias, misclassification, information bias, and confounding, e.g. the Simpson's paradox, and erroneous results (Hammer *et al*., 2009:664). Confounding is said to have occurred when the extent of the relationship between a health status and drug exposure is affected by the effect of external variables that are risk factors for the outcome of interest (Strom *et al*., 2013d:19). Selection bias occurs when there is the selective recruitment of subjects that are do not entirely represent the exposure or outcome pattern in the population, whereas misclassification arises from inaccurate measurement of exposure, outcome or covariate (Ali, 2013:162; ENCePP, 2017:22; Lu, 2015:214). Although various techniques can be applied in the design and analyses stages of pharmacoepidemiological studies to control these, biases still remain a great challenge to pharmacoepidemiology (Lu, 2015:218; Wettermark, 2013:45).

Various techniques have been employed to control such errors in pharmacoepidemiological studies. According to Nørgaard *et al*. (2017:186), the use of restriction, stratification, propensity score matching, self-controlled designs, randomisation and sensitivity analyses have been effective in the control of errors in pharmacoepidemiological studies. Confounding and bias can be controlled at various stages of a pharmacoepidemiological study. During the design phase, randomisation, restriction and matching can be applied to limit these errors (HealthKnowledge, 2018). Randomisation involves an arbitrary selection of study participants whereby every subject in the study population has an equal chance of selection. Restriction is the limitation of study participants to individuals with similar characteristics where the confounder is concerned while matching is the technique of selecting controls that are as similar as possible to cases in terms of the potential confounders. (HealthKnowledge, 2018).

The effects of confounders can also be minimised during the analyses of studies by the use of stratification, propensity scores, multivariate analyses and standardisation, among others (HealthKnowledge, 2018; Nørgaard *et al*., 2017:186). Stratification is applied by dividing the study population into various strata based on predetermined levels of the potential confounder and calculating the relative risks for each stratum (Jager, Zoccali *et al*., 2008:258). A propensity score, the probability of receiving an intervention of interest based on certain observed covariates, can be matched in pharmacoepidemiological studies to reduce bias and confounding and to increase precision (Austin, 2011:402; D'Agostino, 1998:2267). In studies where the number of potential confounders or the levels of their groups are considered large, multivariate analyses are better suited for controlling bias and confounding because they are able to analyses a large number of covariates simultaneously. Several methods of multivariate analyses are used in pharmacoepidemiological studies to control errors and these include logistic regression, linear regression and analysis of covariates (Pourhoseingholi *et al*., 2012:80). Although these techniques can be applied in the design and analyses stages of pharmacoepidemiological studies to control these, biases still remain a great challenge to pharmacoepidemiology (Lu, 2015:218; Wettermark, 2013:45).

Health research requires careful attention to be paid to ethical considerations and pharmacoepidemiology is no exception (Dennis & Lozano, 2001:613). There exist many ethical challenges for pharmacoepidemiology that have been unresolved and have left possibilities of adverse ethical consequences (United Nations (UN), 2004:12). Coughlin (2006:4) asserts that minimising potential harm and risks to study subjects is an ethical role of researchers and rigorously protecting the confidentiality of study subjects is crucial to minimising risks. El Khoury (2013:257) adds that researchers are faced with the challenge of ensuring that privacy and confidentiality are protected to promote accurate and efficient realisation of research goals. Investigators also have the ethical responsibility of providing study participants with information that can be used to decide if to consent to the study or not (Coughlin, 2006:5). Researchers are thus required to inform subjects that their involvement in the study is optional and that subjects can terminate their participation at any time with no consequences (El Khoury, 2013:259).

The challenge to researchers lies in deciding the level of information to provide to participants about the study and the types of information delivery, accounting for the literacy and communication modes that may be preferred, avoiding duress or intimidation and identifying who may provide consent rightfully in the context of set guidelines and these issues are further compounded by a decision to pay participants as this may raise questions about subjects' voluntary consent (UN, 2004:12). The

goal is to have a balance on ethical issues so that procedures are not prohibitively difficult for researchers (El Khoury, 2013:257).

### 2.2.4 Future of pharmacoepidemiology

The future of pharmacoepidemiology appears clear judging from past and current trends and although it appears bright, many challenges remain (Strom *et al*., 2013:407). It is expected that there will be continuous growth in the array of methods available for pharmacoepidemiological studies (Garbe & Suissa, 2014:1914; Lu, 2015:219; Strom *et al*., 2013:409; Wettermark, 2013:48), with potential increase in the application of pharmacoepidemiological insight to the performance of clinical trials and the use of randomised designs to address issues that are handled using observational pharmacoepidemiology (Strom *et al*., 2013:408). An increased use of propensity scores, sensitivity analyses and instrumental variable analyses has been predicted for controlling confounding and time-varying exposures in the future (Garbe & Suissa, 2014:1914; Strom *et al*., 2013:409).

Lu (2015:219) forecast an increase in data availability. It is anticipated that there will be many more databases available for pharmacoepidemiological studies due to the increased computerisation of data, enabling researchers to address questions about drug effects quickly and at a less expensive cost, making use of large sample sizes afforded by these databases (Andrews *et al*., 2014:196; Strom *et al*., 2013:410).

There is a likelihood that all new drugs in the future will be monitored after they are launched (Strom *et al*., 2013:413; Wettermark, 2013:46). Payers and other reimbursement agencies are expected to increasingly request for observational studies that estimate the worth of medicines whereas patients will also insist on improved systems of drug safety and efficacy monitoring (Lu, 2015:219; Wettermark, 2013:46). There is expected to be increased investment in external resources for pharmacoepidemiology by pharmaceutical companies which will lead to increased growth of the field (Strom *et al*., 2013:411).

It is anticipated that pharmacoepidemiology will be linked with the most recent genetic techniques, molecular biology, biochemistry and immunology (Garbe & Suissa, 2014:1914; Strom *et al*., 2013:408). According to Avorn *et al*. (2013:76), pharmacoepidemiology can apply biological evolutions such as pharmacogenomics to augment systems of prescription event monitoring. It is envisioned that pharmacogenomics will also be added to the process of reporting adverse reactions (Strom *et al*., 2013:408) and there will be a surge in the availability of genotypic information for large

populations for whom drug exposures and outcomes have been electronically recorded (Garbe & Suissa, 2014:1914; Strom *et al*., 2013:408).

The World Health Organization's report on Priority Medicines for Europe and the World (Kaplan *et al*., 2013:67) identifies certain diseases for which treatment either does not exist or is inadequate in reaching affected individuals. Pharmacoepidemiology can be applied in the future to improve on treatment in such disease areas by identifying and addressing obstacles to ensure safe and better delivery of medicines to patients (Wettermark, 2013:47). The report highlights the importance of addressing peculiar needs of the elderly, women and children. Data from automated databases can be utilised in future for observational studies on drug use in the elderly, women and children (Wettermark, 2013:47).

Pharmacoepidemiology continues to be a field for collaboration among relevant stakeholders such as payers, regulators, prescribers, patients, manufacturers and the general public. The field is predicted to take up a more extensive multidisciplinary and multiprofessional nature, requiring alliances with decision makers for drug formularies, specialists within specific clinical areas, health economists and health policy researchers (Lu, 2015:219; Wettermark, 2013:48). Although the field of pharmacoepidemiology is rapidly advancing, it is still plagued with challenges and is one of the most dynamic developing fields in research.

### 2.3 Measures of pharmacoepidemiology

Pharmacoepidemiology is concerned with measuring the source, distribution, utilisation and effects of drugs in a population, determining the frequency and distribution of outcomes of drug use and from this information, determining the nature and extent of specific types of drug use (Waning & Montagne, 2001:5). To fully realise this, pharmacoepidemiology makes use of the spectrum of epidemiologic approaches to answer questions on causality and incidence of drug use effects, effectiveness of drugs, prescription patterns, as well as the economic impact of drug use (Kongpatanakul & Strom, 2001:27). According to Balcik and Kahraman (2016:58), pharmacoepidemiology evaluates healthcare systems and patients' health-related behaviours with the aim of increasing individuals' quality-of-life by improving general drug use. Pharmacoepidemiology assesses patients' adherence to medicines (McCaffrey, 2011:136), occurrence, frequency and magnitude of adverse drug events and reactions (Banahan, 2011:156), drug utilisation patterns and economics of drug use (Wettermark *et al*., 2008:161).

### 2.3.1 Incidence

Pharmacoepidemiologists use measures of risk to approximate the likelihood of an event occurring and to measure the extent of a possible relationship between a drug and an event of interest (Quartey *et al*., 2011:549). Incidence is one of the measures of risk that reflect the frequency of a health event (*inter alia*, disease, adverse event, death, births) (Quartey *et al*., 2011:549) and gives a description of disease occurrence (Yang, 2011:24).

Incidence estimates the number of new cases of a disease or other relevant health outcomes (HealthKnowledge, 2012:2; Joubert *et al*., 2014:22; Waning & Montagne, 2001:109; Yang, 2011:19). According to Yang (2011:24), incidence is best applied for the study of disease causes and evaluation of disease prevention programs.

Two types of incidence measures exist: incidence rate and cumulative incidence (HealthKnowledge, 2012:2; Quartey *et al*., 2011:549; Waning & Montagne, 2001:22; Yang, 2011:19). Incidence rate, also referred to as 'incidence density', is defined by Waning and Montagne (2001:23) as "*the measure of rapidity with which a new condition develops in a population; the rate at which newly diagnosed patients are identified over time, measured by actual observation time*" and is expressed as a fraction of the number of new cases to the total person-time of observation in a population at risk (Waning & Montagne, 2001:23; Yang, 2011:20). The Dictionary of Pharmacoepidemiology (Begaud, 2000:74) states that incidence density is particularly useful in expressing an association between a risk and the exposure time and is also a suitable measure when the length of exposure or follow-up differs for the different research subjects.

Cumulative incidence, otherwise called 'incidence proportion' is the proportion of a population who developed the condition of interest in a specified period among those who were initially disease-free (HealthKnowledge, 2012:3; Yang, 2011:19). Incidence proportion indicates the chances of developing the condition of interest within a specified period (HealthKnowledge, 2012:3; Waning & Montagne, 2001:24) and is used primarily in populations which are fixed and in which losses to follow-up are rare (Yang, 2011:19). According to Quartey *et al*. (2011:549), incidence rate is advantageous over cumulative incidence since it accounts for subjects who may become lost to follow-up during the study and would affect the denominator inappropriately.

### 2.3.2 Prevalence

Prevalence measures the proportion of people with a disease or a relevant health outcome in a defined specified population (HealthKnowledge, 2012:2; Joubert *et al*., 2014:22; Waning &

Montagne, 2001:109; Yang, 2011:19) and considers both old and new cases. Prevalence is more indicative of disease burdens in populations (Quartey *et al*., 2011:549), and is useful in predicting health resource needs in a community (Yang, 2011:24).

There are two types of prevalence measures; point prevalence and period prevalence. Whereas point prevalence refers to the number of people with a condition of interest in a population at a point in time, period prevalence is the proportion of persons with a particular disease during a time interval (HealthKnowledge, 2012:6; Waning & Montagne, 2001:21; Yang, 2011:22). Point prevalence gives a snapshot of the population (Yang, 2011:22) whereas period prevalence takes into account both point prevalence and incidence in its computation (Waning & Montagne, 2001:21).

Prevalence is influenced by incidence, duration of disease, interventions and treatments as well as prolongation of life programs in chronic disease management; increasing with the addition of diseased people and the elimination of people who are healthy (Waning & Montagne, 2001:21).

### 2.3.3 Drug utilisation

Kildemoes *et al*. (2012:1027) assert that recent decades have seen dramatic increments in the utilisation of many classes of drugs, possibly caused by the expansion of prescription indications informed by evidence of favourable outcomes in patients other than those for whom the drugs were intended. Drug utilisation is a vital part of pharmacoepidemiology; describing the scale, characteristics as well as determinants of medication exposure (Kumar *et al*., 2013:321; Shalini *et al*., 2010:804) and providing a base from which subsequent qualitative studies can be carried out (Truter, 2008:92). Begaud (2000:53) defines drug utilisation study as "*a study designed to describe — qualitatively and quantitatively — the population of users of a given drug (or class of drugs) and/or the conditions of use (e.g. indications, duration of treatment, dosage, previous or associated treatments, compliance, etc.)*." The main objective of drug utilisation studies is to ensure judicious use of medications in populations (Kaur *et al*., 2014:1; Kumar *et al*., 2013:322; Sachdeva *et al*., 2010:12) and offer avenues for monitoring the use of medications and for acquiring knowledge on prescribing behaviours which can be applied for regulating drug-related costs (Kumar *et al*., 2013:321).

Drug utilisation studies are continuous and systematic quality improvement procedures that assess use of drugs and/or prescribing patterns (Gama, 2008:69; Kumar *et al*., 2013:321; Sashdeva *et al*., 2010:12), alert the irrational use of medicines, give feedback to prescribers, and develop quality standards for describing ideal drug use (Sashdeva *et al*., 2010:12). According to Kaur *et al*. (2014:1),

drug utilisation studies are effective for evaluating hospital formularies for efficiency and cost-effectiveness and critically appraising hospital drug policies, making recommendations for improvement. Varied drug use information such as information regarding overall drug use, drug groups, generic compounds or specific products, obtained from various sources such as the drug-use chain, large databases, regulatory agencies and drug use evaluations are applied in drug utilisation studies depending on the issues being addressed (WHO, 2003:13). Drug utilisation studies are essential in pharmacoepidemiology because of the close relationship with the other relevant fields of public health, pharmacovigilance, pharmacogenetics and pharmacoeconomics (Sashdeva *et al*., 2010:12).

### 2.3.4 Adverse drug reactions

Medicines can pose significant harm to patients and this may be caused by medication errors or by adverse effects of the medications (Leendertse *et al*., 2008:1890). Van den Bemt *et al*. (2000:323) categorised drug-related problems into two: those caused by an error in the drug production and distribution process and those that occur devoid of errors in the process. Drug safety is crucial to overall patient health and is an important aspect addressed by pharmacoepidemiology (Lovely *et al*., 2015:104). Edwards and Aronson (2000:1255) define adverse drug reactions as "*an appreciably harmful or unpleasant reaction, resulting from an intervention related to the use of a medicinal product, which predicts hazard from future administration and warrants prevention or specific treatment, or alteration of the dosage regimen, or withdrawal of the product*". Schatz and Weber (2015:5) added that ADRs occur in a drug's usual clinical use, affect morbidity and mortality and are influenced by the surge in the number of drugs available on the market, the aging population as well as increasing polypharmacy. Adverse drug reactions affect healthcare costs significantly and increase the length of hospital stay of affected patients (Van den Bemt *et al*., 2000:327).

#### 2.3.4.1 Types of adverse drug reactions

Strom (2013a:4) characterises ADRs into two: type-A reactions that are related to doses, common, predictable and less serious compared to the type-B reactions that may be due to hypersensitivity or immunologic reactions and usually require drug cessation and adds that type-B reactions, because of the difficulty in predicting and detecting, are the major focus of many pharmacoepidemiological researches. Pharmacoepidemiological designs such as cohorts, case-controls and meta-analyses can be applied for the identification of ADRs (Edwards & Aronson, 2000:1258; Waning & Montagne, 2001:102) and this affords advantages such as the ability to measure the occurrence rates and to identify related risk factors (Waning & Montagne, 2001:102).

#### 2.3.4.2 Sources of adverse drug reactions

Medication errors and drug-drug interactions are probable causes of ADRs that can be prevented if necessary caution is taken (Lovely *et al.*, 2015:104). The role of medication errors, drug-drug interactions and drug-disease interactions as potential causes of ADRs are addressed in sections 2.3.4.2.1-2.3.4.2.3.

##### 2.3.4.2.1 Medication errors

A medication error is a situation which occurs when a wrong drug is given, or a right drug is wrongly administered to a patient (Lee *et al*., 2013:378). Medication errors are classified according to where they occur in the drug use cycle. These include prescribing errors which occur when an incorrect drug, dose or quantity is prescribed, dispensing errors which occur at the dispensing phase — including wrong medication, strength and quantity of medicines supply as well as administration errors resulting from wrong drugs being administered at wrong doses or through wrong routes to wrong patients at wrong times (Rehan & Bhargava, 2015:1; Van den Bemt *et al*., 2000:327).

Insufficient patient information and lack of knowledge on the drug are contributors to the incidence of medication errors (Williams, 2012:688) and methods such as voluntary reporting by patients, medical rounds, directly observing actual patients on hospital wards, urine testing and tracer analyses have been traditionally used to detect medication errors (Rehan & Bhargava, 2015:2). The high cost of primary data collection in medication error studies limits the sample sizes of such studies, a problem that is addressed using a pharmacoepidemiological approach whereby electronic databases are used to improve on sample sizes in a manner that is cost-effective (Strom, 2013c:120).

##### 2.3.4.2.2 Drug-drug interactions

Drug-drug interactions (DDIs) arise when one drug, or more, influences the pharmacodynamics and pharmacokinetics of other drugs (Hennessy *et al*., 2016:92). A drug-drug interaction is characterised by the effect a drug has on the safety and efficacy of another drug and is an essential consideration to be made in determining the right medication for a patient (Chelkeba *et al*., 2013:144). Pharmacoepidemiology, using large databases with the appropriate data, affords an important means by which the health consequences of potential DDIs can be studied and is critical to ensuring the confirmation or otherwise of the effects of potential DDIs (Hennessy *et al*., 2016:102).

#### 2.3.4.2.3 Drug-disease interactions

When a medication prescribed for the treatment of a pre-existing condition exacerbates comorbid diseases, syndromes or conditions, a drug-disease interaction is said to have occurred (Lindbald *et al*., 2006: 1134). Drug-disease interactions increase in patients who receive multiple medications (Duobova *et al*., 2007:5), being most prevalent in elderly chronic disease patients (Aspinall *et al*., 2015:84) because of their physiological changes and the presence of comorbidities (Lau & Tenney, 2017:1). Several pharmacoepidemiological researches have studied the prevalence of drug-disease interactions and their burden (Aspinall *et al*., 2015:84; Duobova *et al*., 2007:1; Lindbald *et al*., 2006:1133; Nwose & Ye, 2016:1). Pharmacoepidemiological methods are vital in determining these drug-disease interactions and preventing their occurrence among susceptible populations through pharmacovigilance surveillance (Pan *et al*., 2013:103).

### 2.3.5 Economic consequences of medication use

Most developing countries are increasingly becoming concerned about the rising expenditure on healthcare in general and of pharmaceuticals in particular (Hill *et al*., 1997:421; Schulman *et al*., 2013:280) and policy makers are faced with the challenge of managing limited healthcare resources (Schulman *et al*., 2001:37). Cost has thus become a criterion for the evaluation of medications, requiring the use of well-designed research methodologies that guide decision makers (Schulman *et al*., 2013:281).

Pharmacoepidemiology assists in the decision-making process by making data of high quality about the consequences of drug therapy available and assisting in the identification of the most effective and safest treatment options (Briggs & Levy, 2006:1081; Larson & Bjornson, 1996:283). Pharmacoepidemiological studies are applied to evaluate the burden of disease, addressing questions related to direct and indirect expenditure that account for the economic burden associated with diseases and support the field of pharmacoeconomics in the determination of strategies and inputs into economic impact models (Manack *et al*., 2012:165).

Patients are becoming more concerned about the efficacy and safety of medications and yet attentive to the rising costs of these medications while third party payers are concerned with their expenditure on pharmaceuticals and resultant value obtained. Healthcare providers have access to limited resources requiring that they assess cost-effectiveness of pharmaceuticals stringently. A merger of pharmacoepidemiology with pharmacoeconomics is therefore the way to go to meet the needs of all stakeholders (Schulman *et al*., 2013:290).

### 2.3.6 Medication adherence

Adherence to prescribed medicines is critical to the success of pharmacotherapy (Andrade *et al*., 2006:565; Cramer *et al*., 2008:44; Farmer, 1999:1075; Hess *et al*., 2006:1280; Vrijens *et al*., 2012:691).

Adherence is defined as how patients take medications as prescribed (Vrijens *et al*., 2012:696); how patients' drug-taking behaviour coincides with recommendations by the healthcare provider (Grégoire & Moisan, 2016:369).

There are three constructs that make up adherence: initiation or acceptance signifying the patient's agreement with the recommended treatment and subsequent intake of the first dose of the treatment, compliance or implementation which involves following the prescribed schedule of treatment, and discontinuation which marks the end of the therapy, where treatment is stopped permanently. The time frame between acceptance and discontinuation is termed persistence (Grégoire & Moisan, 2016:369; Urquhart, 2001:473; Vrijens *et al*., 2012:696).

Figure 2-1 (adapted from Cramer *et al*., 2008:46; Grégoire & Moisan, 2016:369; Urquhart, 2001:473 and Vrijens *et al*., 2012:696) depicts the three constructs of adherence and how they relate to each other.

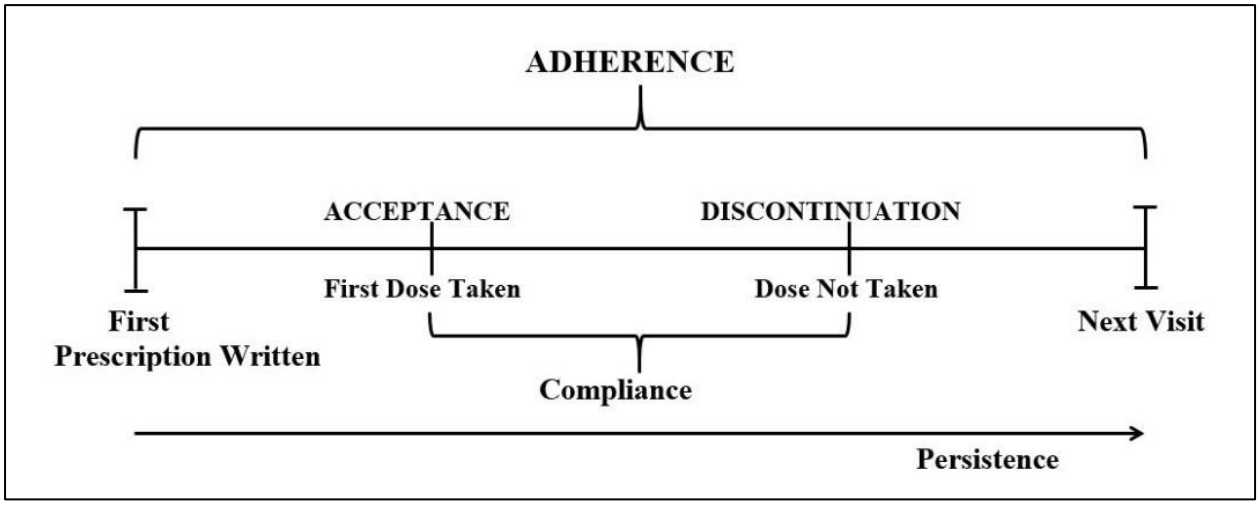



**Figure 2-1: Constructs of adherence**

According to Cramer *et al*. (2008:46), compliance can be defined as the process of abiding with healthcare provider recommendations that pertain to the dosage, timing and frequency of medicines,

reflecting the magnitude of agreement with the prescribed regimen of therapy, while persistence refers to the time from the initiation of therapy to when it is discontinued permanently. Persistence is measured in time, the common unit being days although months and years could also be used (Burnier, 2006:1191; Peterson *et al*., 2007:6).

### 2.4 Study designs used in pharmacoepidemiology

Morroni and Myer (2014:78) refer to a study design as a structured procedure which a researcher follows to address a research question. Study designs used pharmacoepidemiology differ in terms of the number of observations being made, units of observations (whether individual or group), period of time between measurement of exposure and disease, methods by which data are collected and the direction in which exposure is measured relative to outcome (Friis & Sellers, 2009:242). The choice of study design to be employed in pharmacoepidemiology must be considered critically as this has important consequences for the interpretation of results (Sturmer & Brookhart, 2013:21).

Study designs may be experimental or interventional, whereby the investigator manipulates the exposure of interest, or observational in which there is no interaction between the investigator and the exposure being studied (Friis & Sellers, 2009:243; Morroni & Myer, 2014:78; Strom, 2013d:27; Waning & Montagne, 2001:45).

Observational study designs may also be descriptive or analytical. Descriptive observational studies (case reports, case series and cross-sectional studies) give insight into the progress of disease and drug-related problems in a group of people whereas analytical observational studies (ecological studies, case-control studies and cohort studies) rely on the use of new data and are applied to examine causal relationships (Waning & Montagne, 2001:46).

In experimental studies, the researcher aims to influence outcomes by treatment or by altering a disease determinant (Morroni & Myer, 2014:91), compare the advantages of an intervention with current or no treatment or show the existence of a causal relationship (Waning & Montagne, 2001:63). Experimental studies can be distinguished from quasi-experimental studies by the presence of randomization, the process whereby study subjects have equal chances of being assigned to one study group or another (Friis & Sellers, 2009:244; Harpe, 2011:41).

The order in which exposure information is gathered relative to the occurrence of a disease or drug problem is also used to classify study designs (Rothman *et al*., 2008:95). Studies may be prospective, in which the study population is identified at the beginning of the study and observed forward in time, or retrospective, in which the research question is addressed using data that have been previously

collated and recorded (Morroni & Myer, 2014:88; Waning & Montagne, 2001:46). Novel study designs have been postulated and successfully applied to address certain shortfalls in already existing study designs with particular reference to sampling. For example, nested case-control, case-cohort and multi-time case-control designs are new sampling designs within cohort studies which allow accurate estimation of relative risk measures (Rothman *et al*., 2008:122; Schneeweiss & Suissa, 2013:327). There are also designs applicable to pharmacoepidemiology that synthesise knowledge using parts of clinical experience or medical literature and statistical analyses in an attempt to clarify uncertainty and to facilitate policy level decision-making. Meta-analysis and decision-analysis are examples of such quantitative synthesis study designs (Petitti, 2000:2).

Figure 2-2 (compiled from Friis & Sellers, 2009:243; Harpe, 2011:41; Morroni & Myer, 2014:78; Petitti, 2000:2; Rothman *et al*., 2008:95; Strom, 2013d:27; Schneeweiss & Suissa, 2013:327 and Waning & Montagne, 2001:46) depicts the study designs used in pharmacoepidemiology.

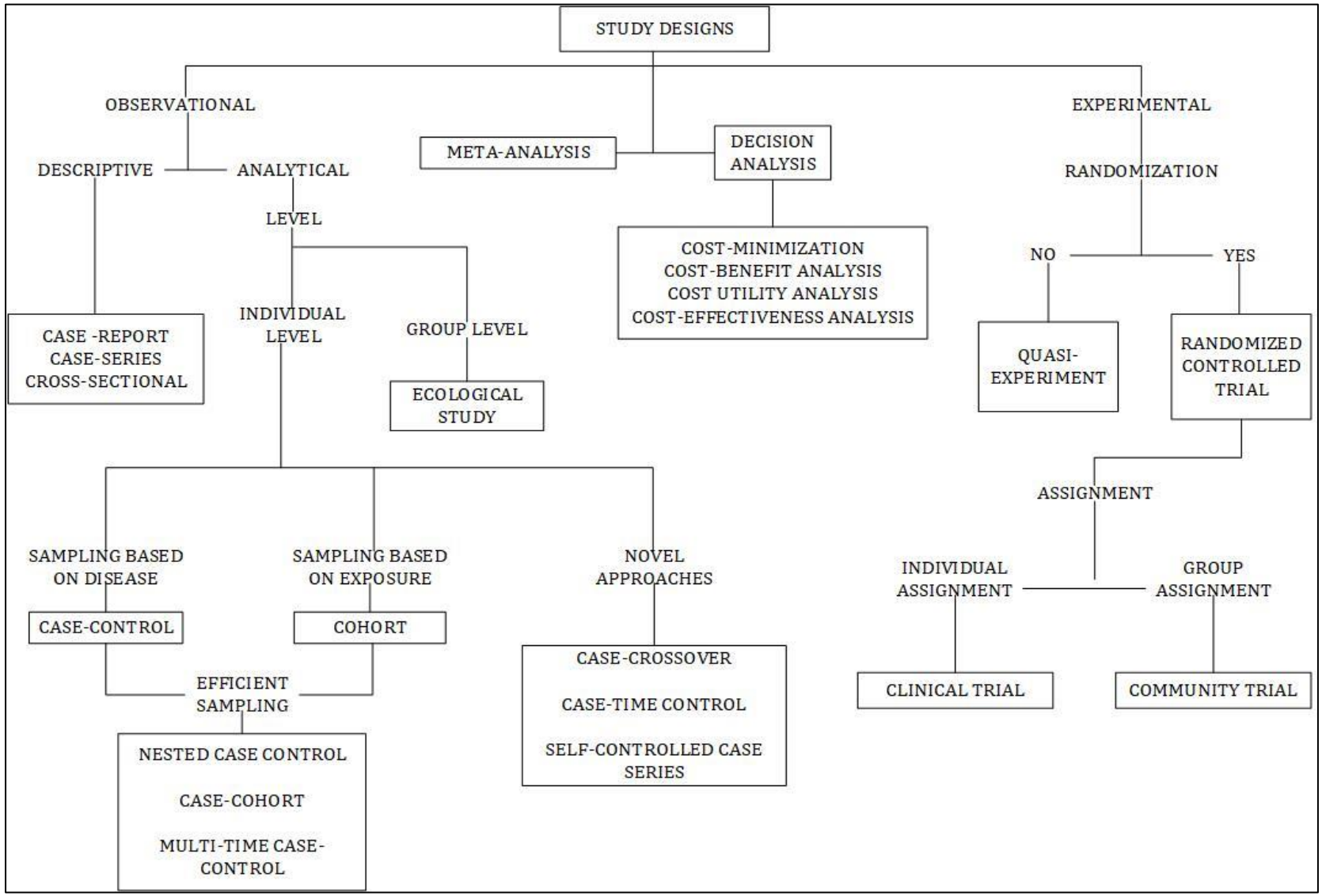



**Figure 2-2 Study designs used in pharmacoepidemiology studies**

Each study design will be discussed further, together with its characteristics, strengths, shortfalls, and applications and summarised in Table 2-1 to highlight the similarities and differences that exist among them.

### 2.4.1 Case reports and case series

A 'case report' is a study design that details an individual patient's clinical experience with a specific medication, and a 'case series' is one that describes the clinical experience of a collection of patients with a common exposure and investigation into their clinical outcomes (Harpe, 2011:44; Strom, 2013d:22). Case reports are deemed the most basic form of descriptive study designs, generating hypothesis about the effects of drug which can be further investigated by more stringent study designs (Friis & Sellers, 2009:146; Strom, 2013d:23), while case series offer a larger population for observation permitting inference of summary measures that assist in drawing firmer conclusions than with case reports (Friis & Sellers, 2009:146).

Despite the relatively simplistic nature of case reports and case series, they are useful study designs as they provide less expensive and easy methods for raising hypotheses (Bhopal, 2002:233; Harpe, 2011:44; Strom, 2013d:22). Bhopal (2002:233) asserts that case series can be applied to study signs and symptoms of diseases as well as mortality and morbidity rates and to define diseases. Verhamme and Sturkenboom (2010:69) add that case series are important for the detection of safety signals of drugs. Case reports and case series, however, do not offer enough proof of causal relationships, limiting their usefulness in pharmacoepidemiological studies (Bhopal, 2002:233; Strom, 2013d:22; Verhamme & Sturkenboom, 2010:69).

### 2.4.2 Cross-sectional studies

Cross-sectional studies, also known as prevalence studies, involve measuring both exposure and outcomes among a group of people at a specified point in time (Bhopal, 2002:242; Harpe, 2011:45, Verhamme & Sturkenboom, 2010:69; Waning & Montagne, 2001:51).

The focus of the cross-sectional study design is on simultaneous collection of information on disease or drug-related problem, characteristics of the population as well as the risk factors (Bhopal, 2002:242; Harpe, 2011:45). According to Morroni and Myer (2014:82), cross-sectional studies are comparatively simple and less expensive to carry out and are applicable for evaluating the healthcare needs of a given population. Prevalence studies can yield valuable details on the burden of a disease and other important variables in a population (Bhopal, 2002:245; Verhamme & Sturkenboom,

2010:69). These studies are also vital in forecasting the future spread of diseases and health-related conditions (Waning & Montagne, 2001:52).

As with the other descriptive studies, cross-sectional studies cannot be used to make inferences about any possible cause-effect relationships (Friis & Sellers, 2009:262; Verhamme & Sturkenboom, 2010:69; Waning & Montagne, 2001:52). Cross-sectional studies are not effective designs to be used for identifying seasonal variations, diseases which have short incubation periods or those for which rates in a population are low (Waning & Montagne, 2001:52).

### 2.4.3 Ecological studies

Ecological studies or analyses of secular trends are designs that study patterns in an exposure and an outcome that are thought to be cause and effect, respectively, and examines the possible overlap of these patterns for a population over time or across geographical boundaries (Strom, 2013d:23; DiPietro, 2010:976). Groups or collections of people are the unit of analysis in ecological studies (Bhopal, 2002:262; Friis & Sellers, 2009:249) and it is required that relevant information on the populations of interest be available for exposure and distributions of diseases to be assessed (Rothman *et al*., 2008:99). Ecological studies involve the application of vital statistics (Strom, 2013d:23), which are usually analysed using correlation analysis (Friis & Sellers, 2009:251).

Analyses of secular trends play an important role in studying variables that can only be measured aggregately, such as environmental factors, and help to study the influence of such variables on health (Morroni & Myer, 2014:91). Strom (2013d:24) and Morroni and Myer (2014:91) agree that ecological studies are simple and rapidly generate hypotheses as well as evidence in support or against already existing hypotheses. Ecological studies have been found to be of particular significance in cases where little variation exists within the population with reference to a relevant exposure (Morroni & Myer, 2014:91). This study design is, however, plagued by the disadvantage of '*ecological fallacy',* a bias in which conclusions are drawn at a single patient level after aggregate data has been used (Friis & Sellers, 2009:248; Harpe, 2011:44; Morroni & Myer, 2014:91). Findings in ecological studies are complex to interpret since the relationship between exposure and outcomes are not measured in each individual and it is not possible to examine whether exposed individuals developed the significant outcome (DiPietro, 2010:976; Morroni & Myer, 2014:91). Confounding variables are difficult to account for in ecological studies since the analyses do not differentiate factors are likely to cause an outcome among exposures whose trends coincide with disease trends (Strom, 2013d:24).

### 2.4.4 Case-control study design

Case-control design involves an analytical, observational investigation that is comparative by nature, contrasting people who have a disease or drug-related problem (cases) with those without it (controls) with reference to precursory exposures (Bhopal, 2002:247; Gordis, 2014:190; Strom, 2013d:24; Szklo & Nieto, 2007:23; Waning & Montagne, 2001:52). According to Waning and Montagne (2001:53), a distinct characteristic of case-control study is the identification of cases based on their disease status and the design entails the formulation of study questions, selection and collation of cases, selection and matching of controls, exposure determination in all study subjects, and the subsequent analysis and interpretation of data

Among the benefits offered by case-control designs is their relatively quick, easy and cost-effective computing (Friis & Sellers, 2009:276). Case-control designs require smaller sample sizes and rely on retrospective data that facilitate their rapid performance at relatively less expensive costs (Strom, 2013d:24; Waning & Montage, 2001:52). In the study of multiple causes for a particular outcome, case-control studies are of particular significance as the same cases and controls can be used to investigate several risk factors and exposures concurrently (Harpe, 2011:47; Morroni & Myer, 2014:86; Strom, 2013d:24). By using case-control designs, rare diseases and diseases with longer latency periods can efficiently be studied since with this design, subjects already have the disease of interest (Morroni & Myer, 2014:86; Strom, 2013d:24; Waning & Montagne, 2001:52).

Limitations of case-control designs include their inability to estimate incidence and risks of diseases, difficulties in identifying cases and controls that are representative and the presence of selection, information and recall biases (Friis & Sellers, 2009:276; Harpe, 2011:47; Morroni & Myer, 2014:86; Strom, 2013d:24; Waning & Montagne, 2001:52). Waning and Montagne (2001:52) and Harpe (2011:48) concur that case-control designs are inefficient in the study of rare exposures such as drugs that are scarcely used.

### 2.4.5 Cohort study design

A cohort is defined as a group of people or a population that can be distinguished by some characteristics they have in common (Bhopal, 2002:251; Friis & Sellers, 2009:252; Rothman & Greenland, 2008:100). A cohort study design is thus one in which the investigator identifies a cohort (consisting of both exposed and unexposed subjects) who are free of a disease of interest and follows them up forward in time, relating them to subsequent development of the disease or outcome (Etminan & Samii, 2004:964; Gordis, 2014:179; Sturmer & Brookhart, 2013:24). Although cohort

studies basically compare exposed and unexposed subjects, they can occasionally be applied to compare different exposures (Strom, 2013d:24) and allow for actual estimation of incidence since they follow up on subjects to identify development of outcomes (Sturmer & Brookhart, 2013:24). Figure 2-3 (adapted from Bhopal, 2002:247; Gordis, 2014:190; Strom, 2013d:25; Szklo & Nieto, 2007:23 and Waning & Montagne, 2001:52) gives a pictorial illustration of the cohort study design.

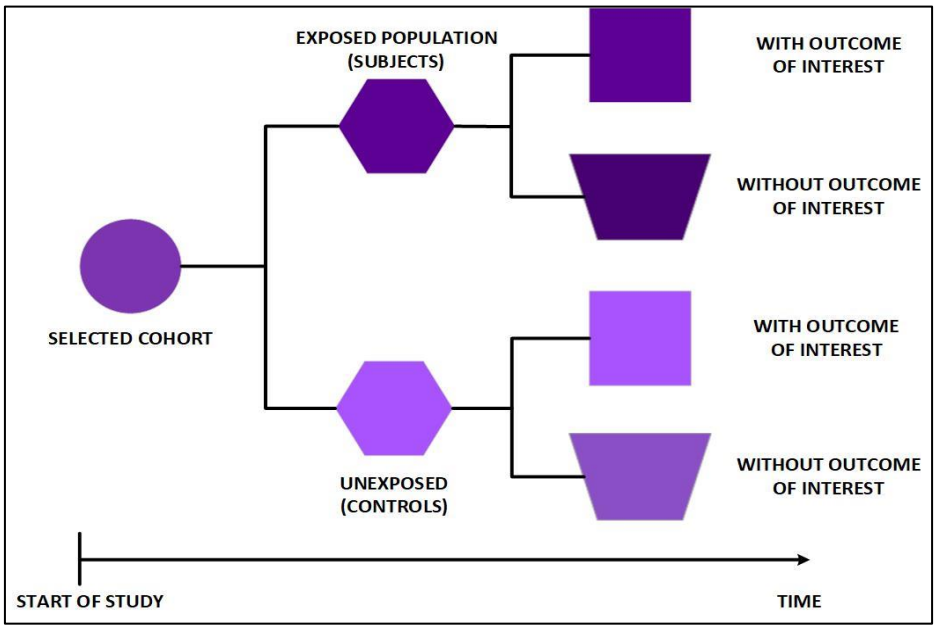



**Figure 2-3 Cohort study design**

Contrary to the assertion made by Waning and Montagne (2001:56) that cohort study designs are prospective by nature, Strom (2013d:25), Szklo and Nieto (2007:22), Gordis (2014:183) and Harpe (2011:49) argue that cohort studies may be prospective, retrospective or of a mixed approach depending on when the cohort is identified comparative to the commencement of the study. In the prospective (concurrent) cohort study, the investigator identifies the cohort at the beginning of the study and follows them forward through calendar time for the occurrence of the outcome of interest (Morroni & Myer, 2014:88). Cohorts are identified in the past using existing information and observed to the present time in retrospective (historical) (Szklo & Nieto, 2007:23) whereas it is possible to perform a cohort study in which exposure data are based on existing information and follow up and outcome measurement continue into the future (Gordis, 2014:183; Morroni & Myer, 2014:88). Cohort studies may also be open or dynamic, in which the size can vary with time with the joining or leaving of subjects, or closed, in which subjects may not enter nor leave the study (Harpe, 2011:49; Waning & Montagne, 2001:57).

The identification and selection of study subjects who are free of an outcome of interest and subsequent follow-up to identify the occurrence of the outcomes make cohort designs suitable for propounding relationships between exposures and outcomes (Harpe, 2011:48; Morroni & Myer, 2014:89; Strom, 2013d:25). Furthermore, from a single exposure, cohort studies can be applied to study several outcomes (Harpe, 2011:49; Morroni & Myer, 2014:89; Waning & Montagne, 2001:56) an advantage that is significant in postmarketing drug surveillance studies which focus on identifying potential effects of a new drug (Strom, 2013d:25). According to Friis and Sellers (2009:318), the cohort study design permits the determination of risks and incidence of diseases and provides evidence of time lapse between exposure and incidence of disease. However, studies that use the cohort design are plagued with problems of being expensive, time-consuming and requiring large sample sizes for conduct (Friis & Sellers, 2009:318; Strom, 2013d:25; Waning & Montagne, 2001:56). Cohort studies are susceptible to limitation to sample size as a result of loss-to-follow up (Friis & Sellers, 2009:318; Morroni & Myer, 2014:89, Waning & Montagne, 2001:56) and are inefficient for studying rare outcomes and diseases (Morroni & Myer, 2014:89; Waning & Montagne, 2001:56).

### 2.4.6 Study designs based on efficient sampling within a cohort study

Recent years in research have seen considerable focus on exploiting features and advantages of cohort and case-control studies by combining elements of both and the development of hybrid or ambi-directional study designs (Gordis, 2014:203; Szklo & Nieto, 2007:27). According to Schneeweiss and Suissa (2013:326), these hybrid designs (nested case-control design, case-cohort design and multi-time case-control design) that depend on sampling within cohorts' counter time, cost and resource restraints of traditional cohort designs and account for multiple confounding factors by matching cases and controls within cohorts. These efficient sampling designs rely on selection of subjects with the event of interest from the cohort but can be distinguished from each other on the basis of the approach for the selection of controls or non-cases (Gordis, 2014:203; Schneeweiss & Suissa, 2013:326; Szklo & Nieto, 2007:27).

In nested case-control study designs, a risk-set sampling approach is used whereby cases are compared with cohort members at chance of the outcome being investigated at the time each case occurs (Etminan, 2004:1106; Schneeweiss & Suissa, 2013:327) and matched on calendar time and follow-up length (Gordis, 2009:205). By this method of control selection, a control may later develop the disease and become a case (Gordis, 2014:205; Szklo & Nieto, 2007:27), and the study subject may thus be a control more than once in the study, necessitating the time-matched nature of cases and controls selection (Schneeweiss & Suissa, 2013:327). The nested case-control study design

allows for better confounding control through its use of matching (Etminan, 2004:1107; Friis & Sellers, 2009:316), better measurement of time-dependent exposures (Etminan, 2004:1107) as well as less complex statistical analysis (Etminan, 2004:1107; Szklo & Nieto, 2007:27). Limitations with this design arise where subjects are not permitted to be selected more than once, leading to the introduction of bias in the relative risk estimation since longer-term cases who do not develop the disease during follow up are more likely to be selected as controls (Schneeweiss & Suissa, 2013:327).

To improve the efficiency of the traditional case-control design when dealing with time-variant exposures and acute effects, the multi-time case-control study design expands the number of observations per control — an important approach when the cost at which additional controls can be obtained, is high (Suissa *et al*., 2010:882). This approach increases odds ratio precision in a case-control study in which there are transient or acute time-varying exposures. Suissa *et al*. (2010:881) assert that this study design is innovative because it estimates the probability of exposure among controls using an expanded time interval approach. The study design works on the assumption that the ratio does not change with duration of exposure and thus the multi-time case-control is appropriate for acute exposures but not chronic exposures or where latencies are being studied (Suissa *et al*., 2010:882).

Each participant in the cohort has an even chance of selection as a control, irrespective of the case or control status, in a case-cohort study design (Gordis, 2014:205; Szklo & Nieto, 2007:28). Unlike in the nested case-control design, the time a subject has contributed to the person-time encounter of the cohort is of no relevance in this study (Gordis, 2014:206; Rothman *et al*., 2008:123). A significant benefit of this design is its ability to facilitate the conduct of different case-control studies from a single cohort because controls are not matched (Gordis, 2014:205; Rothman *et al*., 2008:123; Szklo & Nieto, 2007:28). The case-cohort design has a simple sampling approach, which is an added advantage because it is easily comprehensible and facilitates simpler computer programming and allows for convenient external comparisons to be made (Schneeweiss & Suissa, 2013:327). Risk factor distribution, as well as prevalence rates that are required for the estimation of attributable risks, can be obtained when case-cohort study designs are used where the baseline cohort is representative of its source population (Szklo & Nieto, 2007:29). Schneeweiss and Suissa (2013:327) note that case-cohort designs are limited in analysis which has to account for the overlap of members of the cohort between successive risk sets. Another disadvantage identified by Rothman *et al*. (2008:124) is the need for selection of more controls to manage the imbalance caused by membership overlap and to ensure that statistical precision is maintained.

#### 2.4.6.1 Case-crossover, case time-control and self-controlled case series study designs

Assessment of risks of acute adverse events that result from transient exposures is challenging in pharmacoepidemiology due to the acute nature of the adverse events, confounding by indication and difficulties in determining exposure timing (Schneeweiss & Suissa, 2013:328). Conventional study designs have not been efficient in studying such acute events and this has led to the development of case-crossover, case time-control and self-controlled case series; new study designs that are able to address these complexities.

The following sections discuss novel study designs applied in pharmacoepidemiology to address the limitations of the conventional study designs.

#### 2.4.6.2 Case-crossover study designs

The case-crossover design is used to study acute outcomes in conditions where the exposure is thought to be transient with its effect taking place within a short time (Gordis, 2014:206; Szklo & Nieto, 2007:32). With this design, a case acts as its own control, with the exposure pattern being compared when an outcome occurs (event time) and at some other prior time (control time) (Etminan & Samii, 2004:966; Harpe, 2011:51; Szklo & Nieto, 2007:32). The differences in exposure rates at these times are subsequently used to approximate an odds ratio of the event related to the said exposure (Delaney & Suissa, 2009:54).

Three assumptions are critical to the effective carrying out of case-crossover studies: the study must inevitably involve an acute adverse event suspected to be caused by a temporary drug exposure, the effect period must be known accurately and data on the drug exposure pattern for each case must be reliably obtained over a sufficiently long time period (Schneeweiss & Suissa, 2013:329). Control selection bias is eliminated, and resource savings made in this study design as cases serve as their own controls, ruling out the need to recruit and collect information on another control group (Donnan & Wang, 2001:259; Gordis, 2014:208).

An important strength of this design lies in its inherent ability to account for time-invariant confounders, including indications for drug use (Delaney & Suissa, 2009:59; Etminan & Samii, 2004:966; Schneeweiss & Suissa, 2013:328; Szklo & Nieto, 2007:33). The case-crossover study design, however, cannot be used for the study of drugs used to manage chronic conditions that require constant use of medication (Donnan & Wang, 2001:259) nor for the study of adverse events that happen because of long periods of exposure (Schneeweiss & Suissa, 2013:329). Confounders that change over time can also not be adjusted for by this design (Delaney & Suissa, 2009:57) and

this causes a bias in the risk ratio estimation (Delaney & Suissa, 2009:57). Etminan and Samii (2004:967) and Gordis (2014:208) believe recall bias may be a problem in this design as cases may not be able to recollect accurately their exposure to drugs being studied.

An expansion of the case-crossover study design that deals with confounding by indication, taking into account drug utilisation changes over time, is the case-time-control design (Donnan & Wang, 2001:260; Etminan & Samii, 2004:967; Schneeweiss & Suissa, 2013:329). Confounding by indication is thought to occur when an outcome is associated with a particular exposure when in reality, it is as a result of disease severity (Etminan & Samii, 2004:967). This approach accounts for patterns in drug use over time and provides unbiased odds ratio estimates in the presence of confounding by indication even though drug indication is not measured because within-subject analysis is used (Schneeweiss & Suissa, 2013:329).

#### 2.4.6.3 Self-controlled case series designs

Self-controlled case series designs use data on cases only to study temporal associations between time-variant exposures and adverse events (Whitaker *et al*., 2009:7). The design is thus best suited to transient exposures and acute outcomes in which exposure risk periods are defined clearly (Petersen *et al*., 2016:1; Whitaker *et al.*, 2009:13). This approach is considered bidirectional (Hallas & Pottegård, 2014:584; Maclure *et al.*, 2012:51) because of its inclusion of controlled time from before and after the exposure (Maclure *et al.*, 2012:51).

To be applicable, the self-controlled case series depends on three assumptions: that outcomes occur in a non-homogeneous Poisson process, that when an outcome arises, it does not alter the chances of occurrence of a successive exposure and that an outcome's occurrence neither influences nor censors the period of observation (Whitaker *et al*., 2009:12).

Self-controlled case series are efficient designs in terms of resources and sample size (Layton & Shakir, 2012:326) as well as efforts in data collection (Whitaker *et al*., 2009:10). Adjusting for time-invariant confounding factors is an essential attribute of the self-controlled case series design (Layton & Shakir, 2012:326; Petersen *et al*., 2016:2; Whitaker *et al*., 2009:11). Hallas and Pottegård (2014:587) noted that this design may be used for the study of drugs with indefinite use, adding, however, that the design is limited by its dependence on complex data processing and analysis.

### 2.4.7 Interventional designs

In interventional study designs, the researcher actively exerts some influence on outcomes by changing the determinants of exposure or through treatment (Harpe, 2011:41; Morroni & Myer, 2014:91). Interventional or experimental studies are prospective in nature and usually investigate therapies that minimise disease occurrence or preventive interventions (Waning & Montagne, 2001:64; Morroni & Myer, 2014:91). A vital significance of interventional designs is their ability to determine causal relationships (Harpe, 2011:40).

Although randomisation is a key feature of pure experimental designs that enables control of confounding, it is not always feasible to randomly assign study subjects and for this reason, quasi-experimental designs play a significant role in research (Harpe, 2011:41; Harris *et al*., 2006:16; Morroni & Myer, 2014:91; Thompson & Panacek, 2006:245; White & Sabarwal, 2014:2).

#### 2.4.7.1 Randomised controlled trials (RCTs)

Randomised controlled trials are deemed the most precise pure experimental study design, involving the random assignment of study subjects to an intervention or control group under the control of the investigator (Friis & Sellers, 2009:330; Harpe, 2011:41; Morroni & Myer, 2014:91). The results of an RCT are evaluated by comparing the occurrence of the outcome under investigation between the control and intervention groups (Morroni & Myer, 2014:91). Randomised controlled trials may be divided into clinical trials that assess outcomes in individual subjects or community trials that investigate effects in a large group of people (Friis & Sellers, 2009:330; Morroni & Myer, 2014:91; Rothman *et al*., 2008:87; Waning & Montagne, 2001:46).

According to Lesko and Mitchell (2013:269), RCTs are reckoned the ‘gold standard’ against which other study designs are judged. This standard can be attributed to its strength of being the only real test of causal relationship (Harpe, 2011:40; Lesko & Mitchell, 2013:269; Waning & Montagne, 2001:64).

Blinding, the process in which study subjects are unaware of the group of assignment, is employed in RCTs, an added advantage of RCTs since it reduces biases in assessment (Friis & Sellers, 2009:337; Waning & Montagne, 2001:64). Morroni and Myer (2014:95) add that RCTs provide the most robust information about drug safety and efficacy.

Randomised controlled trials are not devoid of problems. The design is complex, time-consuming, expensive and artificial; not reflecting the natural environment in which drugs are used (Morroni &

Myer, 2014:95; Lesko & Mitchell, 2013:272; Waning & Montagne, 2001:64) and may not be feasible on account of ethical, legal and logistical issues (Friis & Sellers, 2009:340; Morroni & Myer, 2014:95; Rothman *et al*., 2008:88; Waning & Montagne, 2001:64).

#### 2.4.7.2 Quasi-experimental study designs

Quasi-experimental study designs, also known as 'before-after' or 'pre-post intervention' design, aim to assess the effect of an intervention without randomisation (Harris *et al*., 2006:17; Layton & Shakir, 2012:832; Morroni & Myer, 2014:96). The design is important in situations where randomisation is not possible because of logistical, ethical, political or sample size challenges (Harpe, 2011:42; Harris *et al*., 2006:17; White & Sabarwal, 2014:2) but where repeated measurements before and after a change can be made (Harpe, 2011:42). This approach is well-suited to the evaluation of new policies or programs, for the determination of the impact of formulary changes on the prescription behaviours of prescribers (Harpe, 2011:41).

According to White and Sabarwal (2014:10), quasi-experimental designs allow for the conduct of evaluations in settings that reflect the real environment of use of medications. The design can be applied to validate methods of therapy or to establish potential causal relationships, sometimes serving as a precursor study for establishing a rationale for further studies using true experimental designs (Thompson & Panacek, 2006:245). Quasi-experimental study designs are, however, limited by their lack of random assignment which hinders the ability to account for vital confounders, as well as regression to the mean, which can lead to inaccurately concluding that an outcome is as a result of an intervention when it is actually due to chance (Harris *et al*., 2006:18; Layton & Shakir, 2012:832).

### 2.4.8 Quantitative synthesis study designs

Meta-analysis and decision analysis seek to link research and practice (Byers & Stullenbarger, 2003:194) by synthesising knowledge, from clinical experience or medical literature, to provide answers to defined problems (Petitti, 2000:2). These designs are important for disseminating research findings to policy makers and practitioners (Byers & Stullenbarger, 2003:202) and are considered quantitative designs in that they use statistical and numerical analysis to create wholesome evidence that are relevant for clinical and public policy formulation in healthcare (Petitti, 2000:2). Already in 2003, Byers and Stullenbarger (2003:193) noted that meta-analysis has been the benchmark of synthesis research for the past few decades, with decision analysis emerging as an alternative in more recent years.

#### 2.4.8.1 Meta-analyses

Porta (2014:154) defines meta-analysis as "*a statistical analysis of results from separate studies, examining sources of differences in results among studies, and leading to a quantitative summary of the results if results are judged sufficiently similar to support such synthesis*". Where individual studies do not have sufficient power to establish associations, meta-analyses are valuable since they explore heterogeneity in a collection of studies and variations in responses as well as generalisabilities that can precipitate more potent management or modifications in treatment (Attia *et al*., 2002:297; Haidich, 2010:30). Meta-analyses have both qualitative and quantitative components involving classifying studies according to features that have been pre-determined and statistically analysing information (Porta, 2014:154). The design relies on the premise that results combined from a group of investigators permit more precise and balanced effect evaluation than those from individual studies (Doi & Thalib, 2007:19; Haidich, 2010:30).

Applications of meta-analysis in pharmacoepidemiology include, *inter alia*, investigation of adverse effects, assessment of the use of existing therapies for emerging indications, and assessment of divergent outcomes among subgroups of patients (Berlin *et al*., 2012:740).

The benefits of meta-analyses include their ability to shorten time of research and to increase sample size and power for detecting drug effects as well as their resource efficiency (Berlin *et al*., 2012:743; Haidich, 2010:36).

Of critical concern is the issue of publication bias resulting from failure to include studies that are unpublished (Berlin *et al*., 2012:726; Einarson, 2008:397; Haidich, 2010:34; Harpe, 2011:51). Meta-analyses are also subject to limitations such as biases that are inherent in the original researches being combined and difficulties in combining excessively diverse studies (Berlin *et al*., 2012:726).

#### 2.4.8.2 Decision analyses

Decision analysis is a quantitative study design that uses a sequential approach to analyse decisions under varying conditions and incorporates preferences, utility and outcomes (Byers & Stullenbarger, 2003:196). The design presents decision problems of interest as models and offers researchers the avenue to apply evidence-based medicine (EBM) to specific clinical scenarios (Aleem *et al*., 2008:138).

Results from decision analyses are applied to make decisions on patient management, to make policy recommendations and to inform individuals about choices where therapies are concerned

(Petitti, 2000:3). A decision analysis becomes an economic evaluation when it incorporates cost and plays a vital role in funding decisions concerning health (Prosser *et al*., 2012:3). Cost-effectiveness, cost-benefit, cost-minimisation, and cost-utility analyses are types of economic analyses used in pharmacoeconomics (Hoomans & Severens, 2014:2; McGhan, 2010:4). Cost-effectiveness analysis compares medical strategies that produce a common outcome (Hoomans & Severens, 2014:3) with the goal to establish whether the outcomes expected from a preferred strategy are worth the costs (Ryder *et al*., 2009:220) whereas cost-utility is a subset of cost-effectiveness analysis which assesses outcomes using special outcome measure such as quality-adjusted life-years (QALY) and takes patients' preferences into account (Rascati, 2014:72). Cost-minimisation analysis is used to compare the costs of interventions known to have equivalent medical effects with the aim of determining which intervention is the least expensive approach to achieving an outcome for a population (York Health Economics Consortium (YHEC), 2016a). Cost-benefit analysis compares interventions and their outcomes, expressing both costs and benefits in fiscal terms and allowing for the comparison based on a summary metric of net monetary benefit (YHEC, 2016b).

Decision analysis is unique in its ability to model multiple variables under various scenarios and in its ability to be adapted to individuals or groups of patients to facilitate decision-making. In addition, the approach permits the researcher to assess the relative worth of various options and to disseminate findings in a format that is appreciated and better understood by practitioners and policy makers (Byers & Stullenbarger, 2003:201). The most challenging feature of decision analysis is the accurate quantification of utility (Aleem *et al*., 2009:23; Byers & Stullenbarger, 2003:201) which is defined as the worth placed on an outcome based on personal or societal preference (Porta, 2014:61). According to Aleem *et al*. (2008:138), decision analysis is complex, requiring the researcher to take several dimensions of a scenario into consideration, it is time and information intensive and requires rigorous and robust analyses.

Table 2-1 (compiled from the references cited in section 2.4) summarises the study designs used in pharmacoepidemiological research in terms of their characteristics, strengths and weaknesses.

**Table 2-1: Summary of study designs employed in pharmacoepidemiological studies**

| Study design | Properties | Strengths | Weaknesses | Measures of association | Example of applications |
|---|---|---|---|---|---|
| Case reports<br>Case series | Descriptive observational<br>Investigate the natural history of outcomes and case definitions<br>For hypotheses generation | Easy to compute<br>Relatively inexpensive | There is not enough evidence to establish causal relationships | No measure of association | A study of iatrogenic subcutaneous emphysema of endodontic origin (Mishra *et al*., 2014:279) |
| Cross-sectional study design | Descriptive observational<br>Investigate disease distribution in a population and the characteristics of the population | Easy and quick to carry out<br>Relatively inexpensive | Inferences cannot be made about possible cause-effect associations<br>Incidence cannot be identified<br>Seasonal variations cannot be studied | Prevalence ratio | An investigation of the role of inflammation and cardiovascular risk in endothelial function in rheumatoid arthritis patients (Sandoo *et al.*, 2012:1) |
| Ecological studies (Analyses of secular trends) | Observational<br>Groups of people are the units of analysis<br>Results reported on the overlap of trends in exposures and outcomes | Simple and rapid generation of hypotheses<br>Inexpensive | '*Ecological fallacy'*<br>Findings from this design are complex to interpret<br>It is difficult to control confounding when this design is used | Correlation coefficient | Study of the relationship between mortality and suicide in England with reference to antidepressant prescribing (Morgan *et al*., 2004:1) |
| Case-control study design | Analytical observational<br>Study participants are chosen based on outcome status | Requires a relatively smaller sample size<br>Studies multiple exposures<br>Rare outcomes and diseases with long latency periods can be studied<br>Relatively faster and less expensive to carry out | Does not estimate incidence and outcome risks<br>Rare exposures cannot be studied<br>Susceptible to selection, information and recall bias | Odds ratio | A study of acute and clinically relevant drug-induced liver injury (de Abajo *et al*., 2004:71) |

| Study design | Properties | Strengths | Weaknesses | Measures of association | Example of applications |
|---|---|---|---|---|---|
| Cohort | Analytical observational<br>Study subjects selected based on exposure | Several outcomes can be studied<br>Incidence and outcome risks can be estimated<br>Allows for proposal of temporal relationships<br>Investigate rare exposures | A large sample size is required<br>There may be loss to follow up if performed prospectively<br>Inefficient for investigating rare outcomes<br>May be costly and time consuming if performed prospectively | Relative risks | A research into the relationship between non-steroidal anti-inflammatory drugs and the risk of serious coronary heart disease (Ray *et al*., 2002:118) |
| Nested case-control | Based on coherent sampling within a cohort<br>Type of case-control conducted within a defined cohort<br>Involves matching based on time of enrolment of cases and other specified variables | Matching allows for controlling confounding<br>Time-dependent exposures can be measured and studied<br>Requires less complex statistical analysis | Loss-to-follow-up<br>If outcome of interest is rare, the results may not be representative of all controls | Odds ratio | A study of the outcomes of non-steroidal anti-inflammatory drugs on the risks and stages of breast cancer (Sharpe *et al.*, 2000:112) |
| Multi-time Case-control | Based on efficient sampling within a cohort<br>Measures exposure at more than one point in time leading to an increased number of observations per control | Increases the precision of the odds ratio<br>Allows for estimation of probability of exposure among controls<br>Appropriate for the study of acute exposures | Not appropriate for the study of chronic exposures or where latency is being studied | Odds ratio | |
| Case-cohort | Based on efficient sampling within a cohort<br>Controls are selected randomly from the cohort irrespective of time contributed to person-time experience | Different case-control studies can be conducted from a single cohort<br>A simple sampling approach is used | Analysis must account for possible overlap of cohort members<br>There may be a need to select more controls to manage the possible overlap | Odds ratio | An investigation of drug-associated agranulocytosis (Van der Klauw *et al*., 1999:369) |

| Study design | Properties | Strengths | Weaknesses | Measures of association | Example of applications |
|---|---|---|---|---|---|
| Case-crossover | Assesses individual at the point of exposure and at another time when there is no exposure<br>Case serves as own control and is studied at the event and control time<br>Investigate transient exposures and acute outcomes | Eliminates control bias<br>Accounts for time-invariant confounders | Chronic events cannot be studied<br>Time-dependent confounding variables cannot be controlled<br>Susceptible to recall bias<br>Limited use in claims data | Odds ratio | A study of montelukast and the risk of Churg-Strauss syndrome (Hauser *et al*., 2008:677) |
| Case-time-control | Expansion of case-crossover that accounts for drug utilisation over time | Accounts for confounding by indication<br>Can be used to investigate time trends in drug use and exposures | Time-dependent confounders cannot be controlled | Odds ratio | A research into the risk of congenital abnormalities associated with the use of phenobarbital, phenytoin, or diazepam in pregnancy and (Kjær *et al*., 2006:181) |
| Self-controlled case series | Bidirectional study that uses data on cases only to investigate temporal relationship between time-variant exposures and outcomes<br>Investigate transient exposures and acute outcomes | Resource efficient<br>Small sample size required<br>Adjusts for time-invariant confounders<br>Investigate drugs with indefinite use | Data processing and analysis is complex | Odds ratio | An investigation of the relationship between intake of orlistat and the risk of acute liver trauma (Douglas *et al*., 2013:1) |

| Study design | Properties | Strengths | Weaknesses | Measures of association | Example of applications |
|---|---|---|---|---|---|
| Randomized controlled trial | Interventional study design: investigator exerts control<br>Randomisation is used<br>May be carried out on an individual (clinical trial) or on a group of people (Community trial) | Controls for unknown confounders<br>Only real test of causal relationships<br>Blinding is used to reduce bias<br>Provides robust information on drug efficacy and safety | Complex<br>Time consuming<br>Most expensive design<br>May not be feasible due to legal, ethical and logistical constraints | Relative risk | A study of the effect of chronic vitamin E supplementation on cardiovascular events and cancer (Walsh, 2005:1823) |
| Quasi-experiments | Interventional: researcher exerts control<br>No randomisation is used<br>Evaluates new programs and policies | Allows for the validation of various therapies<br>Establishes potential causal relationships | Absence of randomisation introduces bias<br>Regression to the mean | Relative risk | An examination of the differences in the use of antidepressants by young people and suicidal tendencies after FDA warning and media coverage (Lu *et al*., 2014:1) |
| Meta-analyses | Quantitative synthesis design<br>Classifies studies according to predetermined features and statistical analyses to generate conclusions about a body of research | Shortens research time<br>Increases sample size and power<br>Resource efficient<br>Good source of evidence for clinical decision making | Publication bias<br>Susceptible to biases inherent in original studies<br>Susceptible to difficulties in combining excessively diverse studies | If based on RCT, Relative risk is used<br>If based on observational studies,<br>Relative risk or odds ratio used | An analysis of the epidemiology of corticosteroid-induced osteoporosis (Van Staa *et al*., 2002:777) |

| Study design | Properties | Strengths | Weaknesses | Measures of association | Example of applications |
|---|---|---|---|---|---|
| Decision analyses | Quantitative synthesis design<br>Uses chronological methods to analyse decisions<br>Can be economic decision analysis if cost is incorporated:<br>Cost-effectiveness<br>Cost-utility<br>Cost-minimisation<br>Cost-benefit | Allows for the assessment of relative worth of various options<br>Allows results to be presented and disseminated in a model form which is better appreciated by practitioners and policy makers<br>Allows for tailored assessment of decision options based on individual patients<br>Incorporates patients' preferences where utilities are obtained from patients | Process is complex<br>Requires accurate quantification of utilities<br>Time and information sensitive<br>Requires rigorous statistical analyses | | Decision analyses of treatment for patients who had biochemical failure after therapy on prostate cancer<br>(Shimizu *et al*. 2007:763) |

## 2.5 Sources of data for pharmacoepidemiological research

According to Friis and Sellers (2009:204), the findings of a study are only as good as the data upon which it is established, an assertion that is relevant to pharmacoepidemiological studies as well.

Pharmacoepidemiological studies utilise data that are prospectively collected for a study's specific purpose (*de novo* collection of data), known as primary data or secondary data which have previously been collated for some other purposes (Harpe, 2011:55). Primary data, collated in the field for pharmacoepidemiological studies, are called "field studies" or "*ad hoc* studies" and typically comprise of cohort, case-control, cross-sectional designs, post-marketing randomised surveillance, case-control surveillance, and prescription-event monitoring, among others (Kaufman, 2013:178). Secondary data are, however, pre-existing, having been collected for other purposes such as for addressing a previous hypothesis or for some administrative process and offer the researcher limited control over the quality and accuracy of information received (Harpe, 2011:56).

According to Richter *et al*. (2016:84), the use of primary data allows for the effective monitoring of patients and thus ensures high quality of data which, together with the availability of clinical data, offers vital insights that may not be readily available when secondary data is used. Primary data also permit the investigator's influence over the type and amount of data collected as well as the validation of such data (Harpe, 2011:56). Collecting primary data is, however, relatively more expensive and time-consuming than collecting secondary data (Mullner, 2009:368; Harpe, 2011:56) and may be subject to certain biases, such as selection and recall bias, which are often affiliated to data collection procedures (Harpe, 2011:56). Secondary data may, however, be incomplete for research because they are typically collected for other reasons, may lack information on vital confounding variables and may be difficult to validate (Harpe, 2011:56; Torre & Martins, 2012:140).

### 2.5.1 Field or *ad hoc* studies

Field or *ad hoc* studies are of particular significance in instances that require detailed information, such as particulars about over-the-counter medication use, habits and quality of life information that cannot be obtained from secondary data (Kaufman, 2013:178). Strom (2013e:195) categorises data obtained from case-control surveillance, prescription-event monitoring, and registries as *ad hoc* sources of data.

#### 2.5.1.1 Case-control surveillance (CCS)

Case-control surveillance uses a case-control approach to evaluate the consequences of exposures and medications on the risk of severe diseases and includes the monitoring of prescription and non-prescription drugs, biologic components, as well as dietary supplements (Rosenberg *et al*., 2006:138).

Case-control surveillance relies on the interview of patients for long-term regular use of medication and details on potential modifiers and variables that confound the drug-disease association and subsequent comparison between cases with the event of interest and controls with other illnesses (Rosenberg *et al*., 2012:287). The capacity and ability to evaluate prescription and non-prescription medications, as well as dietary supplement use; identify and control potential confounding factors, provide accurate diagnosis classification; assess the effects of exposures using its high statistical power and evaluate if outcomes of long-term drug use are beneficial traits of this source of data. Case-control surveillance is, however, limited by potential selection and reporting biases and exposure misclassification as well as lack or inaccurate dosage information (Rosenberg *et al*., 2012:291).

#### 2.5.1.2 Prescription event monitoring (PEM)

New medicines intended to be used extensively in general practice as well as older drugs for which new indications have been found are targeted for active surveillance by prescription-event monitoring (PEM) and investigated by contacting general practitioners to request for medication use and outcomes information on medications use (Shakir, 2006:152). Prescription event monitoring uses an observational cohort approach and questionnaires as the major data-collection tool (Layton & Shakir, 2012:302).

Among the strengths of the PEM are the ability to provide large sample sizes for research, the representative nature of the data with regard to patients being treated in “real-world” clinical practice, the use of validated dispensed prescriptions to determine exposure which allows for accurate reporting of exposure data and the identification of patients with potential adverse events for further studies. The effectiveness of PEM is hindered by restriction of use to general practice only, selection bias, confounding and the use of a single-group cohort design which prevents the comparison with an unexposed comparator (Layton & Shakir, 2012:302; Shakir, 2006:152).

#### 2.5.1.3 Registries

Gliklich *et al*. (2014:13) define patient registries as "*an organised system that uses observational study methods to collect uniform data (clinical and other) to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure, and that serves one or more predetermined scientific, clinical or policy purposes*".

Registries may be derived from a geographical area or from a clearly-defined underlying population, may be representative of an entire treated population or may depict patients who voluntarily enrol, may have comparators within a registry or may concentrate on a medical intervention or healthcare delivery approach. Conclusions drawn from a study based on a registry will thus differ subject to what is examined, and the sampling approach used (Dreyer & Velentgas, 2012:331).

Registries are essential for elucidating the natural history and course of a disease, investigating determinants of quality of life, assessing patient outcomes, studying variations in interventions and their resultant outcomes, and appraising both clinical and cost effectiveness of interventions (Gliklich *et al*., 2014:17). Dreyer and Velentgas (2012:334) add that registries gather information on a broad range of factors, allowing for the study of exposures, outcomes and potential modifiers and confounders. Registries can be adapted over time for the conduct of new research or lead to new hypotheses that can be further investigated but their diverse nature sometimes render them efficient only for addressing broad questions while the approach of participant recruitment may also subject them to selection bias (Dreyer & Velentgas, 2012:334).

### 2.5.2 Secondary data sources

#### 2.5.2.1 Post-marketing spontaneous pharmacovigilance reporting systems

Spontaneous reporting systems have been an important source of data for pharmacoepidemiological research (Dweik *et al*., 2016:2; Pan *et al*., 2013:101). Waning and Montagne (2001:133) defined spontaneous reporting systems as "*formal reporting systems designed to record, collate, and analyse the occurrence of adverse drug reactions*". The concept of spontaneous reporting systems hinges on voluntary reporting of suspected adverse events to designated centres or to pharmaceutical manufacturers who are subsequently obliged to forward the information to regulators (Pan *et al*., 2013:103).

Data collected from reporting centres are entered into adverse event databases (Pan *et al*., 2013:108) and assessed for patterns that suggest certain drug-outcome associations (Pan *et al*., 2013:108; Waning & Montagne, 2001:133), serving subsequently as a basis for further evaluation,

action by regulatory bodies or both (Pan *et al*., 2013:108). Observations made after marketing of drugs and at the point of care are often vital for enhancing existing knowledge on the safety profile of medications since they offer a channel to generate hypotheses on the safety of such drugs in the absence of other systematic approaches, ensuring the active participation of healthcare practitioners in post-marketing drug safety evaluation (Pan *et al*., 2013:108).

Pan *et al*. (2013:105) suggest that for an adverse event report to be complete, it must contain information on product name, demographics, adverse event description, confounders, start and stop dates of medication intake, as well as date of event onset, dosage regimen for drug use, results from biopsy or autopsy, and de-challenge and re-challenge, since the quality of the report submitted determines the validity of safety issues that are based on it. Spontaneous reporting account for all types of legitimate drugs used in all settings and are designed to be able to identify variables that are relevant to this goal, giving reporters the appropriate means of reporting all details of an adverse event (ENCePP, 2017:14). The strengths of this data source, as applied in pharmacoepidemiology, lie in its ability to facilitate safety signal detection that prompt study into proposed safety issues and, where necessary, action by the regulators; to offer a relatively inexpensive approach to study safety issues on a broad scope of medicines and to engage health professionals in drug safety surveillance, permitting more vigilant clinical observations and reporting of necessary information which may be absent in other sources of data such as the automated databases (Pan *et al*., 2013:112). Rare drug-related events may be identified using spontaneous reporting systems whiles comparisons may also be made among different drug groups (Avorn *et al*., 2013:74).

Limitations of the spontaneous reporting systems comprise its reliance on the quality of individual reports (ENCePP, 2017:15; Pan *et al*., 2013:113); the voluntary nature of reporting in most jurisdictions; the proportionality to duration of time a drug is available on the market, reporting being highest in the initial post-release phase and fading away as practitioners become more familiar with the drug; and the absence or insufficient information on confounding factors (Waning & Montagne, 2001:134). Potential duplication of reports which may misrepresent clinical conclusions and adversely affect statistical findings is also a disadvantage of the spontaneous reporting systems (ENCePP, 2017:15). Another significant problem of spontaneous reporting systems is underreporting which makes it impracticable to estimate population-based rates of adverse events since not all events in the population are submitted and the extent to which particular events are underreported cannot be accurately known (Pan *et al*., 2013:113).

#### 2.5.2.2 Automated databases

Automated databases offer a cost-effective and efficient approach to studying medications after they have been released onto the market and are an essential source of data for pharmacoepidemiological research (Strom, 2013c:120; Torre & Martins, 2012:132). Takahashi *et al*. (2012:124) describe automated databases as electronically stored healthcare utilisation information that have been collected on a regular basis in daily clinical practice and note their increased use in healthcare research in recent years. Hennessy (2006:311) believes that automated databases can be used to study research questions that address issues about medication use and outcomes and assess the effects of programs devised to enhance drug use.

##### 2.5.2.2.1 Types of automated databases

Databases employed in pharmacoepidemiological studies can be generally categorised into medical records databases or electronic health medical records and claims and other administrative databases (Hennessy, 2006:311; Strom, 2013c:119). Medical records databases are documented during outpatient care and are maintained for patients' clinical care while administrative databases designed primarily for administrative purposes such as for reimbursements and offer information for health research as a by-product. Hennessy (2006:311) adds that data in medical records databases take a "*doctor's eye view*" whereas administrative data represents an "*accountant's eye view*".

Claims databases are generated by information submitted to payers by healthcare service providers (physicians, pharmacists, hospitals and laboratories) for reimbursement for services provided (Hennessy, 2006:312). Databases usually contain information that are required for billing purposes such as dates of access of the medical service and associated diagnoses; the type of insurance cover; medical tests that are covered by the insurance; and prescription medications dispensed by the pharmacy (Torre & Martins, 2012:134). Medical record databases, generated by the increased computerisation of medical practice at the care level, primarily contain outpatient data, with inpatient data electronic medical record databases now being introduced and have the unique benefits of the inclusion of diagnosis data which is sometimes absent from the claims databases (Strom, 2013c:120). Hennessy (2006:312) concludes that, whereas medical records databases only account for prescription medications prescribed by the general practitioner, claims data are obtained from pharmacy bills and cover all prescription drugs dispensed from all levels of healthcare and adds that medical records may contain certain vital information on significant confounders, such as smoking status, which may not be present in administrative data.

#### 2.5.2.2.2 Advantages of automated databases

Employing databases for research offer some advantages over the use of other sources (Harpe, 2011:57; Strom, 2013c:120; Torre & Martins, 2012:139). Motheral and Fairman (1997:349) noted that databases have a broad scope, providing information on large populations over long periods, are flexible, providing various options in terms of methodology, and provide significant statistical power at relatively lower costs, advantages which give the researcher insight into whether to use a database to address what is being studied. The use of databases also increases the representativeness of research since it covers large populations in a region under study, its relatively large sizes permit the study of rare drug effects as well as the study of drug safety and utilisation patterns under real-world situations in relatively short time (Park & Stergachis, 2008:533; Schneeweiss & Avorn, 2005:324; Strom, 2013c:120; Torre & Martins, 2012:133). Since the use of databases does not depend on patients' interviews to obtain data, recall and interviewer biases are avoided when databases are used as the source of research. Moreover, databases can expand the scope of research by its abilities to be linked to other external electronic databases such as birth and death records, as well as police accident records (Strom, 2013c:120).

#### 2.5.2.2.3 Limitations of automated databases

Validity of diagnosis data is a major limitation of databases. This is especially true for claims databases that may not contain information on medical records and will therefore require access to medical record data to be used for validation (Park & Stergachis, 2008:534; Strom, 2013c:121).

Databases are also limited with reference to the quality of data since there may be lack of information on relevant confounding factors, such as information on alcohol intake, that may be of value to the questions under study (Park & Stergachis, 2008:534; Strom, 2013c:121; Torre & Martins, 2012:140).

Schneeweiss and Avorn (2005:327) believe that data privacy is another limitation of the use of databases, asserting that new regulations regarding privacy pose a threat to the access to valuable data that may be required for research. Another potential challenge faced using databases is poor outcome definition by the use of diagnostic coding systems (Park & Stergachis, 2008:534). This issue, together with changes in coding practices are significant when the study takes place over time within which switches in coding are likely to take place. Problems with coding may result from changes in documentation practices, reimbursement policies or in the codes themselves, leading to modified research results (Harpe, 2011:58). Due to career changes, health plan changes by employers, and changes in employee and family members' coverage, there is potential high turnover of individuals insured on various health plans and this hinders

longitudinal analyses of data obtained from databases (Harpe, 2011:58; Strom, 2013c:121). Further, since databases only cover cases that have been handled by healthcare service providers and are insured by particular policies, information on diseases that are not considered severe enough for medical attention to be sought and medications obtained over-the-counter cannot be obtained from these databases. Medical services that are not covered will also not be captured by databases (Strom, 2013c:121).

#### 2.5.2.2.4 Application of automated databases in pharmacoepidemiology

Studies on drug utilisation; prescribing habits of physicians; adverse events and risk management, beneficial effects of drugs, as well as health policy can be performed by the application of automated databases (Schneeweiss & Avorn, 2005:324). Strom (2013c:121) adds that databases are useful when uncommon outcomes are being studied, incidence rates are being calculated, short-term drug effects and diagnoses that are objective and laboratory-driven are being studied, interviewer and recall bias could adversely affect the results and when there are time and resource constraints. The choice of automated database to use for a pharmacoepidemiological study is as important as the decision to use a secondary data source in the first place and is dependent on the whether the research question can be appropriately answered using the database instead of an *ad hoc* study (Hennessy, 2006:313). It is imperative to consider the relative size, cost, and representativeness of the database, together with whether it is population-based, offers valid outcome data and an ability to control for confounders when choosing a database for research (Hennessy, 2005:313; Strom, 2013c:122).

Annexure B gives some examples of pharmacoepidemiological studies that have been carried out using automated databases together with the study designs and statistical analyses employed. The databases employed contained medication claims data, medical records or a combination of both and covered large patient populations. Although case-control, cross-sectional and nested case-control designs were among the designs used, cohort studies were the most common. The researchers analysed the data using descriptive statistics, regression methods, chi-squared tests and tailored methods, such as the receiver operating characteristic (ROC) curve and Kaplan-Meier survival analysis for specific analyses (Refer to Annexure B).

Table 2-2 summarises the sources of data used for pharmacoepidemiological research.

**Table 2-2: Summary of data sources employed for pharmacoepidemiological research**

| Classification | Source | Characteristics | Advantages | Disadvantages | References |
|---|---|---|---|---|---|
| *Ad-hoc* sources | Case-control surveillance | This applies a case-control method to investigate the effects of exposures on disease risks using patient interviews | Evaluates both prescription and non-prescription medications together with dietary supplements<br>Has the potential to control for confounding<br>Provides accurate diagnosis classification<br>Evaluates outcomes of long term drug use | Subject to selection and reporting biases<br>There is the possibility of misclassification of exposures<br>There may be lack of or inaccurate dosage information | Rosenberg *et al*. (2012:287) |
| | Prescription event monitoring | It involves the use of observational methods and questionnaires to obtain information on drug use and their associated outcomes from general practitioners | Provides large sample size for research<br>Data are representative since they are obtained from patients being treated in real clinical practice<br>Exposure data are accurate since validated dispensed prescriptions are used<br>Patients with potential adverse events can be identified and further studied | This approach is restricted to general practice only<br>The source is susceptible to selection bias and confounding<br>There is no comparison with an unexposed comparator | Layton and Shakir (2012:302);<br>Shakir (2006:152) |

| Classification | Source | Characteristics | Advantages | Disadvantages | References |
|---|---|---|---|---|---|
| *Ad-hoc* sources | Registries | Organised system that collects clinical and other data to evaluate specific outcomes that can be defined as a specific disease or exposure.<br>Studies based on registries offer conclusions that may vary depending on what is examined, and the technique used for sampling | Registries can be used to elucidate the natural history and course of diseases<br>They are applied for investigating determinants of quality of life<br>They allow for studying variations in interventions and their resultant outcomes<br>They can be used to appraise clinical- and cost-effectiveness of interventions | The diverse nature of registries makes it efficient only for addressing broad questions<br>The use of registries may be subject to selection bias | Dreyer and Velentgas (2012:331);<br>Gliklich *et al*. (2014:13) |
| Secondary data sources | Post-marketing spontaneous pharmacovigilance reporting systems | Reporting systems that have been put in place to record and examine the occurrence of adverse drug events by voluntary reporting of suspected adverse events to assigned centres | It accounts for all types of legal medications that are used in all settings<br>It facilitates the detection of safety signals<br>Offer a relatively inexpensive approach to the study of drug safety<br>It promotes the engagement of health professionals in safety surveillance<br>It allows for identifying rare drug events | Spontaneous reporting systems rely on the quality of individual reports<br>The voluntary nature leads to underreporting<br>There is limited information on confounding | Dweik *et al*. (2016:2);<br>ENCePP (2017:14);<br>Pan *et al*. (2013:101);<br>Waning and Montagne (2001:133) |

| Classification | Source | Characteristics | Advantages | Disadvantages | References |
|---|---|---|---|---|---|
| Secondary data sources | Automated databases | Databases are electronically stored healthcare utilisation information that are gathered regularly in clinical practice. | Use of databases is cost-effective and efficient<br>It offers large sample sizes for studies and increase representativeness of research<br>It allows for study of rare drug effects under real world situations<br>Recall and interviewer biases are avoided | Vital information such as information on confounding may be absent<br>Regulations about data privacy threaten access<br>There is potential poor outcome definition when diagnostic coding systems are used<br>Information on cases that are not considered for medical attention and drugs obtained over the counter may not be captured | |
| | Medical records databases | Medical records databases are generated from information recorded in outpatient care and kept for patient's clinical care. They contain diagnosis data and account for prescription medications from general practice only. | | | |
| | Claims databases | Claims databases are designed for administrative purposes and are generated from information submitted to payers by healthcare service providers. It is compiled from bills and cover all prescription drugs dispensed at all levels of healthcare. | | | |

## 2.6 Research methods in pharmacoepidemiology

Waning and Montagne (2001:159) noted that concepts and methods of pharmacoepidemiology have been broadly applied in pharmacy practice especially for the facilitation of decisions in drug therapy. Pharmacoepidemiology has been applied for the study of drug effects and adverse drug events (Pan *et al*., 2013:116), compliance (Acri & Gross, 2013:314); medication errors (Lee *et al*., 2013:378); drug utilisation (Lee *et al*., 2013:339; West-Strum, 2011:9; Thaker *et al*., 2015:54); and evaluating interventions intended to improve on medication use quality and to contain costs (Wagner *et al*., 2002:299). Several methods have been employed to facilitate these applications of pharmacoepidemiology and in the subsequent paragraphs interrupted time-series analysis, prescription sequence symmetry analysis, survival analysis and measures for the determination of adherence will be discussed.

### 2.6.1 Interrupted time series analysis

Interrupted time series (ITS) analysis is used in pharmacoepidemiology to evaluate the effectiveness of a change in policy or an intervention that is carried out at population-level and is aimed towards improving quality of medication use (Jandoc *et al*., 2015:950; Lagarde, 2012:76; Penfold & Zhang, 2013:38; Wagner *et al*., 2002:299). Interrupted time series analysis is an application of the quasi-experimental study design and is considered the most robust and most effective among such applications (Penfold & Zhang, 2013:44; Shadish *et al*., 2002:171; Wagner *et al*., 2002:308). According to Lagarde (2012:76), interrupted time series makes use of data that have been routinely collected at evenly-spaced time intervals to assess interventions expected to influence outcomes that can accurately be identified and recorded into health information systems. A time series of an event of interest is used to identify a pattern which is “interrupted” by a policy or an intervention at a point in time when an ITS analysis is being carried out (Bernal *et al*., 2017:349).

Time series is defined as time-ordered measurements of observations taken on a population over a period (Bernal *et al*., 2017:349; Granados, 2008:1035; Penfold & Zhang, 2013:39; Wagner *et al*., 2002:299). The methodology relies on the division of a time series into segments that represent rates of an event before and after an intervention (Penfold & Zhang, 2013:39; Wagner *et al*., 2002:299). Where values in the time series deviate from an established pattern as a result of an intervention or policy, change points are recorded (Wagner *et al*., 2002:299). To statistically measure and evaluate the changes and to draw inferences when an ITS analysis is done, segmented regression analysis, which uses modelling to assess the level and trend of a segment, is used (Bernal *et al*., 2017:351; Jandoc *et al*., 2015:953; Lagarde, 2012:79; Penfold & Zhang,

2013:39; Wagner *et al*., 2002:300). The level, value of a series at its beginning and the trend, and the rate at which a measure changes, define a segment of a time series and changes in these represent an effect from the intervention and change in the value of the outcome of interest respectively (Wagner *et al*., 2002:300).

Interrupted time series requires certain conditions to be met with reference to characteristics of the intervention being investigated, the event of interest and data. It is required that the period before and after the introduction of the intervention be clearly differentiated; that outcomes are short-term, expected to show quick and appreciable change after the intervention or after a clearly defined time lag; and that time-ordered measures of the outcome are accessible before and after an intervention or policy is introduced (Bernal *et al*., 2017:349).

An important strength of interrupted time series analysis is its ability to provide intuitive and easily interpreted graphical presentation of results (Penfold & Zhang, 2013:39; Wagner *et al*., 2002:301). This graphical presentation allows for better understanding by people with little or no statistical and epidemiological knowledge (Bernal *et al*., 2017:354) and easy identification of changes that occur together with details of such changes (Penfold & Zhang, 2013:39). The method also permits thorough evaluation of the continuous effects of a policy or an intervention under real-world conditions, increasing its external validity (Bernal *et al*., 2017:354). According to Penfold and Zhang (2013:39), an interrupted time series makes it possible for unintended effects of interventions to be analysed; lends itself to the conduct of analyses that are population-based rather than individual-based; and afford the investigator the opportunity to conduct stratified analyses to assess differences in policy and intervention effects on subpopulation groups. Time-independent confounders generally do not affect the results of interrupted time series studies as they change slowly over time and can be taken into consideration during the modelling (Bernal *et al*., 2017:353).

Certain vital issues limit the effectiveness of an interrupted time series, the most relevant of which may be its inability to be used to infer about outcomes at individual-level; the dependence of power to estimate regression coefficients on the high number of observations before and after the policy or intervention; and the need for adequate time periods between policies or interventions to evaluate their impact separately (Penfold & Zhang, 2013:43). Bernal *et al.* (2017:354) add that interrupted time series is susceptible to imprecise inferences about the effectiveness of a policy because of model specifications and is unable to adjust for time-dependent confounders. In order to enhance the robustness of ITS analysis, issues related to seasonal pattern of diseases and outcomes; the presence of time-varying confounding factors; the possible use of controls and other complex models of ITS; the phenomenon of over-dispersion which leads to imprecise approximation of standard errors; and autocorrelation, which is the correlation between series

values taken at different time points, need to be addressed (Bernal *et al*., 2017:352; Jandoc *et al*., 2015:954; Lagarde, 2012:77; Wagner *et al*., 2002:305).

### 2.6.2 Methods for determining adherence

Adherence to medications therapy can be determined directly and indirectly (Andrade *et al*., 2006:565; Fairman & Motheral, 2000:500; Hess *et al*., 2006:1280; Jimmy & Jose, 2011:157; Lam & Fresco, 2015:2; Lima-Dellamora *et al*., 2017:2; McCaffrey, 2011:139; Osterberg & Blaschke, 2005:488; Vik *et al*., 2004:304). Direct or objective measures of adherence are those that allow for the confirmation of intake of medications (Lima-Dellamora *et al*., 2017:2). Biological assays that measure the concentration of medications and their metabolites in body fluids such as urine, and blood, and the detection of certain biologic markers added to drug formulations, together with directly observed therapy, are examples of direct methods used for the measurement of adherence (Andrade *et al*., 2006:565; Fairman & Motheral, 2000:500; Jimmy & Jose, 2011:157; Lam & Fresco, 2015:2; Lima-Dellamora *et al*., 2017:2; McCaffery, 2011:139; Osterberg & Blaschke, 2005:488; Vik *et al*., 2004:304). Although direct measures of adherence are most accurate and reliable, offering physical proof of patients' medication taking behaviour (Fairman & Motheral, 2000:500; Jimmy & Jose, 2011:157; Lam & Fresco, 2015:2), certain limitations hinder their practicality (Lam & Fresco, 2015:2).

Direct measures can be expensive, labour-intensive (Fairman & Motheral, 2000:500; Jimmy & Jose, 2011:157; Osterberg & Blaschke, 2005:488; Vik *et al*., 2004:304) and intrusive, leading to anxiety and pressure in patients (Lam & Fresco, 2015:2; Vik *et al*., 2004:304). Fairman and Motheral (2000:500), Vik *et al*. (2004:304) and Lam and Fresco (2015:2) agree that trends of adherence cannot be determined using direct measures since these only assess adherence at one point in time and do not allow differentiation in the patterns of adherence nor suggest reasons for non-adherence. Variations in, as well as factors that influence metabolism, affect the results of tests that are used to assess adherence directly (Jimmy & Jose, 2011:157; Lam & Fresco, 2015:2). Drug or food interactions, dosing schedules and drug half-life also influence results and hinder their accuracy (Lam & Fresco, 2015:2; Vik *et al*., 2004:304).

White coat adherence or the "*Hawthorne effect*", i.e. increased adherence to treatment prior to clinic appointments (Driscoll *et al*., 2017:455; Fairman & Motheral, 2000:500; Urquhart & Vrijens, 2006:370), also gives false impressions of patient adherence as patients take medications only before imminent tests and distort measures (Jimmy & Jose, 2011:157; Lam & Fresco, 2015:2; Osterberg & Blaschke, 2005:488). Furthermore, certain direct measures may be limited by ethical concerns (Vik *et al*., 2004:304) and impractical except in impatient settings where there is increased accessibility to patients (Fairman & Motheral, 2000:500).

Indirect measures of adherence are those that result from interactions between users and the instruments that measure adherence (Lima-Dellamora *et al*., 2017:2) and depend on surrogate measures of medication use (McCaffrey, 2011:139). Indirect measures can be classified as self-reports, medication measurements or pill counts, prescription record review and electronic event monitoring (Farmer, 1999:1076; McCaffrey, 2011:140). The most prevalent approach to evaluating adherence is to ask patients about their medication taking behaviours (McCaffrey, 2011:140). Patient interviews involve asking patients to estimate the percentage of their medications they would take according to prescribed regime and may require testing patients' knowledge of drugs' names, indications and prescribed schedules (Lam & Fresco, 2015:5). Standardised survey instruments and validated adherence-specific questionnaires are also employed to estimate adherence (Farmer, 1999:1079; McCaffrey, 2011:140). These questionnaires are usually adapted to accommodate diverse conditions, in various languages and can easily be completed by patients or their caregivers (Lam & Fresco, 2015:6). The Brief Medication Questionnaire, the Medication Adherence Report Scale, Eight-Item Morisky Medication Adherence Scale, Medication Adherence Questionnaire, The Self-Efficacy for Appropriate Medication Use Scale and Hill-Bone Compliance Scale are examples of questionnaires and scales useful for assessing adherence (Farmer, 1999:1080; Lam & Fresco, 2015:6; McCaffrey, 2011:140).

Self-reporting of adherence may also entail the use of patient-kept diaries, in which patients are expected to document medication taking as and when it occurs (Farmer, 1999:1078; Gillisen, 2007:210; Lam & Fresco, 2015:5; McCaffrey, 2011:140). The patient-kept diary is the only self-report approach that allows for a consistent record of how the patient follows a prescribed regime (Lam & Fresco, 2015:5) and gives information on the exact prescribed regime (Farmer, 1999:1078). Although self-report of adherence is fast, inexpensive and simple to carry out (Farmer, 1999:1078; Gillisen, 2007:210; Osterberg & Blaschke, 2005:488) and gives the patients' perspective of their adherence patterns (Fairman & Motheral, 2000:500), it is thought to be unreliable because it is susceptible to recall bias on the part of patients (Fairman & Motheral, 2000:500; Lima-Dellamora, 2017:2; McCaffery, 2011:140; Urquhart & Vrijens, 2006:370). Self-reports are subject to misrepresentations by patients in the form of exaggerations and censoring that alter adherence estimates (Gillisen, 2007:210; Jimmy & Jose, 2011:157; Osterberg & Blaschke, 2005:488; Urquhart & Vrijens, 2006:370). Farmer (1999:1078) asserts that the responses from interviews and questionnaires are influenced by the construction of questions and depend on the expertise of the interviewer as well as the type of instrument used.

Pill counts (i.e. counting the number of pills or capsules that are untaken between clinic visits) (Jimmy & Jose, 2011:157; Lam & Fresco, 2015:5; Urquhart & Vrijens, 2006:369), is the most

prevalent indirect method used to measure adherence (Jimmy & Jose, 2011:157; Osterberg & Blaschke, 2005:488) and offers an objective evaluation of adherence (McCaffrey, 2011:141; Vik *et al*., 2004:305). This tool computes percentage adherence using the number of tablets that have been taken and the number of tablets that should have been taken, represented by the simple formula (Vik *et al*., 2004:305):

$$\%\ adherence = \frac{Number\ of\ tablets\ taken}{Number\ of\ tablets\ that\ should\ have\ been\ taken}\ x\ 100$$

Pill counts, like most tools for determining adherence, are easy to use, straightforward and inexpensive (Farmer, 1999:1078; Gillessen, 2007:210; McCaffery, 2011:141; Osterberg & Blaschke, 2005:488). Despite these attractive features, the effectiveness of pill counts is hindered by their inability to measure whether the medication was taken and on schedule (Fairman & Motheral, 2000:500), their incompatibility to medications taken on when-required basis and their need for patients to maintain medications in original containers with carefully-done labels (Vik *et al*., 2004:305). Patients can also easily alter data from pill counts by switching medicines between containers and worse, discarding them to appear compliant. Furthermore, this method does not offer details on other details of medication taking behaviours such as drug holidays where the patient fails to take the medication on three or more consecutive days (Jimmy & Jose, 2011:157; Osterberg & Blaschke, 2005:489).

Checchi and colleagues (2014:2) describe electronic medication packaging (EMP) as electronic devices that are incorporated into containers in which medications are dispensed and add that these devices have recorders and storage functions installed to facilitate the assessment of adherence to therapy. Microprocessors placed in these devices permit the time and frequency of opening a medication-containing device to be recorded (Farmer, 1999:1083; McCaffrey, 2011:141), giving precise information on patients' medication taking behaviours (Farmer, 1999:1083; Osterberg & Blaschke, 2005:489). Electronic medication packaging devices share certain common features which include: recorded information on dosing events and stored adherence records; audio-visual promptings that signal scheduled dosing; digital displays; instantaneous monitoring; and providing feedback on patient's adherence (Checchi *et al*., 2014:2). Although these are not available in all devices, the ability to record adherence is vital for assessing and deciding on suitable interventions for improving adherence (Lam & Fresco, 2015:4).

Medications Events Monitoring System (MEMS) is an example of an EMP that is widely used for adherence monitoring (Farmer, 1999:1083; Lam & Fresco, 2015:4; Vik *et al*., 2004:305) which not only assesses the medication-taking behaviour but also reports on the suitability of recommended dosage schedules (Vik *et al*., 2004:305). High cost of EMPs coupled with their

limited availability counter their attractive properties of being able to provide precise and easily quantified information on adherence levels and trends (Farmer, 1999:1083; Jimmy & Jose, 2011:158; Lam & Fresco, 2015:4; Osterberg & Blaschke, 2005:489; Vik *et al*., 2004:305). Furthermore, electronic monitoring does not give exact information on if the patient ingests the medication (McCaffrey, 2011:141; Osterberg & Blaschke, 2005:489). Patients may actuate the device without intake of the medication, transfer the medications into different containers, or take multiple doses out at the one time, all of which invalidate the data obtained (Farmer, 1999:1083; Jimmy & Jose, 2011:158; Osterberg & Blaschke, 2005:489).

Researchers are increasingly using administrative claims databases to evaluate adherence to medications (Andrade *et al*., 2006:566; Farmer, 1999:1082; Grégoire & Moisan, 2016:369; Halpern *et al*., 2006:1040; Hess *et al*., 2006:1280; Karve *et al*., 2009:989; Lam & Fresco, 2015:3; Lima-Dellamora *et al*., 2017:2; McCaffrey, 2011:141; Svarstad *et al*., 2001:806; Vik *et al*., 2004:305). Arnet *et al*. (2016:362) noted that the use of medication records for determining adherence is based on the assumptions that records are complete and accurate, the first dose of medication was taken on the day of refill, the drug is ingested as prescribed, medication is not obtained from another source, no dose changes or interruptions occur during the period of observation, and not refilling means no medication is taken during the gap period. Claims data have a significant role in assessing adherence to medications for chronic diseases but are not useful for medications, such as antibiotics, that are taken for short periods and do not require further refill (Vik *et al*., 2004:305).

Administrative datasets produce adherence values based on medication possession and not consumption, estimating the highest level of medication consumption possible (Hess *et al*., 2006:1281) since they are unable to determine patients' intake of prescribed and dispensed medications (Andrade *et al*., 2006:566). Grégoire and Moisan (2016:370) add that dispensing data, which do not indicate discontinuation of medications and are dependent on conditions of reimbursement, may underestimate adherence as patients may be deemed non-adherent if they stopped refilling their prescriptions or pay for their medications out-of-pocket if they are no longer covered by the plan. Moreover, adherence determined solely by claims data may mask periods of over- and under-utilisation of medications (Vik *et al*., 2004:305). Despite these limitations, administrative claims data are non-invasive, convenient (Hess *et al*., 2006:1281), inexpensive (Hess *et al*., 2006:1281; McCaffrey, 2011:141) and efficient if the data used are accurate (Andrade *et al*., 2006:566). Furthermore, claims data permits the researcher to follow up after initiation of therapy and to identify early discontinuation as well as unprescribed medication consumption (Farmer, 1999:1082).

Successful application of claims data for measuring adherence requires that datasets be complete with all relevant information captured accurately (Farmer, 1999:1082), researchers understand the different rules that apply to the insurance plan databases they use (Farmer, 1999:1082; Grégoire & Moisan, 2016:370), and patients be continuously eligible for the medication of interest during study period with enough data for valid conclusions to be drawn about their adherence behaviours (Farmer, 1999:1083; McCaffrey, 2011:146).

While adherence measures computed using claims data differ, they are characterised by three specifications: continuous versus dichotomous variables, single versus multiple dispensing intervals, and examination of medications availability versus treatment gaps (Farmer, 1999:1082; Steiner & Prochazka, 1997:112). The first to last prescription records are considered when adherence is measured as continuous variable, whereas adherence as a dichotomous variable considers patients adherent when they meet certain pre-determined criteria (Farmer, 1999:1082). Adherence as a dichotomous variable requires a cut-off value, distinguishing adherent from non-adherent patients, that must be rationally selected and consistent with existing scientific evidence (Peterson *et al*., 2007:7; Steiner & Prochazka, 1997:114) and is considered to have less reliability and power than continuous adherence measures (DiMatteo *et al*., 2002:806). Dichotomisation does not permit the differentiation of non-adherence into various types such as sporadic and consistent (Lima-Dellamora *et al*., 2017:8). A treatment gap is the time from the anticipated date of depletion of a claim and the date of the subsequent refill (Fairman & Motheral, 2000:502; Vink *et al*., 2009:161) and serves as a proxy for periods of non-adherence (Farmer, 1999:1082) on the assumption that a drug must be in a patient's possession to be consumed (Grégoire & Moisan, 2016:373). Steiner and Prochazka (1997:114) assert that the best measures for assessing the timing and duration of drug exposure are those that evaluate treatment gaps as these can determine intervals in which drug exposure is improbable. Measures based on medication possession are valuable in testing hypotheses associated with the relationship between medications and cumulative dosages and outcomes (Steiner & Prochazka, 1997:114). Treatment gap measures do not consider stockpiling from previous supplies and early refills and may thus underestimate adherence (Grégoire & Moisan, 2016:373). Whereas measures based on single dispensing can be useful in situations where short-term medication use is thought to result in outcomes, multiple interval measures are of benefit where cumulative drug use are essential to outcomes (Lima-Dellamora *et al*., 2017:7; Steiner & Prochazka, 1997:114).

Measures computed from claims data also differ in their suitability for determining the various constructs of adherence i.e. initiation, persistence and compliance (Grégoire & Moisan, 2016:371). While Grégoire and Moisan (2016:374) state that gap measures are suited for studies on persistence and measures based on medication possession for determining compliance, Lima-

Dellamora *et al*. (2017:10) argue that all adherence measures determined using claims data yield results that refer more appropriately to the persistence construct of adherence.

Continuous measure of medication gaps (CMG), continuous multiple interval measure of oversupply (CMOS) and continuous single interval of medication gaps (CSG) are measures that use gaps to calculate adherence. Measures based on medication availability are: medication possession ratio (MPR), proportion of days covered (PDC), continuous single-interval of medication availability (CSA), continuous multiple-interval of medication availability (CMA), medication possession ratio modified (MPRm), refill compliance rate (RCR), compliance rate (CR), days between fills adherence rate (DBR), medication refill adherence (MRA), truncated medication possession ratio (truncated MPR), and proportion of prescribed days covered (PPDC) (Andrade *et al*., 2006:569; Hess *et al*., 2006:1283; Lam & Fresco, 2015:3; Lima-Dellamora *et al*., 2017:6; McCaffrey, 2011:143; Vollmer *et al*., 2012:2;).

Table 2-3 provides an overview of the measures of adherence that can be computed using administrative claims data.

**Table 2-3 Measures of adherence using administrative claims data**

| Measure | Definition/Formula | Advantages | Disadvantages | Comments | Reference |
|---|---|---|---|---|---|
| CMG | Used to compute the total days for which a patient has no medication. It is computed using the formula: $\frac{Total\ days\ of\ gaps\ in\ treatment}{Total\ days\ to\ the\ next\ refill\ or\ to\ end\ of\ study\ period}$ | | Involves lots of calculations and data fields which makes it less attractive | Produces values of non-adherence for cumulative period | Hess *et al*. (2006:1282) |
| | $\frac{Sum\ of\ days\ in\ the\ refill\ gaps\ in\ the\ observation\ period}{Time\ between\ the\ first\ and\ last\ fills}$ | Gives an estimation of the variability in patients' refill behaviour | | | Peterson *et al*. (2007:6) |
| | $\frac{Number\ of\ days\ in\ observation\ period - total\ days'supply}{Number\ of\ days\ in\ observation\ period}$ | Gives a cumulative assessment of treatment gaps | It is imprecise when patient has only one prescription<br>It assumes patient has discontinued therapy after the last dispensing and does not consider stockpiling | Sums the proportion of days patients have no medications across all refill intervals from the first to last dispensing | Raebel *et al*. (2013:17) |
| | $\frac{Days\ on\ which\ patient\ has\ no\ medications\ (gaps)}{days\ in\ the\ observation\ period}$ | | | Gaps may be negative (early refill) or positive (late refill)<br>Requires adjustment for oversupplies by adding negative and positive gaps | Campagna *et al*. (2014:759) |

| Measure | Definition/Formula | Advantages | Disadvantages | Comments | Reference |
|---|---|---|---|---|---|
| CMG | $\frac{Cumulative\ days\ with\ no\ medication\ over\ intervals}{Time\ from\ beginning\ to\ end\ of\ study\ period\ in\ days}$ | | Does not consider the possibility of early refill or overfill | | Lam & Fresco (2015:3) |
| | $\frac{Total\ observation\ period - days'supply}{days\ in\ observation\ period}$ | Simple | Cannot approximate the number of days on which there was no medication thus gives no measure of 'real gap' | A mean is usually used to calculate total adherence and is not based on the dosing history | Lima-Dellamora *et al*. (2017:7) |
| CMOS | $\frac{Total\ days\ of\ medication\ gaps\ (+)or\ surplus\ (-)}{Total\ days\ in\ period\ of\ study}$ | This measure accounts for cases in which the patient was over-supplied medication for days in the observation period | It requires several calculations and data fields, making it less attractive | Value represents nonadherence for a cumulative period | Hess *et al*. (2006:1282) |
| | $\frac{Total\ gaps\ or\ surplus\ days}{Total\ days\ in\ period\ of\ observation\ or\ to\ next\ refill}$ | Easy to calculate | Ignores non-adherence after last refill<br>Can underestimate the gaps that exist between refills | | Raebel *et al*. (2013:17) |
| | $\frac{Total\ observation\ period - days'supply}{days\ in\ the\ observation\ period}$ | | | Like CMG but allows for negative values which represent oversupply of medications | Lima-Dellamora *et al*. (2017:7) |

| Measure | Definition/Formula | Advantages | Disadvantages | Comments | Reference |
|---|---|---|---|---|---|
| CSG | $\frac{Total\ observation\ period - days'supply}{days\ in\ single\ observation\ period}$ | | Observation period is randomly selected thus it is not derived from real usage history<br>In events of oversupply, it can only be used for multiple cases where the negative values can be reset to zero | Value represents adherence for a single dispensing event<br>For more than one dispensing, a mean is used to calculate total adherence | Lima-Dellamora *et al*. (2017:6) |
| MPR | Proportion of days' supply that a patient obtains during a defined time or over a period of refill intervals.<br>Calculated using two formulae:<br>$\frac{number\ of\ days\ supply\ obtained\ in\ the\ study\ period}{number\ of\ days\ in\ the\ study\ period}\ x\ 100$<br>$\frac{number\ of\ days\ supply\ obtained\ without\ the\ last\ refill}{number\ of\ days\ between\ the\ first\ and\ lastfill\ dates}\ x\ 100$ | Calculation and interpretation is easy | Periods of over- and under-supply may not be considered<br>MPR value greater than 1 are difficult to interpret | | Andrade *et al*. (2006:569) |
| | Ratio of total days for which patient receives medication excluding the last prescription to total days in a period represented by:<br>$\frac{Total\ days\ of\ medication\ supplied\ without\ last\ prescription}{Total\ days\ in\ a\ period}$ | Accepted standard for compliance evaluation because it is easy. | Supposes that proportion of days a prescription covers is equivalent to proportion of days the medication is taken. Possession of medication does not mean it is ingested | Most commonly used adherence measure<br>Requires patients receiving 2 or more prescriptions to define the time frame during which therapy is given | Halpern *et al*. (2006:1044) |

| Measure | Definition/Formula | Advantages | Disadvantages | Comments | Reference |
|---|---|---|---|---|---|
| MPR | Ratio of overall days' supply to number of days in the observation period per participant and is represented by $\frac{Days' supply}{Days\ in\ study\ period}$ | | Existence of different measures all referred to as MPR leads to confusion in comparisons across studies | Value represents the ratio of medication available<br>Ratios are divided and averaged to obtain total study adherence value | Hess *et al*. (2006:1282) |
| | This measure is computed by dividing the sum of the number of days supplied for all but the last refill by the number of days between the first and the last refills. It is typically calculated using: $\frac{Number\ of\ days\ of\ medication\ supplied\ in\ refill\ interval}{Number\ of\ days\ in\ refill\ interval}$ | | | Requires at least two refill dates for calculation | Peterson *et al.* (2007:6) |
| | $\frac{Sum\ of\ days' supply\ for\ all\ medications}{Period\ of\ observation}$ | Easy to calculate | Assumes that medications are used within the study period which may be unrealistic<br>Overestimates adherence when there is polytherapy within a drug class | | Martin *et al.* (2009:37) |
| | This estimates the proportion of days' supply dispensed over a specific period or defined refill intervals. There are several variations of this measure but the basic one is calculated using: $\frac{Total\ days' of\ medication\ supply}{Number\ of\ days\ in\ period\ of\ observation}$ | Easy to calculate | There is confusion due to varied methods of calculation<br>Probable adherence overestimation due to early refills | Results are identical to those produced by PDC when assessing adherence to a single drug.<br>Similar to MRA and CMA | Raebel *et al*. (2013:14) |

| Measure | Definition/Formula | Advantages | Disadvantages | Comments | Reference |
|---|---|---|---|---|---|
| MPR | This measure gives continuous adherence values and is calculated by: $\frac{Total\ number\ of\ days\ supply\ for\ all\ fills}{Number\ of\ days\ between\ fills\ plus\ days'supply\ of\ last\ fill}$ | | Sensitive to varied entries in the days' supply field | It may also be calculated by excluding the days' supply of the last fill from both the numerator and denominator<br>Requires at least two dispensing records for calculation. | Campagna *et al*. (2014:759) |
| | $\frac{Days'supply\ obtained\ by\ patient}{Refill\ interval\ or\ fixed\ interval}$ | Simple | Does not consider gaps in refill<br>Variations in denominator make it impossible to use MPR for large population analysis<br>MPR overestimates adherence values | | Lam & Fresco (2015:4) |
| | $\frac{Days\ with\ medication\ supplied}{Days\ in\ time\ interval\ under\ study}$ | | | Can be dichotomized or assessed as a continuous variable | Giardini *et al*. (2016:377) |

| Measure | Definition/Formula | Advantages | Disadvantages | Comments | Reference |
|---|---|---|---|---|---|
| MPR | $\frac{Number\ of\ days\ for\ which\ patient\ has\ medication\ prescribed}{Number\ of\ days\ in\ observation\ period}$ | Measures both single and multiple intervals in which patient has medications | | Usually expressed as a fraction<br>It is suggested that a mean be calculated to obtain overall adherence in the case of polypharmacy | Lima-Dellamora *et al*. (2017:5) |
| PDC | This measure is obtained when the total days' supply is divided by the number of days of in the study according to the formula:<br>$\frac{Total\ days'supply}{Total\ days\ evaluated}\ x\ 100$ | | May underestimate adherence because capping at refill intervals does not account for excess medication carried over from one interval to the next | Value represents percentage of days with medication available<br>Same as MPR but capped at 100% | Hess *et al*. (2006:1282) |
| | It is computed as the number of days with the medication available divided by the number of days in the period of interest and is multiplied presented as a percentage.<br>The formula below is used:<br>$\frac{Number\ of\ days\ with\ drug\ of\ interest}{Number\ of\ days\ in\ period\ under\ investigation}\ x\ 100$ | It can simultaneously provide values for both compliance and persistence constructs of adherence | | Denominator is specifically a number of clinically meaningful days and equal for all patients and intervals<br>May be analysed as a continuous or categorical variable | Peterson *et al*. (2007:6). |

| Measure | Definition/Formula | Advantages | Disadvantages | Comments | Reference |
|---|---|---|---|---|---|
| PDC | It is defined as proportion of days a patient has a drug available during a study period calculated by: $\frac{Number\ of\ days\ patient\ has\ medication}{Total\ number\ of\ days\ in\ the\ specified\ period}$ | Reflects adherence behavior of patients with polytherapy within a drug class accurately<br>It is less sensitive to complex changes in drug therapy.<br>Provides more conservative adherence estimates than MPR | Challenging to calculate | It requires daily examination of medication possession and results in a simple measure that can indicate the presence of medication for each study day | Martin *et al*. (2009:37) |
| | | | Falsely attributes part of prescriber nonadherence to patients since it does not adjust for prescription patterns | Assumes prescribed medication is for chronic daily use which is not always the case Thus, it reflects both patient and physical behaviour where use of medication is concerned | Blais *et al*. (2011:337) |

| Measure | Definition/Formula | Advantages | Disadvantages | Comments | Reference |
|---|---|---|---|---|---|
| PDC | Calculated as:<br>$\frac{Total\ number\ of\ days'supply\ dispensed\ in\ specified\ observation\ peri}{Number\ of\ days\ in\ patient's\ observation\ period}$ | Easy to calculate<br>Provides a more conservative adherence estimate when there are frequent drug changes and multitherapy with drugs within a class | It ignores non-adherence after the last refill<br>Underestimates non-possession of medication in refill intervals if it is followed by early refills | Oversupply is truncated.<br>It should create time series that reflect refill dates within a patient's observation period instead of summing up the days' supply over the period | Raebel *et al*. (2013:15) |
| | Calculated by:<br>$\frac{Number\ of\ days\ with\ medication\ available}{Number\ of\ days\ in\ a\ specified\ time\ interval}$ | This measure is more stable than MPR when a constant denominator is used | | The denominator is typically a number of days that is clinically significant<br>May be continuous or dichotomous | Campagna *et al*. 2014:759 |
| | This measure is defined as the number of days a patient has medication cover over a period of time | | | Different from MPR because it assumes patient finishes current fill before starting a new refill<br>Can be assessed as continuous or dichotomous variable | Giardini *et al*. (2016:377) |

| Measure | Definition/Formula | Advantages | Disadvantages | Comments | Reference |
|---|---|---|---|---|---|
| PDC | The measure is the total number of days of supply dispensed to patient in a specified period divided by the length of that period, computed as: $$\frac{Total\ days'supply\ dispensed\ in\ period}{length\ of\ observation\ period}$$ | Suitable for assessing the compliance construct of adherence | PDC cannot capture periods of non-persistence<br>Difficult to distinguish patients with low compliance who persist with treatment from those who discontinue therapy early<br>It does not consider stockpiling from previous fills resulting in underestimated adherence values<br>It is challenging to calculate PDC for patients taking more than a drug | | Grégoire & Moisan (2016:377) |
| | Calculated by: $$\frac{Number\ of\ days\ patient\ has\ prescribed\ medications}{Number\ of\ days\ in\ observation\ period}\ x\ 100$$ | | | Capped at 100%<br>Expresses single or multiple intervals in which the patient has medications | Lima-Dellamora *et al*. (2017:5) |
| | It is the proportion of days in a fixed study period where the patient has at least one of multiple medications available. | It has a lower risk of overestimating adherence | | Measure preferred in polytherapy | Tang *et al*. (2017:2). |

| Measure | Definition/Formula | Advantages | Disadvantages | Comments | Reference |
|---|---|---|---|---|---|
| CSA | This measure is obtained by dividing the days' supply of medication by the number of days in the study interval from date of fill to, but excluding, the next fill date or through to study completion date.<br>It is represented by the formula:<br>$$\frac{days' supply\ obtained\ at\ the\ beginning\ of\ interval}{days\ in\ the\ interval}$$ | Beneficial for studies in which participation attrition is high. | Bias can occur if participants have more than one refill per day and if the refills are close to the completion of the study<br>The measure does not allow the carryover of drugs from a refill interval to the next since it truncates study intervals | Provides adherence value for each participant between fills and not that of the overall study period<br>Mean of all dispensation values estimates the overall study adherence<br>Participants who fill one prescription do not weigh the same as those with multiple refills when a cumulative analysis is done | Hess *et al*. (2006:1283) |
| | This measure assesses medication availability for a single interval and provides adherence values for each participant between fills using the formula:<br>$$\frac{Number\ of\ days' supply\ dispensed}{Number\ of\ days\ from\ refill\ date\ up\ to, but\ excluding,\ next\ refill\ date}$$ | Easy to compute.<br>Provides overall adherence based on the last refill date<br>Does not require completion of study date | Calculation is not precise when there is only one dispensing record | Mean of all adherence values gives overall study adherence | Raebel *et al*. (2013:16) |

| Measure | Definition/Formula | Advantages | Disadvantages | Comments | Reference |
|---|---|---|---|---|---|
| CSA | The formula used is: $\frac{Number\ of\ days\ patient\ has\ supply\ of\ prescribed\ drugs}{number\ of\ days\ in\ single\ observation\ period}$ | | | Measures single interval in which patient has medication beginning with dispensing<br>Aims to express medication coverage for a given period<br>Values larger than 1 signify oversupply of medication | Lima-Dellamora *et al* *(*2017:5) |
| CMA | This is computed by dividing the days' supply of medication patient obtained throughout the study period by the number of days in the study period by the formula: $\frac{Cumulative\ days'supply\ of\ medication\ obtained}{total\ number\ of\ days\ to\ next\ refill\ or\ end\ of\ study}$ | | | Value represents adherence over a cumulative time period | Hess *et al*. (2006:1281) |
| | Assesses adherence to medication over multiple refill intervals using the equation: $\frac{Total\ days'supply\ obtained\ throughout\ study\ period}{Number\ of\ days\ between\ first\ dispensing\ and\ end\ of\ study}$ | Easy to calculate | Ignores non-adherence after last refill | Mathematically similar to MPR and MRA | Raebel *et al*. (2013:16) |
| | This measure is based on medication availability and calculated according to the formula: $\frac{Cumulative\ days'supply\ over\ series\ of\ intervals}{total\ number\ of\ days\ from\ beginning\ to\ end\ of\ study}$ | | | | Lam & Fresco (2015:3) |

| Measure | Definition/Formula | Advantages | Disadvantages | Comments | Reference |
|---|---|---|---|---|---|
| CMA | CMA is computed by two formulae as follows:<br>$\frac{Total\ number\ of\ days\ with\ medication\ prescribed}{number\ of\ days\ in\ observation\ period}$<br>$\frac{Total\ numberof\ dispensing\ events(amount\ of\ drugs)\ for\ time}{expected\ number\ of\ dispensing\ events\ in\ observation\ time}$ | | | Expresses medication coverage in several continuous periods | Lima-Dellamora *et al*. (2017:5) |
| MPRm | MPRm is the total days' supply of medication divided by the total number of days from the first fill date to, but excluding, the last fill date and the number of days' supply dispensed at the last fill date and is represented by the formula:<br>$\frac{Total\ days'supply}{(last\ claim\ date - first\ claim\ date) + last\ days'supply)}\ x\ 100$ | Reduces possible overestimation as encountered with RCR | Assumes that all participants will have 100% adherence in the last refill period, resulting in adherence values higher than those obtained with other measures | Value represents adherence percentage adjusted to include the last refill period | Hess *et al*. (2006:1282) |
| | Calculated using:<br>$\frac{Number\ of\ days\ on\ which\ patient\ has\ prescribed\ medicati}{(period\ between\ first\ and\ last\ dispensing) + (period\ covered\ by\ last}$ | Removes one day to be sensitive and rules out duplication in the analysis<br>Permits the calculation of coverage, assuming possible gap in information | | Includes the final observation period | Lima-Dellamora *et al*. (2017:5) |

| Measure | Definition/Formula | Advantages | Disadvantages | Comments | Reference |
|---|---|---|---|---|---|
| RCR | Computed by: $$\frac{\left(\frac{sum\ of\ quantity\ dispensed\ over\ interval}{quantity\ to\ be\ taken\ per\ day}\right) x\ 100}{number\ of\ days\ between\ first\ and\ last\ refills}$$ | | | Gives an overall adherence percentage<br>Participants with one dispensing record are not included because it invalidates the denominator | Hess *et al*. (2006:1282) |
| | Defined by the formula: $$\frac{\left(\frac{Sum\ of\ amount\ during\ observation\ period}{Amount\ to\ take\ per\ day\ by\ prescription}\right)}{number\ of\ days\ between\ first\ and\ last\ refills}\ x\ 100$$ | Deals with arbitrary definition of the end of the study period by using dispensing dates as basis | Does not discount duplicate days like MPRm does not compensate for oversupply like MPRm and the truncated MPR | Measures percentage coverage<br>Numerator signifies the number of days on which patients have medications available | Lima-Dellamora *et al*. (2017:6) |
| CR | Computed by dividing the difference between the sum of days' supply for each patient and the days' supply dispensed at the last dispensing event by the number of days from the beginning of the study to, but excluding, the last dispensation.<br>The following equation is used: $$\frac{(total\ days'supplied - last\ days'supply)}{(last\ claim\ date - first\ claim\ date)}\ x\ 100$$ | Calculations are simplified | Does not account for participants who discontinue medication before the end of the study | Provides total adherence rate based on the last refill day without requiring the date of study completion | Hess *et al*. (2006:1281) |

| Measure | Definition/Formula | Advantages | Disadvantages | Comments | Reference |
|---|---|---|---|---|---|
| CR | Calculated using:<br>$\frac{Sum\ of\ days'supplies - days'supply\ for\ last\ refill}{Number\ of\ days\ from\ first\ up\ to, but\ excluding, last\ fill\ date}$ | Gives the overall study adherence based on date of last refill<br>Does not require study completion date | Gives imprecise values when one dispensing record is available | Denominator can also be the difference between the index date and the last claim date | Raebel *et al*. (2013:16) |
| | $\frac{Days'supply\ in\ all\ the\ dispensing\ events\ except\ the\ last}{number\ of\ days\ between\ first\ and\ last\ dispensing}\ x\ 100$ | | | Measures percentage coverage in defined dispensing intervals | Lima-Dellamora *et al*. (2017:6) |
| DBR | The measure is computed according to the formula:<br>$\left(1 - \left[\frac{(last\ claim\ date - first\ claim\ date) - total\ days'supply}{last\ claim\ date - first\ claim\ date}\right]\right) x\ 100$ | | | Value represents overall adherence percentage<br>It evaluates time between dispensing events | Hess *et al*. (2006:1283) |
| | Calculated using the formula:<br>$(1 - \left[\frac{(Period\ between\ first\ and\ last\ fills - days'supply)}{period\ between\ first\ and\ last\ fills}\right] x\ 100$ | | | Uses a device for adjustment to measure percentage coverage | Lima-Dellamora *et al*. (2017:6) |

| Measure | Definition/Formula | Advantages | Disadvantages | Comments | Reference |
|---|---|---|---|---|---|
| MRA | Computed as: $\frac{Total\ days' supply}{Number\ of\ days\ evaluated}\ x\ 100$ | Simple<br>Requires few data | | Value gives overall adherence percentage<br>Results are identical to those of other refill adherence measures | Hess *et al*. (2006:1282) |
| | Computed as: $\frac{Total\ days' supply}{Number\ of\ days\ in\ observation\ period}\ x\ 100$ | Easy to compute | Calculation is imprecise for those with one dispensing | Mathematically similarly to CMA and MPR<br>Mean gives total study adherence value | Raebel *et al*. (2013:15) |
| | Calculated using: $\frac{Days' supply}{number\ of\ days\ in\ observation\ period}\ x\ 100$ | | | Assesses coverage like the CMA but gives value as a percentage | Lima-Dellamora *et al*. (2017:6) |
| Truncated<br><br>MPR | $\frac{Number\ of\ days\ on\ which\ patient\ has\ medications\ available}{Number\ of\ days\ in\ observation\ period}$ | | Information is lost about the medications received in excess of what is used in theory for the period of study | Results must be less than 1<br>Aims to rule out any presumed excess in the measure | Lima-Dellamora *et al*. (2017:6). |

| Measure | Definition/Formula | Advantages | Disadvantages | Comments | Reference |
|---|---|---|---|---|---|
| PPDC | This measure is a modified version of the PDC and accounts for differences in prescription patterns and is calculated according to: $$\frac{Total\ days' supply\ dispensed\ to\ patient}{Total\ days' supply\ prescribed\ in\ the\ follow-up\ period}$$ | Reflects medication adherence to actual prescribed therapy<br>Specifically measures patient behaviour by measuring adherence to therapy as prescribed | | Considers the exact quantity of medications prescribed | Blais *et al*. (2011:336). |

### 2.6.3 Survival analysis

There is a need to assess medical product reliability, drug safety, medical therapies and devices viability, as well as the effects of other interventions and this can be done by assessing the number of people who survive after the implementation of these over a period of time (Goel *et al*., 2010:274).

Survival refers to remaining free of an outcome or event over time (Sullivan, 2016) with events of interest ranging from recovery, development of adverse reactions, incidence of diseases, and relapse, to death or any other specified occurrence an individual may experience (Johnson & Shih, 2007:273; Kleinbaum & Klein, 2005:4; Singh & Mukhopadhyay, 2011:145; Swinscow & Campbell, 2002:126; Zwiener *et al*., 2011:163). Survival analysis is a group of techniques used to model and statistically analyse data, for which the variable of interest is time to which an event occurs (Gardiner & Luo, 2008:1019; Jager, van Djik *et al*., 2008:560; Kleinbaum & Klein, 2005:4; Singh & Mukhopadhyay, 2011:145).

Survival analysis plays an essential role in epidemiology (Tolley *et al*., 2016:262). The method aims to determine and interpret survival and hazard functions using survival data, compare survival and hazard functions and examine the correlation between certain exploratory variables and survival time (Kleinbaum & Klein, 2005:15; Singh & Mukhopadhyay, 2011:148).

A clearly defined starting point of study is required for survival analysis and outcomes must be dichotomous in nature, with each participant having only one outcome during the study (dos Santos Silva, 1999:263). Time to failure or survival time is the time from an initial event or a defined starting time to the onset of an event of interest (dos Santos Silva, 1999:263; Johnson & Shih, 2007:273; Tolley *et al*., 2016:263). Survival data have two key features: lengths of follow up vary among study subjects and target events may never occur in some study subjects at the end of the study (Gardiner & Luo, 2008:1019; Johnson & Shih, 2007:273).

Censoring, a distinguishing phenomenon in survival analysis (Clark *et al*., 2003:232), occurs if the outcome of interest is not observed at the study termination, leaving the survival time unknown for a subgroup of study participants. Censoring may be because of a study participant not yet experiencing the outcome before the end of the study, loss-to-follow-up of patients during the study period or patients experiencing a different outcome, such as death, that makes it impossible to follow up further (Clark *et al*., 2003:232; Gardiner & Luo, 2008:1019; Johnson & Shih, 2007:273; Kleinbaum & Klein, 2005:6; Rich *et al*., 2010:331; Singh & Mukhopadhyay, 2011:146; Tolley *et al*., 2016:263; Zwiener *et al*., 2011:163). Right censoring occurs at the end of the investigation period; left censoring occurs

when the outcome is observed at an unknown time in a subject, whereas interval censoring means the subjects go in and out of observation (Clark *et al*., 2013:232; Sullivan, 2016; Tolley *et al*., 2016:262). Survival analysis assumes that censoring is independent or non-informative and subjects who have censored data would have the same distribution of survival times if they were observed until the events occurred (Gardiner & Luo, 2008:1021; Jager, van Djik *et al*., 2008:561; Johnson & Shih, 2007:273; Kirkwood & Stern, 2003:272; Sullivan, 2016). A key feature of survival analysis is the inclusion of censored data (Virnig *et al*., 2000:86), making it necessary that special analysis be employed (Clark *et al*., 2013:232). In section 2.6.3.1, survival and hazards functions, time-dependent functions that are important in survival analysis (Bewick *et al*., 2004:389, Kleinbaum & Klein, 2005:8), are discussed.

#### 2.6.3.1 Survival and hazard functions

Survival function or probability, *S(t)* is the probability that a subject will survive or will be free of an outcome beyond a specified time, *t,* while the hazard function, *h(t)* is a measure of risk that represents the instantaneous potential per unit time of experiencing the outcome, having survived until time, *t* (Bewick *et al*., 2004:389; Clark *et al*., 2003:233; Gardiner & Luo, 2008:1019; Kirkwood & Stern, 2003:272; Kleinbaum & Klein, 2005:9; Singh & Mukhopadhyay, 2011:146; Tolley *et al*., 2016:264). Survival probability at any time, *t*, is computed by the formula (Hoffman, 2015:632):

$$S(t) = \frac{Number\ of\ subjects\ free\ of\ outcome\ beyond\ time,t}{Total\ number\ of\ study\ subjects}$$

Where:

> *S(t)* denotes the probability that a subject will survive or will be free of an outcome beyond a specified time, *t*

The survival probability which gives vital summary information from time to event data, directly describing the survival experience of a population at various time points (Clark *et al*., 2003:233; Johnson & Shih, 2007:274), can be illustrated graphically by the survival curve, in which the survival function is plotted against time in years (Kirkwood & Stern, 2003:272; Sullivan, 2016). The life table and Kaplan-Meier methods are applied to model survival curves (Hoffman, 2015:621; Kirkwood & Stern, 2003:272; Sullivan, 2016; Tolley *et al*., 2016:269).

The hazard probability is the conditional odds that a person, having survived to time, *t*, will experience the outcome in a short period (Bewick *et al*., 2004:389; Hoffman, 2015:633) and is the slope of the survival curve at time, *t*. The following mathematical formula expresses how hazard function is calculated (Hoffman, 2015:633; Kleinbaum & Klein, 2005:11):

$$h(t) = \lim_{\Delta t \to \infty} \frac{P(t \leq T \leq t+\Delta t \mid T \geq t)}{\Delta t}$$

Where:

*h(t)* denotes the hazard function

$\Delta t$ denotes a small change in time.

$T$ denotes a continuous random variable

This probability gives an understanding of the conditional rates of experiencing an outcome, acts as a means of specifying the survival model, and lends itself for use as a diagnostic tool (Clark *et al*., 2003:233). Unlike the survival function which is concerned with the event of interest not happening, the hazard function focusses on the incidence of the event (Clark *et al*., 2003:233; Kleinbaum & Klein, 2005:10). The hazard function is thus related to the incident event rate whereas the survival probability provides the cumulative non-occurrence (Clark *et al*., 2003:233).

#### 2.6.3.2 Survival models

A survival model analyses time-to-event historical data and is used to generate estimates that show changes in the risk of the event taking place over time and help policy makers better estimate expected timing of certain outcomes (Singh & Mukhopadhyay, 2011:146). Sections 2.6.3.2.1-2.6.3.2.2 discuss the various models used for the analysis of survival data.

##### 2.6.3.2.1 Life tables

According to Kirkwood and Stern (2003:273), life tables display communal survival patterns when the researcher knows the number of individuals who survive at a series of time points but not the exact individual survival times. The life table generally begins with time of entry and a survival of 100% which decreases over time, periodically showing the occurrence of events (Hoffman, 2015:622). Two forms exist: the cohort or follow-up life table which summarises the real-time survival of a group of individuals over a pre-defined follow-up period (Kirkwood & Stern, 2003:273; Sullivan,

2016) and the current life table which shows the expected survival of a hypothetical population to which age-specific rates of death have been applied (Kirkwood & Stern, 2003:273).

To construct a life table, time is usually divided into equally spaced intervals and data are based on the number of subjects at risk of an outcome when each time interval commences, the number of subjects who have the outcome during each interval, and the number censored at the end of each interval. The survival probability is then calculated using the formula (Hoffman, 2015:624; Kirkwood & Stern, 2003:273; Sullivan, 2016:15; Tolley *et al*., 2016:269):

$$P = 1 - \frac{Number\ of\ subjects\ with\ events\ in\ the\ period}{Number\ of\ subjects\ without\ the\ event\ at\ the\ beginning\ of\ the\ period}$$

Where:

P denotes the survival probability

Cumulative survival probability is further calculated based on the principles of conditional probability using the formula (Kirkwood & Stern, 2003:275):

$$S(i) = Chance\ of\ surviving\ to\ month\ (i-1) x\ chance\ of\ surviving\ month\ i.$$

Where:

$S(i)$ denotes

A shortfall of the life table approach is that survival probabilities tend to vary depending on how the researcher organises the time intervals especially where samples are small; an issue the Kaplan-Meier approach addresses by re-estimating survival probabilities each time an event occurs (Sullivan, 2016).

#### 2.6.3.2.2 The Kaplan-Meier approach

The Kaplan-Meier approach uses the exact event and censoring times to estimate survival probabilities known as 'Kaplan-Meier estimates' or product-limit estimates (dos Santos Silva, 1999:270; Kirkwood & Stern, 2003:277). According to Goel *et al*. (2010:275) and Sullivan (2016), when using the Kaplan-Meier approach, it is important to assume that survival probabilities are the same for all subjects irrespective of when they are recruited into the study, patients censored have the same survival chances as those who are not, and events occur at the times specified. In the product-limit approach, time intervals are not defined by fixed lengths but are based on the happening

of an outcome (Tolley *et al*., 2016:270). The survival probability at time, *t*, is estimated using the formula (Kirkwood & Stern, 2003:277):

$$S_t = 1 - r_t = \frac{n_t - d_t}{n_t}$$

Where:

$S_t$ is the survival probability at time *t*

$r_t$ is the estimated risk at time *t*

$n_t$ is the number of individuals at risk

$d_t$ is the number of events that occur at time t.

The survival probability until and including event *j* is then given by (Kirkwood & Stern, 2003:277):

$$S(t_j) = S(t_{(j-1)}) x S_{t_j} = S_{t_1} x S_{t_2} x \ldots x S_{t_j}.$$

When the Kaplan-Meier estimates are used to plot cumulative survival probabilities, a step function is obtained: a horizontal line showing times of no events and a vertical drop signifying the occurrence of an event at a precise time and subsequent change in survival function (dos Santos Silva, 1999:270; Kirkwood & Stern, 2003:278; Rich *et al*., 2010:333). Since the researcher does not know if events of interest would have occurred in censored subjects, the survival curve based on the Kaplan-Meier approach should be considered an estimate once the first patient is censored (Rich *et al*., 2010:336). Kirkwood and Stern (2003:277) noted the relative stability of this approach since loss-to-follow-up of study subjects does not influence the survival probability estimation.

Parametric methods, which assume a defined mathematical formula for calculating the probability of an outcome and the number of survivors in a fixed time interval, are also used to model survival curves (Tolley *et al*., 2016:265). The Weibull, Gompetz, gamma, log-normal and log-logistic functions are parametric survival function distributions that are popularly used (Bradburn *et al*., 2003:434; Kirkwood & Stern, 2003:294; Tolley *et al*., 2016:265).

#### 2.6.3.3 Comparing survival functions

Survival analysis focusses on comparing the survival patterns of different groups (Kirkwood & Stern, 2003:278), the differences of which must be quantified to assess significance statistically (Rich *et al*., 2010:334). Several methods exist for the comparison of survival functions; some of which will be discussed in the following sections.

##### 2.6.3.3.1 Log-rank test

The log-rank test is the most widely-used approach for comparing survivals among independent groups which considers the entire follow-up period (Bland & Altman, 2004:1073; Rich *et al*., 2010:334; Sullivan, 2016) and tests the null hypothesis that their survival curves are equal (Bewick *et al*., 2004:391; Bland & Altman, 2004:1073; Goel *et al*., 2010:276; Jager, van Djik *et al*., 2008:563; Singh & Mukhophadyay, 2011:147; Sullivan, 2016). The test is more likely to identify the differences in the survival of groups when the chances of an outcome occurring is persistently higher for one group than the others (Bland & Altman, 2004:1073; Singh & Mukhopadhyay, 2011:147). Similar to the Kaplan-Meier method, the log rank test assumes that censoring is uninformative, survival probabilities are the same for subjects irrespective of their recruitment time, and events occur at the specified times (Bland & Altman, 2004:1073; Jager, van Dijk *et al*., 2008:563; Singh & Mukhopadhyay, 2011:147). A log rank statistic is calculated with the total number of expected and observed events for each group according to the formula (Bewick *et al*., 2004:391; Goel *et al*., 2010:276; Swinscow & Campbell, 2002:130):

$$Log\ rank\ statistic = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2}$$

Where:

> $O1$ and $O2$ represent the total number of observed events for groups 1 and 2, respectively $E1$ and $E2$ are the total number of expected events for groups 1 and 2, respectively.

The log rank statistic is further compared with a chi-square distribution with 1 degree of freedom to draw significance and to confirm that the survival of the groups being compared are not the same (Bewick *et al*., 2004:391; Goel *et al*., 2010:276).

A significant advantage of this test is that there is no need for prior knowledge of the shape of the survival curve nor the distribution of the survival times (Bland & Altman, 2004:1073). The test is, however, univariate by nature, describing survival with respect to the variable under investigation and ignoring the effect of other factors (Bewick *et al*., 2004:392; Clark *et al*., 2003:431; Goel *et al*., 2010:277; Jager, van Dijk *et al*., 2008:565). The lack of effect size to compliment the *p*-value, together with its inability to estimate the degree of the difference between the survival of groups, limits the efficiency of the test. Bland and Altman (2004:1073) added that the test is unlikely to identify difference in survival when the curves of groups cross.

#### 2.6.3.3.2 Hazard ratio

The hazard ratio is a descriptive measure used to compare survival between independent groups by providing relative event rates in the groups (Hoffman, 2015:634; Rich *et al*., 2010:335; Zwiener *et al*., 2011:167). The ratio is defined as the ratio of hazard occurring in one group to that of another at an exact time (Goel *et al*., 2010:277; Singh & Mukhophadyay, 2011:146) and is computed as a quotient of the hazards of the groups being compared according to the mathematical equation (Zwiener *et al*., 2011:167):

$$Hazard\ ratio = \frac{h_2(t)}{h_1(t)}$$

Where:

$h_1(t)$ and $h_2(t)$ are the hazards of groups 1 and 2, respectively, at time, *t*.

Hazard ratios give an idea of how much higher an event probability is, in one group, than the one to which it is being compared (Zwiener *et al*., 2011:167). In clinical trials, the hazard ratio can be used to understand the extent to which a treatment can shorten disease duration (Spruance *et al*., 2004:2787). Hazard ratios are better estimated using regression modelling techniques such as the Cox regression method (Clark *et al*., 2003:236).

#### 2.6.3.3.3 Cox-proportional hazards model

This model is semi-parametric because it does not specify the base hazard function (Singh & Mukhopadhyay, 2011:147; Sullivan, 2016). The model is a regression, used for analysing survival data, which estimates hazard ratio and confidence interval (Bradburn *et al*., 2003:431; Singh & Mukhopadhyay, 2011:147; Spruance *et al*., 2004:2787) and tests the effect of several risk factors

and independent variables on the survival experience of different groups (Bewick *et al*., 2004:392; Goel *et al*., 2010:277; Sullivan, 2016). It presumes that the hazard ratios of various groups being compared are insensitive to time and thus remain constant over the study period; a phenomenon known as the proportional hazards assumption (Bewick *et al*., 2004:392; Bradburn *et al*., 2003:432; Hoffman, 2015:636; Kirkwood & Stern, 2003:287; Singh & Mukhopadhyay, 2011:147; Spruance *et al*., 2004:2787; Sullivan, 2016; Zwiener *et al*., 2011:167).

The mathematical expression of the model is (Kirkwood & Stern, 2003:288):

$$\log\big(h(t)\big) = \log\big(h_0(t)\big) + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p$$

Where:

*h(t)* is the hazard at time *t*

$h_0(t)$ is the baseline hazard at time *t*

$x_1$ to $x_p$ are the *p* exposure variables.

The Cox-proportional hazards model is robust (Hoffman, 2015:636; Singh & Mukhopadhyay, 2011:147), powerful and flexible (Spruance *et al*., 2004:2791). It can estimate effect sizes (Zwiener *et al*., 2011:167) and calculate hazard ratios (Singh & Mukhopadhyay, 2011:147). The model gives the increased chances of an event incidence in one group compared to another (Rich *et al*., 2010:147) and allows for adjustment of the imbalance between groups (Hoffman, 2015:636). A Cox model can, however, not be applied in instances where the survival curves of populations cross (Hoffman, 2015:636).

### 2.6.4 Prescription sequence symmetry analysis (PSSA)

Observational pharmacoepidemiology addresses the limitations of clinical trials and spontaneous reporting, complementing these methods to provide in-depth knowledge of the safety profile of medications as used in clinical practice (Lesko & Mitchell, 2013:271). Prescription sequence symmetry analysis (PSSA) uses administrative claims data to assess a relationship between medications and potential adverse events (Wahab *et al*., 2016:349) by examining the symmetry in sequence of medication initiation and the onset of an indicator of the adverse event in a specific time frame (Pratt *et al*., 2013:916; Wahab, Pratt, Kalisch, *et al*., 2013:2; Wahab, Pratt, Wiese, *et al*., 2013:496). The approach was proposed by Hallas (1996:478) and was first used to investigate the relationship between cardiovascular medication initiation and depression. Prescription sequence

symmetry analysis has been further used to assess the adverse drug events associated with use of inhaled corticosteroids (Van Boven *et al*., 2013:231), anti-epileptic drugs (Tsiropoulos *et al*., 2009:483), isotretinoin (Hersom *et al*., 2003:424), and sulpiride (Lai *et al*., 2014:1), among others.

In PSSA, the researcher determines the ratio of the sequences of two medications; an index suspected to be the cause of the adverse event and an indicator used to counter the event (Wahab *et al*., 2016:349). The sequence of dispensing the index and indicator medications in the same individual is determined and the number of subjects who were dispensed the indicator medication after the index is divided by the number of those who received the indicator medication before the index medication to determine the crude sequence ratio — a measure of the extent of symmetry (Bytzer & Hallas, 2000:1480; Hashimoto *et al*., 2015:2; Hersom *et al*., 2003:426; Pratt *et al*., 2014:2; Pratt *et al*., 2015:862; Van Boven *et al*., 2013:232; Wahab, Pratt, Wiese, *et al*., 2013:498; Wahab *et al*., 2016:349).

The PSSA is robust towards time-constant variables but sensitive to trends in prescribing over time, requiring a null-effect sequence ratio to control for changing trends. The null-effect sequence ratio estimates the expected sequence ratio due to the trends if the medication and event being studied have no causal relationship. The ratio of the crude sequence ratio to the null-effect sequence ratio produces an adjusted sequence ratio (Hashimoto, 2015:3; Pratt *et al*., 2014:2; Pratt *et al*., 2015:862; Wahab, Pratt, Wiese, *et al*., 2013:498; Wahab *et al*., 2016:349).

One would expect a symmetrical distribution of the indicator drug initiation before and after the index drug is initiated if there is no causal relationship (Lai *et al*., 2017:568). The validity of PSSA depends on the quality and specificity of the indicator of the adverse event (Lai *et al.,* 2017:569) which may be a medication used to treat or hospitalisation that describes the event (Wahab, Pratt, Kalisch, *et al*., 2013:2; Lai *et al*., 2017:569). The method requires the use of an appropriate exposure time frame based on the probable time course of the adverse event development, although one year may be optimal for obtaining adequate sensitivity and positive predictive value in signal detection studies for which there are no specific hypotheses (Lai *et al*., 2017:579).

Use of only three variables — patient identifiers, medication code, and date on which the medication is dispensed — all of which are readily available in administrative claims databases (Pratt *et al*., 2013:916; Pratt *et al*., 2014:1; Lai *et al*., 2017:570) makes the method computationally efficient and simple to use (Hallas, 1996:478; Lai *et al*., 2017:570; Pratt *et al*., 2014:1). The inherent ability to make adjustments for confounding variables that remain stable over time is another important advantage. The method uses study subjects that serve as their own controls and therefore does not depend on

numerical adjustment to control for these time-invariant confounders (Bytzer & Hallas, 2000:1480; Hallas, 1996:478; Pratt *et al*., 2013:918; Pratt *et al*., 2014:1; Van Boven, 2013:232; Wahab, Pratt, Kalisch, *et al*., 2013:2; Wahab *et al*., 2016:349). Graphical output of PSSA allows signals to be better interpreted and a possible temporal association to be observed (Lai *et al*., 2017:570; Pratt *et al*., 2015:863). Wahab, Pratt, Wiese, *et al*. (2013:496) found the method to have a high specificity and moderate sensitivity for identifying adverse events while Pratt *et al*. (2015:858) proved its consistency over time.

Prescription sequence symmetry analysis is only applicable in post-market surveillance studies where medications are prescribed for treating adverse events or the event leads to hospitalisation (Pratt *et al*., 2014:8). There is the tendency to underestimate associations if patients who experience the events discontinue medications or obtain treatment for events over-the-counter since these will not be recorded in the administrative claims datasets (Lai *et al*., 2017:570; Wahab, Pratt, Kalisch, *et al*., 2013:4). Prescribing trends over time, confounding by indication, protopathic bias, as well as time-varying confounders can bias causal associations identified using the PSSA (Bytzer & Hallas, 2000:1483; Hashimoto *et al*., 2015:2; Lai *et al*., 2017:579; Pratt *et al*., 2013:920; Wahab, Pratt, Kalisch, *et al*., 2013:2; Wahab *et al*., 2016:349). Positive signals generated by PSSA do not provide sufficient evidence of causal evidence since it uses only prescription dispensing data and admission data from hospitals without considering the clinical condition of the patients (Wahab, Pratt, Kalisch, *et al*., 2013:3).

## 2.7 Chapter summary

The chapter gave a background of pharmacoepidemiology and its significance in health research. Study designs employed in pharmacoepidemiology, sources of data for pharmacoepidemiological studies, as well as the measures assessed in pharmacoepidemiology were examined. Pharmacoepidemiological studies that have been carried out using secondary data were summarised, together with the methods employed.

The subsequent chapter discusses the results and findings of the empirical study and how the stated objectives were met.

# CHAPTER 3: RESULTS AND DISCUSSION

## 3.1 Introduction

This chapter focuses on the results of the empirical investigation phase of the study. The specific objectives of the empirical study were to:

(i) Determine the time to onset of treatment of hypertension and hyperlipidaemia in patients with type 2 diabetes mellitus using survival analysis.

(ii) Compare different adherence measures, by determining adherence to montelukast among asthma patients, using data from a medicines claims database in South Africa.

Manuscript one, titled “Time to onset of treatment of hypertension and hyperlipidaemia in type 2 diabetes mellitus patients: a survival analysis”, addressed the second objective to compare the time to onset of treatment of hypertension and hyperlipidaemia in patients with type 2 diabetes mellitus using survival analysis. This manuscript was prepared for submission to the *Journal of clinical pharmacy and therapeutics*. The author guidelines for the guideline can be found in Annexure C. Proof of submission of manuscript one to *Journal of clinical pharmacy and therapeutics* can be found in Annexure D.

Manuscript two addressed the third objective, *viz.* to compare different adherence measures, by determining adherence to montelukast among asthma patients, using data from a medicines claims database in South Africa. Manuscript two, titled “Comparison of adherence measures using administrative claims data among asthma patients receiving montelukast” was prepared for submission to *Pharmacoepidemiology and drug safety*. The author guidelines can be found at Annexure E. Proof of submission of manuscript two to *Pharmacoepidemiology and drug safety* can be found in Annexure F.

Each manuscript was written according to the guidelines prescribed for authors by the respective journals. For uniformity, the manuscripts included in this dissertation were formatted to Arial font size 11, with line spacing of 1.5. The reference list of each manuscript was formatted according to journal guidelines and included in the reference list of the dissertation in the Harvard format.

The role of each author in both manuscripts is depicted in Table 3-1.

**Table 3-1: Author's roles and responsibilities**

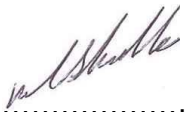
| Author | Role in the study |
|---|---|
| Ms M Obeng-Kusi (Researcher) | Manuscript organisation and design<br>Interpretation of the statistical analyses<br>Writing up of results |
| Prof JR Burger (Supervisor) | Concept of research<br>Manuscript supervision<br>Supervising the results interpretation<br>Revision of the manuscript for important content |
| Prof MS Lubbe (Co-supervisor) | Concept of research<br>Data and statistical analyses<br>Guidance on writing of manuscript<br>Revision of the manuscript for important content. |
| Mrs M Cockeran (Co-supervisor) | Data and statistical analyses<br>Verified results from statistical analyses<br>Revision of the manuscript for important content |

The following statement provided by the co-authors confirms their roles in the study and authorize the inclusion of the manuscripts in the dissertation.

*I declare that I have approved the above-mentioned manuscripts and that my role in this study, as indicated above, is a representation of my actual contribution, and I hereby give my consent that it may be published as part of the MPharm study of M Obeng-Kusi.*

……..

Prof JR Burger



……

Prof MS Lubbe

……

Mrs M Cockeran

## 3.2 Manuscript 1: Determining the time to onset of treatment of hypertension and hyperlipidaemia in patients with type 2 diabetes mellitus

**Manuscript title: Time to onset of treatment of hypertension and hyperlipidaemia in diabetes mellitus patients: a survival analysis**

**Running head: Hypertension and hyperlipidaemia among people with diabetes**

M. Obeng-Kusi[1] BPharm M. S. Lubbe[1] PhD M. Cockeran[2] MSc (Statistics) J. R. Burger[1] PhD

[1] Medicine Usage in South Africa (MUSA), Faculty of Health Sciences, North-West University, Potchefstroom Campus, Potchefstroom.

[2] Statistics, School of Computer, Statistical and Mathematical Sciences, North-West University, Potchefstroom Campus, Potchefstroom

**Correspondence**

J.R. Burger, Medicine Usage in South Africa (MUSA), Faculty of Health Sciences, North-West University, Potchefstroom Campus, Potchefstroom.

Telephone: +27827883319, +27129400091

Fax: +27872314105

Email: Johanita.Burger@nwu.ac.za

**Funding information**

The authors received financial support from the North-West University (Master's bursary 27959716) and the National Research Foundation (Grant number: 85315).

**Acknowledgement**

The authors thank the PBM company for allowing the use of the database in this study. We also appreciate Ms Anne-Marie Bekker for supporting us with data management and Ms. Helena Hoffman for proof-reading and editing this manuscript.

**Conflict of interest**

We report no conflict of interest where this research is concerned.

**Summary**

What is known and Objective

Hypertension and hyperlipidaemia have high prevalence among people with diabetes and increase patients' risk of cardiovascular diseases, ultimately affecting prognosis negatively. Medicines claims data have gained prominence in the study of drug-related events and outcomes. There is paucity of publications on the time to onset of treatment of these conditions among South African patients with diabetes using secondary data. This study aims to apply survival analysis to determine the time to onset of treatment of hypertension and hyperlipidaemia among people with diabetes using a South African medicines claims data.

Methods

Survival analysis was conducted using retrospective data of patients enrolled continuously with a Pharmaceutical Benefit Management (PBM) company in South Africa from 1st January 2008 to 31st December 2016. Information on patients' sex and date of birth were used, with encrypted membership numbers, to follow-up patients' prescriptions over time. We identified patients based on International Classification of Diseases, Tenth Revision (ICD-10) diagnoses codes in conjunction with National Pharmaceutical Product Index (NAPPI) codes provided by the Monthly Index of Medical

Specialities (MIMS) classification codes and extracted their data using SAS® system version 9.4 classification codes. Patients whose sex and age fields were incompletely filled were excluded. The Kaplan-Meier approach, used to compare the survival experience of patients who commenced treatment for hypertension and hyperlipidaemia, was conducted using IBM® SPSS® version 25.

Results and discussion

A total of 494 patients with an average age of 53.5 (SD 11.1) years were included in the study, 34.8% of whom were females. Prevalence of hyperlipidaemia and hypertension among patients were 35.0% and 45.6%, respectively. Average time to onset of treatment of hyperlipidaemia was 2684.4 (SD 42.2) days compared to 2434.2 (SD 47.6) days for hypertension. There was no statistically significant difference in age and sex among patients who developed either of these conditions during the study ($p = 0.404$; Cohen's $d = 0.132$ for hyperlipidaemia and $p = 0.644$, Cohen's $d = 0.059$ for hypertension).

What is new and Conclusion

Within an average of six years after an index period of 1 year free of disease, people with diabetes may commence treatment of hyperlipidaemia, hypertension or both. With all significant data appropriately captured, medicines claims data can be effectively used in survival analysis to determine time to onset of treatment of hyperlipidaemia and hypertension among people with diabetes.

**Summary in three sentences.**

The occurrence of hypertension and hyperlipidaemia in diabetes mellitus patients increase their cardiovascular risk and negatively affect prognosis. To better access and manage these risk factors, it is important that the time to which they occur in people with diabetes be assessed. Survival analysis of data from a privately-owned Pharmaceutical Benefits Management (PBM) company in South Africa

estimated the time to onset of treatment of both hypertension and hyperlipidaemia among diabetes mellitus patients to be approximately 6 years after an index period of one disease free year.

## 1. WHAT IS KNOWN AND OBJECTIVE

Diabetes is associated with high cardiovascular mortality, with the rates of cardiovascular diseases in diabetic patients being about double that in non-diabetic patients.[1-3] Cardiovascular diseases are abnormal conditions of the heart and blood vessels that lead to their dysfunction;[4,5] a group of disorders that include cerebrovascular disease, deep vein thrombosis, congenital heart disease, pulmonary embolism, rheumatic heart disease, peripheral heart disease and coronary heart disease.[6]

Hypertension and hyperlipidaemia are also common in people with diabetes and increase their risk of cardiovascular diseases.[3,6-8] A study among people with diabetes in Nigeria found the prevalence of hypertension to be 54.2%,[9] consistent with another research conducted in the United States of America that established a high hypertension prevalence rate of 59.4% among diabetic patients.[10] According to Shah and Afzal,[11] coexistence of hypertension and diabetes increases the risk of cardiovascular diseases by about 75.0% with a subsequent rise in total morbidity and mortality. Dixit et al[12] found that over 70.0% of people with diabetes had a form of dyslipidaemia, a dominant proportion of who had hyperlipidaemia. Higher dyslipidaemia prevalence rates of 93.5% and 95.0% were recorded among South African[13] and Pakistani diabetic patients,[14] respectively.

The presence of cardiovascular disease significantly affects the prognosis of diabetes.[2] There is the need to study the interaction between these conditions to facilitate better management and ultimately to improve prognosis. Risk management in healthcare is critical for maintaining clinical quality and impacts financial performance,[15] both of which are critical in a managed healthcare environment since managed healthcare organisations influence clinical decisions to increase their cost-effectiveness and productivity.[16] It is important that risk factors be well elucidated for appropriate mitigation measures to be put in place. Several studies have been conducted to investigate these risk factors of cardiovascular disease.[2,3,7,8,17] There is, however, no study that has examined the time to onset of treatment of these comorbidities relative to one another particularly using medicines claims data from the private health sector of South Africa.

The objective of the study was to determine the time to onset of treatment of hypertension and hyperlipidaemia in patients with type 2 diabetes over a ten-year period using South African medicines claims data through the application of survival analysis.

## 2. METHODS

This retrospective cohort study was carried out using medicines claims data for the period 1st January 2008 to 31st December 2016 obtained from a South African Pharmaceutical Benefit Management (PBM) company's central database. The independent PBM company has, over the past 25 years, been processing medicines claims for about 42 medical schemes with 1.6 million beneficiaries in South Africa. Data on about a third of the patients enrolled with private medical aids in South Africa were obtained from this PBM. Information on patients' sex and date of birth were used, with encrypted membership numbers, to follow-up patients' prescriptions over time. Drug trade names, dates on which prescriptions were filled, as well as International Classification of Diseases, Tenth Revision (ICD-10) codes for diagnoses were extracted from the database.

The study population consisted of patients who had a diagnosis code (International Classification of Diseases, Tenth Revision (ICD-10) code E11) for type 2 diabetes mellitus and were receiving antidiabetic medication according to the National Pharmaceutical Product Index (NAPPI) codes provided by the Monthly Index of Medical Specialities (MIMS) classification code 19.1.[18] All antidiabetic medications were considered for this investigation. We selected patients among these who had ICD-10 codes for hypertension (I10, I11, I12, I13, I15, O10, and O11) and hyperlipidaemia (E78.5) and were receiving medications classified according to NAPPI codes provided by MIMS.[18] We excluded patients whose sex and age fields were incompletely filled, as well as those who were not enrolled continuously on the database throughout the period of investigation (N = 2 996). The time to the incidence of hypertension and hyperlipidaemia among the diabetic patients were measured in days. With 2008 serving as the index year, we followed up on patients until 31st December 2016.

The Kaplan-Meier approach was used to analyse the time to onset of treatment of hypertension and hyperlipidaemia among sex and age groups. Patients were divided into age quartiles based on descriptive statistics, viz. Age group 1 (≤51 years), Age group 2 (>51, ≤59 years), Age group 3 (>59, ≤66 years), and Age group 4 (>66 years) for comparison of time to onset of treatment among these. Independent *t*-tests were conducted to compare patients' characteristics. *P*-values < 0.05 were deemed statistically significant. The SAS® system version 9.4[19] was used to extract study population and for baseline analysis whereas IBM® SPSS® version 25[20] was used for the survival analysis.

## 3. RESULTS

A total of 494 patients were involved in this study, 34.8% of which were females. A summary of the patient characteristics is shown in Table 1. The mean age of the study population was 53.5 years (SD 11.4). Males in our study were marginally older than the females at 54.5 (SD 11.1) *vs*. 51.3 (SD 12.1) years ($p = 0.003$; Cohen's $d = 0.3$).

Table 2 summarises the incidence of hyperlipidaemia and hypertension treatments among males and females and the various age groups. Of the study population, 35.0% (n = 173) commenced treatment for hyperlipidaemia during the study compared to 45.6% who started treatment for hypertension. Among men (N = 322), 141 (43.8%) started hypertension treatment whereas 116 (36.0%) commenced treatment for hyperlipidaemia. In comparison, 57 (33.1%) women (N = 172) began treatment of hyperlipidaemia while 84 (48.8%), started receiving treatment for hypertension. Of the 232 patients aged less than 51 years, 73 (31.5%) and 106 (45.7%) started treating hyperlipidaemia and hypertension, respectively. Of the 102 patients between ages 51 and 59 years, 38 (37.3%) patients commenced treatment for hyperlipidaemia while 46 (45.1%) also commenced treatment for hypertension. Among the 82 patients aged between 60 and 66 years, the incidence rates of onset of treatment for hyperlipidaemia and hypertension were 39.0% and 45.1%, respectively. From the 78 patients above age 66 years, 30 (38.5%) developed hyperlipidaemia and 36 (46.2%) developed hypertension. There was no age difference between males and females with

diabetes who received treatment for hyperlipidaemia ($p$ = 0.404; Cohen's $d$ = 0.132) nor hypertension ($p$ = 0.644, Cohen's $d$ = 0.059) (Table 3).

A summary of the time to onset of treatment of hypertension and hyperlipidaemia among patients is presented in Table 4. The mean time to onset of treatment of hyperlipidaemia and hypertension were 2684.4 days (SD 42.2) and 2434.2 (SD 47.6) days, respectively, with no statistically significant differences in the survival experiences according to age categories ($p$ = 0.578 and $p$ = 0.964, respectively) nor sex ($p$ = 0.563 and $p$ = 0.139, respectively). While the mean time to onset of treatment of hypertension among female patients was 2300.8 days (SD 85.6), that among male patients was 2424.2 days (SD 56.5). The average time to onset of treatment of hyperlipidaemia was lower (2673.6 days, SD 52.4) in the male patients than in female patients (2704.6 days, SD 71.1). For hypertension, patients in age range 60-66 years had the highest average survival time (2568.7 days: SD 104.4) while patients aged less than 51 years commenced treatment for hyperlipidaemia later than patients in the other age groups (2718.3, SD 61.9). Table 5 highlights the comparisons among age and sex groups. Figures 1-4 presents the Kaplan-Meier plots for times to onset of treatment of hyperlipidaemia and hypertension among males and females in various age groups.

## 4. DISCUSSION

Our study examined the time to onset of treatment of hypertension and hyperlipidaemia in people with diabetes. The higher incidence of hypertension among females whose average age was 53.7 years compared with their male counterparts of approximately the same age reflects trends that show that prevalence of hypertension is higher among pre-menopausal women than men at the same age.[21,22] This increasing prevalence of hypertension with increasing age in women has been credited to the depleting concentrations of oestrogen, which by vasorelaxation and prevention of vascular remodelling and aortic stiffness, plays a protective cardiovascular role.[23,24]

Consistent with the findings of previous studies,[25-28] male patients were found to have a higher prevalence of hyperlipidaemia than females. It has been proposed that while female sex hormones,

especially oestrogen, protects against rising lipid levels, testosterone may predispose men to elevated levels and thus, account for the high lipid profile of men compared to women.[29] A study conducted among people with diabetes in India reported a high dominance of hyperlipidaemia with female subjects having higher mean lipid concentrations that their male counterparts.[12] In the absence of data on race as well as clinical data from which blood concentrations can be evaluated, the results of our study cannot be evaluated relative to those of the former study. Although previous studies have found that the prevalence of hyperlipidaemia increases with increasing age,[28,30,31] our study reported no significant difference in the prevalence of hyperlipidaemia among the various age stratifications employed in the study. This may be because our study was conducted among a middle-aged South African population (mean age 53.5 years), among whom Maritz[32] had previously found no difference in prevalence of hypercholesterolaemia between males and females in contrast to lower prevalence among younger females than males and higher prevalence in females in older ages.

We found that within about six years after a one-year index period, people with diabetes had the propensity to commence treatment for either hyperlipidaemia, hypertension or both. Insulin has been proposed to play a vital role in the metabolic syndrome,[33] a cluster of conditions and multiplex of predisposing factors for cardiovascular disease to which diabetes, hypertension and hyperlipidaemia, *inter alia*, belong.[34] Insulin stimulates the uptake of glucose in skeletal muscles and the heart, and reduces the synthesis of glucose and triglyceride-rich particles.[35] Insulin resistance or deficiency impairs its ability to facilitate glucose uptake in the muscles and to prevent glucose synthesis leading to hyperglycaemia. According to Goldberg,[36] in people with diabetes, hyperglycaemia as well as defects in the insulin activity cause changes in plasma lipoproteins. Further, insulin — by the endothelial nitric oxide production — exerts anti-inflammatory and vasorelaxation activity,[34] leading to reduced blood pressure. The pathways of oxidative stress, endothelial damage, low grade inflammation and hypercoagulability have been found to be common to these conditions and have also been credited with their co-existence.[34]

To the best of our knowledge, ours is the first study that has employed survival analysis to study the relationship of diabetes with hypertension and hyperlipidaemia in the private health sector of South Africa using claims data. We employed extensive data from the private healthcare sector which is considered representative of the patients in South Africa for which the PBM provides medicines claims services, with a large proportion of privately-owned pharmacies in South Africa being by the PBM. Our study, however, was not devoid of limitations. In the absence of relevant data, the present study did not consider how long our study patients had been receiving treatment for diabetes. This is because although the Chronic Disease List, which contains diabetes along with other conditions for which medications and treatments are covered by PBMs in South Africa, was introduced in 2003,[37] the data from the PBMs on the chronic disease list were only reliable from 2008. It was, therefore, not feasible to go back further to determine how long patients had received treatment for diabetes. Also, the level of glucose control among patients could not be determined because of the lack of clinical data and thus, our results are limited in their ability to estimate the effect of glucose control on the onset of treatment of hypertension and hyperlipidaemia. Our dataset did not permit the classification of hyperlipidaemia as has been done by previous studies. Thus, the prevalence and time to which various forms of lipid alterations occur could not be determined. Finally, fairly wide confidence intervals determined in the time to onset of treatment of diabetes reduced the power to detect significant associations accurately. We recommend further longitudinal studies, in both private and public health sectors of South Africa, into the development of hypertension and the various forms of hyperlipidaemia in patients with type 1 and type 2 diabetes.

## WHAT IS NEW AND CONCLUSION

There was a male dominance of hyperlipidaemia while females were found to have a higher prevalence of hypertension among our study patients. With respect to mean age, there was no difference between the males and females with diabetes who started receiving treatment for both hyperlipidaemia and hypertension. The average time to onset of treatment of both hypertension and

hyperlipidaemia among people with diabetes was estimated to be about 6 years after a one-year index period, with no statistically significant variations between male and female patients of different ages. Although the absence of clinical data and information on other confounders limit their use, medicines claims data present an accessible and large sample size over an extensive period that can be applied in survival analysis for the estimation of time to development of medical outcomes.

**REFERENCES**


1. Centers of Disease Control and Prevention (CDC). National diabetes statistics report, 2017. Estimates of diabetes and its burden in the United States. Atlanta, GA: US Department of Health and Human Services, CDC; 2017. https://www.cdc.gov/diabetes/pdfs/data/statistics/national-diabetes-statistics-report.pdf. Accessed November 9, 2018
2. Laakso M. Cardiovascular disease in type 2 diabetes from population to man to mechanisms: the Kelly West Award Lecture 2008. *Diabetes Care*. 2010;33(2):442-449. doi:10.2337/dc09-0749.
3. Watkins PJ. Cardiovascular disease, hypertension, and lipids. *BMJ*. 2003;326(7394):874-876. doi:10.1136/bmj.326.7394.874.
4. Collins P. *Dictionary of Medical Terms*. 4th ed. London: A&C Black; 2005.
5. O'Toole MT. *Mosby's Medical Dictionary.* 9th ed. St. Louis, MO: Elsevier; 2013.
6. Goldberg RB. Hyperlipidemia and cardiovascular risk factors in patients with type 2 diabetes. *Am J Manag Care.* 2000;6(13 Suppl):S682-S691. https://www.ajmc.com/journals/supplement/2000/2000-08-vol6-n13suppl/aug00-570ps682-s691. Accessed November 10, 2018.
7. World Health Organization (WHO). Cardiovascular disease. http://www.who.int/cardiovascular_diseases/en/. Accessed October 22, 2018.

8. Weycker D, Nichols GA, O'Keeffe-Rosetti M, et al. Risk-factor clustering and cardiovascular disease risk in hypertensive patients. *Am J Hypertens*. 2007;20(6):599-607. doi:10.1016/j.amjhyper.2006.10.013.
9. Unadike BC, Eregie A, Ohwovoriole AE. Prevalence of hypertension amongst persons with diabetes mellitus in Benin City, Nigeria. *Niger J Clin Pract*. 2011;14(3):300-302. doi:10.4103/1119-3077.86772.
10. Gillespie CD, Hurvitz KA; Centers for Disease Control and Prevention (CDC). Prevalence of hypertension and controlled hypertension - United States, 2007-2010. *MMWR Morb Mortal Wkly Rep. Suppl*. 2013;62(3):144-148. https://www.cdc.gov/MMWr/preview/mmwrhtml/su6203a24.htm. Accessed November 10, 2018.
11. Shah A, Afza M. Prevalence of diabetes and hypertension and association with various risk factors among different Muslim populations of Manipur, India. *J Diabetes Metab Disord*. 2013;12:52. doi:10.1186/2251-6581-12-52.
12. Dixit AK, Dey R, Suresh A, Chaudhuri S, Panda AK. The prevalence of dyslipidemia in patients with diabetes mellitus of Ayurveda hospital. *J Diabetes Metab Disord*. 2014;13(58):1-6. doi:10.1186/2251-6581-13-58.
13. Daya R, Bayat Z, Raal FJ. Prevalence and pattern of dyslipidaemia in type 2 diabetes mellitus patients at a tertiary care hospital. *JEMDSA*. 2017;22(3):31-35. doi:10.1080/16089677.2017.1360064.
14. Sarfraz M, Sajid S, Ashraf MA. Prevalence and pattern of dyslipidemia in hyperglycemic patients and its associated factors among Pakistani population. *Saudi J Biol Sci*. 2016; 23(6):761–766. doi:10.1016/j.sjbs.2016.03.001.
15. Singh B, Ghatala MH. Risk management in hospitals. *Int J Innov Manag Technol*. 2012;3(4):417–421. http://www.ijimt.org/papers/266-CM244.pdf. Accessed October 9, 2018.

16. Plumeri PA. Managed care and risk management. *Gastroenterol Clin North Am*. 1997, 26(4):895-910. doi:10.1016/S0889-8553(05)70340-X.
17. Shah AD, Langenberg C, Rapsomaniki E, et al. Type 2 diabetes and incidence of cardiovascular diseases: a cohort study in 1·9 million people. *Lancet Diabetes Endocrinol.* 2015;3(2):105-113. doi:10.1016/S2213-8587(14)70219-0.
18. Snyman JR. Monthly Index of Medical Specialities. Cape Town, CTP Printers; 2017.
19. SAS® [computer program]. Version 9.4. Cary, NC: SAS Institute Inc.; 2002-2012.
20. SPSS® [computer program]. Version 25. Armonk, NY: IBM Corp.; 2017.
21. Akl C, Akik C, Ghattas H, Obermeyer CM. Gender disparities in midlife hypertension: a review of the evidence on the Arab region. *Womens Midlife Health*. 2017;3(1):1-10. doi:10.1186/s40695-017-0020-z.
22. Ghosh S, Mukhopadhyay S, Barik A. Sex differences in the risk profile of hypertension: a cross-sectional study. *BMJ Open*. 2016;6(7):e010085. doi:10.1136/bmjopen-2015-010085.
23. Dubey RK, Oparil S, Imthurn B, Jackson EK. Sex hormones and hypertension. *Cardiovasc Res*. 2002;53(3):688-708. doi:10.1016/S0008-6363(01)00527-2.
24. Fortepiani LA, Zhang H, Racusen L, Roberts LJ 2nd, Reckelhoff JF. Characterization of an animal model of postmenopausal hypertension in spontaneously hypertensive rats. *Hypertension*. 2003;41(3 Pt 2):640-645. doi:10.1161/01.HYP.0000046924.94886.EF.
25. Aguilar-Salinas CA, Olaiz G, Valles V, et al. High prevalence of low HDL cholesterol concentrations and mixed hyperlipidemia in a Mexican nationwide survey. *J Lipid Res*. 2001;42(8):1298-1307. http://www.jlr.org/content/42/8/1298.long. Accessed November 10, 2018.
26. Anil Babu A, Rani S, Revikumar KG, Shafeeque M, Unnikrishnan UG. Assessment of correlation in gender and age with lipoprotein levels in hyperlipidemia patients. *Asian J Pharm Clin Res*. 2016;9(2):197-199.

https://innovareacademics.in/journals/index.php/ajpcr/article/download/10373/4098. Accessed November 10, 2018.

27. Chen GY, Li L, Dai F, Li XJ, Xu XX, Fan JG. Prevalence of and risk factors for type 2 diabetes mellitus in hyperlipidemai in China. *Med Sci Monit*. 2015;21:2476-84. doi:10.12659/MSM.894246.
28. Deng B, Luo T, Huang Y, Shen T, Ma J. Prevalence and Determinants of Hyperlipidemia in Moderate Altitude Areas of the Yunnan-Kweichow Plateau in Southwestern China. *High Alt Med Biol*. 2012;13(1):13-21. doi:10.1089/ham.2011.1037.
29. Samaha AA, Zouein F, Gebbawi M, et al. Associations of lifestyle and dietary habits with hyperlipidemia in Lebanon. *VP*. 2017;1:98-106. doi:10.20517/2574-1209.2017.18.
30. Wang S, Xu L, Jonas JB, You QS, Wang YX, Yang H. Prevalence and associated factors of dyslipidemia in the adult Chinese population. *PLoS One*. 2011;6(3):e17326. doi:10.1371/journal.pone.0017326.
31. Wu JY, Duan XY, Li L, Dai F, Li YY, Li XJ, et al. Dyslipidemia in Shanghai, China. *Prev Med.* 2010;51(5):412-415. doi:10.1016/j.ypmed.2010.08.013.
32. Maritz FJ. Dyslipidaemia in South Africa. In: Steyn K, Fourie J, Temple N, eds. *Chronic diseases of lifestyles in South Africa: 1995-2005. Medical Research Council Technical Report.* Tygerberg, SA: Medical Research Council; 2006:97-108.
33. Cheung BM, Li C. Diabetes and hypertension: is there a common metabolic pathway? *Curr Atheroscler Rep*. 2012;14(2):160-166. doi:10.1007/s11883-012-0227-2.
34. Zhou MS, Wang A, Yu H. Link between insulin resistance and hypertension: What is the evidence from evolutionary biology? *Diabetol Metab Syndr*. 2014;6(1):12. doi:10.1186/1758-5996-6-12.
35. Yki-Järvinen H. Nonglycemic effects of insulin. *Clin Cornerstone*. 2003;5(Suppl 4):S6-S12. doi:10.1016/S1098-3597(03)90061-5.

36. Goldberg IJ. Diabetic dyslipidemia: causes and consequences. *J Clin Endocrinol Metab*. 2001 Mar;86(3):965-971. doi:10.1210/jcem.86.3.7304. Council for Medical Schemes, South Africa. Medical Schemes Act, 131 of 1998. https://www.medicalschemes.com/Content.aspx?130. Accessed November 11, 2018.

**Table 1: Patients' demographics**

| Characteristic | N[1] (%) |
|---|---|
| **Total number of patients, N** | 494 |
| **Age groups (years), n (%)** | |
| **≤ 51** | 232 (47.0) |
| **51-59** | 102 (20.7) |
| **60-66** | 82 (16.6) |
| **>66** | 78 (15.8) |
| **Mean (SD[2])** | 53.47 (11.4) |
| **Sex, n (%)** | |
| **Male** | 322 (65.2%) |
| **Female** | 172 (34.8%) |

[1] N: Number of patients
[2] SD: Standard deviation

**Table 2: Incidence of treatment of hyperlipidaemia and hypertension**

| | | Hyperlipidaemia | | Hypertension | |
|---|---|---|---|---|---|
| | **Total N** | **Number of events, n (%)** | **Number censored, n (%)** | **Number of events, n (%)** | **Number censored, n (%)** |
| **Age group (Years), n (%)** | 494 | 173 (35) | 321 (65.0) | 225 (45.5) | 269 (54.9) |
| <51 | 232 | 73 (31.5) | 159 (68.5) | 106 (45.7) | 126 (54.3) |
| 51-59 | 102 | 38 (37.3) | 64 (37.3) | 46 (45.1) | 56 (54.9) |
| 60-66 | 82 | 32 (39.0) | 50 (61.0) | 37 (45.1) | 45 (54.9) |
| >66 | 78 | 30 (38.5) | 48 (61.5) | 36 (46.2) | 42 (53.8) |
| **Sex, n (%)** | 494 | 173 (35.0) | 321 (65.0) | 225 (45.5) | 269 (54.5) |
| Male | 322 | 116 (36.0) | 206 (64.0) | 141 (43.8) | 181 (56.2) |
| Female | 172 | 57 (33.1) | 115 (66.9) | 84 (48.8) | 88 (51.2) |

**Table 3: *T*-test for equality of mean ages between sex groups**

| Condition | | Mean age in years (SD)[3] | *T*-test for equality of means | |
|---|---|---|---|---|
| | | | ***p*-value** | **Cohen's *d*-value** |
| **Hyperlipidaemia** | Males | 55.0 (10.6) | 0.404 | 0.132 |
| | Females | 53.6 (11.2) | | |
| **Hypertension** | Males | 54.4 (9.9) | 0.644 | 0.059 |
| | Females | 53.7 (11.3) | | |

[3] SD: Standard deviation

**Table 4: Mean time to onset of treatment of hyperlipidaemia and hypertension**

| | Hyperlipidaemia | | Hypertension | |
|---|---|---|---|---|
| | **Estimate (days)** | **95% CI[4]** | **Estimate (days)** | **95% CI** |
| **Age groups** | | | | |
| <51 | 2718.3 | 2597.0-2839.6 | 2426.7 | 2289.6-2563.9 |
| 51-59 | 2654.1 | 2475.5-2832.7 | 2373.1 | 2160.3-2586.0 |
| 60-66 | 2685.9 | 2493.6-2878.3 | 2568.7 | 2364.0-2773.3 |
| >66 | 2620.2 | 2400.5-2839.8 | 2394.6 | 2154.6-2634.6 |
| ***p*-values** | 0.578 | | 0.964 | |
| **Sex** | | | | |
| Male | 2673.6 | 2570.9-2776.3 | 2505.3 | 2394.6-2616.1 |
| Female | 2704.6 | 2565.1-2844.0 | 2300.8 | 2133.0-2468.7 |
| ***p*-values** | 0.563 | | 0.139 | |

[4] CI: Confidence interval

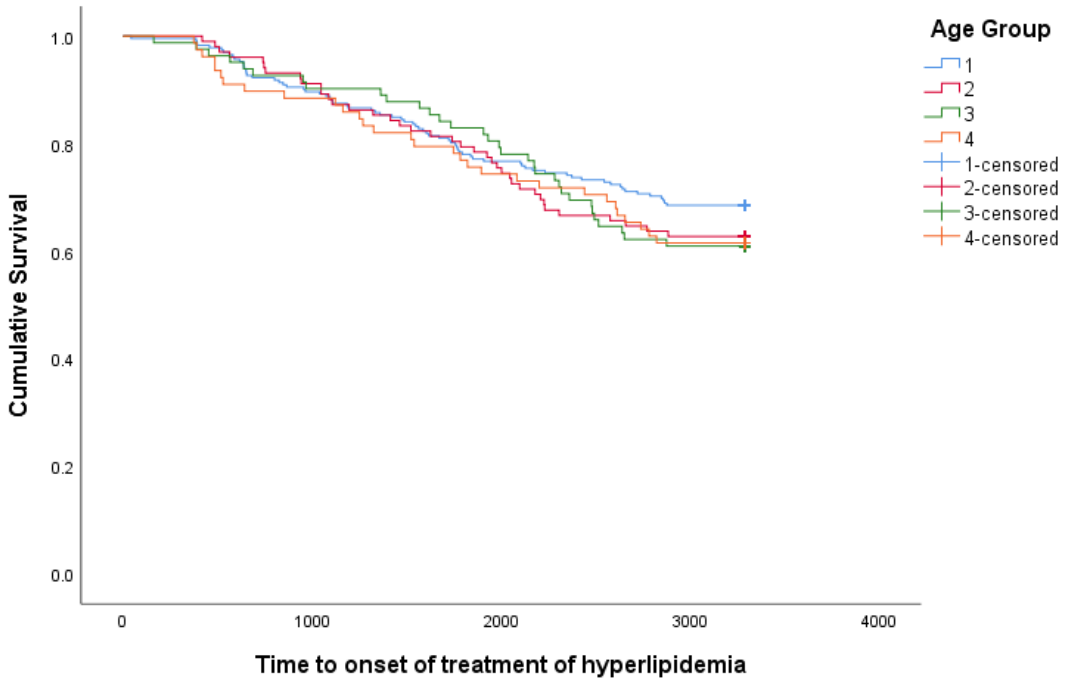



**Figure 1: Kaplan-Meier plot for time to onset of treatment of hyperlipidaemia among various age groups**[5]

[5] Age groups: 1: ≤ 51 years; 2: 51-59 years; 3: 60-66 years; 4: > 66 years

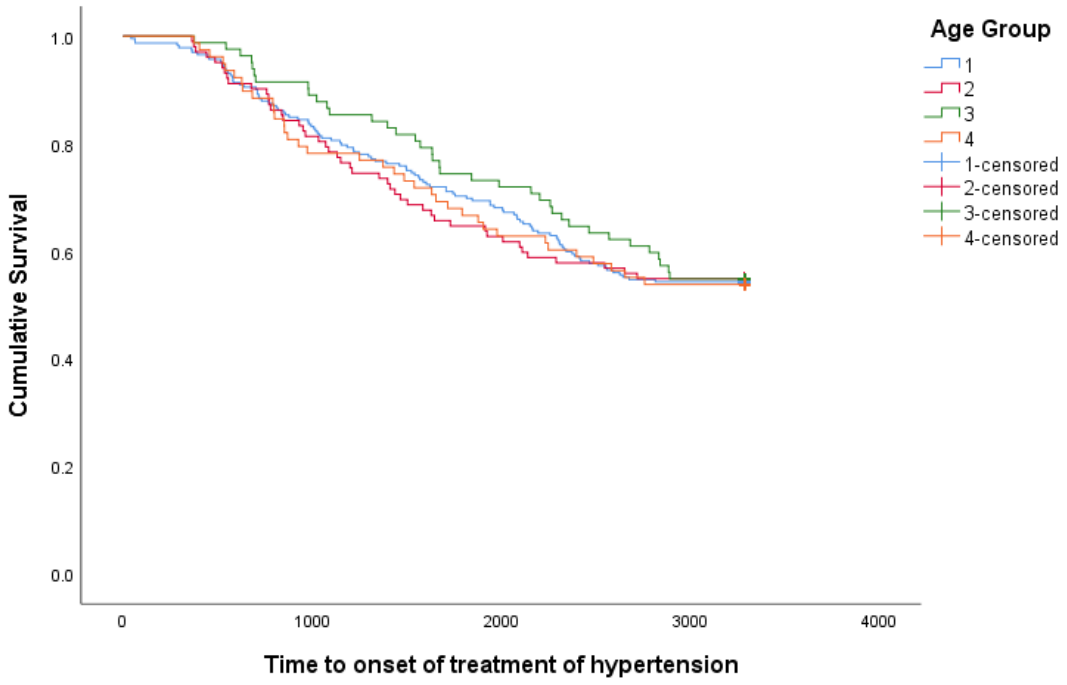



**Figure 2: Kaplan-Meier plot for time to onset of treatment of hypertension among various age groups**[6]

[6] Age groups: 1: ≤ 51 years; 2: 51-59 years; 3: 60-66 years; 4: > 66 years

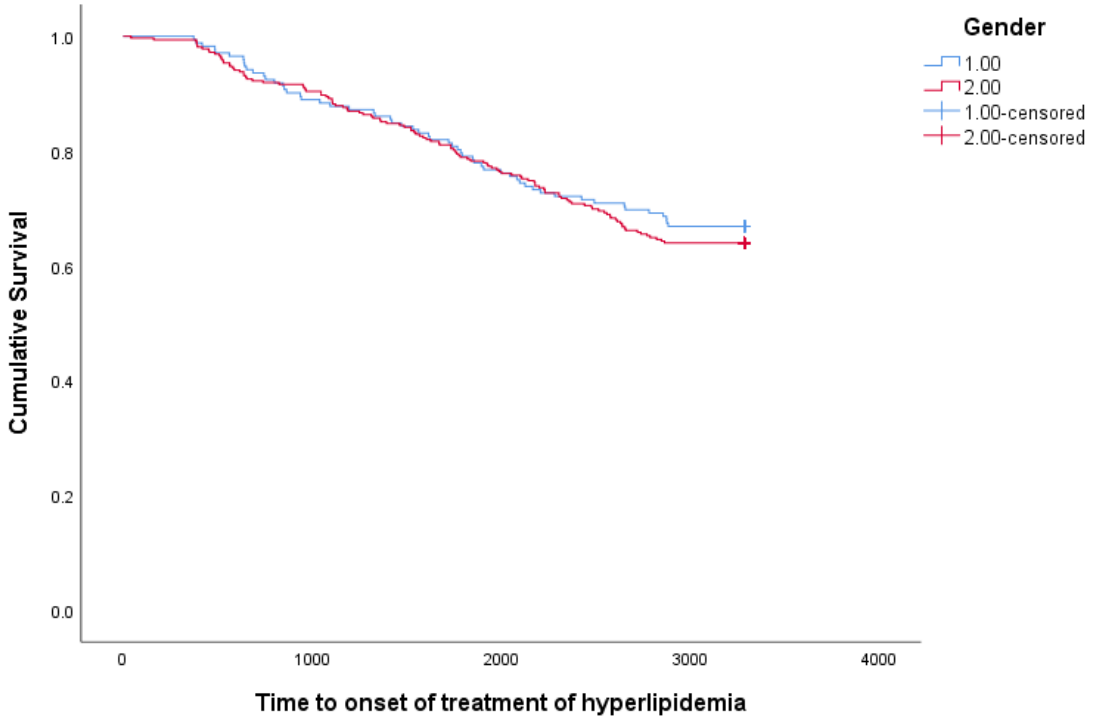



**Figure 3: Kaplan-Meier plot for time to onset of treatment of hyperlipidemia among male and female patients**

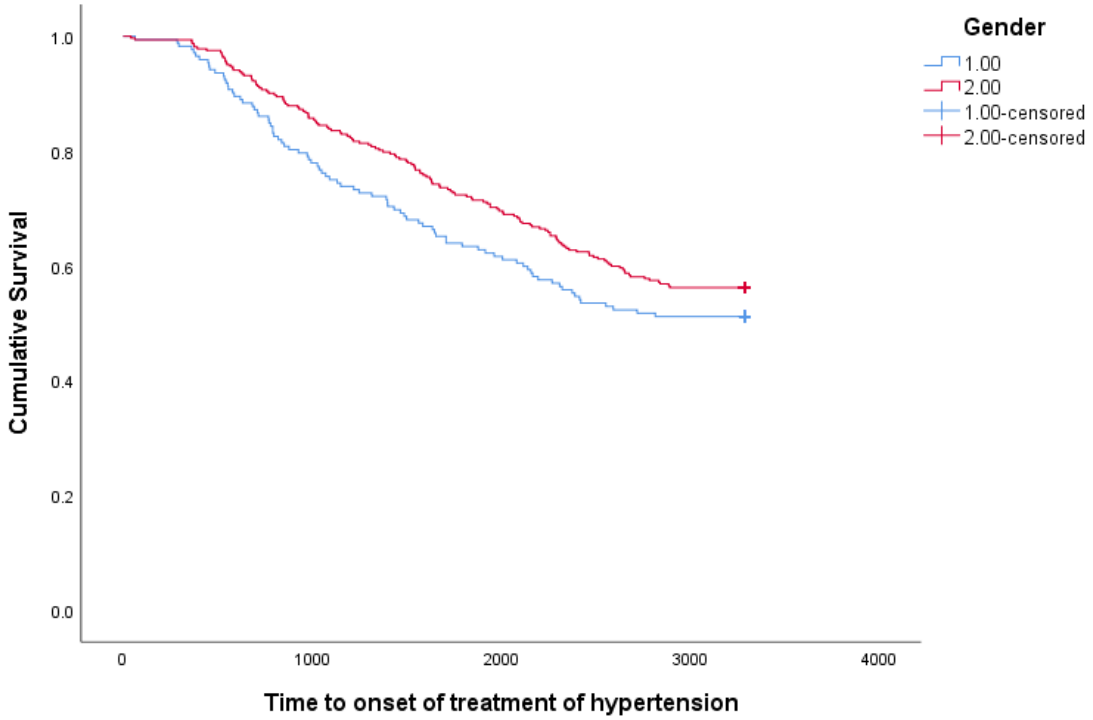



**Figure 4: Kaplan-Meier plots for time to onset of treatment of hypertension among male and female patients**

### 3.3 Manuscript 2: Comparison of different adherence measures

**Manuscript title: Comparison of adherence measures using medicine claims data among asthma patients receiving montelukast**

**Running head:**

Measuring adherence with medicines claims data

**Authors:**

Mavis Obeng-Kusi[1], Martie S Lubbe [1], Marike Cockeran[2], Johanita R Burger [1*]

[1] Medicine Usage in South Africa, Faculty of Health Sciences, North-West University, Potchefstroom Campus, Potchefstroom, Republic of South Africa.

[2] Statistics, School of Computer, Statistical and Mathematical Sciences, North-West University, Potchefstroom Campus, Potchefstroom, Republic of South Africa.

**Correspondence:**

J.R. Burger, North-West University, Potchefstroom Campus, Private Bag X6001, Potchefstroom, 2520.

Telephone: +27827883319, +27129400091

Fax: +27872314105

Email: Johanita.Burger@nwu.ac.za

**Key words:**

Adherence, Medicines claims data, South Africa, Cross-sectional analysis

**Take home points/ key points:**

- Compliance ratio (CR) and Continuous multiple interval measure of medication acquisition (CMOS) are adherence measures that can be considered equivalent to MPR because of the similarity in their mathematical formulae and the resultant adherence values.
- Capped proportion of days covered (PDC capped) produces lower adherence values compared to medicine possession ratio (MPR) because it considers the entire study period and does not allow for carry-over of excess medication.
- MPR is equivalent to capped PDC for adherence values less than 1, after which adherence values from MPR are consistently higher.
- Refill compliance rate (RCR) overestimates adherence because it evaluates the period between refills instead of the entire study period, leading to reduced denominators.
- Modified medicine possession ratio (MPRm) attempts to correct the shortfall of the RCR by adding the number of days of last refill, thus assuming patients to be 100% compliant after last refill. This leads to higher adherence estimates than the MPR.
- Adherence measures that present the entire study period as denominator produce uniform and consistent adherence values compared to those that only consider the time between dispensations.

**Prior postings and presentations, names of sponsors and grant numbers:**

The findings of this study was presented as a poster at the First Conference of Biomedical and Natural Sciences and Therapeutics (CoBNeST), held at the Spier Estate, Stellenbosch, South Africa from October 7, 2018 to October 10, 2018. The authors received financial support from the North-West University (Master's bursary 27959716) and the National Research Foundation (Grant number: 85315).

## ABSTRACT

### Purpose

This study aimed to compare different adherence measures, by determining adherence to montelukast among asthma patients, using data from a medicines claims database in South Africa.

### Methods

This retrospective, cross-sectional research employed data from 1st January 2006 to 31st December 2015 from a privately-owned Pharmaceutical Benefits Management (PBM) in South Africa. Claims for montelukast were identified and adherence determined using the continuous multiple interval measure of oversupply (CMOS), compliance ratio (CR), modified medication possession ratio (MPRm), refill compliance rate (RCR), continuous single interval measure of medication acquisition (CSA) and proportion of days covered (PDC) capped at 1. The measures were compared with the medication possession ratio (MPR) as the reference.

### Results

The MPR, CMOS and CR were equivalent, each yielding an adherence value of 86%. The MPRm, RCR and average CSA yielded higher adherence values of 96.9%, 117.2% and 129.0%, respectively whereas the PDC produced a lower adherence value of 76.0%. The measures that used the entire study period as the denominator produced consistent results compared to the measures that used the difference between claims dates as denominator.

### Conclusions

Although some degree of equivalence was observed among the adherence measures evaluated, the variables applied in their computation resulted in varied adherence values. The definitions of days' supply and the study period contributed significantly to the difference in adherence values obtained.

Selection of adherence measures should, therefore, be determined by the data available and the accurate definition of the variables to be used for computing the measures.

## INTRODUCTION

Adherence to medication is the degree to which the drug-taking behaviours of patients agree with recommendations by the prescriber.[1] The measurement of adherence is becoming increasingly important since it is critical to the success of pharmacotherapy.[2-5] There are several methods of computing adherence and these include, *inter alia*, the use of biological assays and markers, directly observed therapy, self-reports, pill counts, use of surveys, questionnaires and electronic medication packaging.[2,4,6-11] These methods, although beneficial, may be limited by cost,[6,10,11] ethical concerns[11] and self-reported bias.[6,7,10] Administrative claims data offer an inexpensive, efficient and non-invasive means by which adherence can be measured.[2,4,9]

In addition, administrative claims databases provide access to large populations for study under real clinical practice situations, and in a timely and effective manner.[12-15] Data obtained from such databases are less susceptible to recall and interviewer bias and can be linked to other databases,[14] such as medical records databases, to facilitate the determination of adherence. The use of administrative databases for measuring adherence, however, has some inherent disadvantages. Administrative datasets estimate adherence based on medication possession and not consumption[4] since they are unable to determine patients' intake of prescribed and dispensed medications.[2] Grégoire and Moisan[1] add that dispensing data depend on conditions of reimbursement and will, thus, not measure adherence to medications purchased over-the-counter and those not covered by a healthcare scheme. Adherence determined solely by claims data may mask periods of over- and under-utilisation of medications.[11] Successful use of administrative claims data for estimating adherence requires that all relevant information is recorded accurately, and patients are eligible for the medication of interest during the period of study to allow for valid conclusions to be made.[9]

Although several measures have been proposed to estimate adherence using medicines claims data and validated using other methods such as patient reports and pill counts, there are no specifications

for their mathematical calculation.[16] With the wide range of adherence measures available, researchers are often faced with the decision of choosing which is appropriate.[16,17] Some adherence measures have been identified to be mathematically equivalent, yielding similar adherence values.[16] Hess et al[4] suggests that it may not be necessary to have the variety of measures currently employed to assess adherence when administrative claims are used. An assessment of the various methods is, thus, important for better understanding and to facilitate future adherence studies using administrative data.

Previous studies have utilised secondary data to evaluate adherence to hypoglycemic medications,[18-20] lipid-lowering medications,[20,21] antihypertensives,[20] medications for managing multiple sclerosis,[22,23] and oral bone sparing medications,[24] among others. Several studies have also employed secondary data from South Africa for the assessment of adherence to anti-epileptics,[25] antidepressants,[26,27] and other medications. To the best of our knowledge, however, no study has been published on the use of secondary data for the comparison of different adherence measures in South Africa. This study aimed to compare different adherence measures, by determining adherence to montelukast among asthma patients, using data from a medicines claims database in South Africa.

## METHODS

### Study design

We performed a quantitative cross-sectional study analysing medicines claims data that is nationally representative for a ten-year period (1st January 2006 to 31st December 2015).

### Data source

We employed nationally-representative medicines claims data obtained from a privately-owned South African Pharmaceutical Benefit Management (PBM) company. This PBM is a large independent company that has been providing medicine claims processing services to about 1.6

million beneficiaries of about 42 medical schemes in South Africa for over 25 years. The data obtained represent about one-third of the South African patients registered with private medical aid schemes.[28]

Demographic information, such as sex and date of birth, together with encrypted membership numbers were used to the follow-up patients' prescriptions over a period of time. Information on prescribed medications dispensed to patients included drug trade names, quantities dispensed, days' supply, the prescription fill date, as well as International Classification of Diseases, Tenth Revision (ICD-10) codes for diagnoses.

## STUDY MEASUREMENTS

### Patient characteristics

All patients who had a diagnosis code (ICD-10 code J45) for asthma in conjunction with at least two consecutive claims for montelukast based on the National Pharmaceutical Product Index (NAPPI) code 10.4.2, provided by MIMS,[29] during the study period, were included in the research. Patients had to be enrolled continuously with the PBM throughout the study period.

### Measuring adherence

The medication possession ratio (MPR) is one of the most extensively-used measures of adherence based on claims data.[2,17,30,31] Karve et al[31] propose that researchers consider the MPR first for the calculation of adherence since several studies have discovered it to be valuable as an adherence measure.[2,17,32] Although MPR is easy to compute and interpret, it may mask periods of oversupply.[2,17] There is presently no perfect adherence measure using claims data[33] but the MPR serves as an acceptable standard against which other measures can be assessed. For this study, the MPR was used as a reference to which other measures of adherence were compared.

Six adherence measures[4] namely the proportion of days covered (PDC) capped at 1, refill compliance rate (RCR), compliance ratio (CR), modified medication possession ratio (MPRm), continuous multiple interval of oversupply (CMOS) and the continuous single interval measure of medication acquisition (CSA) averaged over the period of observation were determined, three of which (capped PDC, MPRm and average CSA) were compared to the MPR using Bland-Altman plots. In interpreting the results, smaller bias levels, together with narrow limits of agreements were considered representative of more equivalent measures than larger wide values with wider margins. The mathematical formulae for the determination of these adherence measures are described in Table 1.[4]

<<Place Table 1 here>>

**Statistical analysis**

The SAS® system version 9.4[34] was used to compute the means and standard deviations of the various measures. Bland-Altman plots, plotted using IBM® SPSS® version 25,[35] were used to compare selected adherence measures against the medication possession ratio as reference category.

**RESULTS**

The cohort consisted of 9 141 patients with a median age of 13.3 (5.2 - 49.2) years. In total, 52.8% (n = 4 825) were females. Females in the cohort were significantly older than their male counterparts at cohort entry ($p < 0.1$; Cohen's $d = 0.5$).

<<Place Table 2 here>>

Table 3 depicts the values for each adherence measure. The MPR and CR resulted in adherence of 86.4%. Adherence values for PDC, modified MPR, RCR and CSA, averaged were 74.0%, 96.6%, 117.2% and 129.1%, respectively. The value for CMOS, the gap measure evaluated, was 0.136.

<<Place Table 3 here>>

Results from Bland-Altman plots for MPR as reference category against MPRm, Average CSA and PDC are presented in Table 4. Varied degrees of equivalence were observed between MPR and each of the measures to which it was compared. Against PDC capped, there was a stronger agreement as evidenced by the small bias (0.02) value as well as narrow limits of agreement (-0.11-0.15). This was, however, not the case with the MPRm from which there was a wide limit of agreement (-42.97-2.06).

<<Place Table 4 here>>

<<Place Figures 1-3 here>>

**DISCUSSION**

This large cross-sectional study produced three key findings. Firstly, the study showed that two measures (CR and CMOS) produced equivalent mean adherence as the MPR. The observation, on one hand, confirms the results of the study by Hess et al[4] in which CMOS and MPR produced the same mean values but is contrary to the same study's finding of significant difference between CR and MPR. This variation in findings can be attributed to the different ways in which the variables where defined for computing the measures. The formulae by which the MPR, CMOS and the CR were calculated were almost the same, requiring the same variables and thus resulting in adherence values that were equivalent. The CMOS, being a gap measure, produced a value of 0.14 which is equivalent to adherence levels of 0.86 in non-gap methods. For gap measures, adherence values closer to zero represent better adherence levels.[31]

Secondly, it was also observed that besides the capped PDC which produced adherence values lower than the MPR, all the other measures (Table 3) resulted in adherence values greater than the MPR (Average CSA>>>RCR>>MPRm>MPR), a trend consistent with results of earlier studies.[4,29]

The PDC capped resulted in adherence values lower than those from MPR because this measure uses the total study period and does not consider excess medication on hand at the termination of the study. We observed a strong agreement between MPR and PDC capped for adherence values below 1 (Figure 1) as most of the data points that correspond with mean values below 1 showed zero difference between these measures. This is also evidenced by a small bias value of 0.02 and the narrow range between the limits of agreement (-0.11-0.15). Consequently, our results show that MPR and PDC capped are equivalent when adherence is below 1, where there is no oversupply or excess medication on hand. In the event of adherence values above 1 (oversupply) as shown in Fig. 2, however, there is increasing difference with increasing mean. Thus, MPR consistently produces values higher than PDC capped with increasing adherence. The presence of patients with some degree of medication oversupply in our investigation may, therefore, be the reason for the higher mean MPR value relative to the mean value of capped PDC.

The RCR evaluates the time period between dispensations instead of the entire study period, leading to a denominator which is smaller compared to that for measures that consider the entire period. This results in overestimation of adherence reflected by the high mean RCR value. To overcome the drawback of RCR in overestimating adherence, the MPRm, which also assesses the period between fills, includes number of days equal to the number of days' supply of medication acquired at the last dispensation. This addition of days assumes that the patient is completely adherent in the period to be covered by the last dispensation, accounting for its relatively high value compared to the MPR. There is some degree of agreement between MPR and MPRm as most data points fall within the limits of agreement (Figure 2). A bias of -20.46 together with the wide range of limits of agreement between -42.97 and 2.06, however, suggests that MPR and MPRm are not equivalent. The difference between the MPRm and MPR can be attributed to their difference in denominator and the consideration of the last dispensation in the computation of the MPRm.

In the calculation of the CSA, adherence for each dispensation period is calculated independently and averaged. The number of refills affects the weight of adherence in the cumulative analysis thus patients with single refills will not have the same weight as those with multiple refills. Patients' receipt of medications close to the end of the study period leads to bias as this is reflected as an oversupply and accounts for the measure's high adherence values. This measure does not allow a carry-over of excess medication from a refill interval to another; a very likely event in practical settings.[4,31] The MPR and average CSA also exhibit an agreement as most data points fall within the limits of agreement (Fig. 3). With a bias of -0.59 and range of limits of agreement between -3.06 and 1.88, the two measures can be said to be fairly equivalent. For mean adherence above 2, it can be noticed that there is decreasing difference with increasing mean; MPR consistently producing smaller values than the average CSA.

Finally, it was interesting to note that the definition of the denominator was critical to the results obtained from measure computations. Similar to what has been found by Hess et al[4] and Karve et al[29] adherence measures that presented the entire study period as denominator (MPR, capped PDC and CMOS) in our study ensured uniformity and consistency as opposed to measures that used the time between dispensations as the denominator. The measures that estimated the denominator as the difference between refills do not provide a uniform denominator. These measures may overestimate adherence since they may not be able to account for patients who discontinued medications early resulting in overestimation of adherence for such patients.[29]

A notable limitation of this study was that the data used were intended for reimbursement purposes and not for research purposes and may, therefore, lack certain data fields that would be relevant for the calculation of adherence methods such as the exact number of days for which drugs were supplied. Again, there was no definitive means by which conclusions could be drawn as to the best adherence measures to be used when administrative claims data are employed.

In conclusion, determination of the variables for computation of adherence is critical to the accuracy of results in assessing adherence using these measures.[3,4,23] Days' supply must be cautiously estimated. Depending on the available data, this can be determined as the quotient of the prescribed dose and the number of units of the drug dispensed.[3,4] The number of evaluation days is also essential in the calculation of adherence measures. It is therefore important to determine *a priori* the time frame for the evaluation of adherence to obtain consistent values. Understanding the data available for determining adherence is important for a decision to be made on the appropriate measures.[8] The variables required as well as the complexity of calculations are features that need to be considered.

## ETHICS STATEMENT

We sought permission from the board of directors of the PBM as well as the Health Research Ethics Committee (HREC) of the North-West University (NWU-00179-14-A1-06) to conduct the study.


## ACKNOWLEDGEMENT

The authors thank the PBM Company for permitting the use of the database in this study. We are also grateful to Ms Anne-Marie Bekker for her administrative support with regard to the database and Ms. Helena Hoffman for her help in proof-reading and editing this article. We acknowledge the North-West University and the National Research Foundation (Grant number 85315) for financial support.


## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## DISCLAIMER

The views expressed in this article are the personal views of the authors. Our funding agencies played no part in the study's design; data collection, analysis and interpretation; write-up of findings nor in the decision to publish this manuscript.

**REFERENCES**


1. Grégoire JP, Moisan J. Assessment of adherence to drug treatment in database research. In: Elseviers M, Wettermark B, Almarsdóttir AB, Morten A, Benko R, Bennie M, et al, eds. Drug utilization research. Hoboken, NJ: Wiley; 2016:369-380.
2. Andrade SE, Kahler KH, Frech F, Chan KA. Methods for evaluation of medication adherence and persistence using automated databases. *Pharmacoepidemiol Drug Saf.* 2006*;*15(8):565-574. https://doi.org/10.1002/pds.1230. Accessed November 10, 2018.
3. Cramer JA, Roy A, Burrell A, et al. Medication compliance and persistence: terminology and definitions. *Value Health.* 2008;11(1):44-47. https:// doi.org/10.1111/j.1524-4733.2007.00213.x. Accessed November 10, 2018.
4. Hess LM, Raebel MA, Conner DA, Malone DC. Measurement of adherence in pharmacy administrative databases: a proposal for standard definitions and preferred measures. *Ann Pharmacother*. 2006;40(7-8):1280-1288. https://doi.org/10.1345%2Faph.1H018. Accessed November 10, 2018.
5. Vrijens B, Geest SD, Hughes DA, Przemyslaw K, Demonceau J, Ruppar T, et al. A new taxonomy for describing and defining adherence to medications. *Br J Clin Pharmacol*. 2012;73(5):691-705. https://doi.org/10.1111/j.1365-2125.2012.04167.x. Accessed November 10, 2018.
6. Jimmy B, Jose J. Patient medication adherence: measures in daily practice. *Oman Med J*. 2011;26(3):155-159. https://doi.org/10.5001/omj.2011.38. Accessed November 10, 2018.
7. Lam WY, Fresco P. Medication adherence measures: an overview. *Biomed Res Int*. 2015; 2015:217047. https://doi.org/10.1155/2015/217047. Accessed November 10, 2018.

8. Lima-Dellamora ED, Osorio-de-Castro CG, Madruga LG, Azeredo, TB. Use of pharmacy records to measure treatment adherence: a critical review of the literature. *Cad Saude Publica*. 2017;33(3):e00136216. https://doi.org/10.1590/0102-311X00136216. Accessed November 10, 2018.
9. McCaffrey DJ. Medication utilization patterns. In: Yang Y, West-Strum D, eds. Understanding pharmacoepidemiology. Manhattan, NY: McGraw-Hill; 2011:369-380.
10. Osterberg L, Blaschke T. Adherence to medication. *N Engl J Med*. 2005;353(5):487-497. https://doi.org/10.1056/NEJMra050100. Accessed November 10, 2018.
11. Vik SA, Maxwell CJ, Hogan DB. Measurement, correlates, and health outcomes of medication adherence among seniors. *Ann Pharmacother*. 2004;38(2):303-312. https://doi.org/10.1345%2Faph.1D252. Accessed November 10, 2018.
12. Park BJ, Stergachis A. Automated databases in pharmacoepidemiologic studies. In: Hartzema AG, Tilson HH, Chan KA, eds. Pharmacoepidemiology and therapeutic risk management. Cincinnati, OH: Harvey Whitney; 2008:519-544.
13. Schneeweiss S, Avorn J. A review of uses of health care utilisation databases for epidemiologic research on therapeutics. *J Clin Epidemiol*. 2005;58(4):323-337. https://doi.org/10.1016/j.jclinepi.2004.10.012. Accessed November 10, 2018.
14. Strom BL. Overview of automated databases in pharmacoepidemiology. In: Strom BL, Kimmel SE, Hennessy S, eds. Pharmacoepidemiology. 5th ed. Chichester: Wiley; 2012:158-162.
15. Torre C, Martins AP. Overview of pharmacoepidemiological databases in the assessment of medicines under real-life conditions. In: Lunet N, ed. Epidemiology – current perspectives on research and practice. Rijeka: InTech; 2012:130-154.
16. Karve S, Cleves MA, Helm M, Hudson TJ, West DS, Martin BC. Prospective validation of eight different adherence measures for use with administrative claims data among patients

with schizophrenia. *Value Health*. 2009;12(6):989-959. https://doi.org/10.1111/j.1524-4733.2009.00543.x. Accessed November 10, 2018.

17. Martin BC, Wiley-Exley EK, Richards S, Domino ME, Carey TS, Sleath BL. Contrasting measures of adherence with simple drug use, medication switching, and therapeutic duplication. *Ann Pharmacother*. 2009;43(1):36-44. https://doi.org/ 10.1345/aph.1K671. Accessed November 10, 2018.
18. Donnan PT, MacDonald TM, Morris AD. Adherence to prescribed oral hypoglycaemic medication in a population of patients with Type 2 diabetes: a retrospective cohort study. *Diabet Med*. 2002;19(4):279-284. https://doi.org/10.1046/j.1464-5491.2002.00689.x. Accessed November 10, 2018.
19. Hertz RP, Unger AN, Lustik MB. Adherence with pharmacotherapy for type 2 diabetes: a retrospective cohort study of adults with employer-sponsored health insurance. *Clin Ther*. 2005; 27(7):1064-1073. https://doi.org/10.1016/j.clinthera.2005.07.009. Accessed November 10, 2018.
20. Vink NM, Klungel OH, Stolk RP, Denig P. Comparison of various measures for assessing medication refill adherence using prescription data. *Pharmacoepidemiol Drug Saf*. 2009;18(2):159-165. https://doi.org/10.1002/pds.1698. Accessed November 10, 2018.
21. Jackevicius CA, Mamdani M, Tu JV. Adherence with statin therapy in elderly patients with and without acute coronary syndromes. *JAMA*. 2002;288(4):462-467. https://doi.org/ 10.1001/jama.288.4.462. Accessed November 10, 2018.
22. Hansen K, Schüssel K, Kieble M et al. Adherence to Disease Modifying Drugs among Patients with Multiple Sclerosis in Germany: A Retrospective Cohort Study. *PLoS One*. 2015;10(7):e0133279. https://doi.org/10.1371/journal.pone.0133279. Accessed November 10, 2018.
23. Kozma CM, Dickson M, Phillips AL, Meletiche DM. Medication possession ratio: implications of using fixed and variable observation periods in assessing adherence with disease-

modifying drugs in patients with multiple sclerosis. *Patient Prefer Adherence*. 2013;7:509-516. https://doi.org/10.2147/PPA.S40736. Accessed November 10, 2018.

24. Van der Zwaard BC, Van Hout W, Hugtenburg JG, Van der Horst HE, Elders PJM. Adherence and persistence of patients using oral bone sparing drugs in primary care. *Fam Pract*. 2017;34(5):525-531. https://doi.org/10.1093/fampra/cmw120. Accessed November 10, 2018.
25. Jacobs K, Julyan M, Lubbe MS, Burger JR, Cockeran M. Medicine possession ratio as proxy for adherence to antiepileptic drugs: prevalence, associations, and cost implications. *Patient Prefer Adherence*. 2016;10:539-547. https://doi.org/10.2147/PPA.S98940. Accessed November 10, 2018.
26. Slabbert FN, Harvey BH, Brink CB, Lubbe MS. The impact of HIV/AIDS on compliance with antidepressant treatment in major depressive disorder: A prospective study in a South African private healthcare cohort. *AIDS Res Ther.* 2015;12:9. https://doi.org/10.1186/s12981-015-0050-2. Accessed November 10, 2018.
27. Slabbert FN, Harvey BH, Brink CB, Lubbe MS. Prospective analysis of the medicine possession ratio of antidepressants in the private health sector of South Africa, 2006 - 2011. *S Afr Med J*. 2015;105(2):139-144. https://doi.org/10.7196/samj.8394. Accessed November 10, 2018.
28. Council for Medical Schemes (CMS). About us. 2017. https://www.medicalschemes.com/Content.aspx?1. Accessed: June 25. 2018.
29. Ho PM, Bryson CL, Rumsfeld JS. Medication adherence: its importance in cardiovascular outcomes. *Circulation.* 2009;119(23):3028-3035. https://doi.org/ 10.1161/CIRCULATIONAHA.108.768986. Accessed November 10, 2018.
30. Snyman JR. Monthly Index of Medical Specialities. Cape Town, CTP Printers; 2017.
31. Karve S, Cleves MA, Helm M, Hudson TJ, West DS, Martin BC. An empirical basis for standardizing adherence measures derived from administrative claims data among diabetic

patients. *Med Care*. 2008;46(11):1125-1133. https://doi.org/10.1097/MLR.0b013e31817924d2. Accessed November 10, 2018.

32. Arnet I, Kooij MJ, Messerli M, Hersberger KE, Heerdink ER, Bouvy M. Proposal of standardisation to assess adherence with medication records: methodology matters. *Ann Pharmacother*. 2016;50(5):360-368. https://doi.org/10.1177/1060028016634106. Accessed November 10, 2018.
33. Sattler EL, Lee JS, Perri M, 3rd. Medication (re)fill adherence measures derived from pharmacy claims data in older Americans: a review of the literature. *Drugs Aging*. 2013;30(6):383-399. https://doi.org/10.1007/s40266-013-0074-z. Accessed November 10, 2018.
34. SAS® [computer program]. Version 9.4. Cary, NC: SAS Institute Inc.; 2002-2012.
35. SPSS® [computer program]. Version 25. Armonk, NY: IBM Corp.; 2017.

**TABLES**

**Table 1. Mathematical formulas for adherence measures**

| Adherence measure | Formula |
|---|---|
| **Medication possession ratio (MPR)** | $\frac{Number\ of\ days'\ supply\ in\ index\ period}{number\ of\ days\ in\ the\ study\ period}$ |
| **Proportion of days covered capped (PDC capped)** | $\frac{Number\ of\ days'\ supply\ in\ index\ period}{number\ of\ days\ in\ study\ period}$ |
| **Refill compliance rate (RCR)** | $\frac{Number\ of\ days'\ supply}{last\ claim\ date - index\ date} x\ 100$ |
| **Compliance ratio (CR)** | $\frac{Number\ of\ days'\ supply\ in\ index\ period - last\ days\ supply}{last\ claim\ date - index\ date}$ |
| **Medication possession ratio, modified (MPRm)** | $\frac{Number\ of\ days'\ supply}{last\ claim\ date - index\ date + last\ days'\ supply}\ x\ 100$ |
| **Continuous multiple interval measure of oversupply (CMOS)** | $\frac{Total\ days\ of\ treatment\ gaps\ (+) or\ surplus\ (-)}{total\ days\ to\ next\ fill\ or\ end\ of\ observation\ period}$ |
| **Continuous single interval measure of medication acquisition, average (Average CSA)** | $\frac{Days'\ supply\ obtained\ at\ the\ beginning\ of\ the\ interval}{days\ in\ the\ interval}$ |

**Table 2. Baseline characteristics of the cohort (N = 9 141)**

| Characteristic | n (%) |
|---|---|
| **Total number of patients, N** | 9141 |
| **Age (years)** | |
| **Median (IQR)[7]** | 13.3 (5.2-49.2) |
| **Sex** | |
| **Male** | 4316 (47.22) |
| **Female** | 4825 (52.78) |

[7] IQR: Interquartile range

**Table 3. Adherence values for the cohort (N = 9 141)**

| Adherence measure | Mean | Standard deviation |
|---|---|---|
| **Medication possession ratio (MPR)** | 0.86 | 1.44 |
| **Proportion of days covered capped (PDC capped)** | 0.76 | 0.25 |
| **Refill compliance rate (RCR), %** | 117.15 | 281.08 |
| **Compliance ratio (CR)** | 0.86 | 1.44 |
| **Medication possession ratio, modified (MPRm), %** | 96.60 | 35.91 |
| **Continuous multiple interval measure of oversupply (CMOS)** | 0.14 | 1.44 |
| **Continuous single interval measure of medication acquisition, average (Average CSA)** | 1.29 | 2.11 |

**Table 4. Comparison of adherence measures**

| | Bias | Standard deviation | Lower limit of agreement | Upper limit of agreement |
|---|---|---|---|---|
| **MPRm** | -20.46 | 11.49 | -42.97 | 2.06 |
| **PDC capped** | 0.02 | 0.07 | -0.11 | 0.15 |
| **Average CSA** | -0.59 | 1.26 | -3.06 | 1.88 |

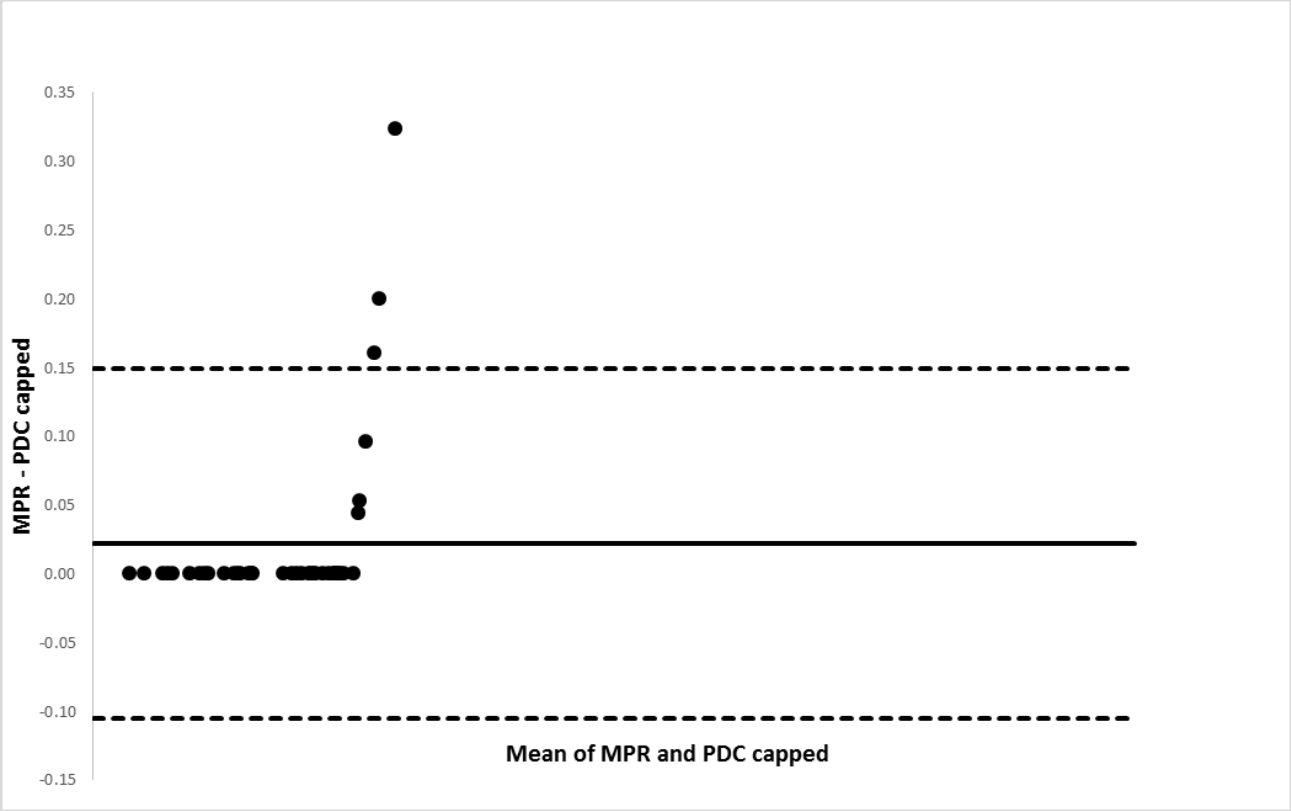



**Figure 1: Bland-Altman plot of medication possession ratio (MPR) vs proportion of days covered capped at 1 (PDC capped)**

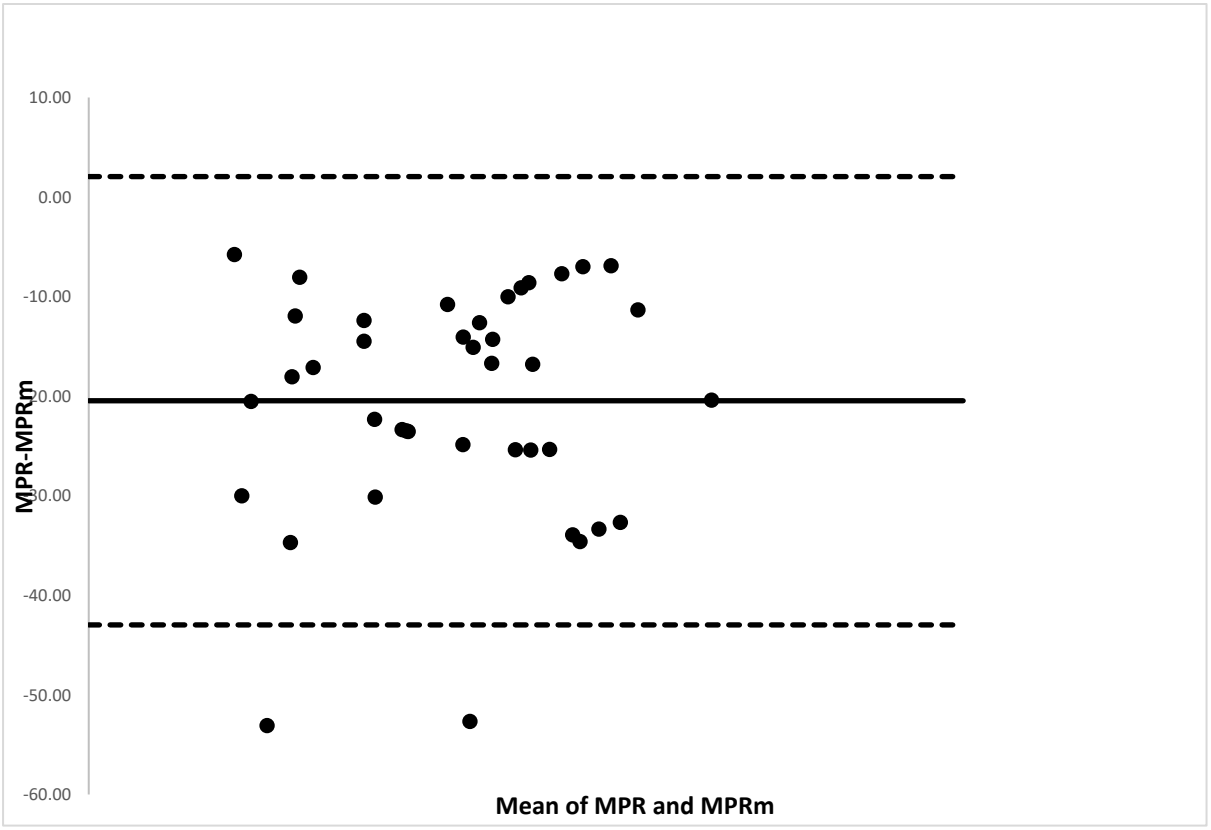



**Figure 2. Bland-Altman plot of medication possession ratio (MPR) vs modified medication possession ratio (MPRm)**

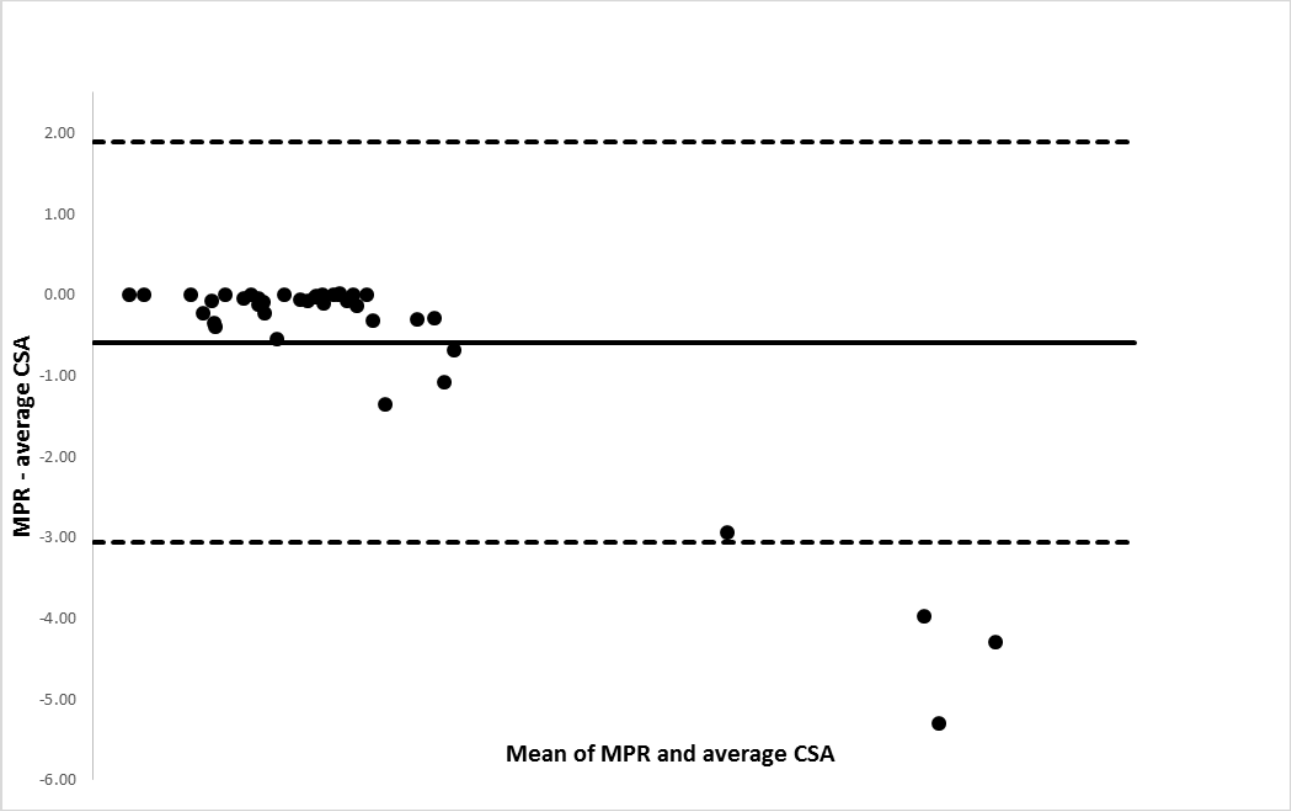



**Figure 3. Bland-Altman plot of medication possession ratio (MPR) vs average continuous single interval measure of medication acquisition (average CSA)**

**Figure Legends**

Figure 1: Bland-Altman plot showing good agreement in the adherence determined using the medication possession ratio (MPR) and the proportion of days covered capped at 1 (PDC capped). The plot depicts data from 40 patients randomly selected from the study population. The mean of the adherence values determined using both MPR and PDC capped for each patient (x-axis) was plotted against the difference between these same adherence values (y-axis). Solid lines represent bias while the dashed lines represent the lower and upper limits of agreement.

Figure 2: Bland-Altman plot showing poor agreement in adherence values determined using the medication possession ratio (MPR) and the modified medication possession ratio (MPRm). The plot depicts data from 40 patients randomly selected from the study population. The mean of the adherence values determined using both MPR and MPRm for each patient (x-axis) was plotted against the difference between these same adherence values (y-axis). Solid lines represent bias while the dashed lines represent the lower and upper limits of agreement.

Figure 3: Bland-Altman plot showing agreement in adherence computed using the medication possession ratio (MPR) and the average continuous single interval of medication acquisition (Average CSA). Data from 40 patients randomly selected from the study population was employed in this plot. The mean of the adherence values determined using both MPR and average CSA for each patient (x-axis) was plotted against the difference between these same adherence values (y-axis). Solid lines represent bias while the dashed lines represent the lower and upper limits of agreement.

## 3.4 Chapter summary

In this chapter, two manuscripts that address the objectives of the empirical investigation were presented. The next chapter, which concludes the study, discusses the study's conclusions, strengths and limitations, and proposes recommendations for future studies.

# CHAPTER 4: CONCLUSIONS, STRENGTHS, LIMITATIONS AND RECOMMENDATIONS

## 4.1 Introduction

The focus of this final chapter is to draw conclusions from the study with respect to the objectives specifically outlined in Chapter 1. The chapter highlights the strengths and limitations of the study and proposes recommendations for future studies.

## 4.2 Conclusions from the study

The study aimed to assess the appropriate use of different research methods in pharmacoepidemiological studies using a South African medicines claims database. Two main approaches, a literature review and an empirical study, were employed towards attaining this goal. Conclusions drawn after meeting the aims of the study follow in the subsequent paragraphs.

### 4.2.1 Conclusions from literature study

The specific objectives of the literature review were achieved in Chapter 2 of this dissertation and the following paragraphs discuss the conclusions drawn from the findings.

#### 4.2.1.1 Conceptualisation of pharmacoepidemiology; its uses and relevance

Pharmacoepidemiology was defined as the study of the distribution and determinants of drug-related outcomes in large populations and how this is applied to efficacious treatments (Porta, 2014:184). It was also established that the discipline applies epidemiologic methods to the field of pharmacology to study the use and effects of medications among groups of people (Evans, 2012:973; Kongpatanakul & Strom, 2001:27; Spitzer, 1999:352; Varanasi, 2012:11; Wettermark, 2013:43; Wise, 2011:95). The history and growth of pharmacoepidemiology were summarised in section 2.2.1.

The literature study revealed that pharmacoepidemiology can be applied in all phases of drug development, making and use and at all levels of healthcare (Thaker *et al*. 2015:53) to obtain useful insight into the use of drugs with resultant outcomes and contribute to decision-making based on scientific evidence (Rodriguez & Gutthann, 1998:421). Pharmacoepidemiology was found to be useful for understanding target populations, indications and relationships that exist among these and the drugs under investigation, increasing product name recognition and market penetration during drug marketing, serving as 'legal prophylaxis' in anticipation of future medication ligations, supporting rational and cost-effective medication use, monitoring

medication-taking behaviours, among others (refer to sections 2.2.2.1, 2.2.2.2, 2.2.2.3, 2.2.2.4 and 2.2.2.5).

From the literature, it was determined that pharmacoepidemiology addresses questions on incidence and causality of the outcomes of medications, medication effectiveness, economic impact of medications as well as prescription and utilisation patterns (Kongpatanakul & Strom, 2001:27). To do this, specific measures of pharmacoepidemiology are evaluated in this discipline. These include, *inter alia*, adherence to medicines (McCaffrey, 2011:136), occurrence, frequency and magnitude of adverse drug events and reactions (Banahan, 2011:156), drug utilisation patterns and economics of drug use (Wettermark *et al*., 2008:161). The measures of pharmacoepidemiology were briefly discussed in section 2.3.1.

Bias, inadequate data quality and ethical concerns were among the challenges of pharmacoepidemiology identified in this literature review (refer to section 2.2.3). The anticipated future of pharmacoepidemiology was briefly outlined in section 2.2.3.

From this summary, it can be concluded that the emergence of new diseases, the need for new pharmaceuticals for the management of existing conditions and the increased utilisation of medications among populations have led to a steady rise in the activities of pharmaceutical manufacturers. This activity surge consequently increases the exposure of communities to medications with their associated outcomes, necessitating their study to enhance their use and minimise potential untoward effects. Pharmacoepidemiology is thus an important field of study which is rapidly gaining significance in science in general and specifically in medical practice.

#### 4.2.1.2 Conceptualisation of various study designs; their advantages and disadvantages as well as the statistical analyses applicable to them

Study designs used for the conduct of pharmacoepidemiological studies were extensively discussed to achieve this objective. A study design was defined as a well-designed procedure followed to answer a research question (Morroni & Myer, 2014:78). Diverse study designs exist for the conduct of pharmacoepidemiological studies (Refer to Figure 2-2), differing with respect to the number and units of observations, data collection methods, time between measurement of exposure and outcomes, as well as the direction in which these are measured relative to one another (Friis & Sellers, 2009:242). Their marked effect on results and interpretations requires researchers to critically consider the choice of study designs for pharmacoepidemiological studies.

An experimental study design, requiring the investigator to interact with the exposure of interest, or observational study design in which the investigator plays no interactive role, may be used in

pharmacoepidemiological studies (Friis & Sellers, 2009:243; Morroni & Myer, 2014:78; Strom, 2013d:27; Waning & Montagne, 2001:45). It was also deduced from literature that pharmacoepidemiological studies may be based on designs that are prospective (where the study population is observed forward in time) or retrospective (in which the research is based on previously recorded data) (Morroni & Myer, 2014:88; Rothman *et al*., 2008:95; Waning & Montagne, 2001:46).

The most commonly-used study designs were found to be the randomized controlled trials, quasi-experimental studies, cohorts, case-controls, cross-sectional studies, case-series and case-reports (Friis & Sellers, 2009:243; Harpe, 2011:41; Morroni & Myer, 2014:78; Petitti, 2000:2; Rothman *et al*., 2008:95; Strom, 2013d:22 and Waning & Montagne, 2001:46). The nested case-control, case-cohort and multi-time case-control designs were identified as novel study designs which allow relative risks to be measured accurately (Rothman *et al*., 2008:122; Schneeweiss & Suissa, 2013:326) while meta- and decision-analyses were determined to be designs used to synthesize knowledge using medical literature and statistical analyses (Petitti, 2000:2). The characteristics of the various studies designs together with their strengths and limitations were summarised (Refer to Table 2-1).

It can be concluded thus that all study designs employed in pharmacoepidemiology, like in most scientific research, have both advantages and disadvantages and have specific conditions under which they can be effectively used. The choice of study design for any pharmacoepidemiological study will be dependent on the objectives of the study and availability of resources in terms of finance, human resources, data and time.

#### 4.2.1.3 Conceptualisation of data sources and databases used in pharmacoepidemiological studies

The literature showed that pharmacoepidemiological studies can utilise primary data collected *de novo* for a specific study's purposes or secondary data previously collated for other purposes ((Harpe, 2011:56; Kaufman, 2013:178). It was found that although primary data ensure high quality data with insights that may not be available when secondary data are used, they are more expensive, time-consuming and susceptible to biases that are often associated with the processes of data collection (Harpe, 2011:56; Richter *et al*., 2016:84). Secondary data, on the other hand, were found to be more economical in terms of time and resources but potentially lacking information on important confounders and difficult to validate since they are generally collated for purposes other than research (Harpe, 2011:56; Torre & Martins, 2012:140).

*Ad hoc* or field studies were identified as studies based on primary data and are particularly relevant when detailed information that may not be available from secondary data are required (Kaufman, 2013:178). The following are categorised as *ad -hoc* data sources (Strom, 2013e:191):

- Case-control surveillance
- Prescription event monitoring
- Registries

Post-marketing spontaneous pharmacovigilance reporting systems and automated databases were the secondary data sources identified from the literature study. Spontaneous reporting systems were described as formal systems that record and analyse the incidence of adverse drug reactions (Waning & Montagne, 2001:133). The resultant data are entered into adverse event databases and become a source of data for further studies or regulatory action (Pan *et al*., 2013:141). Pharmacovigilance reporting offers an inexpensive approach to studying drug safety by promoting the detection of safety signals that subsequently lead to further studies and action by regulators where necessary (Pan *et al*., 2013:141). Spontaneous reporting is, however, voluntary, leading to underreporting and is proportional to the time a drug is available on the market and these limit its usefulness in pharmacoepidemiology (Pan *et al*., 2013:150; Waning & Montagne, 2001:134).

Automated databases, defined as information from utilisation of healthcare that have been stored electronically and collated on regular basis in daily clinical practice (Takahashi *et al*., 2012:124), were found to be efficient and cost-effective for studying issues related to post-marketing drug use and effects and evaluating programmes targeted at enhancing medication use (Hennessy, 2006:31; Torre & Martins, 2012:132). Automated databases for pharmacoepidemiological studies may be electronic health medical records; kept for patients' clinical management or administrative claims databases, which are intended for reimbursement purposes (Hennessy, 2006:311; Strom, 2013c:119). Databases were found to provide large sample sizes for study increasing the statistical power at relatively less expensive cost, with increased representativeness while allowing for the study of rare medication outcomes as well as utilisation and safety profiling (Motheral & Fairman, 1997:349; Park & Stergachis, 2008:533; Schneeweiss & Avorn, 2005:324; Strom, 2013c:121; Torre & Martins, 2012:133). It was, however, learnt that databases are limited in the area of data quality due to lack of important confounder information (Park & Stergachis, 2008:534; Strom, 2013c:121; Torre & Martins, 2012:140). The use of diagnostic coding systems in databases may also lead to poor outcomes definition that hinder the usefulness of databases in pharmacoepidemiological studies (Harpe, 2011:58; Park & Stergachis, 2008:534).

Furthermore, information on conditions not considered severe enough for medical attention, medications not insured by particular policies and those purchased over-the-counter will not be captured by databases and can therefore not be studied by the use of these sources of data (Strom, 2013c:120).

The various data sources were expounded on extensively in Chapter 2 with a summary of their characteristics, advantages and disadvantages outline in Table 2-2. It was deduced that various data sources have their respective roles in pharmacoepidemiological investigations. Randomized controlled trials, although still deemed the 'gold standard', may not always be the most appropriate choice especially in cases where the study is hindered by resource and ethical constraints. The use of automated databases for pharmacoepidemiological research is, however, gaining prominence and this may be attributed to their ability to provide readily available data in large volumes at relatively less expensive rates.

#### 4.2.1.4 Determination, from literature, of the application of various research methods in pharmacoepidemiological studies

Several methods have been employed in pharmacoepidemiology to study drug utilisation (Lee *et al*., 2013:339; West-Strum, 2011:9; Thaker *et al*., 2015:54); drug effects and adverse drug events, medication errors (Lee *et al*., 2013:339); and compliance (Acri & Gross, 2013:314) and to evaluate interventions intended to improve on medication use quality and to contain costs (Wagner *et al*., 2002:299). In this study, interrupted time series analysis, methods for determining adherence and survival analysis and were discussed in sections 2.6.1, 2.6.2 and 2.6.3.

Interrupted time series assess the effectiveness of interventions or policy changes conducted to improve medication use quality at a population level (Jandoc *et al*., 2015:950; Lagarde, 2012:76; Penfold & Zhang, 2013:38; Wagner *et al*., 2002:299). The method divides time-ordered observations into segments representing rates of occurrence of an event before and after an intervention (Wagner *et al*., 2002:299; Penfold & Zhang, 2013:39). The changes in the level, value of a series and the rate at which a measure changes represent an effect from the intervention or policy change (Wagner *et al*., 2002:300). The ability to provide easy-interpretable results coupled with a graphical presentation makes time series analysis a preferred method by people with little or no statistical and epidemiological knowledge (Bernal *et al*., 2017:354; Penfold & Zhang, 2013:39; Wagner *et al*., 2002:301). It was also deduced from the literature that the method thoroughly evaluates the continuous effects of interventions under real-life conditions and analyses the unintended consequences of these interventions, allowing for stratified assessment of differences in effect among varying population groups. Interrupted time series cannot be used to make definitive conclusions at individual level and require a large sample size before and after

a policy with ample time periods between interventions; limitations to their use for pharmacoepidemiological studies (Penfold & Zhang, 2013:39).

It can be concluded from this summary that interrupted time series can be applied in policy formulation and assessment in the health sciences to facilitate recommendations to adopt new policies or to modify existing policies to improve outcomes.

From literature, adherence can be determined directly, using biological assays and markers, as well as directly observed therapy that allow for confirmation of medication intake, or indirectly, using surrogate measures of medication utilisation (Andrade *et al*., 2006:565; Fairman & Motheral, 2000:500; Hess *et al*., 2006:1280; Jimmy & Jose, 2011:157; Lam & Fresco, 2015:2; Lima-Dellamora *et al*., 2017:2; McCaffrey, 2011:139; Osterberg & Blaschke, 2005:488; Vik *et al*., 2004:304). Direct measures were found to be the most reliable and accurate measures because they offer proof of medication consumption but are expensive, could lead to anxiety in patients (Fairman & Motheral, 2000:500; Jimmy & Jose, 2011:157; Osterberg & Blaschke, 2005:488; Vik *et al*., 2004:304) without allowing also for the study of trends in adherence as they assess adherence at a point in time (Fairman & Motheral, 2000:500; Lam & Fresco, 2015:2; Vik *et al*., 2004:304). Indirect methods of measuring adherence include patient interviews, self-reporting of adherence, pill counts, the use of electronic medication packages, and the use of administrative databases. It was established that the administrative claims data are increasingly being used for determining adherence (Andrade *et al*., 2006:566; Farmer, 1999:1082; Grégoire & Moisan, 2016:369; Halpern *et al*., 2006:1040; Hess *et al*., 2006:1280; Karve *et al*., 2009:989; Lam & Fresco, 2015:3; Lima-Dellamora *et al*., 2017:2; McCaffrey, 2011:141; Svarstad *et al*., 2001:806; Vik *et al*., 2004:305). This approach is non-invasive, convenient and inexpensive (Hess *et al*., 2006:1281; McCaffrey, 2011:141) and offers the researcher the opportunity to identity events of early discontinuation and use of medications by patients in a manner that is different from what was prescribed (Farmer, 1999:1082). The adherence values obtained using administrative claims data are, however, reliant on the accuracy and completeness of the datasets used. The use of administrative claims databases for determining adherence is also based on medication possession and not consumption and may mask periods of under- and over-supply of medication (Hess *et al*., 2006:1280; Vik *et al*., 2004:305). Several measures of adherence can be computed from administrative claims data (Refer to Table 2-2) and a choice of the appropriate measure should be determined by the data available and the objective of the study.

From the literature investigation, it was gathered that the methods used for determining adherence to medications vary depending on the source of data employed for the determination. Although the direct methods may be more precise in this measurement, their usefulness is hampered by cost and the potential for patients to alter their medication-taking before to produce

desirable adherence findings. These challenges are addressed by the indirect methods which are themselves not devoid of limitations, the prominent being their potential inaccuracy compared to their direct counterparts. It can be concluded that adherence measures determined from administrative claims data have yet not been standardised and must, thus, be computed and interpreted cautiously with well-defined variables and formulae.

Survival analysis models and statistically analyses data to determine the time until a given outcome occurs (Gardiner & Luo, 2008:1019; Jager, van Dijk *et al*., 2008:560; Kleinbaum & Klein, 2005:4; Singh & Mukhopadhyay, 2011:145). Survival time, also known as time to failure is the time between a defined starting time to the onset of an event of interest (dos Santos Silva, 1999:263; Johnson & Shih, 2007:273; Tolley *et al*., 2016:263). In survival analysis, survival and hazard functions are determined, analysed and compared to determine any correlation with exploratory variables and survival time (Kleinbaum & Klein, 2005:15; Singh & Mukhopadhyay, 2011:148). The survival probability describes the survival experience of a population at various time points (Clark *et al*., 2003:233; Johnson & Shih, 2007:274) and can be illustrated graphically by a survival curve (Kirkwood & Stern, 2003:272; Sullivan, 2016). The life table and Kaplan-Meier methods are applied to model survival curves (Hoffman, 2015:621; Kirkwood & Stern, 2003:272; Sullivan, 2016; Tolley *et al*., 2016:269) (Refer to sections 2.6.3.2.1 and 2.6.3.2.2). The survival patterns of different populations groups can be compared and quantified using the log rank test, hazard ratio and the cox-proportional hazards model (Refer to sections 2.6.3.3.1, 2.6.3.3.2 and 2.6.3.3.3).

From literature it was established that the prescription sequence symmetry analysis (PSSA) was proposed and first employed in 1999 to study the relationship between cardiovascular medication initiation and the incidence of depression (Hallas, 1996:478). This method, by means of administrative claims data, analyses the symmetry in sequence of medication initiation and the onset of an indicator of the adverse event in a specific time frame (Pratt *et al*., 2013:916; Wahab, Pratt, Kalisch, *et al*., 2013:2; Wahab, Pratt, Weise, *et al*., 2013:496) and uses this to assess a relationship between medications and potential adverse events (Wahab *et al*., 2016:349). PSSA has high specificity, is computationally efficient and simple to use (Hallas, 1996:478; Lai *et al*., 2017:570; Pratt *et al*., 2014:1) because it relies on only three variables — patient identifiers, medication code, and date on which the medication is dispensed — all of which are readily available in administrative claims databases (Pratt *et al*., 2013:916; Pratt *et al*., 2014:1; Lai *et al*., 2017:570). It has an inherent ability to control for confounding variables that remain stable over time and uses study subjects that serve as their own controls; requiring no numerical adjustment to control for time-invariant confounders (Bytzer & Hallas, 2000:1480; Hallas, 1996:478; Pratt *et al*., 2013:918; Pratt *et al*., 2014:1; Van Boven, 2013:232; Wahab, Pratt, Kalisch, *et al*., 2013:2;

Wahab *et al*., 2016:349). Graphical output of PSSA allows signals to be better interpreted and a possible temporal association to be observed (Lai *et al*., 2017:570; Pratt *et al*., 2015:863). The method, however, was found to be applicable only in post-market surveillance studies where medications are prescribed for treating untoward medication events or the event leads to hospitalisation (Pratt *et al*., 2014:8) and susceptible to confounding by indication and protopathic bias that can alter causal associations identified using the PSSA (Bytzer & Hallas, 2000:1483; Hashimoto *et al*., 2015:2; Lai *et al*., 2017:579; Pratt *et al*., 2013:920; Wahab, Pratt, Kalisch, *et al*., 2013:2; Wahab *et al*., 2016:349). The conclusion drawn was that PSSA is an important method which can be widely used for determination diverse adverse events especially with the increased development of various healthcare databases that support pharmacoepidemiological research. This method carries the potential to cause an upsurge in the number of studies that investigate causal relationships and thus can lead to increased knowledge in and prevention of ADRs.

From the methods studied in this literature search, it can be concluded that the prevalent use of automated databases as a source of data for pharmacoepidemiology brings with it a host of research methods that can be applied to address various questions in pharmacoepidemiology.

The empirical study involved the illustration of the use of some methods in pharmacoepidemiological studies when medication claims data are employed.

### 4.2.2 Conclusions from empirical investigation

Conclusions from the empirical investigation are explored in the paragraphs that follow.

#### 4.2.2.1 Determination of the time to onset of treatment of hypertension and hyperlipidaemia in patients with type 2 diabetes mellitus using survival analysis

Medicines claims data can be effectively employed in survival analysis to determine time-to-events. From the investigation, it was established that there was a male dominance of hyperlipidaemia while females were found to have a higher prevalence of hypertension among people with diabetes. The average time-to-onset-of-treatment of both hypertension and hyperlipidaemia among people with diabetes was estimated to be about six years after a one-year disease-free period, with no statistically significant variations between male and female patients of different ages (Refer to section 3.3).

Although the absence of clinical data and information on other confounding variables limits their use, medicines claims data are readily accessible and sufficiently large, covering an extensive period and can be efficiently applied in survival analysis for the estimation of time to development of medical outcomes. The extent of analysis, however, depends on the types of data fields

available and the accuracy of patients' records in the database. Survival analysis using retrospective data requires sufficient statistical aptitude for accurate data analysis and subsequent interpretation.

#### 4.2.2.2 Comparison of different adherence measures, by determining adherence to montelukast among asthma patients, using data from a medicines claims database in South Africa

Although medicines claims data have been used extensively for the determination of adherence, the measures for assessing adherence have still not been standardised. Several mathematical formulae exist for various measures, some of which overlap and lead to mathematically equivalent adherence measures. From the study, it was found that although some degree of equivalence exists among the adherence measures investigated and the medication possession ratio (MPR)- which was used as a reference measure - there were some variations in the adherence values obtained using these measures. The continuous multiple interval measure of oversupply (CMOS) and the compliance ratio (CR) were both found to produce the same adherence values as the MPR. When adherence was computed using the proportion of days covered (PDC), adherence was determined to be lower than adherence obtained using MPR. Modified medication possession ratio (MPRm), refill compliance rate (RCR) and average continuous single interval measure of medication acquisition (CSA) produced adherence measures that were higher that adherence produced using MPR. It can also be concluded from the study that the measures that employ the entire study period as the denominator, in their computation, produce consistent adherence values compared to those that use the period between dispensations (Refer to section 3.4).

It can be concluded that the definition of the variables for computation of adherence is critical to obtaining accurate results when assessing adherence using these measures. Variables, such as days' and the time frame for the evaluation, directly affect the nature of results obtained and must be clearly defined to be used in the respective calculations of the measures. Understanding the data available for determining adherence is also important for a decision to be made on the appropriate measures. The choice of measure to be used for computing adherence as well as the interpretation of study findings, thus, depends on the nature of data available from the database and clear definition of variables to be used in computing the adherence measures.

### 4.3 Strengths and limitations of the study

The database does not provide information on important confounders such as disease duration, presence of comorbidities, alcohol consumption, place of birth and smoking status. In this study,

therefore, the methods did not consider how such confounders can be accounted for in pharmacoepidemiological studies using databases.

In spite of this limitation, this is the first study to be carried out in the South African private health sector that assesses different methods, in one study, employed in pharmacoepidemiological studies using medicines claims data and provides more insight into the use of claims data for such studies. Furthermore, the PBM ensures the reliability and validity of the data by eligibility, gate-keeping, utilisation management and clinical management services, with timely benefit pricing and management (refer to paragraph 1.3.4.1). This ensured that information received for this study, and subsequently its results, were sufficiently accurate.

In addition to these, amendments made to the Medical Schemes Act 131 of 1998 mandate the payment for diagnoses, treatment and medications for 27 selected chronic diseases that constitute the Chronic Disease List (CDL), to which diabetes, hypertension, hyperlipidaemia, asthma, *inter alia*, belong (Council for Medical Schemes (CMS), 2017). Registration of patients' chronic conditions with their medical aid schemes using valid ICD-10 codes, that correspond with codes specified by the Council for Medical Schemes, and a report of severity status are required for access to treatment and medications for these chronic conditions. This provision ensures that claims for medications for these conditions are verified for eligibility and are approved prior to reimbursement, resulting in relative accuracy of data contained in medicines claims data for the CDL conditions and thus reducing the possibility of a 'false' disease-free year.

## 4.4 Recommendations

The following recommendations are proposed from the study:

(i) A database that links claims data from all medical providers in South Africa could be established to help obtain larger sample sizes and a more nationally representative data for conducting pharmacoepidemiological studies.

(ii) Further studies should be conducted to streamline and standardise research methods that employ secondary data for pharmacoepidemiological studies.

## 4.5 Chapter summary

This final chapter correlates the achievements of the study with the specifically outlined study objectives. The strengths and limitations were highlighted and discussed, and recommendations proposed for further research.

# BIBLIOGRAPHY


Abbing-Karahagopian, V., Kurz, X., de Vries, F., Van Staa, T.P., Alvarez, Y., Hesse, U., Hasford, J., Van Dijk, L., de Abajo, F.J., Weil, J.G., Grimaldi-Bensouda, L., Egberts, A.C.G., Reynolds, R.F. & Klungel, O.H. 2014. Bridging differences in outcomes of pharmacoepidemiological studies: design and first results of the PROTECT project. *Current clinical pharmacology*, 9(2):130-138.

Acri, T. & Gross, R. 2013. Studies of medication adherence. (*In* Strom, B.L., Kimmel, S.E. & Hennessy, S., *eds*. Textbook of pharmacoepidemiology. 2nd ed. Chichester: Wiley. p. 314-323).

Aguilar-Salinas, C.A., Olaiz, G. Valles, V., Torres, J.M., Gómez Pérez, F.J., Rull, J.A., Rojas, R., Franco, A. & Sepulveda, J. 2001. High prevalence of low HDL cholesterol concentrations and mixed hyperlipidemia in a Mexican nationwide survey. *Journal of lipid research*, 42(8):1298-307.

Akl, C., Akik, C., Ghattas, H & Obermeyer, C.M. 2017. Gender disparities in midlife hypertension: a review of the evidence on the Arab region. *Women's Midlife Health*, 3(1):1-10.

Aleem, I.S., Schemitsch, E.H. & Hanson, B.P. 2008. What is a clinical decision analysis study? *Indian journal of orthopaedics*, 42(2):137-139.

Ali, A.K. 2013. Methodological challenges in observational research: a pharmacoepidemiological perspective. *British journal of pharmaceutical research*, 3(2):161-175.

Andrade, S.E., Kahler, K.H., Frech, F. & Chan, K.A. 2006. Methods for evaluation of medication adherence and persistence using automated databases. *Pharmacoepidemiology and drug safety*, 15(8):565-574.

Andrews, E.B., Margulis, A.V., Tennis, P. & West, S.L. 2014. Opportunities and challenges in using epidemiologic methods to monitor drug safety in the era of large automated health databases. *Current epidemiology reports*, 1(4):194-205.

Anil Babu, A., Rani, S., Revikumar, K.G., Shafeeque, M. & Unnikrishnan, U.G. 2016. Assessment of correlation in gender and age with lipoprotein levels in hyperlipidemia patients. *Asian Journal of pharmaceutical and clinical research*, 9(2):197-199.

Arnet, I., Kooij, M.J., Messerli, M., Hersberger, K.E., Heerdink, E.R. & Bouvy, M. 2016. Proposal of standardisation to assess adherence with medication records: methodology matters. *Annals of pharmacotherapy*, 50(5):360-368.

Aronson, J.K. 2009. Medication errors: what they are, how they happen and how to avoid them. *International journal of medicine*, 102(8):513-521.

Asghari, S., Courteau, J., Carpentier, A.C. & Vanasse, A. 2009. Optimal strategy to identify incidence of diagnostic of diabetes using administrative data. *BMC medical research methodology*, 9(62):1-7.

Aspinall, S.L., Zhao, X., Semla, T.P., Cunningham, F.E., Paquin, A.M., Pugh, M.J., Schmader, K.E., Stone, R.A., Hanlon, J.T. & Veterans Affairs Community Living Center Pharmacotherapy Research Group. 2015. Epidemiology of drug-disease interactions in older veteran nursing home residents. *Journal of the American Geriatrics Society*, 63(1):77-84.

Attia, J., Thakkinstian, A. & D'Este, C. 2003. Meta-analyses of molecular association studies: methodologic lessons for genetics epidemiology. *Journal of clinical epidemiology*, 56(4):297-303.

Austin, P.C. 2011. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate behavioural research*, 46(3):399-424.

Avorn, J. 2004. The role of pharmacoepidemiology and pharmacoeconomics in promoting access and stimulating innovation. *Pharmacoeconomics*, 22(2):81-86.

Avorn, J., Mo, J., Reynolds, R.F., Pan, G.J.D., Arlett, P. & Kesselheim, A.S. 2013. Views from academia, industry, regulatory agencies and the legal system. (*In* Strom, B.L., Kimmel, S.E. & Hennessy, S., *eds*. Textbook of pharmacoepidemiology. 2nd ed. Chichester: Wiley. p. 63-97).

Balcik, P.Y. & Kahraman, G. 2016. Pharmacoepidemiology. *International Organization of Scientific Research (IOSR) journal of pharmacy*, 6(2):57-62.

Banahan, B.F. 2011. Medication safety and pharmacovigilance. (*In* Yang, Y. & West-Strum, D., *eds*. Understanding pharmacoepidemiology. Manhattan, NY: McGraw-Hill. p. 155-168).

Bateman, E.D., Hurd, S.S., Barnes, P.J., Bousquet, J., Drazen, J.M., FitzGerald, M., Gibson, P., Ohta, K., O'Byrne, P., Pedersen, S.E., Pizzichini, E., Sullivan, S.D., Wenzel, S.E. & Zar, H.J. 2008. Global strategy for asthma management and prevention: GINA executive summary. *The European respiratory journal*, 31(1):143-178.

Begaud, B. 2000. Dictionary of pharmacoepidemiology. Chichester: Wiley.

Berger, M.L., Mamdani, M., Atkins, D. & Johnson, M.L. 2009. Good research practices for comparative effectiveness research: defining and interpreting non-randomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Taskforce report – part 1. *Value in health*, 12(8):1044-1052.

Bergman, U. 2001. Pharmacoepidemiology — from description to quality assessment: a Swedish perspective. *Norwegian journal of epidemiology*, 11(1):31-36.

Berlin, J.A., Cepeda, M.S. & Kim, C.J. 2012. The use of meta-analysis in pharmacoepidemiology. (*In* Strom, B.L., Kimmel, S.E. & Hennessy, S., *eds*. Pharmacoepidemiology. 5th ed. Chichester: Wiley. p. 723-756).

Berlin, J.A., Glasser, S.C. & Ellenberg, S.S. 2008. Adverse events detection in drug development: recommendations and obligations beyond phase 3. *American journal of public health*, 98(8):1366-1371.

Bernal, J.L., Cummins, S. & Gasparrini, A. 2017. Interrupted time series regression for the evaluation of public health interventions: a tutorial. *International journal of epidemiology*, 46(1):348-355.

Bewick, V., Cheek, L. & Ball, J. 2004. Statistics review 12: survival analysis. *Critical care*, 8(5):389-394.

Bhopal, R. 2002. Concepts of epidemiology, an integrated introduction to the ideas, theories, principles and methods of epidemiology. New York, NY: Oxford.

Bhurke, S.M., Martin, B.C., Li, C., Franks, A.M., Bursac, Z. & Said, Q. 2012. Effect of the clopidogrel-proton pump inhibitor drug interaction on adverse cardiovascular events in patients with acute coronary syndrome. *Pharmacotherapy*, 32(9):809-818.

Blais, L., Kettani, F.-Z., Beauchesne, M.-F., Lemière, C., Perreault, S. & Forget, A. 2011. New measure of adherence adjusted for prescription patterns: the case of adults with asthma treated with inhaled corticosteroid monotherapy. *The annals of pharmacotherapy*, 45(3):335-341.

Bland, J.M & Altman, D.G. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet*, 327(8476):307-310.

Bland, J.M. & Altman, D.G. 2004. The logrank test. *British medical journal (Clinical research ed.)*, 328(7447):1073.

Bonita, R., Beaglehole, R. & Kjellstrom, T. 2006. Basic epidemiology. 2nd ed. Geneva: World Health Organization.

Bradburn, M.J., Clark, T.G., Love, S.B. & Altman, D.G. 2003. Survival analysis part II: Multivariate data analysis- an introduction to concepts and methods. *British journal of cancer*, 89(3):431-436.

Briggs, A.H. & Levy, A.R. 2006. Pharmacoeconomics and pharmacoepidemiology, curious bedfollows or a match made in heaven? *Pharmacoeconomics*, 24(11):1079-1086.

Brink, H., Van Rensburg, G. & Van der Walt, C. 2012. Fundamentals of research methodology for health care professionals. Cape Town: Juta.

Burnier, M. 2006. Medication adherence and persistence as the cornerstone of effective antihypertensive therapy. *American journal of hypertension,* 19(11):1190-1196.

Byers, J.B. & Stullenbarger, E. 2003. Meta-analysis and decision analysis bridge research and practice. *Western journal of nursing research*, 25(2):193-204.

Bytzer, P. & Hallas, J. 2000. Drug-induced symptoms of functional dyspepsia and nausea: a symmetry analysis of one million prescriptions. *Alimentary pharmacology & therapeutics*, 14(11):1479-1484.

Campagna, E.J., Muser, E., Parks, J. & Morrato, E.H. 2014. Methodological considerations in estimating adherence and persistence for a long-acting injectable medication. *Journal of managed care and specialty pharmacy*, 20(7):756-66.

CDC (Centers for Disease Control and Prevention). 2017. National diabetes statistics report, 2017. Estimates of diabetes and its burden in the United States. Atlanta, GA: US Department of Health and Human Services, CDC.

CDC (Centers for Disease Control and Prevention). 2012. Principles of epidemiology in public health. 3rd ed. Atlanta, GA. https://www.cdc.gov/ophss/csels/dsepd/ss1978/ss1978.pdf. Date of access: 29 Mar. 2017.

Checchi, K.D., Huybrechts, K.F., Avorn, J. & Kesselheim, A.S. 2014. Electronic medication packaging devices and medication adherence: a systematic review. *The journal of the American Medical Association*, 312(12):1237-1247.

Chelkeba, L., Alemseged, F. & Bedada, W. 2013. Assessment of potential drug-drug interactions among outpatients receiving cardiovascular medications at Jimma University specialised hospital, South West Ethiopia. *International journal of basic & clinical pharmacology*, 2(2):144-152.

Chen, G.Y., Li, L., Dai, F., Li, X.J., Xu, X.X, & Fan, J.G. 2015. Prevalence of and risk factors for type 2 diabetes mellitus in hyperlipidemia in China. *Medical science monitor*, 21:2476-84.

Cheung, B.M. & Li, C. 2012. Diabetes and hypertension: is there a common metabolic pathway? *Current atherosclerosis reports*, 14(2):160-6.

Clark, T.G., Bradburn, M.J., Love, S.B. & Altman, D.G. 2003. Survival analysis part I: concepts and first analyses. *British journal of cancer*, 89(2):232-238.

CMS (Council for Medical Schemes). 2017. About us. https://www.medicalschemes.com/Content.aspx?1. Date of access: 25 Jun. 2018.

Cohen, J. 1988. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum.

Cohen, B.J. & Lea, R.B. 2004. Essentials of statistics for the social and behavioural sciences. Hoboken, NJ: Wiley.

Collins, P. 2005. Dictionary of medical terms. 4th ed. London: A&C Black.

Coughlin, S.S. 2006. Ethical issues in epidemiologic research and public health practice. *Emerging themes in epidemiology*, 3(1):3-16.

Cramer, J.A., Roy, A., Burrell, A., Fairchild, C.J., Fuldeore, M.J., Ollendorf, D.A. & Wong, P.K. 2008. Medication compliance and persistence: terminology and definitions. *Value in health*, 11(1):44-47.

D'Agostino, R.B. 1998. Tutorial on biostatistics: propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in medicine*, 17:2265-2281.

Daya, R., Bayat, Z. & Raal, F.J. 2017. Prevalence of pattern of dyslipidaemia in type 2 diabetes mellitus patients at a tertiary care hospital. *Journal of endocrinology, metabolism and diabetes of South Africa*, 22(3):31-35.

De Abajo, F.J., Montero, D., Madurgo, M. & Garcia Rodriguez, L.A. 2004. Acute and clinically relevant drug-induced liver injury: a population-based case-control study. *British journal of clinical pharmacology*, 58(1):71-80.

Delaney, J.A., Opatrny, L., Brophy, J.M. & Suissa, S. 2007. Drug-drug interactions between antithrombotic medications and the risk of gastro-intestinal bleeding. *Canadian Medical Association journal*, 177(4):347-351.

Delaney, J.A.C. & Suissa, S. 2009. The case-crossover study design in pharmacoepidemiology. *Statistical methods in medical research*, 18(1):53-65.

Delgado-Rodríguez, M. & Llorca, I. 2004. Bias. *Journal of epidemiology and community health*, 58(8):635-641.

Deng, B., Luo, T., Huang, Y., Shen, T. & Ma, J. Prevalence and determinants of hyperlipidemia in moderate altitude areas of the Yunnan-Kweichow plateau in Southwestern China. *High altitude medicine and biology*, 13(1):13-21.

Dennis, R.J. & Lozano, J.M. 2001. Ethical challenges in pharmacoepidemiological research in Colombia. *Pharmacoepidemiology and drug safety*, 10(7):613-618.

DiMatteo, M.R., Giordani, P.J., Lepper, H.S. & Croghan, T.W. 2002. Patient adherence and medical treatment outcomes: a meta-analysis. *Medical care*, 40(9):794-811.

DiPietro, N.A. 2010. Methods in epidemiology: observational study designs. *Pharmacotherapy*, 30(10):973-984.

Dixit, A.K., Dey, R, Suresh, A. Chaudhuri, S. & Panda, A.K. 2014. The prevalence of dyslipidemia in patients with diabetes mellitus of Ayurveda hospital. *Journal of diabetes and metabolic disorders*, 13(58):1-6.

Doi, S. & Thalib, L. 2007. A quality-effects model for meta-analysis. *Epidemiology*, 19(1):94-100.

Donnan, P.T., MacDonald, T.M. & Morris, A.D. 2002. Adherence to prescribed oral hypoglycaemic medication in a population of patients with Type 2 diabetes: a retrospective cohort study. *Diabetic medicine*, 19(4):279-284.

Donnan, P.T. & Wang, J. 2001. The case-crossover and case-time-control designs in pharmacoepidemiology. *Pharmacoepidemiology and drug safety,* 10(3):259-262.

Dos Santos Silva, I. 1999. Introduction to survival analysis. (*In* Dos Santos Silva, I., *ed*. Cancer epidemiology: principles and methods. Lyon: International Agency for Research on Cancer (IARC). p. 263-276).

Douglas, I.J., Langham, J., Bhaskaran, K., Brauer, R. & Smeeth, L. 2013. Orlistat and the risk of acute liver injury: self-controlled case series study in UK Clinical Research Datalink. *The British medical journal*, 346(f1936):1-24.

Dreyer, N.A. & Velentgas, P. 2012. Registries. (*In* Strom, B.L., Kimmel, S.E. & Hennessy, S., *eds*. Pharmacoepidemiology. 5th ed. Chichester: Wiley. p.331-346).

Driscoll, K.A., Wang, Y., Johnson, S.B., Gill, E., Wright, N. & Deeb, L.C. 2017. White coat adherence occurs in adolescents with type 1 diabetes receiving intervention to improve insulin pump adherence behaviours. *Journal of diabetes science and technology*, 11(3):455-460.

Dubey, R.K., Oparil, S., Imthurn, B. & Jackson, E.K. 2002. Sex hormones and hypertension. *Cardiovascular research*, 53(3):688-708.

Doubova, S.V., Reyes-Morales, H., Torres-Arreola Ldel, P. & Suárez-Ortega, M. 2007. Potential drug-drug and drug-disease interactions in prescriptions for ambulatory patients over 50 years of age in family medicine clinics in Mexico City. *BMC Health services research*, 7(147):1-8.

Dweik, R.A., Yaya, S. Stacey, D. & Kohen, D. 2016. Spontaneous adverse drug reaction reporting by patients in Canada: a multi-method study-study protocol. *Springerplus*, 5:213. https://springerplus.springeropen.com/articles/10.1186/s40064-016-1838-9. Date of access: 3 Aug. 2017.

Edwards, I.R. & Aronson, J.K. 2000. Adverse drug reactions: definitions, diagnosis and management. *The lancet*, 356(9237):1255-1259.

Eggen, E.A. & Straand, J. 2001. Notes from the guest editors: pharmacoepidemiology — a bridge science. *Norwegian journal of epidemiology*, 11(1):3-4.

Einarson, T.R. 2008. Meta-analysis of pharmacoepidemiologic studies. (*In* Hartzema, A.G., Tilson, H.H. & Chan, K.A., *eds.* Pharmacoepidemiology and therapeutic risk management. Cincinnati, OH: Harvey Whitney. p. 397-437).

El Khoury, A.C. 2013. Bioethical issues in pharmacoepidemiologic research. (*In* Strom, B.L., Kimmel, S.E. & Hennessy, S., *eds*. Textbook of pharmacoepidemiology. 2nd ed. Chichester: Wiley. p. 256-268).

EMA (European Medicines Agency and Heads of Medicines Agency). 2011. Guideline on good pharmacovigilance. (Revision 3). EMA/813938/2011. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500129137.pdf. Date of access: 27 Feb. 2017.

ENCePP (European Network of Centres for Pharmacoepidemiology and Pharmacovigilance). 2017. Guide on methodological standards in pharmacoepidemiology (Revision 6). EMA/95098/2010. http://www.encepp.eu/standards_and_guidances Date of access: 27 Jul. 2017.

Etminan, M. 2004. Pharmacoepidemiology II: The nested case-control study - a novel approach in pharmacoepidemiologic research. *Pharmacotherapy*, 24(9):1105-1109.

Etminan, M. & Samii, M.D. 2004. Pharmacoepidemiology I: a review of pharmacoepidemiologic study designs. *Pharmacotherapy*, 24(8):964-969.

Etminan, M., Gill, S., FitzGerald, M. & Samii, A. 2006. Challenges and opportunities for pharmacoepidemiology in drug-therapy decision making. *Journal of clinical pharmacology*, 46(1):6-9.

Evans, S.J.W. 2012. An agenda for UK clinical pharmacology, pharmacoepidemiology. *British journal of clinical pharmacology*, 73(6):973-978.

Faillie, J.-L. 2015. Indication bias or protopathic bias? *British journal of clinical pharmacology*, 80(4):779-780.

Fairman, K. & Motheral, B. 2000. Evaluating medication adherence: which measure is right for your program? *Journal of managed care pharmacy*, 6(6):499-504.

Farmer, K.C. 1999. Methods for measuring and monitoring medication regimen adherence in clinical trials and clinical practice. *Clinical therapeutics*, 21(6):1074-1090.

Fautrel, B. 2004. Pharmacoepidemiology: lessons from real life. *Joint bone spine*, 71(3):175-177.

Fletcher, R.H. & Fletcher, S.W. 2005. Clinical epidemiology: the essentials. 4th ed. Philadelphia, PA: Williams & Wilkins.

Fortepiani, L.A., Zhang, H., Racusen, L., Roberts, L.J. & Reckelhoff, J.F. 2003. Characterization of an animal model of postmenopausal hypertension in spontaneously hypertensive rats. *Hypertension*, 41(3 pt 2):640-645.

Freeman, J.L., Zhang, D., Freeman, D.H., Jr. & Goodwin, J.S. 2000. An approach to identifying incident breast cancer cases using Medicare claims data. *Journal of clinical epidemiology*, 53(6):605-614.

Friis, R.H. & Sellers, T.A. 2009. Epidemiology for public health practice. 4th ed. Sudbury, MA: Jones and Bartlett.

Gama, H. 2008. Drug utilisation studies. *Arquivos de medicina*, 22(2/3):69-74.

Garbe, E. & Suissa, S. 2014. Pharmacoepidemiology. (*In* Ahrens, W. & Pigeot, I., *eds*. Handbook of epidemiology. 2nd ed. New York, NY: Springer. p. 1875-1925).

Gardiner, J.C. & Luo, Z. 2008. Survival analysis. (*In* Boslaugh, S., *ed*. Encyclopedia of epidemiology 1&2. Los Angeles, CA: SAGE. p. 1019-1024).

Gelfand, J.M., Weisten, R., Porter, S.B., Neimann, A.L., Berlin, J.A. & Margolis, D.J. 2005. Prevalence and treatment of psoriasis in the United Kindgom. *Archives of dermatology journal*, 141(12):1537-1541.

Ghosh, S., Mukhopadhyay, S. & Barik, A. 2016. A sex differences in the risk profile of hypertension: a cross-sectional study. *BMJ open*, 6(7): e010085.

Giardini, A., Martin, M.T., Cahir, C., Lehane, E., Menditto, E., Strano, M., Pecorelli, S., Monaco, A. & Marengoni, A. 2016. Toward appropriate criteria in medication adherence assessment in older persons: position paper. *Aging clinical and experimental research*, 28(3):371-381.

Giavarina, D. 2015. Understanding Bland Altman analysis. *Biochemia medica*, 25(2):141-151.

Gillespie, C.D., Hurvitz, K.A. & CDC (Centers for Disease Control and Prevention). Prevalence of hypertension and controlled hypertension- United States, 2007-2010. *MMWR Supplements*, 62(3):144-148.

Gillisen, A. 2007. Patients' adherence in asthma. *Journal of physiology and pharmacology*, 58(Suppl. 5):S205-S222.

Gliklich, P.E., Dreyer, N.A. & Leavy, M.B. 2014. Registries for evaluating patient outcomes: a user's guide. 3rd ed. Rockville, MD: Agency for Healthcare Research and Quality (AHRQ).

Go, A.S., Hylek, E.M., Phillips, K.A., Chang, Y.C., Henault, L.E., Selby, J.V. & Singer, D.E. 2001. Prevalence of diagnosed atrial fibrillation in adults. National implications for rhythm management and stroke prevention: the AnTicoagulation and Risk factors on atrial fibrillation (ATRIA) study. *The journal of the American Medical Association*, 285(10):2370-2375.

Goel, M.K., Khanna, P. & Kishore, J. 2010. Understanding survival analysis: Kaplan-Meier estimate. *International ayurveda research*, 1(4):274-278.

Goldberg, I.J. 2001. Diabetic dyslipidemia: causes and consequences. *The Journal of clinical endocrinology and metabolism*, 86(3):965-71

Goldberg, R.B. 2000. Hyperlipidemia and cardiovascular risk factors in patients with type 2 diabetes. The American journal of managed care, 6(13 Suppl):S682-S691.

Gordis, L. 2014. Epidemiology. 5th ed. Philadelphia, PA: Elsevier.

Granados, J.A.T. 2008. Time series. (*In* Boslaugh, S., *ed*. Encyclopedia of epidemiology 1&2. Los Angeles, CA: SAGE. p. 1035-1038).

Gray, A., Riddin, J. & Jugathpal, J. 2016. Health care and pharmacy practice in South Africa. *The Canadian journal of hospital pharmacy*, 69(1):36-41.

Grégoire, J.-P. & Moisan, J. 2016. Assessment of adherence to drug treatment in database research. (*In* Elseviers, M., Wettermark, B., Almarsdóttir, A.B., Morten, A., Benko, R., Bennie, M., Eriksson, I., Godman, B., Krska, J., Poluzzi, E., Taxis, K., Vlahović-Palčevski, V. & Vander-Stichele, R.H., *eds*. Drug utilization research. Hoboken, NJ: Wiley. p. 369-380).

Grzeskowiak, L.E., Gilbert, A.L. & Morrison, J.L. 2012. Investigating outcomes associated with medication use during pregnancy: a review of methodological challenges and observational study designs. *Reproductive toxicology*, 33(3):280-289.

Gulmez, S.E., Holm, A., Frederriksen, H., Jensen, T.G., Pederson, C. & Hallas, J. 2007. Use of proton pump inhibitors and the risk of community-acquired pneumonia, a population-based case-control study. *Archives of internal medicine*, 167(9):950-955.

Gurtwitz, J.H. & Field, T.S. 2008. Medication errors. (*In* Hartzema A.G., Tilson, H.H., Chan, K.A., *eds*. Pharmacoepidemiology and therapeutic risk management. Cincinnati, OH: Harvey Whitney. p. 653-670).

Haidich, A.B. 2010. Meta-analysis in medical research. *Hippokratia*, 12(S1):29-37.

Hall, G.C., Sauer, B., Bourke, A., Brown, J.S., Reynolds, M.W. & Casale, R.L. 2011. Guidelines for good database selection and use in pharmacoepidemiological research. *Pharmacoepidemiology and drug safety*, 21(1):1-10.

Hallas, J. 1996. Evidence of depression provoked by cardiovascular medication: a prescription sequence symmetry analysis. *Epidemiology*, 7(5):478:484.

Hallas, J. 2001. Pharmacoepidemiology- current opportunities and challenges. *Norwegian journal of epidemiology*, 11(1):7-12.

Hallas, J. & Pottegård, A. 2014. Use of self-controlled designs in pharmacoepidemiology. *Journal of internal medicine*, 275(6):581-589.

Halpern, M.T., Khan, Z.M., Schmier, J.K., Burnier, M., Caro, J.J., Cramer, J., Daley, W.L., Gurwitz, J. & Hollenberg, N.K. 2006. Recommendations for evaluating compliance and persistence with hypertension therapy using retrospective data. *Hypertension*, 47(6):1039-1048.

Hammer, G.P., Du Prel, J.-B. & Blettner, M. 2009. Avoiding bias in observational studies. *Deutsches ärzteblatt international*, 106(41):664-668.

Hansen, K., Schüssel, K., Kieble, M., Werning, J., Schulz, M., Friis, R., Pöhlau, D., Schmitz, N. & Kugler, J. 2015. Adherence to Disease Modifying Drugs among Patients with Multiple Sclerosis in Germany: A Retrospective Cohort Study. *PLoS One,* 10(7):1-12.

Harpe, S.E. 2009. Using secondary data sources for research. *Pharmacotherapy*, 28(2):138-153.

Harpe, S. 2011. Study designs for pharmacoepidemiology. (*In* Yang, Y. & West-Strum, D., *eds*. Understanding pharmacoepidemiology. New York, NY: McGraw-Hill. p. 39-54).

Harris, A.D., McGregor, J.C., Perencevich, E.N., Furuno, J.P., Zhu, J., Peterson, D.E. & Finkelstein, J. 2006. The use and interpretation of quasi-experimental studies in medical informatics. *Journal of the American Medical Informatics Association*, 13(1):16-23.

Hartzema, A.G., Tilson, H.H. & Chan, K.A. 2008. The contribution of pharmacoepidemiology to the study of drug uses and effects, and risk management. (*In* Hartzema, A.G., Tilson, H.H., Chan, K.A., *eds*. Pharmacoepidemiology and therapeutic risk management. Cincinnati, OH: Harvey Whitney. p. 4-6).

Hashimoto, M., Hashimoto, K., Ando, F., Kimura, Y, Nagase, K. & Arai, K. 2015. Prescription rate of medications potentially contributing to lower urinary tract symptoms and detection of adverse reactions by prescription sequence symmetry analysis. *Journal of pharmaceutical health care and sciences*, 1(7):1-9.

Hauser, T., Mahr, A., Metzler, C., Coste, J., Sommerstein, R., Gross, W.L., Guillevin, L. & Hellmich, B. 2008. The leukotriene receptor antagonist Montelukast and the risk of Churg-Strauss syndrome: a case-crossover study. *Thorax*, 63(8):677-682.

HealthKnowledge. 2012. Measures of disease frequency and disease burden. https://www.healthknowledge.org.uk/e-learning/epidemiology/practitioners/measures-disease-frequency. Date of access: 17 Aug. 2017.

HealthKnowledge. 2018. Biases and confounding. https://www.healthknowledge.org.uk/public-health-textbook/research-methods/1a-epidemiology/biases. Date of access: 21 Jan. 2019.

Heiman, G.W. 2014. Basic statistics for behavioural sciences. 7th ed. Belmont, CA: Wadsworth.

Heit, J.A., Silverstein, M.D., Mohr, D.N., Petterson, T.M., O'Fallon, M. & Melton, L.J., 3rd. 2000. Risk factors for deep vein thrombosis and pulmonary embolism; a population-based case-control study. *Archives of internal medicine*, 160(6):809-815.

Hennessy, S. 2006. Use of health care databases in pharmacoepidemiology. *Basic & clinical pharmacology & toxicology*, 98(3):311-313.

Hennessy, S., Leonard, C.E., Gagne, J.J., Flory, J.H., Han, X, Brensinger, C.M & Bilker, W.B. 2016. Pharmacoepidemiologic methods for studying the health effects of drug-drug interactions (DDIs). *Clinical pharmacology and therapeutics*, 99(1):92-100.

Hersom, K., Neary, M., Levaux, H.P., Klaskala, W. & Strauss, J.S. 2003. Isotretinoin and antidepressant pharmacotherapy: a prescription sequence symmetry analysis. *Journal of the American Academy of Dermatology*, 49(3):424-432.

Hertz, R.P., Unger, A.N. & Lustik, M.B. 2005. Adherence with pharmacotherapy for type 2 diabetes: a retrospective cohort study of adults with employer-sponsored health insurance. *Clinical therapeutics*, 27(7):1064-1073.

Hess, L.M., Raebel, M.A., Conner, D.A. & Malone, D.C. 2006. Measurement of adherence in pharmacy administrative databases: a proposal for standard definitions and preferred measures. *The annals of pharmacotherapy*, 40(7-8):1280-1288.

Hill, S., Henry, D., Pekarsky, B. & Mitchell, A. 1997. Economic evaluation of pharmaceuticals: what are reasonable standards for clinical evidence: the Australian experience. *British journal of clinical pharmacology*, 44(5):421-425.

Ho, P.M., Bryson, C.L. & Rumsfeld, J.S. 2009. Medication adherence: its importance in cardiovascular outcomes. *Circulation*, 119(23):3028-35.

Hoffman, J.I.E. 2015. Biostatistics for medical and biomedical practitioners. San Diego, CA: Academic Press.

Hoomans, T. & Severens, J.L. 2014. Economic evaluation of implementation strategies in health care. *Implementation science*, 9(168):1-6.

Howren, M.B. 2013. Quality-adjusted life years (QALYs). (*In* Gellman, M.D. & Turner, J.R. *eds.* Encyclopedia of behavioural medicine. New York, NY: Springer. p. 104-113).

Hughes, J.P., Haley, D.F., Frew, P.M., Golin, C.E., Adimora, A.A., Kuo, I., Justman, J., Soto-Torres, L., Wang, J. & Hodder, S. 2015. Regression to the mean and changes in risk behaviour following study enrolment in a cohort of U.S. women at risk for H.I.V. *Annals of epidemiology*, 25(6):439-444.

Hux, J.E., Ivis, F., Flintoft, V. & Bica, A. 2002. Diabetes in Ontario: determination of prevalence and incidence using a validated administrative data algorithm. *Diabetes care*, 25(3):512-516.

IBM Corp. 2017. IBM SPSS Statistics for Windows, Version 25. Armonk, NY:IBM Corp.

ISPE (International Society of Pharmacoepidemiology). 2016. Guidelines for good pharmacoepidemiology practices (GPP*). Pharmacoepidemiology and drug safety*, 25(1):2-10.

Ivankova, N.V., Creswell, J.W. & Plano Clark, V.L. 2016. Foundations and approaches to mixed methods research. (*In* Maree, K., *eds*. First steps in research. Pretoria: Van Schaik. p. 305-328).

Jackevicius, C.A., Mamdani, M. & Tu, J.V. 2002. Adherence with statin therapy in elderly patients with and without acute coronary syndromes. *The journal of the American Medical Association*, 288(4):462-467.

Jacobs, K., Julyan, M., Lubbe, M.S., Burger, J.R. & Cockeran, M. 2016. Medicine possession ratio as proxy for adherence to antiepileptic drugs: prevalence, associations, and cost implications. *Patient preference and adherence*, 10:539-47.

Jager, K.J., van Djik, P.C., Zoccali, C. & Dekker, F.W. 2008. The analysis of survival data: The Kaplan-Meier method. *Kidney international*, 74(5):560-565.

Jager, K.J., Zoccali, C., MacLeod, A. & Dekker, F.W. 2008. Confounding: what it is and how to deal with it. *Kidney international*, 73(3):256-260.

Jandoc, R., Burden, A.M., Mamdani, M., Levesque, L.E. & Cadarette, S.M. 2015. Interrupted time series analysis in drug utilisation research is increasing: systematic review and recommendations. *Journal of clinical epidemiology*, 68(8):950-956.

Jick, H., Vasilakis, C., Weinrauch, L.A., Meier, C.R., Jick, S.S. & Derby, L.E. 1998. A population-based study of appetite-suppressant drugs and the risk of cardiac-valve regurgitation. *The New England journal of medicine*, 339(11):719-724.

Jimmy, B. & Jose, J. 2011. Patient medication adherence: measures in daily practice. *Oman medical journal*, 26(3):155-159.

Johnson, L.L. & Shih, J.H. 2007. An introduction to survival analysis. (*In* Gallin, J.I. & Ognibene, F.P., *eds*. Principles and practice of clinical research. 2nd ed. Burlington, MA: Academic press. p. 273-282).

Joubert, G. 2014. Exploring, summarising and presenting data. (*In* Ehrlich, R. & Joubert, G., *eds*. Epidemiology: a research manual for South Africa. Cape Town: Oxford. p. 132-146).

Joubert, G., Ehrlich, R., Katzenellenbogen, J. & Hoffman, M. 2014. Key concepts in epidemiology. (*In* Ehrlich, R. & Joubert, G., *eds*. Epidemiology: a research manual for South Africa. 3rd ed. Cape Town: Oxford. p. 13-32).

Juurlink, D.N., Mamdani, M., Kopp, A., Laupacis, A. & Redermeier, D.A. 2003. Drug-drug interactions among elderly patients hospitalized for drug toxicity. *Journal of the American Medical Academy*, 289(13):1652-1658.

Juurlink, D.N., Gomes, T., Ko, D.T., Szmitko, P.E., Austin, P.C., Tu, J.V., Henry, D.A., Kopp, A. & Mamdani, M.M. 2009. A population-based study of the drug interaction between proton pump inhibitors and clopidogrel. *Canadian Medical Association journal*, 180(7):713-718.

Kaplan, W., Wirtz, V.J., Mantel-Teeuwisse, A., Stolk, P., Duthey, B. & Laing R. 2013. Priority medicines for Europe and the world: 2013 update. Geneva: WHO.

Karve, S., Cleves, M.A., Helm, M., Hudson, T.J., West, D.S. & Martin, B.C. 2008. An empirical basis for standardizing adherence measures derived from administrative claims data among diabetic patients. *Medical care*, 46(11):1125-33.

Karve, S., Cleves, M.A., Helm, M., Hudson, T.J., West, D.S. & Martin, B.C. 2009. Prospective validation of eight different adherence measures for use with administrative claims data among patients with schizophrenia. *Value in health*, 12(6):989-995.

Kaufman, D.W. 2013. Field studies. (*In* Strom, B.L., Kimmel, S.E. & Hennessy, S., *eds*. Textbook of pharmacoepidemiology. 2nd ed. Chichester: Wiley. p. 178-189).

Kaur, S., Rajagopalan, S., Kaur, N., Shafiq, N., Bhalla, A., Pandhi, P. & Malhotra, S. 2014. Drug utilisation study in medical emergency unit of a tertiary care hospital in North India. *Emergency medicine international*, 2014(973578):1-5.

Kennedy, M.S.N. 2012. Pancreatic hormones & antidiabetic drugs. (*In* Katzung, B.G., Masters, S.B. & Trevor, A., *eds*. Basic & clinical pharmacology. 12th ed. Manhattan, NY: McGraw-Hill. p. 743-768).

Kildemoes, H.W., Hendriksen, C. & Morten, A. 2012. Drug utilisation according to reason for prescribing: a pharmacoepidemiologic method based on an indication hierarchy. *Pharmacoepidemiology and drug safety*, 21(10):1027-1035.

Kirkwood, B.R. & Sterne, J.A.C. 2003. Essential medical statistics. Malden, MA: Blackwell.

Kjær, D., bet Horvath-Puho, E., Christensen, J., Vestergaard, M., Czeizel, A.E., Sørensen, H.T. & Olse, J. 2006. Use of phenytoin, phenobarbital, or diazepam during pregnancy and risk of congenital abnormalities: a case-time-control study. *Pharmacoepidemiology and drug safety*, 16(2):181-188.

Kleinbaum, D.G. & Klein, M. 2005. Survival analysis: a self-learning text. 2nd ed. New York, NY: Springer.

Kongpatanakul, S. & Strom, B.L. 2001. Pharmacoepidemiology and drug evaluation. (*In* Van Boxtel, C.J., Santoso, B. & Edwards, I.R., *eds*. Drug benefits and risks, International textbook for clinical pharmacology. Chichester: Wiley. p. 27-36).

Kozma, C.M., Dickson, M., Phillips, A.L. & Meletiche, D.M. 2013. Medication possession ratio: implications of using fixed and variable observation periods in assessing adherence with disease-modifying drugs in patients with multiple sclerosis. *Patient preference and adherence*, 7:509-16.

Kumar, A., Sattigeri, B.M. & Chauhan, S. 2013. Drug utilisation in clinical conditions: an update on its essentials. *International journal of research in medical sciences*, 1(4):320-325.

Laakso, M. 2010. Cardiovascular disease in type 2 diabetes from population to man to mechanism. *Diabetes care*, 33(2):442-449.

Lagarde, M. 2012. How to do (or not to do)…Assessing the impact of a policy change with routine longitudinal data. *Health policy and planning*, 27(1):76-83.

Lai, E.C.-C., Hsieh, C.Y., Yang, Y.H.K. & Lin, S.J. 2014. Detecting potential adverse reactions of sulpiride in schizophrenic patients by prescription sequence symmetry analysis. *PLoS one*, 9(2):1-8.

Lai, E.C.-C., Pratt, N., Hsieh, C.-Y., Lin, S.-J., Pottegård, A., Roughead, E.E., Yang, Y.-H.K. & Hallas, J. 2017. Sequence symmetry analysis in pharmacovigilance and pharmacoepidemiologic studies. *European journal of epidemiology*, 32(7):567-582.

Lam, W.Y. & Fresco, P. 2015. Medication adherence measures: an overview. *BioMed research international*, 2015(217047):1-12.

Larson, L.N. & Bjornson, D.C. 1996. Interface between pharmacoepidemiology and pharmacoeconomics in managed care pharmacy. *Journal of managed care pharmacy*, 2(3):282-289.

Lau, M.H.M. & Tenney, J.W. 2017. Evaluation of drug-disease interactions and their association with unplanned hospital readmission utilizing STOPP Version 2 Criteria. *Geriatrics*, 2(3):1-10.

Layton, D. & Shakir, S.A.W. 2012. Prescription-event monitoring. (*In* Strom, B.L., Kimmel, S.E. & Hennessy, S., *eds*. Pharmacoepidemiology. 5th ed. Chichester: Wiley. p. 301-330).

Lee, D., Majumdar, S.R, Lipton, H.L., Soumerai, S.B., Vellozzi, C, Chen, R.T., Glanz, J., Marinac-Dabic, D., Normand, S.L.T., Sedrakyan, A., Mitchell, A.A., Pan, G.J.D., Blackburn, S., Manzo, C., Seidling, H.M., Bates, D.W., Robb, M.A., Sherman, R.E., Strom, B.L., Schinnar, R. & Hennessy, S. 2013. Special applications of pharmacoepidemiology. (*In* Strom, B.L., Kimmel,

S.E. & Hennessy, S., *eds*. Textbook of pharmacoepidemiology. 2nd ed. Chichester: Wiley. p. 339-406).

Leendertse, A.J., Egberts, A., Stoker, L.J. & Van den Bemt, P.M.L.A. 2008. Frequency of and risk factors for preventable medication-related hospital admissions in the Netherlands. *Archives of internal medicine*, 168(17):1890-1896.

Lesko, S.M. & Mitchell, A.A. 2013. The use of randomized controlled trials for pharmacoepidemiologic studies. (*In* Strom, B.L., Kimmel, S.E. & Hennessy, S., *eds*. Textbook of pharmacoepidemiology. 2nd ed. Chichester: Wiley. p. 269-279).

Lighter, D.E. 2004. Process orientation in healthcare. (*In* Lighter, D.E. & Fair, D.C., *eds*. Quality management in healthcare: principles and methods. Mississauga: Jones & Bartlett. p. 43-101).

Lima-Dellamora, E.d.C., Osorio-de-Castro, C.G., Madruga, L.G.d.S.L. & Azeredo, T.B. 2017. Use of pharmacy records to measure treatment adherence: a critical review of the literature. *Cadernos de Saude publica*, 33(3):1-16.

Lindblad, C.I., Hanlon, J.T., Gross, C.R., Sloane, R.J., Pieper, C.F., Hajjar, E.R., Ruby, C.M., Schmader, K.E. & Multidisciplinary Consensus Panel. 2006. Clinically important drug-disease interactions and their prevalence in older adults. *Clinical therapeutics*, 28(8):1133-1143.

Lovely, J.A., Esper, S., Hutchens, M.P., Nicholson, W.T. & Marcucci, C. 2015. The pharmacoepidemiology of drug interactions: why and how they are important. (*In* Marcucci, C., Hutchens, M.P., Wittwer, E.D., Weingarten, T.N., Sprung, J., Nicholson, W.T., Lalwani, K., Metro, D.G., Dull, R.O., Swide, C.E., Seagull, F.J., Kirsch, J.R. & Sandson, N.B., *eds*. A case approach to perioperative drug-drug interactions. New York, NY: Springer. p. 103-108).

Lu, C.Y. 2009. Observational studies: a review of study designs, challenges and strategies to reduce confounding. *International journal of clinical practice*, 63(5):691-697.

Lu, C.Y. 2015. Pharmacoepidemiological approaches in health care. (*In* Babar, Z.-U.-D., *ed*. Pharmacy practice research methods. Cham: Springer. p. 197-222).

Lu, C.Y., Zhang, F., Lakoma, M.D., Madden, J.M., Rusinak, D., Penfold, R.B., Simon, G., Ahmedani, B.K., Clarke, G., Hunkeler, E.M., Waitzfelder, B., Owen-Smith, A., Raebel, M.A., Rossom, R., Coleman, K.J., Copeland, L.A. & Soumerai, S.B. 2014. Changes in antidepressant use by young people and suicidal behavior after FDA warnings and media coverage: quasi-experimental study. *British medical journal (clinical research edition),*

348(g3596). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4062705/. Date of access: 7 Mar. 2017.

Maclure, M., Fireman, B., Nelson, J.C., Hua, W., Shoaibi, A., Paredes, A. & Madigan, D. 2012. When should case-only designs be used for safety monitoring of medical products? *Pharmacoepidemiology and drug safety*, 21(Suppl.1):S50-S61.

Malloy, M.J. & Kane, J.P. 2012. Agents used in dyslipidemia. (*In* Katzung, B.G., Masters, S.B. & Trevor, A., *eds*. Basic & clinical pharmacology. 12th ed. Manhattan, NY: McGraw-Hill. p. 619-633).

Malone, D.C., Hutchins, D.S., Haupert, H., Hansten, P., Duncan, B., Van Bergen, R.C., Solomon, S.L. & Lipton, R.B. 2005. Assessment of potential drug-drug interactions with a prescription claims database. *American journal of health-system pharmacy*, 62(19):1983-1991.

Manack, A., Turkel, C.C. & Kaplowitz, H. 2012. Role of epidemiological data within the drug development lifecycle: a chronic migraine case study. (*In* Lunet, N., *ed*. Epidemiology-current perspectives on research and practice. Rijeka: InTech. p. 155-170).

Maree, K. & Pietersen, J. 2016. The quantitative research process. (*In* Maree, K., *eds*. First steps in research. Pretoria: Van Schaik. p. 162-172).

Maritz, F.J. Dyslipidaemia in South Africa. (*In* Steyn K, Fourie J, Temple N, *eds*. Chronic diseases of lifestyles in South Africa: 1995-2005. Medical Research Council Technical. Tygerberg, SA: Medical Research Council. P. 97-108).

Martin, B.C., Wiley-Exley, E.K., Richards, S., Domino, M.E., Carey, T.S. & Sleath, B.L. 2009. Contrasting measures of adherence with simple drug use, medication switching, and therapeutic duplication. *The annals of pharmacotherapy*, 43(1):36-44.

Masic, I., Miokovic, M. & Muhamedagic, B. 2008. Evidence-based medicine-new approaches and challenges. *Journal of the Society of Medical Informatics of Bosnia & Herzegovina*, 16(4):219-225.

Masoli, M., Fabian, D., Holt, S., Beasley, R. 2004. The global burden of asthma: executive summary of the GINA Dissemination Committee report. *Allergy*, 59(5):469-478.

Matshidze, P. & Hanmer, L. 2007. Health information systems in the private health sector. (*In* Harrison, S., Bhana, R.& Ntuli, A., *eds*. South Africa Health Review 2007. National Department of Health: Health Systems Trust. p. 89-102).

McCaffrey, D.J. 2011. Medication utilization patterns. (*In* Yang, Y. & West-Strum, D., *eds.* Understanding pharmacoepidemiology. Manhattan, NY: McGraw-Hill. p. 369-380).

McCarthy, S., Asherson, P., Coghill, D., Hollis, C., Murray, M., Potts, L., Sayal, K., de Soysa, R., Taylor, E., Williams, T. & Wong, I.C.K. 2009. Attention-deficit hyperactivity disorder: treatment discontinuation in adolescents and young adults. *The British journal of psychiatry*, 194(3):273:277.

McCowan, C., Shearer, J., Donnan, P.T., Dewar, J.A., Crilly, M., Thompson, A.M. & Fahey, T.P. 2008. Cohort study examining tamoxifen adherence and its relationship to mortality in women with breast cancer. *British journal of cancer*, 99(11):1763-1768.

McGhan, W.F. 2010. Introduction to pharmacoeconomics. (*In* Arnold, R.J., *eds*. Pharmacoeconomics from theory to practice. Boca Raton, FL: Taylor and Francis Group. p. 1-16).

McMahon, A.D. & MacDonald, T.M. 2000. Design issues for drug epidemiology. *British journal of clinical pharmacology*, 50(5):419-425.

Mishra, L., Patnaik, S., Patro, S., Debnath, N. & Mishra, S. 2014. Iatrogenic subcutaneous emphysema of endodontic origin- case report with literature review. *Journal of clinical and diagnostic research*, 8(1):279-281.

Monane, M., Bohn, R.L., Gurwitz, J.H., Glynn, R.J., Levin, R. & Avorn, J. 1996. Compliance with antihypertensive therapy among elderly Medicaid enrollees: the roles of age, gender, and race. *American journal of public health*, 86(12):1805-1808.

Morgan, O.W.C., Griffiths, C. & Majeed, A. 2004. Association between mortality from suicide in England and antidepressant prescribing: an ecological study. *BMC public health*, 4(63):1-6.

Morroni, C. & Myer, L. 2014. Study design. (*In* Erhlich, R. & Joubert, G., *eds*. Epidemiology. a research manual for South Africa. 3rd ed. Cape Town: Oxford. p. 78-97).

Motheral, B.R. & Fairman, K.A. 1997. The use of claims databases for outcomes research: rationale, challenges, and strategies. *Clinical therapeutics*, 19(2):346-366.

Motheral, B., Brooks, J., Clark, M.A., Crown, W.H., Davey, P., Hutchins, D., Martin, B.C. & Stang, P. 2003. A checklist for retrospective database studies-report of the ISPOR Task Force on Retrospective Databases. *Value in health*, 6(2):90-97.

Mullner, R.M. 2009. Encyclopedia of health services research. Thousand Oaks, CA: SAGE.

Narayana, S.K., Woods, D.R. & Boos, C.J. 2011. Management of amiodarone-related thyroid problems. *Therapeutic advances in endocrinology and metabolism,* 2(3):115-126.

NCC MERP (National Coordination Council for Medication Error Reporting and Prevention). 2017. About medication errors. http://www.nccmerp.org/about-medication-errors. Date of access: 9 Mar. 2017.

Neimann, A.L., Shin, D., Wang, X., Margolis, D.J., Troxel, A.B. & Gelfand, J.M. 2006. Prevalence of cardiovascular risk factors in patients with psoriasis. *Journal of the American Academy of Dermatology,* 55(5):829-835.

Nørgaard, M., Ehrenstein, V. & Vandenbroucke, J.P. 2017. Confounding in observational studies based on large health care databases: problems and potential solutions – a primer for the clinician. *Clinical epidemiology*, 9:185-193.

Nwose, E.U. & Yee, K.C. 2016. Drug-disease interactions: narrative review of aspirin in gastric ulcer. *Expert opinion on drug metabolism and toxicology*, 12(9):1081-7.

Ong, K.L., Cheung, B.M.Y., Man, Y.B., Lau, C.P. & Lam, K.S.L. 2007. Prevalence, awareness, treatment and control of hypertension among United States Adults 1999-2004. *Hypertension*, 49(1):69-75.

Osheroff, J. 2009. Improving medication use and outcomes with clinical decision support: a step-by-step guide. Chicago, IL: Healthcare Information and Management Systems Society (HIMSS).

Osterberg, L. & Blaschke, T. 2005. Adherence to medication. *The New England journal of medicine*, 353(5):487-497.

O'Toole, M.T. 2013. Mosby's medical dictionary. 9th ed. St. Louis, MO: Elsevier.

Pan, G.J.D., Lindquist, M. & Gelperin, K. 2013. Postmarketing spontaneous pharmacovigilance reporting systems. (*In* Strom, B.L., Kimmel, S.E. & Hennessy, S., *eds*. Textbook of pharmacoepidemiology. 2nd ed. Chichester: Wiley. p. 101-117).

Park, B.J. & Stergachis, A. 2008. Automated databases in pharmacoepidemiologic studies. (*In* Hartzema A.G., Tilson, H.H., Chan, K.A., *eds*. Pharmacoepidemiology and therapeutic risk management. Cincinnati, OH: Harvey Whitney. p. 519-544).

Pearl, J. 2013. Understanding Simpson's Paradox. http://ftp.cs.ucla.edu/pub/stat_ser/r414.pdf Date of access: 29 Jun. 2017.

Penfold, R.B. & Zhang, F. 2013. Use of interrupted time series analysis in evaluating health care quality improvements. *Academic paediatrics*, 13(Suppl. 6):S38-S44.

Petersen, I., Douglas, I & Whitaker, H. 2016. Self-controlled case series methods: an alternative to standard epidemiological study designs. *British medical journal (clinical research edition)*, 354:i4515. https://www.bmj.com/content/bmj/354/bmj.i4515.full.pdf. Date of access: 31 May 2018.

Peterson, A.M., Nau, D.P., Cramer, J.A., Benner, J., Gwadry-Sridhar, F. & Nichol, M. 2007. A checklist for medication compliance and persistence studies using retrospective databases. *Value in health*, 10(1):3-12.

Petitti, D.B. 2000. Meta-analysis, decision analysis and cost-effectiveness analysis: methods for quantitative synthesis in medicine. New York, NY: Oxford.

Plumeri, P.A. 1997. Managed care and risk management. *Gastroenterology clinics of North America*, 26(4):895-910.

Porta, M. 2014. A dictionary of epidemiology. 6th ed. Manhattan, NY: Oxford.

Pourhoseingholi, M.A., Baghestani, A.R. & Vahedi, M. 2012. How to control confounding effects by statistical analysis. *Gastroenterology and hepatology from bed to bench*, 5(2):79-83.

Pratt, N., Andersen, M., Bergman, U., Choi, N.-K., Gerhard, T., Huang, C., Kimura, M., Kimura, T., Kubota, K., Lai, E.C.-C., Man, K.K.C., Ooba, N., Park, B.-J., Sato, T., Shin, J.-Y., Sundström, A., Yang, Y.-H.K. & Roughead, E.E. 2013. Multi-country rapid adverse drug event assessment: the Asian Pharmacoepidemiology Network (AsPEN) antipsychotic and acute hyperglycaemia study. *Pharmacoepidemiology and drug safety*, 22(9):915-924.

Pratt, N., Chan, E.W., Choi, N.-K., Kimura, M., Kimura, T., Kubota, K., Lai, E.C-C., Man, K.K.C., Ooba, N., Park, B.-J., Sato, T., Shin, J.-Y, Wong, I.C.K., Yang, Y.-H.K. & Roughead, E.E. 2015. Prescription sequence symmetry analysis: assessing risk, temporality, and consistency for adverse drug reactions across datasets in five countries. *Pharmacoepidemiology and drug safety*, 24(8):858-864.

Pratt, N.L., Ilomäki, J., Raymond, C. & Roughead, E. 2014. The performance of sequence symmetry analysis as a tool for post-market surveillance of newly marketed medicines: a simulation study. *BMC medical research methodology*, 14(66):1-9.

Prosser, L.A., Grosse, S.D., Kemper, A.R., Tarini, B.A. & Perrin, J.M. 2012. Decision analysis, economic evaluation, and newborn screening: challenges and opportunities. *Genetics in medicine*, 10(1038):1-18.

Pugh, M.B., Werner, B., Filardo, T.W., Binns, P.W., Francis, L.G., Lukens, R. & Montgomery, B. 2000. Stedman's medical dictionary. 27th ed. Baltimore, MD: Lippincott Williams & Wilkins.

Quartey, G., Wang, J. & Kim, J. 2011. A review of risk measures in pharmacoepidemiology with tips for statisticians in the pharmaceutical industry. *Pharmaceutical statistics*, 10(6):548-553.

Qureshi, Z., Seoane-Vazquez, E., Rodriguez-Monguio, R., Stevenson, K.B. & Szeinbach, S.L. 2011. Market withdrawal of new molecular entities approved in the United States from 1980 to 2009. *Pharmacoepidemiology and drug safety*, 20(7):772-777.

Raebel, M.A., Schmittdiel, J., Karter, A.J., Konieczny, J.L. & Steiner, J.F. 2013. Standardising terminology and definitions of medication adherence and persistence in research employing electronic databases. *Medical care*, 51(8 Suppl. 3):S11-S21.

Rascati, K.L. 2014. Essentials of pharmacoeconomics. 2nd ed. Philadelphia, PA: Lippincott Williams & Wilkins.

Ray, W.A., Stein, C.M., Daugherty, J.R., Hall, K., Arbogast, P.G. & Griffin, M.R. 2002. COX-2 selective non-steroidal anti-inflammatory drugs and risks of serious coronary heart disease. *The lancet*, 360(9339):1071-1073.

Reardon, G., Schwartz, G.F. & Mozaffari, E. 2004. Patient persistency with topical ocular hypotensive therapy in a managed care population. *American journal of ophthalmology*, 137(1 Suppl.):S3-S12.

Rehan, H.S. & Bhargava, S. 2015. Medication errors are preventable. *Journal of pharmacovigilance*, 2(5):1-3.

Rich, J.T., Neely, G., Paniello, R.C., Voelker, C.C.J., Nussenbaum, B. & Wang, E.W. 2010. A practical guide to understanding Kaplan-Meier curves. *Otolaryngology-head and neck surgery*, 143(3):331-336.

Richter, A., Meißner, Y., Strangfeld, A. & Zink, A. 2016. Primary and secondary patient data in contrast: the use of observational studies like RABBIT. *Clinical and experimental rheumatology*, 34(5 Suppl. 101):S79-S86.

Rodriguez, L.A.G. & Gutthann, S.P. 1998. Use of the UK General Practice Research Database for pharmacoepidemiology. *British journal of clinical pharmacology*, 45(5): 419-425.

Rosenberg, L., Coogan, P.F. & Palmer, J.R. 2006. Case-control surveillance. (*In* Strom, B.L. & Kimel, S.E., *eds*. Textbook of pharmacoepidemiology. Chichester: Wiley. p. 137-150).

Rosenberg, L., Coogan, P.F. & Palmer, J.R. 2012. Case-control surveillance. (*In* Strom, B.L., Kimmel, S.E. & Hennessy, S., *eds*. Pharmacoepidemiology. 5th ed. Chichester: Wiley. p.287-300).

Rothman, K.J. & Greenland, S. 2008. Cohort studies. (*In* Rothman, K.J., Greenland, S. & Lash, T.L., *eds*. Modern epidemiology. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins. p. 100-127).

Rothman, K.J., Greenland, S. & Lash, T.L. 2008. Types of epidemiologic studies. (*In* Rothman, K.J., Greenland, S. & Lash, T.L., *eds*. Modern epidemiology. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins. p. 87-99).

Roujeau, J.-C., Kelly, S.P., Naldi, L., Rzamy, B., Stern, R.S., Anderson, T., Auquier, A., Bastyi-Garin, S., Correia, O., Locati, F., Mockenhaupt, M., Paoletti, C., Shapiro, S., Shear, N., Schopf, E. & Kaufman, D.W. 1995. Medication use and the risk of Steven-Johnsons syndrome or toxic epidermal necrosis. *The New England journal of medicine,* 333(24):1600-1607.

Ryder, H.F., McDonough, C., Tosteson, A.N.A. & Lurie, J.D. 2009. Decision analysis and cost effectiveness analysis. *Seminars in spine surgery*, 21(4):216-222.

Sabate, E. 2003. Adherence to long-term therapies: evidence for action. Geneva: World Health Organization (WHO).

Sachdeva, P.D. & Patel, B.G. 2010. Drug utilisation studies- scope and future perspectives. *International journal on pharmaceutical and biological research*, 1(1):11-17.

Samaha, A.A., Zouein, F., Gebbawi, M., Fawaz, M., Houjayri, R., Samaha, R., Baydoun, S. & Eid, A.H. 2017. Associations of lifestyle and dietary habits with hyperlipidemia in Lebanon. *Vessel plus*, 1:98-106.

Sandoo, A., Kitas, G.D., Carroll, D. & Van Zanten, J.J.C.S.V. 2012. The role of inflammation and cardiovascular disease risks on microvascular and macrovascular endothelial function in patients with rheumatoid arthritis: a cross-sectional and longitudinal study. *Arthritis research and therapy*, 14(3):1-10.

Sarfraz, M., Sajid, S. & Ashraf, M.A. Prevalence and patter of dyslipidemia in hyperglycemic patients and its associated factors among Pakistani population. *Saudi journal of biological sciences*, 23(6):761-766.

SAS 9.4® (Statistical Analysis System®). 2002-2012. SAS for Windows 9.4. Cary, North Carolina.

Sattler, E.L., Lee, J.S. & Perri, M. 2013. Medication (re)fill adherence measures derived from pharmacy claims data in older Americans: a review of the literature. *Drugs and aging*, 30(6):383–399.

Saurat, J.H., Guerin, A. & Yu, A.P. 2010. High prevalence of potential drug-drug interactions for psoriasis patients prescribed methotrexate or cyclosporine for psoriasis: associated clinical and economic outcomes in real-world practice. *Dermatology*, 220(2):128-137.

Schatz, S.N. & Weber, R.J. 2015. Adverse drug reactions. (*In* Murphy, J.E. & Wun-Len, L.M., *eds.* PSAP 2015 Book 2 CNS/Pharmacy practice: pharmacotherapy self-assessment program. Lenexa, KS: American College of Clinical Pharmacy. p. 5-26).

Schneeweiss, S. & Avorn, J. 2005. A review of uses of health care utilisation databases for epidemiologic research on therapeutics. *Journal of clinical epidemiology*, 58(4):323-337.

Schneeweiss, S. & Suissa, S. 2013. Advanced approaches to controlling confounding in pharmacoepidemiologic studies. (*In* Strom, B.L., Kimmel, S.E. & Hennessy, S., *eds*. Textbook of pharmacoepidemiology. 2nd ed. Chichester: Wiley. p. 324-336).

Schulman, K.A., Glick, H.A., Polsky, D. & Reed, S.D. 2013. Pharmacoeconomics: economic evaluation of pharmaceuticals. (*In* Strom, B.L., Kimel, S.E. & Hennessy, S., *eds*. Textbook of pharmacoepidemiology. 2nd ed. Chichester: Wiley. p. 280-290).

Schulman, K.A., Glick, H.A., Polsky, D., John, K.R. & Eisenberg, J.M. 2001. Pharmacoeconomics. (*In* Van Boxtel, C.J., Santoso, B. & Edwards, I.R., *eds*. Drug benefits and risks-International textbook for clinical pharmacology. Chichester: Wiley. p. 37-53).

Shadish, W.R., Cook, T.D. & Campbell, D.T. 2002. Experimental and quasi-experimental designs for generalised causal inference. Boston, MA: Houghton.

Shah, A. & Afzal, M. Prevalence of diabetes and hypertension and association with various risk factors among different Muslim populations of Manipur, India. *Journal of diabetes and metabolic disorders,* 13(58):1-6.

Shah, A.D., Langenberg, C., Rapsomaniki, E., Denaxas, S., Pujades-Rodriguez, M., Gale, C.P., Deanfield, J., Smeeth, L., Timmis A. & Hemingway, H. 2015. Type 2 diabetes and incidence of cardiovascular diseases: a cohort study in 1.9 million people. *The lancet*, 3(2):105-113.

Shakir, S.A.W. 2006. Prescription-event monitoring. (*In* Strom, B.L. & Kimel, S.E., *eds*. Textbook of pharmacoepidemiology. Chichester: Wiley. p. 151-167).

Shalini, S., Ravichandran, V., Mohanty, B.K., Dhanaraj, S.K. & Saraswathi, R. 2010. Drug utilisation studies: an overview. *International journal of pharmaceutical sciences and nanotechnology*, 3(1):803-810.

Sharpe, C.R., Collet, J.P., McNutt, M., Belzille, E., Boivin, J.F. & Hanley, J.A. 2000. Nested case-control study of the effects of non-steroidal anti-inflammatory drugs on breast cancer risk and stage. *British journal of cancer*, 83(1):112-120.

Shimizu, F., Igarashi, A. Fukuda, T., Kawachi, Y., Minowada, S., Ohashi, Y. & Fujime, M. 2007. Decision analyses in consideration of treatment strategies for patients with biochemical failure after curative therapy on clinically localized prostate cancer in the prostate-specific antigen era. *Japanese journal of clinical oncology*, 37(10):763–774.

Silverberg, J.I. 2015. Study designs in dermatology: a review for the clinical dermatologist. *Journal of the American Academy of Dermatology*, 73(5):721-731.

Silverstein, M.D., Heit, J.A., Mohr, D.N., Petterson, T.M., O'Fallon, W. & Melton, L. 1998. Trends in the incidence of deep vein thrombosis and pulmonary embolism: a 25-year population-based study. *Archives of internal medicine*, 158(6):585-593.

Simpson, S.H., Corabian, P., Jacobs, P. & Johnson, J. 2003. The cost of major comorbidity in people with diabetes mellitus. *Canadian Medical Association journal*, 168(13):1661-1667.

Singh, B. & Ghatala, M.H. 2012. Risk management in hospitals. International journal of innovation, management and technology, 3(4):417-21.

Singh, R. & Mukhopadhyay, K. 2011. Survival analysis in clinical trials: basics and must know areas. *Perspectives in clinical research*, 2(4):145-148.

Sjöqvist, F. & Birkett, D. 2004. Drug utilization. (*In* Souich, P.D., Orme, M., Erill, S., *eds*. The IUPHAR Compendium of basic principles for pharmacological research in humans. Irvine, CA: International Union of Basic and Clinical Pharmacology (IUPHAR). p. 76-84).

Slabbert, F.N., Harvey, B.H., Brink, C.B. & Lubbe, M.S. 2015. The impact of HIV/AIDS on compliance with antidepressant treatment in major depressive disorder: A prospective study in a South African private healthcare cohort. *AIDS Research and therapy*, 12(9):1-10.

Slabbert, F.N., Harvey, B.H., Brink, C.B. & Lubbe, M.S. 2015. Prospective analysis of the medicine possession ratio of antidepressants in the private health sector of South Africa, 2006 - 2011. *South African medical journal*, 105(2):139-44.

Snyman, J.R. 2017. Monthly Index of Medical Specialities. Cape Town:CTP Printers.

South Africa. 1998. Medical Schemes Act 131 of 1998.

Sowers, J.R., Epstein, M. & Frohlich, E.D. 2001. Diabetes, hypertension and cardiovascular disease: an update. *Hypertension*, 37(4):1053-1059.

Spitzer, W. 1999. Pharmacoepidemiology: barriers and bridges, tribute to Irene Guggenmoos-Holzmann. *Journal of clinical epidemiology*, 52(4):351-354.

Spruance, S.L., Reid, J.E. Grace, M. & Samore, M. 2004. Hazard ratio in clinical trials. *Antimicrobial agents and chemotherapy*, 48(8):2787-2792.

Steiner, J.F. & Prochazka, A.V. 1997. The assessment of refill compliance using pharmacy records: methods, validity, and applications. *Journal of clinical epidemiology*, 50(1):105-116.

Stergachis, A., Hazlet, T.K. & Boudreau, D. 2008. Pharmacoepidemiology. (*In* DiPiro, J.T., Talbert, R.L., Yee, G.C., Matzke, G.R., Wells, B.G. & Posey, L., *eds*. Pharmacotherapy: a pathophysiological approach. 7th ed. New York, NY: McGraw Hill. p. 67-68).

Steyn, H.S. 2009. Manual for the determination of effect size indices and practical significance. Potchefstroom: NWU.

Strom, B.L. 2013a. What is pharmacoepidemiology? (*In* Strom, B.L., Kimmel, S.E. & Hennessy, S., *eds*. Textbook of pharmacoepidemiology. 2nd ed. Chichester: Wiley. p. 3-16).

Strom, B.L. 2013b. When should one perform pharmacoepidemiologic studies? (*In* Strom, B.L., Kimmel, S.E. & Hennessy, S., *eds*. Textbook of pharmacoepidemiology. 2nd ed. Chichester: Wiley. p.54-62).

Strom, B.L. 2013c. Overview of automated databases in pharmacoepidemiology. (*In* Strom, B.L., Kimmel, S.E. & Hennessy, S., *eds*. Textbook of pharmacoepidemiology. 2nd ed. Chichester: Wiley. p. 118-122).

Strom, B.L. 2013d. Study designs available for pharmacoepidemiologic studies. (*In* Strom, B.L., Kimmel, S.E. & Hennessy, S., *eds*. Textbook of pharmacoepidemiology. 2nd ed. Chichester: Wiley. p. 17-29).

Strom, B.L. 2013e. How should one perform pharmacoepidemiologic studies? Choosing among the available alternatives. (*In* Strom, B.L., Kimmel, S.E. & Hennessy, S., *eds*. Textbook of Pharmacoepidemiology. 2nd ed. Chichester: Wiley. p.190-202).

Strom, B.L., Kimmel, S. & Hennessy, S. 2013. The future of pharmacoepidemiology. (*In* Strom, B.L., Kimmel, S.E. & Hennessy, S., *eds*. Textbook of pharmacoepidemiology. 2nd ed. Chichester: Wiley. p. 407-416).

Sturmer, T. & Brookhart, M.A. 2013. Study design considerations. (*In* Velentgas, P., Dreyer, N.A. & Nourjah, P., *eds*. Developing a protocol for observational comparative effectiveness research: A user's guide. AHRQ Publication 12(13)-EHC099. Rockville, MD: Agency for Healthcare Research and Quality. p. 21-34).

Suissa, S. 2009. Statistical methods in pharmacoepidemiology: advances and challenges *Statistical methods in medical research*, 18(1):3-6.

Suissa, S., Dell'Aniello, S & Martinez, C. 2010. The multitime case-control design for time-varying exposures. *Epidemiology*, 21(6):876-883.

Sullivan, L. 2016. Survival analysis. http://sphweb.bumc.bu.edu/otlt/MPH-Modules/BS/BS704_Survival/BS704_Survival_print.html. Date of access: 27 Oct. 2017.

Sultana, J., Cutroneo, P.&Trifirò, G. 2013. Clinical and economic burden of adverse drug reactions. *Journal of pharmacology and pharmacotherapeutics*, 4(1):73-77.

Suruki, R.Y. & Chan, K.A. 2008. Basic pharmacoepidemiology methods. (*In* Hartzema A.G., Tilson, H.H. & Chan, K.A., *eds*. Pharmacoepidemiology and therapeutic risk management. Cincinnati, OH: Harvey Whitney. p. 219-238).

Svarstad, B.L., Shireman, T.I. & Sweeney, J.K. 2001. Using drug claims data to assess the relationship of medication adherence with hospitalisation and costs. *Psychiatric services*, 52(6):805-811.

Swanepoel, J.W.H., Swanepoel, C.J., Van Graan, F.C., Allison, J.S. & Santana, L. 2010. Elementary statistical methods. 3rd ed. Potchefstroom: AndCork.

Swinscow, T.D.V. & Campbell, M.J. 2002. Statistics at square one. 10th ed. London: BMJ Books.

Szklo, M. & Nieto, F.J. 2007. Epidemiology beyond the basics. Sudbury, MA: Jones & Bartlett.

Szumilas, M. 2010. Explaining odds ratios. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 19(3):227-229.

Takahashi, Y., Nishida, Y. & Asai, S. 2012. Utilisation of health care databases for pharmacoepidemiology. *European journal of clinical pharmacology*, 68(2):123-129.

Tanaka, S., Seto, K. & Kwakami, K. 2015. Pharmacoepidemiology in Japan: medical databases and research achievements. *Journal of pharmaceutical health care and sciences*, 1(16):1-4.

Tang, K.L., Quan, H. & Rabi, D.M. 2017. Measuring medication adherence in patients with incident hypertension: a retrospective cohort study. *British medical journal health services research*, 17(1):1-16.

Thaker, S.J., Gogtay, N.J. & Thatte, U.M. 2015. Pharmacoepidemiology: the essentials. *Clinical epidemiology and global health*, 3(2):52-57.

Thompson, C.B. & Panacek, E.A. 2006. Research study designs: experimental and quasi-experimental. *Air medical journal*, 25(6):242-246.

Tolley, H.D., Barnes, J.M. & Freeman, M.D. 2016. Survival analysis. (*In* Freeman, M.D. & Zeegars, M.P., *eds*. Forensic epidemiology: principles and practice. San Diego, CA: Academic Press. p. 261-284).

Torre, C. & Martins, A.P. 2012. Overview of pharmacoepidemiological databases in the assessment of medicines under real-life conditions. (*In* Lunet, N., *ed*. Epidemiology: current perspectives on research and practice. Rijeka: InTech. p. 130-154).

Truter, I. 2008. A review of drug utilisation studies and methodologies. *Jordan journal of pharmaceutical sciences*, 1(2):91-104.

Tsiropoulos, I., Andersen, M. & Hallas, J. 2009. Adverse events with use of antiepileptic drugs: a prescription and event symmetry analysis. *Pharmacoepidemiology and drug safety*, 18(6):483-491.

Tzimis, L. 2001. Ethical issues arising in pharmacoepidemiologic research in a community hospital in Crete, Greece. *Pharmacoepidemiology and drug safety*, 10(7):631-633.

UN (United Nations). 2004. Ethical issues in drug epidemiology: issues, principles and guidelines. https://www.unodc.org/documents/GAP/GAP%20toolkit%20module%207%20ENGLISH%2003-89740_ebook.pdf. Date of access: 14 Aug. 2017.

Unadike, B.C., Eregie, A. & Ohwovoriole, A.E. 2011. Prevalence of hypertension amongst persons with diabetes mellitus in Benin City, Nigeria. *Nigerian Journal of clinical practice*, 14(3):300-302.

Urquhart, J. 2001. Some economic consequences of noncompliance. *Current hypertension reports*, 3(6):473-480.

Urquhart, J. & Vrijens, B. 2006. Patient adherence to prescribed drug dosing regimens in ambulatory pharmacotherapy. (*In* Strom, B.L. & Kimmel, S.E., *eds*. Textbook of pharmacoepidemiology. Chichester: Wiley. p. 368-382).

University of Washington. 2016. Data resources in the health sciences. http://guides.lib.uw.edu/hsl/data/findclin. Date of access: 29 Mar. 2017.

Van Boven, J.F.M., de Jong-Van den Berg, L.T.W. & Vegter, S. 2013. Inhaled corticosteroids and the occurrence of oral candidiasis: a prescription sequence symmetry analysis. *Drug safety*, 36(4):231-236.

Van den Bemt, P., Egberts, T. & Brouwers, J.R.B.J. 2000. Drug-related problems in hospitalised patients. *Drug safety*, 22(4):321-333.

Van der Klauw, M.M., Goudsmit, R., Halie, M.R., van't Veer M.B., Herings, R.M., Wilson, J.H. & Stricker, B.H. 1999. A population-based case-cohort study of drug-associated agranulocytosis. *Archives of internal medicine*, 159(4):369-374.

Van der Linden, P.D., Sturkenboom, M.C.J.M., Herings, R.M.C., Leufkens, H.M.G., Rowlands, S. & Stricker, B.H.C. 2003. Increased risk of achilles tendon rapture with quinolone antibacterial use, especially in elderly patients taking oral corticosteroids. *Archives of internal medicine*, 163(15):1801-1807.

Van der Zwaard, B.C., Van Hout, W., Hugtenburg, J.G., Van der Horst, H.E. & Elders, P.J.M. 2017. Adherence and persistence of patients using oral bone sparing drugs in primary care. *Family practice*, 34(5):525-531.

Van Staa, T.P., Leufkens, H.G. & Cooper, C. 2002. The epidemiology of corticosteroid-induced osteoporosis: a meta-analysis. *Osteoporosis international*, 13(10):777-787.

Varanasi, V.K. 2012. Pharmacoepidemiology versus clinical pharmacology and epidemiology. *Advances in pharmacoepidemiology and drug safety*, 1(2):1-2.

Verhamme, K. & Sturkenboom, M. 2010. Study designs in paediatric pharmacoepidemiology. *European journal of clinical pharmacology*, 67(Suppl. 1):S67-S74.

Vik, S.A., Maxwell, C.J. & Hogan, D.B. 2004. Measurement, correlates, and health outcomes of medication adherence among seniors. *The annals of pharmacotherapy*, 38(2):303-312.

Vink, N.M., Klungel, O.H., Stolk, R.P. & Denig, P. 2009. Comparison of various measures for assessing medication refill adherence using prescription data. *Pharmacoepidemiology and drug safety*, 18():159-165.

Virnig, B.A., Ash, A., Kind, S. & Mesler, D.E. 2000. Survival analysis using medicare data: example and methods. *Health services research*, 35(5 Pt. 3):86-101.

Vollmer, W.M., Xu, M., Feldstein, A., Smith, D., Waterbury, A. & Rand, C. 2012. Comparison of pharmacy-based measures of medication adherence. *British medical journal health services research*, 12(155):1-8. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3413584/pdf/1472-6963-12-155.pdf Date of access: 31 May 2018.

Vrijens, B., Geest, S.D., Hughes, D.A., Przemyslaw, K., Demonceau, J., Ruppar, T., Dobbels, F., Fargher, E., Morrison, V., Lewek, P., Matyjaszczyk, M., Mshelia, C., Clyne, W., Aronson, J.K. & Urquhart, J. 2012. A new taxonomy for describing and defining adherence to medications. *British journal of clinical pharmacology*, 73(5):691-705.

Wagner, A.K., Soumerai, S.B., Zhang, M.S. & Ross-Degnan, D. 2002. Segmented regression analysis of interrupted time series studies in medication use research. *Journal of clinical pharmacy and therapeutics*, 27(4):299-309.

Wahab, I.A., Pratt, N.L., Ellett, L.K. & Roughead, E.E. 2016. Sequence symmetry analysis as a signal detection tool for potential heart failure adverse events in an administrative claims database. *Drug safety*, 39:347-354.

Wahab, I.A., Pratt, N., Kalisch, L.M. & Roughead, E.E. 2013. Sequence symmetry analysis and disproportionality analyses: what percentage of adverse drug reaction do they signal? *Advances in pharmacoepidemiology & drug safety*, 2(4):1-5.

Wahab, I.A., Pratt, N.L., Wiese, M.D., Kalisch, L.M. & Roughead, E.E. 2013. The validity of sequence symmetry analysis (SSA) for adverse drug reaction signal detection. Pharmacoepidemiology and drug safety, 22(5):496-502.

Waller, P. 2001. Pharmacoepidemiology: a tool for public health. *Pharmacoepidemiology and drug safety*, 10(2):165-172.

Walsh, P.C. 2005. Effects of long-term vitamin E supplementation on cardiovascular events and cancer: a randomised controlled trial. *The journal of urology*, 174(5):1823-1824.

Wang, S., Xu, L., Jonas, J.B., You, Q.S., Wang, Y.X. & Yang, H. 2011. Prevalence and associated factors of dyslipidemia in the adult Chinese population. *PLoS One*, 6(3):e17326.

Waning, B. & Montagne, M. 2001. Pharmacoepidemiology principles and practice. New York, NY: McGraw-Hill.

Watkins, P.J. 2003. Cardiovascular disease, hypertension and lipids. *BMJ (Clinical research ed.)*, 326(7394):874-879.

Wells, B.G., DiPiro, J.T., Schwinghammer, T.L. & DiPiro, C.V. 2015. Pharmacotherapy handbook. 9th ed. Manhattan, NY: McGraw-Hill.

West-Strum, D. 2011. Introduction to pharmacoepidemiology. (*In* Yang, Y. & West-Strum, D., *eds*. Understanding pharmacoepidemiology. Manhattan, NY: McGraw-Hill. p. 1-14).

Wettermark, B. 2013. The intriguing future of pharmacoepidemiology. *European journal of clinical pharmacology*, 69(1):43-51.

Wettermark, B., Vlahovic-Palcevski, V., Salvesen Blix, H., Rønning, M. & Vander Stichele, R.H. 2008. Drug utilization research. (*In* Hartzema A.G., Tilson, H.H., Chan, K.A., *eds*. Pharmacoepidemiology and therapeutic risk management. Cincinnati, OH: Harvey Whitney. p. 159-174).

Weycker, D., Nichols, G.A., O'Keefe-Rosetti, M., Edelsberg, J., Khan, Z.M., Kaura, S. & Oster, S. 2007. Risk factor clustering and cardiovascular disease risk in hypertensive patients. *American journal of hypertension*, 20:599-607.

Whitaker, H.J., Hocine, M.N. & Farrington, C.P. 2009. The methodology of self-controlled case series studies. *Statistical methods in medical research*, 18(1):7-26.

White, H. & Sabarwal, S. 2014. Quasi-experimental design and methods. Florence, Italy: United Nations Children's Emergency Fund (UNICEF).

WHO (World Health Organization). 2018. Cardiovascular diseases. Oslo. https://www.who.int/cardiovascular_diseases/en/. Accessed 22 Oct. 2018.

WHO (World Health Organization). 2016. Global report on diabetes. Oslo. http://apps.who.int/iris/bitstream/handle/10665/204871/9789241565257_eng.pdf;jsessionid=B42038D0F9D4A6F073EFD0F7E2ECC7E2?sequence=1. Date of access: 15 Jul. 2017.

WHO (World Health Organization). 2003. Introduction to drug utilization studies. Oslo. http://www.who.int/medicine/areas/quality_safety/safety_efficacy/Dru%20utilisation%research.pdf. Date of access: 2 Mar. 2017.

WHO (World Health Organization). 2017b. Women, gender and health. Oslo. apps.who.int/gender/whatisgender/en/. Date of access: 29 Mar. 2017.

Williams, D. 2012. Monitoring medicines use: the role of the clinical pharmacologist. *British journal of clinical pharmacology*, 74(4):685-690.

Wise, L. 2011. Risks and benefits of (pharmaco)epidemiology. *Therapeutic advance in drug safety*, 2(3):95-102.

Wolozin, B., Kellman, W., Rousseau, P., Celesia, G.G. & Siegel, G. 2000. Decreased prevalence of alzheimer disease association with 3-hydroxy-3-methyglutaryl coenzyme A reductase inhibitors. *Archives of neurology*, 57(10):1439-1443.

Wong, N.D., Lopez, V, Tang, S. & Williams, G.R. 2006. Prevalence, treatment, and control of combined hypertension and hypercholesterolemia in the United States. *The American journal of cardiology*, 98(2):204-206.

Wu, J.Y., Duan, X.Y., Li, L., Dai, F., Li, Y.Y., Li, X.J. & Fan, J.G. 2010. Dyslipidemia in Shanghai, China. *Preventive medicine*, 51(5):412-5.

Yasmina, A., Deneer, V.H. M., Maitlaand-van der Zer, A.H., Van Staa, T.P., De Boer, A. & Klungel, O.A. 2014. Application of routine electronic health record databases for pharmacogenetic research. *Journal of internal medicine*, 275(6):590- 604.

Yang, Y. 2011. Principles of epidemiology applied to the study of medication use. (*In* Yang, Y. & West-Strum, D., *eds*. Understanding pharmacoepidemiology. Manhattan, NY: McGraw-Hill. p. 15-37).

YHEC (York Health Economics Consortium). 2016a. Cost minimisation analysis. http://www.yhec.co.uk/glossary/cost-minimisation-analysis/. Date of access: 2 Sep. 2017.

YHEC (York Health Economics Consortium). 2016b. Cost benefit analysis. http://www.yhec.co.uk/glossary/cost-benefit-analysis/. Date of access: 2 Sep. 2017.

Yki-Järvinen, H. 2003. Nonglycemic effects of insulin. *Clinical cornerstone*, Suppl 4:S6-12.

Zornberg, G.L. & Jick, H. 2000. Antipsychotic drug use and risk of first-time idiopathic venous thromboembolism: a case-control study. *The lancet*; 356(9237):1219-1223.

Zhou MS, Wang A, Yu H. 2014. Link between insulin resistance and hypertension: What is the evidence from evolutionary biology? *Diabetology and metabolic syndrome*, 6(1):12.

Zwiener, I., Blettner, M. & Hommel, G. 2011. Survival analysis. *Deutsches ärzteblatt international*, 108(10):163-169.

# ANNEXURE A: CERTIFICATE OF ETHICAL APPROVAL

NORTH-WEST UNIVERSITY
YUNIBESITI YA BOKONE-BOPHIRIMA
NOORDWES-UNIVERSITEIT

Private Bag X6001, Potchefstroom,
South Africa, 2520

Tel: (018) 299-4900
Faks: (018) 299-4910
Web: http://www.nwu.ac.za

**Research Ethics Regulatory Committee**
Tel: +27 18 299 4849
Email : Ethics@nwu.ac.za

**ETHICS APPROVAL CERTIFICATE OF STUDY**

Based on approval by **Health Research Ethics Committee (HREC)** on **11/10/2017**, the North-West University Research Ethics Regulatory Committee (NWU-RERC) hereby **approves** your study as indicated below. This implies that the NWU-RERC grants its permission that provided the special conditions specified below are met and pending any other authorisation that may be necessary, the study may be initiated, using the ethics number below.

**Study title: Research methods for conducting pharmacoepidemiological studies using medicines claims data**

**Study Leader/Supervisor: Prof JR Burger**
**Student: M Obeng-Kusi (27959716)**

**Ethics number**:

| N | W | U | - | 0 | 0 | 1 | 7 | 9 | - | 1 | 4 | - | A | 1 | - | 0 | 6 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Institution | | | | Study Number | | | | | | Year | | | Status | | | | |

Status: S = Submission; R = Re-Submission; P = Provisional Authorisation; A = Authorisation

**Application Type: Sub-study**
**Commencement date**: 11/10/2017 **Risk:** **Minimal**

**Approval of the study is initially provided for a year, after which continuation of the study is dependent on receipt of the annual (or as otherwise stipulated) monitoring report and the concomitant issuing of a letter of continuation.**

**General conditions:**

*While this ethics approval is subject to all declarations, undertakings and agreements incorporated and signed in the application form, please note the following:*

- *The study leader (principle investigator) must report in the prescribed format to the NWU-IRERC via HREC:*
  - *annually (or as otherwise requested) on the monitoring of the study, and upon completion of the study*
  - *without any delay in case of any adverse event or incident (or any matter that interrupts sound ethical principles) during the course of the study.*
- *Annually a number of studies may be randomly selected for an external audit.*
- *The approval applies strictly to the proposal as stipulated in the application form. Should any changes to the proposal be deemed necessary during the course of the study, the study leader must apply for approval of these amendments at the HREC, prior to implementation. Should there be any deviations from the study proposal without the necessary approval of such amendments, the ethics approval is immediately and automatically forfeited.*
- *The date of approval indicates the first date that the study may be started.*
- *In the interest of ethical responsibility the NWU-IRERC and HREC retains the right to:*
  - *request access to any information or data at any time during the course or after completion of the study;*
  - *to ask further questions, seek additional information, require further modification or monitor the conduct of your research or the informed consent process.*
  - *withdraw or postpone approval if:*
    - *any unethical principles or practices of the study are revealed or suspected,*
    - *it becomes apparent that any relevant information was withheld from the HREC or that information has been false or misrepresented,*
    - *the required amendments, annual (or otherwise stipulated) report and reporting of adverse events or incidents was not done in a timely manner and accurately,*
    - *new institutional rules, national legislation or international conventions deem it necessary.*
- *HREC can be contacted for further information or any report templates via Ethics-HRECApply@nwu.ac.za or 018 299 1206.*

The IRERC would like to remain at your service as scientist and researcher, and wishes you well with your study. Please do not hesitate to contact the IRERC or HREC for any further enquiries or requests for assistance.

Yours sincerely,

**Prof. Refilwe Phaswana-Mafuya**
*Chair NWU Research Ethics Regulatory Committee (RERC)*

# ANNEXURE B: EXAMPLES OF APPLICATION OF METHODS FOR PHARMACOEPIDEMIOLOGICAL STUDIES

| Measure | Database | Description of database | Study design | Statistical analyses | Reference |
|---|---|---|---|---|---|
| Adverse drug reactions | Prescription database | This contains pharmacy claims for approximately 90 million patients nationwide | Cohort | Logistic regression | Hersom *et al*. (2003) |
| | Odense University Pharmacoepidemiological Database | The database contains a complete list of all reimbursed prescriptions in Odense | Case-control | Unconditional logistic regression | Gulmez *et al*. (2007) |
| | General Practice Research Database | Contains computerized medical information of approximately 8 million inhabitants in the UK | Case-control | Multivariate logistic regression | Van der Linden *et al*. (2003) |
| | General Practice Research Database | Contains computerized medical information of approximately 8 million inhabitants in the UK | Case-control | Conditional logistic regression | Zornberg & Jick (2000) |
| | General Practice Research Database | Contains computerized medical information of approximately 8 million inhabitants in the UK | Nested case-control | Kaplan-Meier survival analysis | Jick *et al*. (1998) |

| Measure | Database | Description of database | Study design | Statistical analyses | Reference |
|---|---|---|---|---|---|
| Drug-drug interactions | IMS Lifelink Health Plan Administrative Database | Contains commercial health plan information for 98 health care plans across USA for about 61 million patients | Cohort | Cox regression | Bhurke *et al*. (2012) |
| | Ingenix Impact National Managed Care Database | General medical and pharmacy history of 60 million patients from 46 health plans from the USA. It contains inpatient information, medical services utilization and pharmacy claims | Cohort | Wilcoxon rank sum test<br>Multivariate regression<br>Linear model regression | Saurat *et al*. (2010) |
| | Ontario Public Drug Program<br><br>Canadian Institute for Health Information Discharge Abstract Database | Has comprehensive records of prescription medications dispensed to residents.<br><br>Has detailed diagnostic and procedural information about hospital admission.<br><br>Identifies claims for inpatients and | Nested case-control | Logistic regression | Juurlink *et al*. (2009) |

| Measure | Database | Description of database | Study design | Statistical analyses | Reference |
|---|---|---|---|---|---|
| | Ontario Health Insurance Plan Database<br><br>Ontario Registered Persons' Database | outpatient physician services.<br><br>Contains basic demographic information such of date of death, age, etc. | | | |
| | General Practice Research Database | Contains computerized medical information of approximately 8 million inhabitants in the UK | Case-control | Conditional logistic regression | Delaney *et al*. (2007) |
| | Remote Electronic Claim Adjudication (RECAP) system | Contains prescription claims and eligibility records for 54000 pharmacy and mail-service facilities | Cross-sectional | Chi-square analysis | Malone *et al*. (2005). |
| | Ontario Public Drug Program<br><br>Canadian Institute for Health Information Discharge Abstract Database | Has comprehensive records of prescription medications dispensed to residents.<br><br>Has detailed diagnostic and procedural information about hospital admission. | Nested case-control | Conditional logistic regression | Juurlink *et al*. (2003) |

| Measure | Database | Description of database | Study design | Statistical analyses | Reference |
|---|---|---|---|---|---|
| | Ontario Health Insurance Plan Database<br><br>Ontario Registered Persons' Database | Identifies claims for inpatients and outpatient physician services.<br><br>Contains basic demographic information such of date of death, age, etc. | | | |
| Prevalence | General Practice Research Database | Contains computerized medical information of approximately 8 million inhabitants in the UK | Cohort | Logistic regression analysis<br>Chi-square tests | McCarthy *et al*. (2009) |
| | General Practice Research Database | Contains computerized medical information of approximately 8 million inhabitants in the UK | Cross-sectional | Conditional logistic regression | Neimann *et al*. (2006) |
| | General Practice Research Database | Contains computerized medical information of approximately 8 million inhabitants in the UK | Cross-sectional | Descriptive statistics | Gelfand *et al*. (2005) |
| | Clinical database patients who received care within Kaiser Permanente | Contains diagnoses and pharmacy claims for all patients who received care within | Cross-sectional | Chi-square | Go *et al*. (2001) |

| Measure | Database | Description of database | Study design | Statistical analyses | Reference |
|---|---|---|---|---|---|
| | | the Kaiser Permanente, a large health maintenance organization | | | |
| | Relational databases from, Edward Hines Jr Veterans Affairs Hospital, Loyola University and Carl T. Hayden Veterans Affairs Medical Centers | These contain patients' medical records, diagnoses and medication history | Cross-sectional | Chi-square | Wolozin *et al*. (2000) |
| Incidence | Régie de l'Assurance Maladie du Québec (RAMQ) | Contains physician claims, admission and discharge information of patients | Cohort | McNemar matched pair chi-square test<br>Kappa statistics | Asghari *et al*. (2009) |
| | Ontario Diabetes database | Contains diagnoses, patient demography and medication claims | Cohort | Use of algorithms | Hux *et al*. (2002) |
| | Surveillance, Epidemiology and End Results (SEER) Program | Medical data of persons who are 65 years and older and have been diagnosed with any form of cancer | Case-control | Logistic regression | Freeman *et al*. (2000) |
| | Rochester Epidemiology Project | Contains details of inpatient and outpatient information and medical care provided to residents of Olmsted and Rochester | Cohort | Poisson regression | Silverstein *et al*. (1998) |

| Measure | Database | Description of database | Study design | Statistical analyses | Reference |
|---|---|---|---|---|---|
| Adherence | Prescription data from the Groningen Initiative to Analyze type 2 diabetes treatment (GIANTT) | Contains full prescriptions recorded in electronic medical records by engaged physicians | Cohort | Receiver operating characteristic (ROC) curve | Vink *et al*. (2009) |
| | Cancer registry in Tayside | Contains diagnoses records, hospital admissions and prescription information of cancer patients in Tayside | Cohort | Shapiro-Wilks test<br>Cox proportional hazards models<br>Propensity scoring | McCowan et al. (2008) |
| | National Managed Care Benchmark Database | Contains administrative data of benefeciaries enrolled in 30 employer-sponsored managed care plans across 7 US census areas | Cohort | Weibull parametric survival model<br>Logistic regression | Hertz *et al*. (2005) |
| | Medicines Monitoring Unit (MEMO) | Contains information on 18 million dispensed prescriptions in Tayside as well as demographic data and diagnostic information on all patients admitted to hospitals | Cohort | Multiple logistic regression | Donnan *et al*. (2002) |

| Measure | Database | Description of database | Study design | Statistical analyses | Reference |
|---|---|---|---|---|---|
| | Ontario Drug Benefit (ODB) prescription claims database<br><br>Canadian Institutes of Health Information (CIHI) hospital discharge abstract database<br><br>Ontario Health Insurance Plan (OHIP) database<br><br>Ontario Registered Persons Database (RPDB) | Has comprehensive records of prescription medications dispensed to Ontario residents.<br><br>Has detailed diagnostic and procedural information about hospital admission.<br><br>Identifies claims for inpatients and outpatient physician services.<br><br>Contains basic demographic information such of date of death, age, etc. | Cohort | Chi-square analysis<br>Kruskal-Wallis test<br>Cox proportional hazards model | Jackevicius *et al*. (2002) |
| Compliance | New Jersey Medicaid Program | Identifies persons eligible to receive Medicaid benefits, dates of coverage, demographical information, information on nursing home residency and death | Cohort | Logistic regression | Monane *et al*. (1996) |

| Measure | Database | Description of database | Study design | Statistical analyses | Reference |
|---|---|---|---|---|---|
| Persistence | Ingenix LabRx administrative claims database | Contains data from medical, hospital and pharmacy claims submitted to a large, single insurance company | Cohort | Linear regression | Reardon *et al*. (2004) |
| | Rochester Epidemiology Project | Contains details of inpatient and outpatient information and medical care provided to residents of Olmsted and Rochester | Nested case-control | Conditional logistic regression | Heit *et al*. (2000) |

# ANNEXURE C: AUTHOR GUIDELINES FOR JOURNAL OF CLINCAL PHARMACY AND THERAPEUTICS

Journal of Clinical Pharmacy and Therapeutics

**Author Guidelines**

**1. General**
The *Journal of Clinical Pharmacy and Therapeutics (JCPT)* provides a forum for clinicians, pharmacists and pharmacologists to explore and report on issues of common interest. It welcomes five main types of articles

- Editorials
- Original research
- Review articles (including Mini-reviews)
- Commentaries
- Case reports

As our main interest is on novelty, irrespective of the type of contribution, the sub-headings should identify what is known and what is new. A clear description of these aspects is important as they are used by us to filter submissions at the very first stage. This helps us to return manuscripts quickly to authors for submission elsewhere.

Please read the instructions below carefully for details on the submission of manuscripts, the Journal's requirements and standards as well as information concerning the procedure after a manuscript has been accepted for publication in *JCPT*.

**2. Ethical Guidelines**
*JCPT* has adopted the following ethical guidelines for publication and research.

**2.1 Authorship and Acknowledgements**
**Authorship:** Authors submitting a paper do so on the understanding that the manuscript has been read and approved by all authors and that all authors agree to the submission of the manuscript to the Journal.

*JCPT* adheres to the definition of authorship set up by The International Committee of Medical Journal Editors (ICMJE). According to the ICMJE authorship criteria, all named authors should meet the following conditions: 1) substantial contributions to conception and design of, or acquisition of data or analysis and interpretation of data, 2) drafting the article or revising it critically for important intellectual content and 3) final approval of the version to be published. Participation solely in the acquisition of funding or the collection of data does not justify authorship. All people who fulfil the criteria for authorship should be listed as authors. Contributors who do not qualify as authors should be mentioned in Acknowledgements.

The Editors recognise that complex, large-scale and multi-centre research will often result in a significant number of people fulfilling the authorship criteria. However, they reserve the right to ask the lead author to justify the inclusion of more than six authors.

**Acknowledgements:** Under Acknowledgements please specify contributors to the research/article other than the authors accredited. Please note that research funders are now listed separately under Source of Funding.

**2.2 Conflict of Interest and Source of Funding**
*JCPT* requires that sources of financial support for the work reported within the manuscript are fully acknowledged, and any potential conflicts of interest noted.

**Conflict of Interest:** All manuscripts submitted to the Journal require a statement about authors' conflicts of interest. Please disclose any possible conflict of interest under the heading 'Conflicts of Interest' on the title page of your manuscript. Any reported conflicts of interest will be published in a highlighted box as part of the article. If no conflicts of interest are reported, the box will include the statement "No conflicts of interest have been declared". Possible conflicts of interest include financial interests relating to issues discussed in the manuscript (e.g. patent ownership, stock ownership, consultancies and speaker's fees).

**Source of Funding:** Authors are required to specify the sources of funding for their research when submitting a manuscript. These include the individuals and organisations that supplied resources for interventions as well as those that funded researcher time and other research costs. All sources of funding should be named and their location (town, state/county, country) included. The information should be provided on the title page of the manuscript and will be disclosed in the published article.

**2.3 Appeal of Editorial Decisions**
The Editors make careful judgements about the selection of manuscripts for publication, taking into account the extent to which the manuscript is consistent with the aims and scope of the Journal and their own and referees' assessments of the quality of the work and the contribution it is likely to make to knowledge, policy and practice. We are able to accept only a proportion of the manuscripts that are submitted to the Journal, and recognise that authors are often disappointed when we decline to publish their manuscripts. We strongly discourage routine appeals against such decisions. Authors who believe there were serious flaws in our editorial judgement may appeal decisions by e-mailing the editorial office with a detailed explanation of their concerns.

**2.4 Permissions**
If all or parts of previously published illustrations are used, permission must be obtained from the copyright holder concerned. It is the author's responsibility to obtain these permissions in writing and provide copies to the Publishers.

**2.5 Copyright Assignment**
If your paper is accepted, the author identified as the formal corresponding author for the paper will receive an email prompting them to login into Author Services; where via the Wiley Author Licensing Service (WALS) they will be able to complete the license agreement on behalf of all authors on the paper.

**For authors signing the copyright transfer agreement**
If the OnlineOpen option is not selected the corresponding author will be presented with the copyright transfer agreement (CTA) to sign. The terms and conditions of the CTA can be previewed in the samples associated with the Copyright FAQs below:
CTA Terms and Conditions http://authorservices.wiley.com/bauthor/faqs_copyright.asp.

**For authors choosing OnlineOpen**
If the OnlineOpen option is selected the corresponding author will have a choice of the following Creative Commons License Open Access Agreements (OAA):
Creative Commons Attribution Non-Commercial License OAA
Creative Commons Attribution Non-Commercial -NoDerivs License OAA

To preview the terms and conditions of these open access agreements please visit the Copyright FAQs hosted on Wiley Author Services
http://authorservices.wiley.com/bauthor/faqs_copyright.asp and visit
http://www.wileyopenaccess.com/details/content/12f25db4c87/Copyright--License.html.

If you select the OnlineOpen option and your research is funded by The Wellcome Trust and members of the Research Councils UK (RCUK) you will be given the opportunity to publish your article under a CC-BY license supporting you in complying with Wellcome Trust and Research Councils UK requirements. For more information on this policy and the Journal's compliant self-archiving policy please visit: http://www.wiley.com/go/funderstatement.

For RCUK and Wellcome Trust authors click on the link below to preview the terms and conditions of this license: Creative Commons Attribution License OAA.

To preview the terms and conditions of these open access agreements please visit the Copyright FAQs hosted on Wiley Author Services
http://authorservices.wiley.com/bauthor/faqs_copyright.asp and visit
http://www.wileyopenaccess.com/details/content/12f25db4c87/Copyright--License.html.

**3. Submission of Manuscripts**
Manuscripts should be submitted electronically via http://mc.manuscriptcentral.com/jcpt. Authors may track the status of their own manuscripts. Complete instructions for submitting papers are available online and a user ID and password can be obtained from the first visit. Further assistance can be obtained from: support@scholarone.com. If you cannot submit online or have a general query, please contact Professor Alain Li Wan Po (Editor-in-Chief) at alainliwanpo@yahoo.com

**Papers do not attract page charges.** OnlineOpen is available to authors of primary research articles who wish to make their article available to non-subscribers on publication, or whose funding agency requires grantees to archive the final version of their article. With OnlineOpen, the author, the author's funding agency, or the author's institution pays a fee (currently $3000) to ensure that the article is made available to non-subscribers upon publication via Wiley Online Library, as well as deposited in the funding agency's preferred archive. For the full list of terms and conditions, see http://olabout.wiley.com/WileyCDA/Section/id-406241.html.

Prior to acceptance you should not inform the Editorial Office that you intend to publish your paper OnlineOpen. All OnlineOpen articles are treated in the same way as any other article. They go through the Journal's standard peer-review process and will be accepted or rejected based on

their own merit.

### 3.1 Graphical abstract

A Graphical abstract is a single, concise, pictorial and visual summary of the main findings of the article. This could either be the concluding figure from the article or a figure that is specially designed for the purpose, which captures the content of the article for readers at a single glance. The Graphical abstract will be displayed in online search result lists, the online contents list and the online article, but will not (yet) appear in the article PDF file or print. Please click **here** to view how the graphical abstract appears on journal home page.

If you would like to have the graphical abstract for your paper, please submit the image and caption in a separate document along with the manuscript to editorial office.

## 4. Manuscripts Types Accepted

**Original research:**Reports in this section should have a structured summary and a main text, both of which must have the following sub-headings: What is known and Objective; Methods; Results and discussion; What is new and Conclusion.

The maximum word-length for reports of original research is 3000 words excluding tables, figures, references and summary. We encourage submission of additional supporting material for online-only publication but this should be clearly identified and labeled as ‘Online appendix A1’ etc. within the text.

**Review articles:** These contributions should have a structured summary and a main text both of which must have the following sub-headings: What is known and Objective; Methods; Results and discussion; What is new and Conclusion. **If your review is not a systematic review, then it should be submitted as a commentary. A mini-review can be submitted either as a commentary or as a systematic review depending on the methodology used.**

The maximum word-length for a Review is 5000 words excluding tables, figures, references and summary. A mini-review is by definition shorter than this but we impose no specific word-length. We encourage submission of additional supporting material for online-only publication but this should be clearly identified and labeled as ‘Online appendix A1’ etc. within the text.

**Commentaries:** A commentary should have:
(i) a structured summary of no more than 150 words with the following subheadings: What is known and Objective; Comment; What is new and Conclusion.
(ii) a main text with the same sub-headings as the summary but with a maximum of 2000 words excluding references.
In both the summary and the main text, the Comment section should make up the bulk of the contribution (> 90%).

**Editorials:** Generally these are contributed by our own Editors to describe specific developments at the Journal but may also include invited contributions from leading experts on highly topical subjects for which the novelty is obvious. These expert contributions may vary considerably in length and style so as to ensure particularly rapid publication.

**Case reports:** A case report should have:
(i) a summary of not more than 100 words
(ii) a main text of not more than 1500 words excluding references.

Both sections should have the following sub-headings: What is known and objective; Case description; What is new and Conclusion. In both sections the case-description should make up the bulk (> 90%) of the contribution. We encourage submission of additional supporting material for online-only publication but this should be clearly identified and labelled as 'Online appendix A1' etc. within the text.

**Letters:** Correspondence is invited. Letters will only be considered if they contain constructive comments on published articles and if they are received in time to allow the authors a right of reply. Publication of correspondence is at the discretion of the Editor.

**5. Manuscript Format and Structure**

**5.1 Format**

**Language:** The language of publication is English. Authors for whom English is a second language should have their manuscript professionally edited by an English speaking person before submission to make sure the English is of high quality. A list of independent suppliers of editing services can be found at http://authorservices.wiley.com/bauthor/english_language.asp. All services are paid for and arranged by the author, and use of one of these services does not guarantee acceptance or preference for publication.

**Units and Spellings:** Système International (SI) units should be used, as given in *Units, Symbols and Abbreviations* (4th edition, 1988), published by the Royal Society of Medicine Services Ltd, 1 Wimpole Street, London W1M 8AE, UK. Other abbreviations should be used sparingly and only if a lengthy name or expression is repeated throughout the text. Spelling should conform to that used in *The Concise Oxford Dictionary*, published by Oxford University Press. Authors should strenuously avoid the use of jargon or obscure technical terms.

The typescript should be on A4 paper on one side only, double spaced with a wide margin on each side. The title and short title (to be printed at the head of alternate pages), authors' names, qualifications and the department(s) where the work was carried out, and the name and full postal address of the author to whom all correspondence should be sent, should be typed on a separate sheet. Please include a telephone, a fax number and an e-mail address.

**5.2 References**

All references should be numbered consecutively in order of appearance and should be as complete as possible. In text citations should cite references in consecutive order using Arabic superscript numerals. Sample references follow:

**Journal article:**

**1**. King VM, Armstrong DM, Apps R, Trott JR. Numerical aspects of pontine, lateral reticular, and inferior olivary projections to two paravermal cortical zones of the cat cerebellum. J Comp Neurol 1998;390:537-551.

**Book:**

2. Voet D, Voet JG. Biochemistry. New York: John Wiley & Sons; 1990. 1223 p.

Please note that journal title abbreviations should conform to the practices of Chemical Abstracts.

For more information about AMA reference style, please click here

**5.3 Figures and Tables**
**Figures:** All graphs, drawings and photographs are considered figures and should be numbered in Arabic numerals e.g. Fig. 1, Fig. 2, etc. in order of appearance. Each figure should have a legend and all legends should be typed together on a separate sheet and numbered correspondingly. If all or parts of previously published illustrations are used, permission must be obtained from the copyright holder concerned. It is the author's responsibility to obtain these in writing and provide copies to the Publisher.

In the full-text online edition of the Journal figure legends may be truncated in abbreviated links to the full screen version. Therefore the first 100 characters of any legend should inform the reader of key aspects of the figure.

**Tables:** Clear tables presenting relevant data are welcome. If tables of important data are particularly lengthy (e.g. tables reporting details of a large number of studies included in a systematic review), the Editors may suggest that some are published as supporting online material. Each table must be typewritten on a separate sheet and should be numbered consecutively with Arabic numerals, e.g. Table 1, and given a short caption. No vertical rules should be used. Units should appear in parentheses in the column headings and not in the body of the table. All abbreviations should be defined in a footnote.

**Electronic Artwork:** We would like to receive your artwork in electronic form. Please save vector graphics (e.g. line artwork) in Encapsulated Postscript format (EPS), and bitmap files (e.g. half-tones) in Tagged Image File Format (TIFF). Ideally, vector graphics that have been saved in metafile (.WMF) or pict (.PCT) format should be embedded within the body of the text file.

Further information can be obtained at Wiley Blackwell's guidelines for figures: http://authorservices.wiley.com/bauthor/illustration.asp.

**Guidelines for Cover Submissions**
If you would like to send suggestions for artwork related to your manuscript to be considered to appear on the cover of the journal, please follow these general guidelines.

**5.4 Colour Artwork**
It is the policy of the JCPT for authors to pay the full cost for the reproduction of their colour artwork. Therefore, please note that if there is colour artwork in your manuscript when it is accepted for publication, Wiley Blackwell require you to complete and return a Colour Work Agreement form before your paper can be published.

Any article received by Wiley Blackwell with colour work will not be published until the form has been returned. If you are unable to access the Internet, or are unable to download the form, please contact the Production Editor for a form. In compliance with the Payment Card Industry Data Security Standard, the Colour Work Agreement Form should be sent by post or courier to the address below, should you require colour:

Customer Services (OPI)
John Wiley & Sons Ltd
European Distribution Centre
New Era Estate

Oldlands Way
Bognor Regis
West Sussex PO22 9NQ8

**5.5 Supporting Information**
Supporting Information can be a useful way for an author to include important but ancillary information with the online version of an article. Examples of Supporting Information include additional tables, data sets, figures, movie files, audio clips, 3D structures, and other related nonessential multimedia files. Supporting Information should be cited within the article text, and a descriptive legend should be included. It is published as supplied by the author, and a proof is not made available prior to publication; for these reasons, authors should provide any Supporting Information in the desired final format.

For further information on recommended file types and requirements for submission, please visit: http://authorservices.wiley.com/bauthor/suppinfo.asp.

**6. After Acceptance**
Upon acceptance of a paper for publication, the manuscript will be forwarded to the Production Editor who is responsible for the production of the Journal.

**Online Production Tracking:** Online production tracking is available for your article through Wiley Blackwell's Author Services. Author Services enables authors to track their article — once it has been accepted — through the production process to publication online and in print. Authors can check the status of their articles online and choose to receive automated e-mails at key stages of production. Authors will receive an e-mail with a unique link that enables them to register and have their article automatically added to the system. Please ensure that a complete e-mail address is provided when submitting the manuscript. Visit http://authorservices.wiley.com/ for more details on online production tracking and for a wealth of resources including FAQs and tips on article preparation, submission and more.

**Proof Corrections:** The corresponding author will receive an e-mail alert containing a link to a website. A working e-mail address must therefore be provided for the corresponding author. The proof can be downloaded as a PDF (portable document format) file from this site.

Acrobat Reader will be required in order to read this file. This software can be downloaded (free of charge) from the following website: www.adobe.com/products/acrobat/readstep2.html. This will enable the file to be opened, read on screen, and printed out in order for any corrections to be added. Further instructions will be sent with the proof. Hard copy proofs will be posted if no e-mail address is available; in your absence, arrange for a colleague to access your e-mail to retrieve the proofs.

**Early View Publication:** *JCPT* is covered by Wiley Blackwell's Early View service. Early View articles are complete full-text articles published online in advance of their publication in a printed issue. Early View articles are complete and final. They have been fully reviewed, revised and edited for publication, and the authors' final corrections have been incorporated. Because they are in final form, no changes can be made after online publication. The nature of Early View articles means that they do not yet have volume, issue or page numbers, so these articles cannot be cited in the traditional way. They are given a Digital Object Identifier (DOI), which allows the article to be cited and tracked before it is allocated to an issue. After print publication, the DOI remains valid and can continue to be used to cite and access the article.

**Search Engine Optimization for Your Paper:** By optimizing your article for search engines, you will increase the chance of someone finding it. This in turn will make it more likely to be viewed and/or cited in another work. Consult our SEO Tips for Authors page in order to maximize online discoverability for your published research. Included are tips for making your title and abstract SEO-friendly, choosing appropriate keywords, and promoting your research through social media.

**Offprints:** Free access to the final PDF offprint of your article will be available via Author Services only. Please therefore sign up for Author Services if you would like to access your article PDF offprint and enjoy the many other benefits the service offers. Additional paper offprints may be ordered online. Please click on this link, fill in the necessary details and ensure that you type information in all of the required fields.

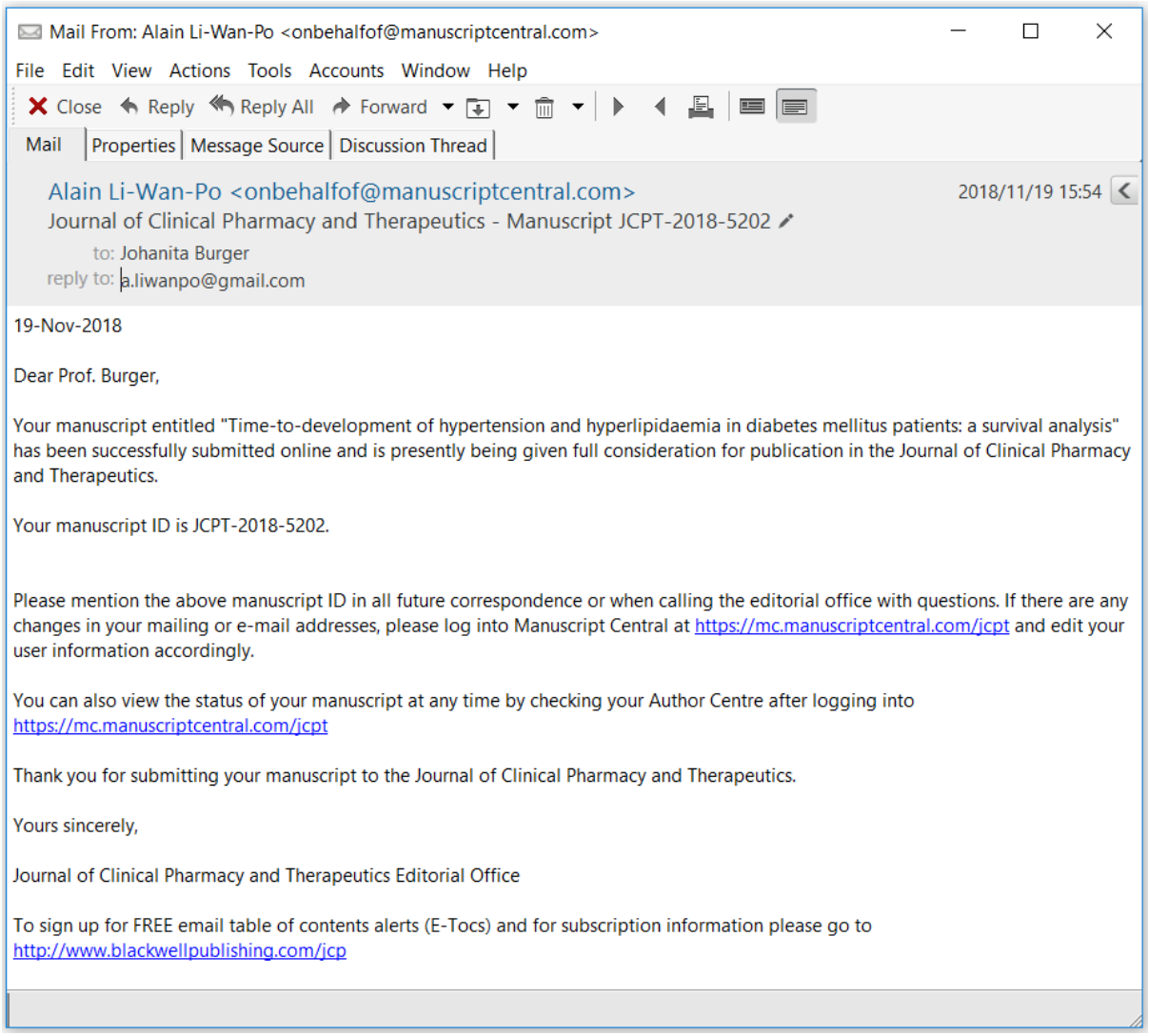
If you have queries about offprints, please email offprint@cosprinters.com.

**Note to NIH Grantees:** Pursuant to NIH mandate, Wiley Blackwell will post the accepted version of contributions authored by NIH grant-holders to PubMed Central upon acceptance. This accepted version will be made publicly available 12 months after publication. For further information, see www.wiley.com/nihmandate.

**Author Material Archive Policy:** Please note that unless specifically requested, Wiley Blackwell will dispose of all hardcopy or electronic material submitted two months after publication. If you require the return of any material submitted, please inform the Editorial Office or Production Editor as soon as possible if you have not yet done so.

**Clinical Case Reports:** We work together with Wiley's Open Access Journal, *Clinical Case Reports*, to enable rapid publication of good quality case reports that we are unable to accept for publication in our Journal. Authors of case reports rejected by our Journal will be offered the option of having their case report, along with any related peer reviews, automatically transferred for consideration by the *Clinical Case Reports* Editorial Team. Authors will not need to reformat or rewrite their manuscript at this stage, and publication decisions will be made a short time after the transfer takes place. *Clinical Case Reports* will consider case reports from every clinical discipline including Medicine, Nursing, Dentistry, and Veterinary Science and may include clinical images or clinical videos. *Clinical Case Reports* is a Wiley Open Access journal and article publication fees apply. For more information please go to www.clinicalcasesjournal.com.

# ANNEXURE D: PROOF OF SUBMISSION OF MANUSCRIPT 1



Mail From: Alain Li-Wan-Po <onbehalfof@manuscriptcentral.com>

File Edit View Actions Tools Accounts Window Help

Close Reply Reply All Forward

Mail | Properties | Message Source | Discussion Thread

Alain Li-Wan-Po <onbehalfof@manuscriptcentral.com> 2018/11/19 15:54

Journal of Clinical Pharmacy and Therapeutics - Manuscript JCPT-2018-5202

to: Johanita Burger

reply to: a.liwanpo@gmail.com

19-Nov-2018

Dear Prof. Burger,

Your manuscript entitled "Time-to-development of hypertension and hyperlipidaemia in diabetes mellitus patients: a survival analysis" has been successfully submitted online and is presently being given full consideration for publication in the Journal of Clinical Pharmacy and Therapeutics.

Your manuscript ID is JCPT-2018-5202.

Please mention the above manuscript ID in all future correspondence or when calling the editorial office with questions. If there are any changes in your mailing or e-mail addresses, please log into Manuscript Central at https://mc.manuscriptcentral.com/jcpt and edit your user information accordingly.

You can also view the status of your manuscript at any time by checking your Author Centre after logging into https://mc.manuscriptcentral.com/jcpt

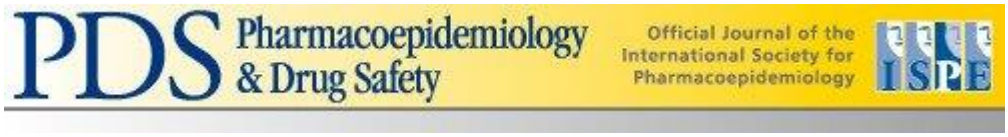
Thank you for submitting your manuscript to the Journal of Clinical Pharmacy and Therapeutics.

Yours sincerely,

Journal of Clinical Pharmacy and Therapeutics Editorial Office

To sign up for FREE email table of contents alerts (E-Tocs) and for subscription information please go to http://www.blackwellpublishing.com/jcp

## ANNEXURE E: AUTHOR GUIDELINES FOR JOURNAL OF PHARMACOEPIDEMIOLOGY AND DRUG SAFETY




# 1. SUBMISSION

Thank you for your interest in *Pharmacoepidemiology and Drug Safety*. Note that submission implies that the content has not been published or submitted for publication elsewhere except as a brief abstract in the proceedings of a scientific meeting or symposium.

**Once you have prepared your submission in accordance with the Guidelines, manuscripts should be submitted online at https://mc.manuscriptcentral.com/pds**

Please ensure you have read and understood section 5 before submitting your manuscript. Completed Conflict of Interest Disclosure Forms must accompany your submission.

The submission system will prompt you to use an ORCID iD (a unique author identifier) to help distinguish your work from that of other researchers. Click here to find out more.

Click here for more details on how to use ScholarOne.

For help with submissions, please contact: pds@wiley.com

We look forward to your submission.

# 2. AIMS AND SCOPE

The aim of *Pharmacoepidemiology and Drug Safety* is to provide an international forum for the communication and evaluation of data, methods and opinion in the discipline of pharmacoepidemiology, defined broadly.

Particular areas of interest include:

• design, analysis, results, and interpretation of studies looking at the benefit or safety of specific pharmaceuticals, biologics, or medical devices, including studies in pharmacovigilance, postmarketing surveillance, pharmacoeconomics, patient safety, drug utilization, molecular pharmacoepidemiology, or any other study within the broad field of pharmacoepidemiology;

• comparative effectiveness research relating to pharmaceuticals, biologics, and medical devices. Comparative effectiveness research is the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition, as these methods are truly used in the real world;

• methodologic contributions of relevance to pharmacoepidemiology, whether original contributions, reviews of existing methods, or tutorials for how to apply the methods of pharmacoepidemiology;

• assessments of harm versus benefit in drug therapy;

• relationships between pharmacoepidemiology and the formulation and interpretation of regulatory guidelines; contributions about regulatory science

• evaluations of risk management plans and programmes relating to pharmaceuticals, biologics and medical devices.

# 3. MANUSCRIPT CATEGORIES AND REQUIREMENTS

*Pharmacoepidemiology and Drug Safety* publishes a number of different article types including:

**Original Reports**

Original reports are the Journal's primary mode of scientific communication. Original reports typically do not exceed **3,000** words of body text, excluding abstract, tables, figures, and references. All original reports being considered for publication will be subject to peer review.

**Reviews**

Reviews of 'hot topics, controversies, and pharmacoepidemiologic methods are welcome. Reviews should be of a critical nature, discussing all sides of a question in a balanced manner. Experts considering offering such a review should feel free to contact one of the Regional Editors, as appropriate, in order to avoid unnecessary effort. All reviews will be peer-reviewed. Reviews typically should not exceed **3,000** words of body text (excluding abstract, figures, tables and references), and be limited to **150** references.

**Brief Reports**

Succinct data papers, and in highly unusual situations case reports (*Pharmacoepidemiol Drug Saf* 2007; **16** :473), will be considered for publication as Brief Reports and subjected to peer review. Brief Reports should not exceed **1,500** words excluding abstract, and be limited to 1 table, 1 figure and **15** references.

**Commentaries**

Commentaries cover a variety of topics of current interest in pharmacoepidemiology and pharmacovigilance, and the intersection between these disciplines and society. The Journal welcomes submissions and proposals. Commentaries are generally limited to **1,500** words and **15** references. No abstract is required for a commentary.

**Letters to the editor**

Letters to the Editor are encouraged, and may be in response to issues arising from recently published articles, or short, free-standing pieces expressing an opinion. No abstract is required,

and text should be formatted in one continuous section. Letters are limited to **1,000** words.

**Research Protocol**

*Pharmacoepidemiology and Drug Safety* does not ordinarily publish study protocols without results. Rather, we strongly recommend that investigators post their research protocols in a publicly available archive such as ClinicalTrials.gov (http://clinicaltrials.gov/) or ENCePP (http://www.encepp.eu/encepp/studiesDatabase.jsp) and ask that they describe that posting in their manuscripts submitted to *Pharmacoepidemiology and Drug Safety*. However, in unusual circumstances, *Pharmacoepidemiology and Drug Safety* will consider publishing descriptions of the design and rationale of pharmacoepidemiologic studies, before study results are available. Characteristics of such descriptions that support consideration for publication include:

• the study is of unusually high public health importance and interest to the readership of *Pharmacoepidemiology and Drug Safety*

• the study is of a scale that is likely to lead to multiple different subsequent results-oriented publications, each then able to refer to this original methods paper, rather than having to repeat the methods in detail

• the rationale for important aspects of the research design is discussed in more depth than could be accommodated in a paper reporting the results, and in more detail than would usually be included in the protocol that would be posted on ClinicalTrials.gov or the ENCePP database.

• the description will serve as an instructive teaching example

The format for the manuscript should be: Introduction, Design and Research Plan, Results (optional), and Discussion. Data describing the study population recruited can be included, if available, in the Results section of the publication. Please select 'Research Protocol' as the category for submission of the manuscript. The remainder of the format should be the same as that of Original Articles.

**Other**

Reviews of books and other media may be submitted only at the invitation of the Editors. However, suggestions are welcome.

# 4. EDITORS AND PEER REVIEW

The Editor-in-Chief (Brian Strom) will apportion manuscripts to a Regional Editor based on the location of the corresponding Author, unless there are conflicts of interest between the paper's authors and that regional office.

Papers from The Americas will be handled by:

**Vincent Lo Re**
University of Pennsylvania Perelman School of Medicine, Center for Pharmacoepidemiology Research and Training, 803 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104-6021, USA
Email: vincentl@pennmedicine.upenn.edu

Papers from Europe and UK will be handled by:

**Hubert G.M. Leufkens**
Utrecht University, Faculty of Science, Utrecht Institute for Pharmaceutical Sciences (UIPS), Division of Pharmacoepidemiology and Clinical Pharmacology, PO Box 80 082, 3508 TB Utrecht, The Netherlands
E-mail: H.G.M.Leufkens@uu.nl

Papers from Asia, Africa, the Middle East, and Oceania will be handled by:

**Byung Joo Park**
Seoul National University College of Medicine, Department of Preventive Medicine
103 Daehakno, Chongno-Gu, Seoul 110-799, Korea
Tel: 82 2 740 8325. E-mail: bjpark@snu.ac.kr

A fast-track review and publication process is in place for particularly time-sensitive findings of urgent public health importance. The Editor-in-Chief should be contacted to request this process.

Authors are encouraged to propose reviewers who have special competence to review their work. Authors may also ask that, due to a possible conflict of interest, named members of the Editorial Board or other individuals should not be selected to review a particular submission. The Editors will pay close attention to such requests, but must reserve the final choice of reviewers.

# 5. PREPARING YOUR SUBMISSION

Manuscripts must be submitted as a Word or rtf file and should be written in English. The guidelines below specify the format and layout required for submissions to *Pharmacoepidemiology and Drug Safety*.

We do appreciate that authors spend a lot of time reformatting their manuscripts for submission and therefore, *Pharmacoepidemiology and Drug Safety* does not require authors to reformat their manuscript to the journal's layout upon initial submission. We ask only that you consider your reviewers by supplying your manuscript in a clear, generic and readable layout (including abstract), and that a reasonable scientific standard is adhered to. However, subsequent revised versions of the manuscript must conform to the requirements of the journal as specified below.

**Text File**

The text file should be presented in the following order:

(i) Title page;

a) full title

b) a short running title of less than 50 characters;

c) the full names of the authors;

d) the authors' institutional affiliations at which the work was carried out, (footnote for authors' present address if different from where the work was carried out);

e) name, address, and email address of corresponding author;

f) key words;

g) 5 ‘take home’ messages or key points;

h) word count excluding abstract, tables, figures and references;

i) a statement about prior postings and presentations, name(s) of any sponsor(s) of the research contained in the paper, along with grant number(s);

(ii) abstract;

(iii) main text;

(iv) acknowledgments (including funding information);

(v) references;

(vi) tables (each table complete with title and footnotes)

(vii) figure legends

(viii) figures provided on a separate page

Appendices and supporting information should be supplied as separate files under supplementary material for review.

**Title**

The title should be a short informative title that contains the major key words. The title should not contain abbreviations (see Wiley's best practice SEO tips).

**Authorship**

Please refer to the journal’s authorship policy in the Editorial Policies and Ethical Considerations section for details on eligibility for author listing.

**Acknowledgements**

Contributions from anyone who does not meet the criteria for authorship should be listed, with permission from the contributor, in an Acknowledgements section. Financial and material support should also be mentioned. Thanks to anonymous reviewers are not appropriate.

## Conflict of Interest Statement

Authors are responsible for disclosing all financial and personal relationships between themselves and others that might appear to bias their work. To prevent ambiguity, authors must state explicitly in the Conflict of interest form whether potential conflicts do or do not exist. Authors should describe the role of the study sponsor(s), if any, in study design, in the collection, analysis and interpretation of data, in the writing of the report and in the decision to submit the report for publication. If the supporting source(s) had no such involvement, the authors should so state.

Completed Conflict of Interest Disclosure Form(s). Conflict of Interest (COI) disclosure forms must be uploaded with your manuscript files at submission. Please choose the file designation 'Conflict of Interest form' when submitting each of your forms. **Please note:** a separate COI form **must** be completed by the corresponding author **and** each co-author. If you do not submit separate COI forms for each of the authors, your manuscript will be un-submitted back to you. If your manuscript has not been formatted specifically for *Pharmacoepidemiology and Drug Safety*, authors can supply the international COI form upon initial submission but must complete the *Pharmacoepidemiology and Drug Safety* COI form if invited to provide a resubmission.

## Keywords

Please provide 2-6 keywords and list them in alphabetical order. Keywords should be taken from those recommended by the US National Library of Medicine's Medical Subject Headings (MeSH) browser list at https://www.nlm.nih.gov/mesh/.

## Key points

Please include up to five bullet points, of around 100 words only, explaining the importance of the paper's findings. These points will be published with article in a box entitled 'Key Points'.

## Abstract

Authors should note that structured abstracts (maximum **250** words) are required. The structured abstract should adopt the format: Purpose, Methods, Results, and Conclusions. Abstracts should not contain citations to other published work.

Letters and Commentaries do not require abstracts.

## Main Text

Where possible, the text should be divided into the following sections: Introduction, Methods, Results, Discussion, Acknowledgements, and References. Tables and Figures follow where applicable. The approximate placement of the tables and/or figures should be noted in the text.

Letters should be formatted in one continuous section. Commentaries should be formatted as appropriate to content.

**References**

All references should be numbered consecutively in order of appearance and should be as complete as possible. In text citations should cite references in consecutive order using Arabic superscript numerals. Sample references follow:

**Journal article:**

1. King VM, Armstrong DM, Apps R, Trott JR. Numerical aspects of pontine, lateral reticular, and inferior olivary projections to two paravermal cortical zones of the cat cerebellum..*J Comp Neurol* 1998;390:537-551.

**Book:**

2. Voet D, Voet JG. *Biochemistry*. New York: John Wiley & Sons; 1990. 1223 p.

Please note that journal title abbreviations should conform to the practices of Chemical Abstracts.

For more information about AMA reference style - AMA Manual of Style

**Tables**

Tables should be self-contained and complement, but not duplicate, information contained in the text. They should be supplied as editable files, not pasted as images. Legends should be concise but comprehensive – the table, legend and footnotes must be understandable without reference to the text. All abbreviations must be defined in footnotes. Footnote symbols: †, ‡, §, ¶, should be used (in that order) and *, **, *** should be reserved for P-values. Statistical measures such as SD or SEM should be identified in the headings.

**Figure Legends**

Legends should be concise but comprehensive – the figure and its legend must be understandable without reference to the text. Include definitions of any symbols used and define/explain all abbreviations and units of measurement.

**Preparing Figures**

Although we encourage authors to send us the highest-quality figures possible, for peer-review purposes we are happy to accept a wide variety of formats, sizes, and resolutions. Click here for the basic figure requirements for figures submitted with manuscripts for initial peer review, as well as the more detailed post-acceptance figure requirements.

**Colour Figures**

Figures submitted in colour may be reproduced in colour online free of charge. Please note, however, that it is preferable that line figures (e.g. graphs and charts) are supplied in black and white so that they are legible if printed by a reader in black and white. If you wish to have figures printed in colour in hard copies of the journal, a fee will be charged by the Publisher.

**Guidelines for Cover Submission**

If you would like to send suggestions for artwork related to your manuscript to be considered to appear on the cover of the journal, please follow these general guidelines.

**Appendices**

Appendices will be listed before the references in the online version of the article, and a link to download the appendices provided. Appendices will also be listed in the print version of the article, but readers will need to access the online version to view these. For submission they should be supplied as separate files (supplementary material for review) but referred to in the text.

**Supporting information**

Supporting information is information that is not essential to the article but that provides greater depth and background. It is hosted online, and appears without editing or typesetting. It may include tables, figures, videos, datasets, etc. Click here for Wiley's FAQs on supporting information. Note: if data, scripts or other artefacts used to generate the analyses presented in the paper are available via a publicly available data repository, authors should include a reference to the location of the material within their paper.

**General Style Points**

The following links provide general advice on formatting and style.

• **Abbreviations:** In general, terms should not be abbreviated unless they are used repeatedly and the abbreviation is helpful to the reader. Initially use the word in full, followed by the abbreviation in parentheses. Thereafter use the abbreviation only.

• **Units of measurement:** Measurements should be given in SI or SI-derived units. Visit the Bureau International des Poids et Mesures (BIPM) website at http://www.bipm.fr for more information about SI units.

• **Trade names:**Chemical substances should be referred to by the generic name only. Trade names should not be used. Drugs should be referred to by their generic names. If proprietary drugs have been used in the study, refer to these by their generic name, mentioning the proprietary name, and the name and location of the manufacturer, in parentheses.

**Wiley Author Resources**

Wiley has a range of resources for authors preparing manuscripts for submission available here. In particular, authors may benefit from referring to Wiley's best practice tips on Writing for Search Engine Optimization.

### Editing, Translation and Formatting Support

Wiley Editing Services can improve the chances of your manuscript being accepted. Offering expert help in English language editing, translation, manuscript formatting and figure preparation, Wiley Editing Services ensures that your manuscript is ready for submission.

# 6. EDITORIAL POLICIES AND ETHICAL CONSIDERATIONS

### Editorial Review and Acceptance

The acceptance criteria for all papers are the quality and originality of the research and its significance to our readership. Except where otherwise stated, manuscripts are single-blind peer reviewed. Papers will only be sent to review if the Editors determine that the paper meets the appropriate quality and relevance requirements. Wiley's policy on confidentiality of the review process is available here.

### Manuscript Transfer Programme

This journal works together with the American Society for Pharmacology and Experimental Therapeutics (ASPET), the British Pharmacological Society (BPS) and Wiley's Open Access Journal, *Pharmacology Research & Perspectives*, to enable rapid publication of good quality research that is unable to be accepted for publication by our journal. Authors may be offered the option of having the paper, along with any related peer reviews, automatically transferred for consideration by the Editors of *Pharmacology Research & Perspectives*. Authors will not need to reformat or rewrite their manuscript at this stage, and publication decisions will be made a short time after the transfer takes place. The Editors of *Pharmacology Research & Perspectives* will accept submissions that report well-conducted research that reaches the standard acceptable for publication. *Pharmacology Research & Perspectives* is a Wiley Open Access journal and article publication fees apply. For more information please go to http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)2052-1707.

### Data Sharing and Accessibility

The journal encourages authors to share the data and other artefacts supporting the results in the paper by archiving it in an appropriate public repository. Authors should include a data accessibility statement, including a link to the repository they have used, in order that this statement can be published alongside their paper.

### Human Studies and Subjects

For manuscripts reporting medical studies involving human participants, we require a statement identifying the ethics committee that approved the study, and that the study conforms to recognized standards, for example: Declaration of Helsinki; US Federal Policy for the Protection of Human Subjects; or European Medicines Agency Guidelines for Good Clinical Practice. All human participants must have given written informed consent.

If patients were involved in the development of the study question or its outcome measures, or were involved in the conduct of the study, a statement to this effect should be included in the Methods section.

Images and information from individual participants will only be published where the authors have obtained the individual's free prior informed consent. Authors do not need to provide a copy of the consent form to the publisher. However in signing the author license to publish authors are required to confirm that consent has been obtained. Wiley has a standard patient consent form available.

Authors are expected to follow the Guidelines for Good Pharmacoepidemiology Practice (GPP) as described in *Pharmacoepidemiol Drug Saf.* DOI: 10.1002/pds.3891 (Volume 25, Issue 1, January 2016, Pages: 2–10). *Pharmacoepidemiology and Drug Safety* also recommends that authors use STROBE as a guideline for the reporting of observational studies; RECORD when using observational routinely-collected health data; and CONSORT as a guideline for the reporting of randomized controlled clinical trials.

**Clinical Trial Registration**

The Journal requires clinical trials to be prospectively registered in a publicly accessible database. Clinical trial registration numbers should be included in all papers that report their results. Please include the name of the trial register and your clinical trial registration number at the end of your abstract. If your trial is not registered, or was registered retrospectively, please explain the reasons for this.

**Research Reporting Guidelines**

Accurate and complete reporting enables readers to fully appraise research, replicate it, and use it. We encourage authors to adhere to the following research reporting standards.

• CONSORT
• SPIRIT
• PRISMA
• PRISMA-P
• STROBE
• CARE
• COREQ
• STARD and TRIPOD
• CHEERS
• the EQUATOR Network
• Future of Research Communications and e-Scholarship (FORCE11)
• ARRIVE guidelines
• National Research Council's Institute for Laboratory Animal Research guidelines: the Gold Standard Publication Checklist from Hooijmans and colleagues
• Minimum Information Guidelines from Diverse Bioscience Communities (MIBBI) website; Biosharing website
• REFLECT statement

**Genetic Nomenclature**

Sequence variants should be described in the text and tables using both DNA and protein designations whenever appropriate. Sequence variant nomenclature must follow the current HGVS guidelines; see http://varnomen.hgvs.org/, where examples of acceptable nomenclature are provided.

Nucleotide sequence data can be submitted in electronic form to any of the three major collaborative databases: DDBJ, EMBL or GenBank. It is only necessary to submit to one database as data are exchanged between DDBJ, EMBL and GenBank on a daily basis. The suggested wording for referring to accession-number information is: 'These sequence data have been submitted to the DDBJ/EMBL/GenBank databases under accession number U12345'. Addresses are as follows:

DNA Data Bank of Japan (DDBJ) http://www.ddbj.nig.ac.jp

EMBL Nucleotide Sequence Submissions http://www.ebi.ac.uk

GenBank http://www.ncbi.nlm.nih.gov

**Conflict of Interest**

Authors are responsible for disclosing all financial and personal relationships between themselves and others that might appear to bias their work. To prevent ambiguity, authors must state explicitly in the Conflict of Interest form whether potential conflicts do or do not exist. Authors should describe the role of the study sponsor(s), if any, in study design, in the collection, analysis and interpretation of data, in the writing of the report and in the decision to submit the report for publication. If the supporting source(s) had no such involvement, the authors should so state.

Completed Conflict of Interest Disclosure Form(s). Conflict of Interest (COI) disclosure forms must be uploaded with your manuscript files at submission. Please choose the file designation 'Conflict of Interest form' when submitting each of your forms. Please note: a separate COI form must be completed by the corresponding author and each co-author. If you do not submit separate COI forms for each of the authors, your manuscript will be un-submitted back to you. If your manuscript has not been formatted specifically for Pharmacoepidemiology and Drug Safety, authors can supply the international COI form upon initial submission but must complete the Pharmacoepidemiology and Drug Safety COI form upon resubmission.

Pharmacoepidemiology and Drug Safety's Editor-in-Chief, Regional Editors and Associate Editors are required to disclose any possible conflict of interests on an annual basis. These can include financial conflicts of interest e.g. patent ownership, stock ownership, employment by dialysis/pharmaceutical companies, consultancies or speaker's fees from pharmaceutical companies related to the research/topic area. These disclosures can be found here.

Reviewers are also requested to disclose any possible conflict of interests which may preclude them from reviewing a manuscript to the Editorial Office.

The Editors are not involved in any decisions regarding submissions to the Journal where they are author/co-author, where the author is a collaborator of the Editor, or the research originates

from the Editor's institution. There is a system in place for dealing with this when such occasions arise.

**Funding**

Authors should list all funding sources in the Acknowledgments section. Authors are responsible for the accuracy of their funder designation. If in doubt, please check the Open Funder Registry for the correct nomenclature: http://www.crossref.org/fundingdata/registry.html

**Authorship**

The list of authors should accurately illustrate who contributed to the work and how. All those listed as authors should qualify for authorship according to the following criteria:

1) Have made substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data;
2) Been involved in drafting the manuscript or revising it critically for important intellectual content;
3) Given final approval of the version to be published. Each author should have participated sufficiently in the work to take public responsibility for appropriate portions of the content; and
4) Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Contributions from anyone who does not meet the criteria for authorship should be listed, with permission from the contributor, in an Acknowledgments section (for example, to recognize contributions from people who provided technical help, collation of data, writing assistance, acquisition of funding, or a department chairperson who provided general support). Prior to submitting the article all authors should agree on the order in which their names will be listed in the manuscript.

**Additional authorship options**

Joint first or senior authorship: In the case of joint first authorship a footnote should be added to the author listing, e.g. 'X and Y should be considered joint first author' or 'X and Y should be considered joint senior author.'

**ORCID**

As part of our commitment to supporting authors at every step of the publishing process, *Pharmacoepidemiology and Drug Safety* requires the submitting author (only) to provide an ORCID iD when submitting a manuscript. This takes around 2 minutes to complete. Find more information.

**Publication Ethics**

*Pharmacoepidemiology and Drug Safety* is a member of the Committee on Publication Ethics (COPE). Note this journal uses iThenticate's CrossCheck software to detect instances of overlapping and similar text in submitted manuscripts. Read our Top 10 Publishing Ethics Tips for Authors here. Wiley's Publication Ethics Guidelines can be found at https://authorservices.wiley.com/ethics-guidelines/index.html

# 7. AUTHOR LICENSING

If your paper is accepted, the author identified as the formal corresponding author will receive an email prompting them to login into Author Services; where via the Wiley Author Licensing Service (WALS) they will be able to complete the license agreement on behalf of all authors on the paper.

Authors may choose to publish under the terms of the journal's standard copyright agreement, or OnlineOpen under the terms of a Creative Commons License.

To find out which Creative Commons Licences are available for the journal, click here. To learn more about Creative Commons Licenses and to preview terms and conditions of the agreements, please click here. Note that certain funders mandate that a particular type of CC license has to be used; to check these please click here.

General information regarding licensing and copyright is available here. To review the Creative Commons License options offered under OnlineOpen, please click here. (Note that certain funders mandate that a particular type of CC license has to be used; to check this please click here.)

**Self-Archiving definitions and policies:** Note that the journal's standard copyright agreement allows for self-archiving of different versions of the article under specific conditions. Please click here for more detailed information about self-archiving definitions and policies.

**Open Access fees:** If you choose to publish using OnlineOpen you will be charged a fee. A list of Article Publication Charges for Wiley journals is available here.

**Funder Open Access:** Please click here for more information on Wiley's compliance with specific Funder Open Access Policies.

# 8. PUBLICATION PROCESS AFTER ACCEPTANCE

### Accepted article received in production

When your accepted article is received by Wiley's production team, you (corresponding author) will receive an email asking you to login or register with Author Services. You will be asked to sign a publication license at this point.

### Proofs

Once your paper is typeset you will receive an email notification of the URL from where to download a PDF typeset page proof, associated forms and full instructions on how to correct and return the file.

Please note that you are responsible for all statements made in your work, including changes made during the editorial process and thus you must check your proofs carefully. Note that proofs should be returned 48 hours from receipt of first proof.

**Publication Charges**

Colour figures may be published online free of charge. However the journal charges for publishing figures in colour in print. If you supply colour figures you will be sent a Colour Work Agreement once your accepted paper moves to the production process. If your Colour Work Agreement is not returned by the specified date figures will be converted to black and white for print publication.

**Early View**

The journal offers rapid publication via Wiley's Early View service. Early View (Online Version of Record) articles are published on Wiley Online Library before inclusion in an issue. Once your article is published on Early View no further changes to your article are possible. Your Early View article is fully citable and carries an online publication date and DOI for citations.

# 9. POST PUBLICATION

**Access and sharing**

When your article is published online:

• You receive an email alert (if requested).

• You can share a link to your published article through social media.

• As the author, you will have free access to your paper (after accepting the Terms & Conditions of use, you can view your article).

• The corresponding author and co-authors can nominate up to ten colleagues to receive a publication alert and free online access to your article.

You can now order print copies of your article (instructions are sent at proofing stage or use the below contact details).

Email offprint@cosprinters.com

Now is the time to start promoting your article. Find out how to do that here. Wiley also helps you measure the impact of your research through our specialist partnerships with Kudos and Altmetric.

# 10. DATA PROTECTION AND PRIVACY

By submitting a manuscript to, or reviewing for, this publication, your name, email address, institutional affiliation, and other contact details the publication might require, will be used for the regular operations of the publication, including, when necessary, sharing with the publisher (Wiley) and partners for production and publication. The publication and the publisher recognize the importance of protecting the personal information collected from users in the operation of these services, and have practices in place to ensure that steps are taken to maintain the security,

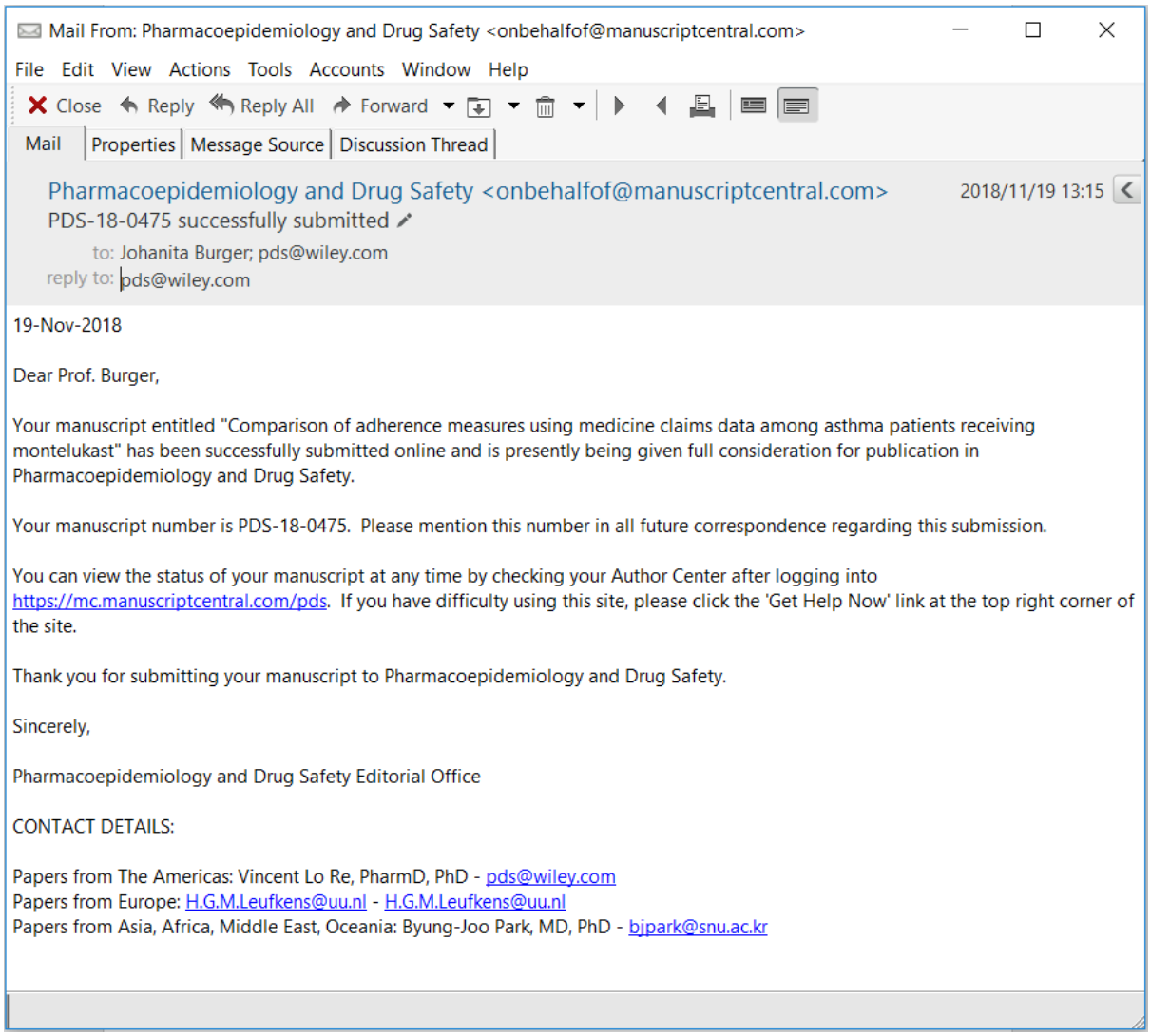
integrity, and privacy of the personal data collected and processed. You can learn more at https://authorservices.wiley.com/statements/data-protection-policy.html.

# 11. EDITORIAL OFFICE CONTACT DETAILS

pds@wiley.com

*Author Guidelines updated 23rd May 2018*

# ANNEXURE F: PROOF OF SUBMISSION OF MANUSCRIPT 2



Mail From: Pharmacoepidemiology and Drug Safety <onbehalfof@manuscriptcentral.com>

File Edit View Actions Tools Accounts Window Help

Close Reply Reply All Forward

Mail | Properties | Message Source | Discussion Thread

Pharmacoepidemiology and Drug Safety <onbehalfof@manuscriptcentral.com> 2018/11/19 13:15

PDS-18-0475 successfully submitted

to: Johanita Burger; pds@wiley.com

reply to: pds@wiley.com

19-Nov-2018

Dear Prof. Burger,

Your manuscript entitled "Comparison of adherence measures using medicine claims data among asthma patients receiving montelukast" has been successfully submitted online and is presently being given full consideration for publication in Pharmacoepidemiology and Drug Safety.

Your manuscript number is PDS-18-0475. Please mention this number in all future correspondence regarding this submission.

You can view the status of your manuscript at any time by checking your Author Center after logging into https://mc.manuscriptcentral.com/pds. If you have difficulty using this site, please click the 'Get Help Now' link at the top right corner of the site.

Thank you for submitting your manuscript to Pharmacoepidemiology and Drug Safety.

Sincerely,

Pharmacoepidemiology and Drug Safety Editorial Office

CONTACT DETAILS:

Papers from The Americas: Vincent Lo Re, PharmD, PhD - pds@wiley.com
Papers from Europe: H.G.M.Leufkens@uu.nl - H.G.M.Leufkens@uu.nl
Papers from Asia, Africa, Middle East, Oceania: Byung-Joo Park, MD, PhD - bjpark@snu.ac.kr

## ANNEXURE G: LANGUAGE EDITOR'S LETTER

*Declaration*

*This is to declare that I, Annette L Combrink, accredited language editor and translator of the South African Translators' Institute, have language-edited the dissertation by*

**M Obeng-Kusi**

*With the title*

**Research methods for conducting pharmacoepidemiological studies using medicines claims data**

*Prof Annette L Combrink*
*Accredited translator and language editor*
*South African Translators' Institute*
*Membership No. 1000356*
*Date: 18 November 2018*

# ANNEXURE H: CERTIFICATE OF TECHNICAL EDITING

WHOM IT MAY CONCERN

I hereby declare that I have done the technical editing of the mini-dissertation with the title:

**Research methods for conducting pharmacoepidemiological studies using medicines claims data**
by

**M Obeng-Kusi**

**27959716**

Technical editing includes all tables, figures as well as the layout of the text.

E Oosthuizen

November 2018