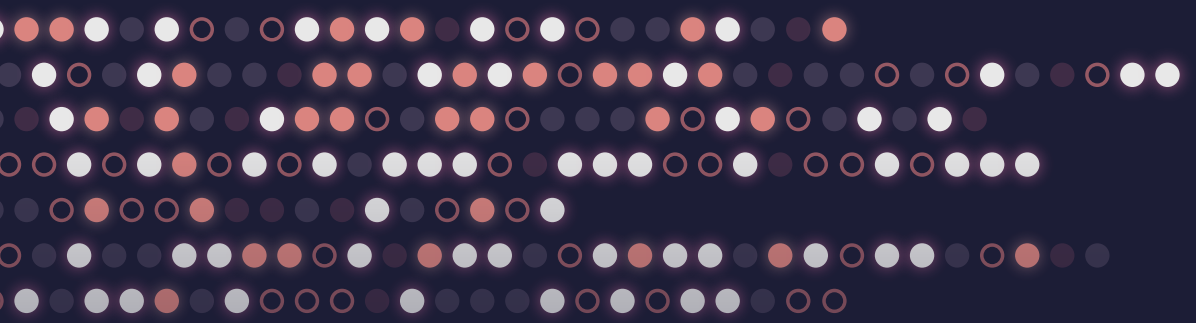
Drug exposure assessment in pharmacoepidemiological database studies. Reporting and impact of exposure misclassification.



Mirjam Hempenius

Drug exposure assessment in pharmacoepidemiological database studies

Reporting and impact of exposure misclassification

Mirjam Hempenius

Drug exposure assessment in pharmacoepidemiological studies: reporting and impact of exposure misclassification.
PhD thesis, Utrecht University, The Netherlands
Mirjam Hempenius

Printed by Ipskamp Printing | proefschriften.net
Cover and illustration design | Rieko Medendorp, studiorieko.nl
Layout and design | Birgit Vredenburg, persoonlijkproefschrift.nl

ISBN: 978-90-393-7445-0
DOI: https://doi.org/10.33540/687

Financial support for printing this thesis by the Utrecht Institute for Pharmaceutical Sciences (UIPS) is greatly acknowledged.

Drug exposure assessment in pharmacoepidemiological database studies

Reporting and impact of exposure misclassification

Bepaling van blootstelling aan geneesmiddelen in farmacoepidemiologische database studies

Het rapporteren en de impact van misclassificatie van geneesmiddelblootstelling

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

maandag 7 maart 2022 des middags te 2.15 uur

door

Mirjam Hempenius
geboren op 28 november 1988
te Kampen

Promotoren:

Prof. dr. O.H. Klungel
Prof. dr. A. De Boer

Copromotor:

Dr. H. Gardarsdottir

TABLE OF CONTENTS

1.

GENERAL INTRODUCTION

Regulatory authorities assess and approve drugs based on evidence generated from randomized clinical trials (RCTs), among other things. During the drug approval process, there are often unanswered questions about the safety and efficacy of the drug, and new questions may arise after the approval. For example, the effectiveness of a drug is not always known in subjects who do not meet the RCT eligibility criteria, such as the elderly, pregnant woman, and children, or in subjects with multiple or specific comorbidities or comedication. Moreover, rare adverse events or delayed effects are seldom detected in RCTs, due to the limited sample size and time span of such trials. In addition, the effectiveness of the drug in question in comparison with other treatments for the same indication is often not known at the time of marketing authorization.[1–3]

To answer those questions, post-authorization efficacy studies (PAESs) and post-authorization safety studies (PASSs) - together called post-authorization studies (PASs) - are conducted that generate evidence regarding the effects of a drug when used in clinical practice. These PASs are often observational studies that use real-world data (RWD) that are based on routine electronic healthcare records. The results from these studies inform regulators and health care providers about the risks and benefits of the drug, for example the expected effects in a specific subgroup, the comedications with which it can or cannot be combined, and the adverse events that could occur.[4] Another advantage of using RWD is that these studies can be conducted quickly and inexpensively. For example, at the start of the COVID-19 pandemic, there was an urgent need for evidence about possible treatments. In those circumstances, observational studies could provide answers quicker than RCTs and can offer initial guidance for physicians.

However, given their observational nature, these studies are vulnerable to biases. In contrast to an RCT, treatments in the real world are not randomly allocated, but are the outcome of the physician's decision, which can be influenced by factors that also relate to the outcome, causing confounding or selection bias. In addition, there is no standard protocol for how data should be registered in clinical practice as is the case for highly controlled RCTs. Therefore, the design of an observational study is crucial to avoid or at least minimize biases. There are currently a number of guidelines to support methodological choices in designing an observational study, such as the ISPE Guidelines for Good Pharmacoepidemiology Practices,[5] the ENCePP Guide on Methodological Standards in Pharmacoepidemiology,[6] the FDA's Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data,[7] and the User's Guide for Developing a Protocol for Observational Comparative Effectiveness Research by the Agency for Healthcare Research and Quality.[8]

In each of these guidelines, a section is devoted to drug exposure assessment. These sections describe how exposure information retrieved from routinely collected health data can be defined and modeled for use in estimating the association between drug exposure and outcome. The exposure assessment starts with assumptions and choices about the type of association between the drug exposure and the outcome. This can be, among other things, a direct effect, a delayed effect, a cumulative effect, or a combination of different effects.[9,10] Depending on the assumed relationship and the clinical setting in which the drug is being used, a data source should be chosen to measure the drug exposure, for example a questionnaire or a database containing information on prescriptions, dispensings, or reimbursement claims of the drug under study. This exposure information must then be translated into exposure episodes to which the occurrence of the outcome can be related. This process is summarized in Figure 1.

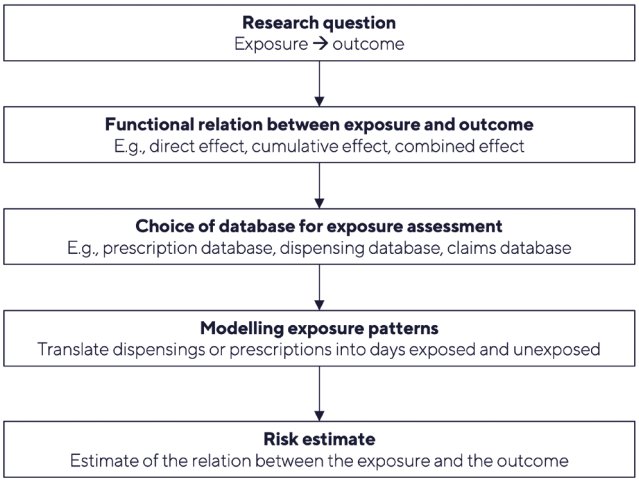


FIGURE 1. Process of drug exposure assessment

THE FUNCTIONAL RELATION BETWEEN DRUG AND OUTCOME

The first decision concerns the expected type of relation between exposure and outcome.[11] The effect can be related only to instantaneous use, for example when studying an anaphylactic drug reaction.[12] Other adverse events are related to the dose that is taken, such as the relation between benzodiazepines and fractures, where the fracture risk increases with higher doses of benzodiazepines.[13,14] However, not all effects are related to direct exposure, and some effects only occur when a certain amount of exposure is reached (cumulative exposure), such as the preventive effect of statins, which becomes detectable after two years of use.[15] This cumulative effect is also seen in the relation between the use of thiazolidinediones and the outcome of use, where prolonged use of thiazolidinediones increases the risk of fractures more than short-term use.[16]

Prior knowledge from case reports or other studies can be used to determine the temporal relationship between exposure and outcome and thus which model to use. However, if the incorrect model is used, there is a chance that an existing effect will not be observed.[12] Alternatively, observational studies can also be used to study the time relation, by applying different time-dependent models for the same exposure–outcome relationship. This was, for example, applied in a study of the preventive effect of statins. In this study, the effect estimate of cumulative statin use (≥ 2 years) was more in line with the effect estimates from trials than when exposure to statins was assessed only at baseline.[15]

Furthermore, multiple effects could also play a role, such as in the relationship between the use of selective serotonin reuptake inhibitors (SSRIs) and the risk of a hip fracture, where there is both a direct effect and a cumulative effect. On the one hand, SSRIs increase the risk of falling, due to day-time drowsiness,[17] but on the other hand, they have an effect on bone mineral density, causing a higher risk of fractures in the long term.[18,19] These different effects in time have also been described for the relation between the use of cox-2 inhibitor and the risk of myocardial infarction, where the rate of cardiovascular events increased rapidly within the first month due

to a direct thrombogenic effect; decreased thereafter; and increased again eight months after the start of rofecoxib exposure, due to delayed atherogenic effects.[11,20]

Since simple models, such as those in current use at a specific point in time (yes/no) or cumulative dose models, do not do justice to the complexity of these relations, more sophisticated models can be used, such as the weighted cumulative exposure model.[12,21,22] In this model, different weights are allocated to different timings of exposure, quantifying the relative importance of past doses on current risk.[22]

CHOICE OF DATA SOURCE

The next decision pertains to the data source in which the exposure can be measured. Today, routinely collected data from administrative or health care databases are often used for exposure assessment. However, the recording of information in these databases is not intended for research purposes, but for correct administration and billing of delivered care or for the monitoring and continuity of care – so-called secondary data. This contrasts with information about drug exposure that is primarily collected for a specific research question, for example via questionnaires. Yet, exposure assessment using routinely collected health data is seen as a more reliable source for exposure assessment than the use of questionnaires because these databases do not suffer from recall bias, which can be a problem when using questionnaires.[23–25]

Different types of databases are available for the assessment of drug exposure, including claims databases from health insurers, out-patient pharmacy databases, and electronic health records with information from general practitioners (GPs) or hospitals and nationwide registries.[4,23,26,27] Since these databases differ in their purpose and represent different aspects of health care delivery, they also differ in the information they hold about drug exposure.

Hospital databases contain electronic medical records of in-hospital treated patients, and they are often single center databases or sometimes a cluster of different hospitals. Electronic medical records contain information on drug use during hospitalization but often lack information on drug use outside the time window of hospitalization.[28] GP databases, such as the UK Clinical Practice Research Datalink® (CPRD®) contain prescriptions issued by GPs, but whether these prescriptions are filled at the pharmacy is not registered.[29] Out-patient pharmacy dispensing databases, such as the PHARMO Out-patient Pharmacy Database or the Nordic Prescription registries, contain information about prescriptions – both from GPs and hospital specialists – that are dispensed by the pharmacy.[30,31] In this type of database, information about drug use during hospitalization is lacking. Insurance claims databases, such as the US government claims database Medicaid, contain information about reimbursed drugs but lack information about drugs that are not reimbursed or about drugs that are only reimbursed under certain circumstances (restrictive coverage policy), dependent on the reimbursement system.[32–36]

In all these databases, no information is available about the actual drug use by the patient, except in in-hospital databases, where drug administration is registered, next to the prescriptions or dispensings. In addition, information about drugs that can be bought by the patient (over-the-counter [OTC] use) is often not available in all these databases,[37] nor is information about free

drug samples, which are in some countries distributed as promotional tools for pharmaceutical manufacturers,[38,39] or drug use in RCTs.

As a result, no database contains a complete record of all drug exposures. It is therefore important to understand how data in these specific databases are generated for proper use and application in pharmacoepidemiologic studies.

MODELING EXPOSURE PATTERNS

Depending on the assumed functional relation between the drug exposure and the outcome, the exposure information available in the databases must be modeled for analysis. For time-fixed exposure definitions, where exposure is assessed only once, this is straightforward (e.g., an intention to treat analysis, assessing drug exposure only at baseline). However, as most drugs are used for longer periods or at different time points, these drug exposures should be modeled as time-varying exposures, for which different steps and assumptions are needed.

For instance, a database contains information about a particular prescription, which was provided on Day x. For exposure to chronically used drugs, the information about the amount dispensed and the daily dose can be used to calculate the period in which the subject is assumed to be exposed. However, information about the daily dose of a drug and the amount dispensed is not always available in the database. In such cases, a fixed time period can be applied to all prescriptions based on standard duration for a prescription in clinical practice. The treatment duration can also be estimated using more advanced methods such as the waiting time distribution, which determines the treatment duration by estimating when a specified percentage of prevalent users had renewed their prescription.[40]

In addition, patients do not always take their medicines as prescribed – they can take more or less than prescribed, skip days, or stop using the drug earlier than prescribed.[41] For the use of antidepressants, for example, it is shown that only 30% of patients use the drug for more than 80% of all prescribed days,[42,43] and the mean proportion of prescribed daily doses that is actually taken is about 50%.[44,45] To account for these intake patterns when modeling exposure, a grace period is often applied after each prescription, which bridges gaps between two prescriptions that are caused by non-adherence.[46] Furthermore, most patients will fill their next prescription a few days before they have fully used up the previous prescription so that they have sufficient supplies at home. In the data, this is expressed as overlapping periods of two prescriptions, and this "stockpiling" must be taken into account when assessing the time period of exposure.[47,48]

While the construction of exposure episodes for chronically used drugs requires many assumptions, this is even more true when it concerns the use of "take-as-needed" drugs, such as non-steroidal anti-inflammatory drugs, or for drugs where dose instructions vary from day to day, such as coumarins. In these cases, sophisticated methods are needed to create episodes of assumed exposure and non-exposure, such as the waiting time distribution.[49,50]

BIASES DUE TO EXPOSURE ASSESSMENT

All of the aforementioned methodological choices in the process of exposure assessment (Figure 1) may have an impact on the estimated effect or risk of medicine use. First, assumptions regarding the relation between exposure and outcome have an impact on the effect estimate and the interpretation of this estimate. In addition, using the incorrect model for the relation between exposure and outcome may leave adverse events undetected. Second, exposure status may be misclassified as non-exposed when information about the exposure is not captured in the database, for example due to OTC use or drug use during hospitalization. Third, despite all efforts to model the information on drug exposure to episodes of use and non-use, the information retrieved from these databases can only serve as a proxy for actual use (i.e., the patient ingesting the drug). Exposure misclassification can consequently occur in pharmacoepidemiologic research, which might cause bias.

This misclassification can be either non-differential or differential. Non-differential exposure misclassification is not related to the outcome and occurs at random.[4] This type of misclassification can be problematic because it dilutes the effect of interest. If this effect is important but small, then there is a risk that this signal will not be detected in the study. In contrast, differential misclassification is misclassification that does not occur randomly, but whose probability is associated with the risk of the outcome.[4] Differential misclassification can lead to bias in different directions. On the one hand, protective or harmful effects can be observed when in fact there is no relationship between the exposure and the outcome. On the other hand, effects that do exist may go undetected. Hence, both non-differential and differential exposure misclassification can lead to biased estimates and invalid conclusions.[51]

It is not always possible to fully avoid biases due to exposure misclassification. Assumptions are needed about the time window in which subjects are considered to be exposed and how exposure relates to the outcome. In addition, subjects do not always adhere to drug use as prescribed, and none of the databases captures exposure fully.

REPORTING OF EXPOSURE ASSESSMENT

Methodological choices underlying the exposure assessment may have an impact on the estimate of the exposure–outcome relationship.[46,52–55] For example, different exposure definitions lead to different effects being estimated (e.g., short-term vs long-term effects).[53] Furthermore, design elements, such as how gaps and overlaps are handled, may have an impact on the estimates.[46] Therefore, transparent reporting is important not only for the interpretation of published study results but also for reproducibility and validity assessment.

Reporting guidelines support researchers to describe their research in a transparent and complete manner, like CONSORT,[56] for reporting on clinical trials, STROBE,[57] for reporting on observational studies, and STROBE-RECORD,[58] for reporting on observational studies using routinely collected health data. These guidelines do however not capture the operational details required for the conduct of pharmacoepidemiological research, including complex exposure ascertainment definitions.[59] Recently, two guidelines were published that do focus on these complex operational details.[59,60] The first is a guideline of a joint ISPE-ISPOR Task Force,

published in 2017 and the second an extension of the RECORD statement, the RECORD-PE, published in 2018.[59,60] Both guidelines include a separate section on the reporting of the drug exposure definition, with the former being specific to the operational details of exposure ascertainment. Adhering to these reporting guidelines will enable the correct interpretation of results, reproducibility, and validity assessments.

Adherence to reporting guidelines is, however, not always optimal, as shown, for example, for guidelines such as CONSORT and STROBE.[61] No research has yet been conducted into the current state of reporting in pharmacoepidemiology, nor what is reported about exposure assessment according to reporting guidelines for pharmacoepidemiologic studies. It is therefore important to obtain insight into how exposure is currently reported and where improvements may be needed.

OBJECTIVE AND OUTLINE OF THIS THESIS

While the sources of misclassification and bias are known to exist, less is known about their impact. Insights into the impact of these biases on effect estimates will guide researchers in the use of valid methods. They will also help decision-makers, such as regulators, to assess the validity of drug effectiveness and safety studies, which in turn leads to better decisions by physicians and safer use of medicines by patients. The aims of this thesis are hence a) to investigate the current reporting of exposure assessment and the risk of bias in those reports and b) to explore the extent and impact of misclassification of exposure on risk estimates in pharmacoepidemiology.

Chapter 2 provides a review of the reporting of exposure assessment and biases in published pharmacoepidemiologic studies. In *Chapter 2.1,* the reporting of exposure assessment in peer reviewed articles is reviewed in six pharmacoepidemiologic journals to identify the current quality of reporting. *Chapter 2.2* provides insights into the risk of bias in a period in which rapid evidence generation was urgently needed by applying the seven ROBINS-I tool domains to assess study design of in-hospital studies regarding potential COVID-19 treatments.

In **Chapter 3**, different sources of exposure misclassification are studied, as well as the impact thereof. *Chapter 3.1* provides an estimate of the amount of newly prescribed drugs that are not dispensed at the pharmacy and therefore will lead to misclassification in a prescription database. The impact of choosing different exposure definitions is then studied in *Chapter 3.2*, which includes a case study of the risk of amiodarone in amiodarone users. *Chapter 3.3* explores the impact of exposure misclassification caused by using a GP prescription database to assess outcomes in users of direct oral anticoagulants, which are prescribed by both GPs and specialists. In *Chapter 3.4,* simulated data are used to provide an overview of the different sources of exposure misclassification and their relative impact on pharmacoepidemiologic studies.

Finally, **Chapter 4** provides a general discussion of the findings within a broader perspective.

Author's contribution

The idea and set-up of the general introduction are by MH; she conducted the short literature search and wrote the general introduction. During the whole process she implemented input and feedback from her PhD supervisors.

REFERENCES

1. Stricker BH, Psaty BM. *Detection, Verification, and Quantification of Adverse Drug Reactions*. Vol 329. BMJ Publishing Group; 2004:44-47. doi:10.1136/bmj.329.7456.44
2. Strom BL, Kimmel SE, Hennessy S. *Pharmacoepidemiology*. 5th ed. Wiley-Blackwell; 2011.
3. Glasser SP, Salas M, Delzell E. Importance and challenges of studying marketed drugs: What is a phase IV study? common clinical research designs, registries, and self-reporting systems. *J Clin Pharmacol*. 2007;47(9):1074-1086. doi:10.1177/0091270007304776
4. Strom BL ed. Pharmacoepidemiology. Third edition. *Eur J Public Health*. 2002;12(2):155-156. doi:10.1093/eurpub/12.2.155-a
5. ISPE. Guidelines for good pharmacoepidemiology practices (GPP). *Pharmacoepidemiol Drug Saf*. 2008;17(2):200-208. doi:10.1002/pds.1471
6. European Medicines Agency. *The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology, 9th Revision.*; 2010. Accessed June 1, 2021. http://www.encepp.eu/standards_and_guidances
7. U.S. Department of Health and Human Services, Administration Food and Drug, Research (CDER) Center for Drug Evaluation and, Research (CBER) Center for Biologics Evaluation and. Guidance for Industry and FDA Staff Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data. *Drug Saf*. Published online 2013. Accessed June 18, 2021. http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm
8. Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. Published online 2013. Accessed September 3, 2021. https://www.ncbi.nlm.nih.gov/books/NBK126190/
9. Edwards IR, Aronson JK. Adverse drug reactions: Definitions, diagnosis, and management. *Lancet*. 2000;356(9237):1255-1259. doi:10.1016/S0140-6736(00)02799-9
10. Davies D. *Textbook of Adverse Drug Reactions*. 4th ed. (Davies D, ed.). Oxford University Press; 1991.
11. Suissa M, LeLorier J. Pharmacologically pertinent period of effect (PPPE). *Pharmacoepidemiol Drug Saf*. 2020;29(S1):61-67. doi:10.1002/pds.4714
12. van Gaalen RD, Abrahamowicz M, Buckeridge DL. The impact of exposure model misspecification on signal detection in prospective pharmacovigilance. *Pharmacoepidemiol Drug Saf*. 2015;24(5):456-467. doi:10.1002/pds.3700
13. Zint K, Haefeli WE, Glynn RJ, Mogun H, Avorn J, Stürmer T. Impact of drug interactions, dosage, and duration of therapy on the risk of hip fracture associated with benzodiazepine use in older adults. *Pharmacoepidemiol Drug Saf*. 2010;19(12):1248-1255. doi:10.1002/pds.2031
14. Requena G, Huerta C, Gardarsdottir H, et al. Hip/femur fractures associated with the use of benzodiazepines (anxiolytics, hypnotics and related drugs): A methodological approach to assess consistencies across databases from the PROTECT-EU project. *Pharmacoepidemiol Drug Saf*. 2016;25:66-78. doi:10.1002/pds.3816
15. Maitland-Van Der Zee AH, Klungel OH, Stricker BHC, et al. Comparison of two methodologies to analyze exposure to statins in an observational study on effectiveness. *J Clin Epidemiol*. 2004;57(3):237-242. doi:10.1016/j.jclinepi.2003.07.011
16. Bazelier MT, de Vries F, Vestergaard P, et al. Risk of fracture with thiazolidinediones: An individual patient data meta-analysis. *Front Endocrinol (Lausanne)*. 2013;4(11):1-9. doi:10.3389/fendo.2013.00011
17. Darowski A, Chambers SACF, Chambers DJ. Antidepressants and falls in the elderly. *Drugs and Aging*. 2009;26(5):381-394. doi:10.2165/00002512-200926050-00002
18. Zhou C, Fang L, Chen Y, Zhong J, Wang H, Xie P. Effect of selective serotonin reuptake inhibitors on bone mineral density: a systematic review and meta-analysis. *Osteoporos Int*. 2018;29(6):1243-1251. doi:10.1007/s00198-018-4413-0

19. Saraykar S, John V, Cao B, Hnatow M, Ambrose CG, Rianon N. Association of Selective Serotonin Reuptake Inhibitors and Bone Mineral Density in Elderly Women. *J Clin Densitom*. 2018;21(2):193-199. doi:10.1016/j.jocd.2017.05.016
20. Suissa S, Garbe E. Primer: Administrative health databases in observational studies of drug effects - Advantages and disadvantages. *Nat Clin Pract Rheumatol*. 2007;3(12):725-732. doi:10.1038/ncprheum0652
21. Sylvestre MP, Abrahamowicz M. Flexible modeling of the cumulative effects of time-dependent exposures on the hazard. *Stat Med*. 2009;28(27):3437-3453. doi:10.1002/sim.3701
22. Abrahamowicz M, Beauchamp M-EE, Sylvestre M-PP. Comparison of alternative models for linking drug exposure with adverse effects. *Stat Med*. 2012;31(11-12):1014-1030. doi:10.1002/sim.4343
23. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol*. 2005;58(4):323-337. doi:10.1016/j.jclinepi.2004.10.012
24. Beiderbeck AB, Sturkenboom MCJM, Coebergh JWW, Leufkens HGM, Stricker BHC. Misclassification of exposure is high when interview data on drug use are used as a proxy measure of chronic drug use during follow-up. *J Clin Epidemiol*. 2004;57(9):973-977. doi:10.1016/j.jclinepi.2003.12.017
25. Cohen JM, Wood ME, Hernandez-Diaz S, Nordeng H. Agreement between paternal self-reported medication use and records from a national prescription database. *Pharmacoepidemiol Drug Saf*. 2018;27(4):413-421. doi:10.1002/pds.4411
26. Pacurariu A, Plueschke K, McGettigan P, et al. Electronic healthcare databases in Europe: Descriptive analysis of characteristics and potential for use in medicines regulation. *BMJ Open*. 2018;8(9). doi:10.1136/bmjopen-2018-023090
27. ENCePP Resources Database. Accessed October 5, 2021. http://www.encepp.eu/encepp/resourcesDatabase.jsp
28. Fisher BT, Lindenauer PK, Feudtner C. In-Hospital Databases. In: Strom BL, Kimmel SE, Hennessy S, eds. *Pharmacoepidemiology*. 5th ed. John Wiley & Sons, Ltd; 2012:244-258. doi:10.1002/9781119959946.CH16
29. Ogdie A, Langan SM, Parkinson J, Dattani H, Kostev K, Gelfand JM. Medical Record Databases. In: Strom BL, Kimmel SE, Hennessy S, eds. *Pharmacoepidemiology*. 5th ed. John Wiley & Sons, Ltd; 2012:224-243. doi:10.1002/9781119959946.CH15
30. Out-patient Pharmacy Database - Pharmo. Accessed June 10, 2021. https://pharmo.nl/what-we-have/pharmo-database-network/out-patient-pharmacy-database/
31. Herings RM, Pedersen L. Pharmacy-based medical record linkage systems. In: Strom B, Kimmel S, Hennessy S, eds. *Pharmacoepidemiology*. 5th ed. Wiley; 2012:270–86. Accessed October 5, 2021. www.wiley.com/
32. Pauly NJ, Talbert JC, Brown J. Low-cost generic program use by medicare beneficiaries: Implications for medication exposure misclassification in administrative claims data. *J Manag Care Spec Pharm*. 2016;22(6):741-751. doi:10.18553/jmcp.2016.22.6.741
33. Gamble JM, McAlister FA, Johnson JA, Eurich DT. Restrictive Drug Coverage Policies can Induce Substantial Drug Exposure Misclassification in Pharmacoepidemiologic Studies. *Clin Ther*. 2012;34(6):1379-1386. doi:10.1016/j.clinthera.2012.04.009
34. Gagne JJ. Restrictive Reimbursement Policies: Bias Implications for Claims-Based Drug Safety Studies. *Drug Saf*. 2014;37(10):771-776. doi:10.1007/s40264-014-0220-5
35. Seeger J, Daniel GW. Commercial Insurance Databases. In: Strom BL, Kimmel SE, Hennessy S, eds. *Pharmacoepidemiology*. 5th ed. John Wiley & Sons, Ltd; 2012:189-208. doi:10.1002/9781119959946.CH13
36. Hennessy S, Freeman CP, Cunningham F. US Government Claims Databases. In: Strom BL, Kimmel SE, Hennessy S, eds. *Pharmacoepidemiology*. 5th ed. John Wiley & Sons, Ltd; 2012:209-223. doi:10.1002/9781119959946.CH14

37. Yood MU, Campbell UB, Rothman KJ, et al. Using prescription claims data for drugs available over-the-counter (OTC). *Pharmacoepidemiol Drug Saf.* 2007;16(9):961-968. doi:10.1002/pds.1454
38. Jacobus S, Schneeweiss S, Chan KA. Exposure misclassification as a result of free sample drug utilization in automated claims databases and its effect on a pharmacoepidemiology study of selective COX-2 inhibitors. *Pharmacoepidemiol Drug Saf.* 2004;13(10):695-702. doi:10.1002/pds.981
39. Hampp C, Greene P, Pinheiro SP. Use of prescription drug samples in the USA: A descriptive study with considerations for pharmacoepidemiology. *Drug Saf.* 2016;39(3):261-270. doi:10.1007/s40264-015-0382-9
40. Støvring H, Pottegård A, Hallas J. Determining prescription durations based on the parametric waiting time distribution. *Pharmacoepidemiol Drug Saf.* 2016;25(12):1451-1459. doi:10.1002/pds.4114
41. Vrijens B, De Geest S, Hughes DA, et al. A new taxonomy for describing and defining adherence to medications. *Br J Clin Pharmacol.* 2012;73(5):691-705. doi:10.1111/j.1365-2125.2012.04167.x
42. Degli Esposti L, Piccinni C, Sangiorgi D, Fagiolini A, Buda S. Patterns of Antidepressant Use in Italy: Therapy Duration, Adherence and Switching. *Clin Drug Investig.* 2015;35(11):735-742. doi:10.1007/s40261-015-0332-4
43. Keyloun KR, Hansen RN, Hepp Z, Gillard P, Thase ME, Devine EB. Adherence and Persistence Across Antidepressant Therapeutic Classes: A Retrospective Claims Analysis Among Insured US Patients with Major Depressive Disorder (MDD). *CNS Drugs.* 2017;31(5):421-432. doi:10.1007/s40263-017-0417-0
44. Krivoy A, Balicer RD, Feldman B, et al. The impact of age and gender on adherence to antidepressants: a 4-year population-based cohort study. *Psychopharmacology (Berl).* 2015;232(18):3385-3390. doi:10.1007/s00213-015-3988-9
45. Mårdby A-C, Schiöler L, Sundell KA, Bjerkeli P, Lesén E, Jönsson AK. Adherence to antidepressants among women and men described with trajectory models: a Swedish longitudinal study. *Eur J Clin Pharmacol.* 2016;72(11):1381-1389. doi:10.1007/s00228-016-2106-1
46. Gardarsdottir H, Souverein PC, Egberts TCGG, Heerdink ER. Construction of drug treatment episodes from drug-dispensing histories is influenced by the gap length. *J Clin Epidemiol.* 2010;63(4):422-427. doi:10.1016/j.jclinepi.2009.07.001
47. Pazzagli L, Brandt L, Linder M, et al. Methods for constructing treatment episodes and impact on exposure-outcome associations. *Eur J Clin Pharmacol.* 2020;76(2):267-275. doi:10.1007/s00228-019-02780-4
48. Sinyavskaya L, Matteau A, Johnson S, Durand M. Methodological challenges in assessment of current use of warfarin among patients with atrial fibrillation using dispensation data from administrative health care databases. *Pharmacoepidemiol Drug Saf.* 2018;27(9):979-986. doi:10.1002/pds.4570
49. Laugesen K, Støvring H, Hallas J, et al. Prescription duration and treatment episodes in oral glucocorticoid users: Application of the parametric waiting time distribution. *Clin Epidemiol.* 2017;9:591-600. doi:10.2147/CLEP.S148671
50. Hallas J, Pottegård A, Støvring H. Using probability of drug use as independent variable in a register-based pharmacoepidemiological cause-effect study—An application of the reverse waiting time distribution. *Pharmacoepidemiol Drug Saf.* 2017;26(12):1520-1526. doi:10.1002/pds.4326
51. De Smedt T, Merrall E, Macina D, Perez-Vilar S, Andrews N, Bollaerts K. Bias due to differential and non-differential disease- And exposure misclassification in studies of vaccine effectiveness. *PLoS One.* 2018;13(6). doi:10.1371/journal.pone.0199180
52. Petrone AB, DuCott A, Gagne JJ, Toh S, Maro JC. The Devil's in the details: Reports on reproducibility in pharmacoepidemiologic studies. *Pharmacoepidemiol Drug Saf.* 2019;28(5):671-679. doi:10.1002/pds.4730

53. de Vries F, de Vries C, Cooper C, Leufkens B, van Staa T-P. Reanalysis of two studies with contrasting results on the association between statin use and fracture risk: the General Practice Research Database. *Int J Epidemiol*. 2006;35(5):1301-1308. doi:10.1093/ije/dyl147
54. Pye SR, Sheppard T, Joseph RM, et al. Assumptions made when preparing drug exposure data for analysis have an impact on results: An unreported step in pharmacoepidemiology studies. *Pharmacoepidemiol Drug Saf*. 2018;27(7):781-788. doi:10.1002/pds.4440
55. Eskin M, Simpson SH, Eurich DT. Impact of drug exposure definitions on observed associations in pharmacoepidemiology research. *J Popul Ther Clin Pharmacol*. 2018;25(1):e39-e52. doi:10.22374/1710-6222.25.1.4
56. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c332. doi:10.1136/bmj.c332
57. von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ*. 2007;335(7624):806-808. doi:10.1136/bmj.39335.541782.AD
58. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLOS Med*. 2015;12(10):e1001885. doi:10.1371/journal.pmed.1001885
59. Langan SM, Schmidt SA, Wing K, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). *BMJ*. 2018;363:k3532. doi:10.1136/bmj.k3532
60. Wang S V., Schneeweiss S, Berger ML, et al. Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies V1.0. *Pharmacoepidemiol Drug Saf*. 2017;26(9):1018-1032. doi:10.1002/pds.4295
61. Agha RA, Fowler AJ, Limb C, et al. Impact of the mandatory implementation of reporting guidelines on reporting quality in a surgical journal: A before and after study. *Int J Surg*. 2016;30:169-172. doi:10.1016/j.ijsu.2016.04.032

2.

QUALITY OF DRUG EXPOSURE REPORTING AND STUDY DESIGN

2.1

QUALITY OF REPORTING OF DRUG EXPOSURE IN PHARMACOEPIDEMIOLOGICAL STUDIES.

Mirjam Hempenius, Kim Luijken, Anthonius de Boer, Olaf H. Klungel, Rolf H. H. Groenwold, Helga Gardarsdottir

Pharmacoepidemiol Drug Saf, 2020 Sep; 29(9): 1141–1150

ABSTRACT

Purpose

Exposure definitions vary across pharmacoepidemiological studies. Therefore, transparent reporting of exposure definitions is important for interpretation of published study results. We aimed to assess the quality of reporting of exposure to identify where improvement may be needed.

Method

We systematically reviewed observational pharmacoepidemiological studies that used routinely collected health data, published in 2017 in six pharmacoepidemiological journals. Reporting of exposure was scored using 11 items of the ISPE-ISPOR guideline on reporting of pharmacoepidemiological studies.

Results

Of the 91 studies included, all studies reported the type of exposure (100%), while most reported the exposure risk window (85%) and the exposure assessment window (98%). Operationalization of the exposure window was described infrequently: 16% (14/90) of the studies explicitly reported the presence or absence of an induction period if applicable, 11% (5/47), and 35% (17/49) reported how stockpiling and gaps between exposure episodes were handled, respectively, and 35% (17/49) explicitly mentioned the exposure extension. Switching/add-on was reported in 62% (50/81). How switching between drugs was dealt with and specific drug codes were reported in 52 (57%) and 24 (26%) studies, respectively.

Conclusion

Publications of pharmacoepidemiological studies frequently reported the type of exposure, the exposure risk window, and the exposure assessment window. However, more details on exposure assessment are needed, especially when it concerns the operationalization of the exposure risk window (eg, the presence or absence of an induction period or exposure extension, handling of stockpiling and gaps, and specific codes), to allow for correct interpretation, reproducibility, and assessment of validity.

1. INTRODUCTION

Transparent reporting is important for interpretation of published study results, but also for reproducibility and validity assessment. Reporting guidelines support researchers to describe their research in a transparent and complete manner, like CONSORT,[1] for reporting on clinical trials, STROBE,[2] for reporting on observational studies, and STROBE-RECORD,[3] for reporting on observational studies using routinely collected health data. These guidelines however do not capture the complex operational details required for the conduct of pharmacoepidemiological research, including complex exposure ascertainment algorithms.[4] Recently, two guidelines were published that focus on these complex operational details. The first is a guideline of a joint ISPE-ISPOR Task Force,[5] published in 2017 and the second an extension of the RECORD statement, the RECORD-PE,[4] published in 2018.

Both RECORD-PE and the joint ISPE-ISPOR Task-Force guidelines include a separate section on the reporting of the drug exposure definition, with the latter most specific on operational details of exposure ascertainment. Reporting details about the drug exposure definition is important, since drug exposure can be defined in various ways in observational research, including time-fixed, time-varying, and cumulative drug exposure definitions. In particular regarding time-varying definitions, researchers must make choices how the drug exposure risk window is defined and how gaps or overlapping periods between drug prescriptions or dispensings are being addressed when constructing drug use episodes. As different choices may lead to different effects being estimated,[6–9] it is important that researchers report transparently how exposure was defined to aid correct interpretation of results.

It takes a substantial amount of time to see the effects of published guidelines on transparent reporting in practice. In the case of CONSORT, reporting has improved in the 20 years after the first version was published, but remains suboptimal, with on average 18 of 37 items being reported over the period 2010-2014.[10,11] Also in case of STROBE, reporting has improved after publishing of the guideline, but there is still room for improvement as the median compliance with the 22 items is 77% in 2016, 9 years after STROBE was published.[12] We therefore assessed the quality of exposure assessment reporting according to ISPE-ISPOR Task Force Guidelines, in studies published around the time frame where these guideline were published, to provide a baseline exposure assessment and to determine where improvement may be needed.

2. METHODS

To assess the quality of reporting of pharmacoepidemiological research, we systematically reviewed observational pharmacoepidemiological studies that used routinely collected health data. We used the guideline by the ISPE-ISPOR Task Force to evaluate quality of reporting of exposure, because this guideline aims to facilitate not only validity assessment but also (direct) reproducibility, and therefore is most specific about operationalization of the exposure risk window.

2.1. Journal selection and eligibility of studies

We selected six pharmacoepidemiological journals: *Annals of Pharmacotherapy, British Journal of Clinical Pharmacology, Drug Safety, European Journal of Clinical Pharmacology, Pharmacoepidemiology and Drug Safety, and Pharmacotherapy.* This selection of journals was based on predefined criteria: at least 20 hits in 2017 that met the search criterion in search of routinely collected health data. These 20 hits had to cover at least 5% of the total publications of that specific journal and the journal had to be classified in the category "Pharmacology and pharmacy" at InCites Journal Citation Reports, with an impact factor of at least 2 (Figure S1).[13]

We included all of the studies published in 2017 in these six journals that used routinely collected health data for exposure assessment, such as prescription data, dispensing data or claims data. All studies needed to include at least 250 subjects, to ensure that the exposure assessment was not performed manually. Studies that used questionnaires for exposure assessment were excluded. Studies assessing vaccines were also excluded, as our interest was in reporting of exposure that is used over a certain period of time, whereas vaccines are administered as single administrations.

2.2. Extraction of study characteristics

The following general items were extracted: journal name, word count limit of the article (≤1500 [short report], 1500-3000, 3000-4000 or ≥4000 words), study design (cohort, case-control, case-crossover or other study design), the route of administration of the drug (oral/inhaled or intravenous/subcutaneous), the type of outcome (beneficial or adverse effect), the number of included subjects (categorized as 250-1000, 1001-10 000, 10 001-100 000 and >100 000 subjects), the type of database used for assessment of exposure (claims, General Practitioner [GP], pharmacy or hospital database), and the geographical area where the study was conducted, defined as continents.

Because the ISPE-ISPOR guideline focuses on time-varying exposure, we categorized the studies according to exposure definition in five predefined categories (illustrated in Figure 1):

1. Intention to treat: drug exposure at baseline was included as a time-fixed variable in the model.
2. The presence of ≥1 prescription during a certain period, for example during pregnancy or during the last 12 months prior to the outcome event of interest.
3. Time-varying: episodes of (non)exposure were constructed based on duration of each prescription, without distinguishing between different dosages.

4. Measures of adherence: for example, level of drug exposure was measured as proportion of days a subject has drug in possession divided by the total number of days of follow-up.
5. Dose and cumulative dose: drug exposure was modeled as a continuous or ordinal variable and the effects of different dosages at index date were compared.

This categorization was carried out to notice any difference in reporting between studies with different types of exposure definitions. Characteristics of each included study were extracted independently by two reviewers (M.H. and K.L.).

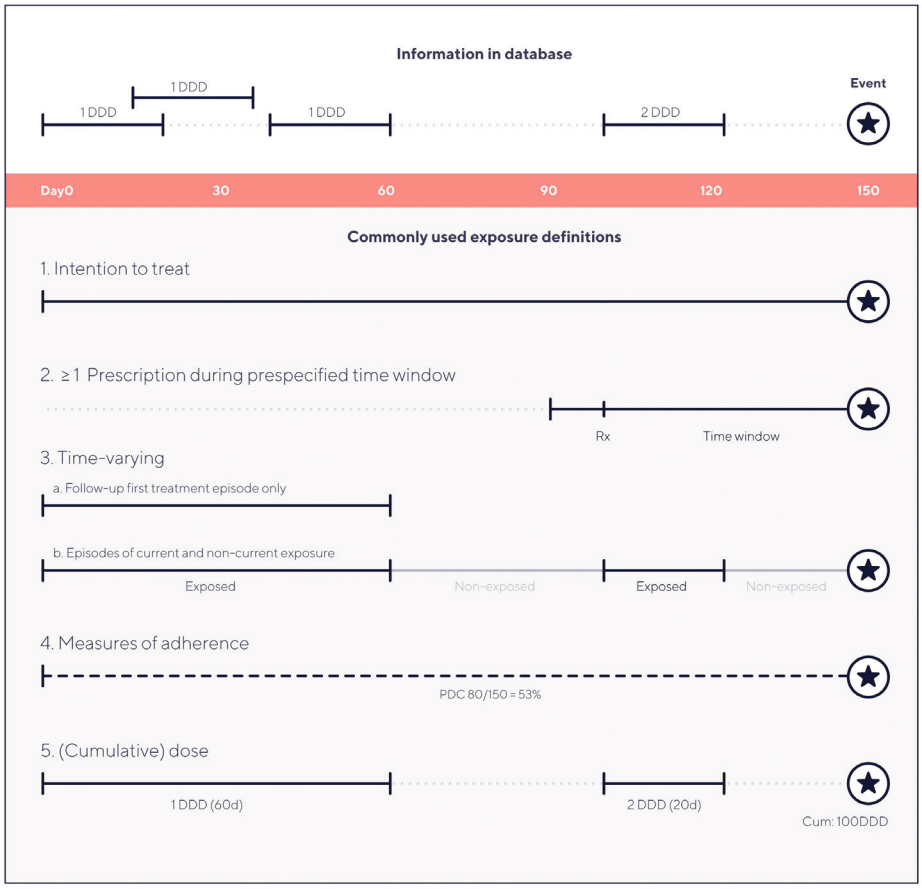


FIGURE 1. Categorization of commonly used exposure definitions in pharmacoepidemiological studies. Different types of exposure definition are applied in pharmacoepidemiological research. We divided these in five categories for further analysis: 1. intention to treat: exposure at baseline is included as a time-fixed variable in the model; 2. the presence of ≥1 prescriptions during a certain time period, for example during pregnancy or during the last 12 months prior to the event; 3. time-varying: episodes of (non)exposure are constructed based on duration of each prescription; 4. measures of adherence: for example, level of exposure is measured as proportion of days covered and 5. dose and cumulative dose: exposure is modeled as a continuous or ordinal variable and the effects of different dosages are compared (time-fixed or time-varying). DDD, daily defined dose; PDC, percentage of days covered; Rx, prescription

2.3. Evaluation of reporting quality

Quality of reporting of exposure was assessed according to the ISPE-ISPOR guideline[5]. All items listed under "Section D – Reporting on exposure definition" were assessed. Item D4 of the guideline contains four elements ("Codes, frequency and temporality of codes, diagnosis position and care setting") and is linked with guideline section C ("Inclusion and exclusion criteria") for further clarification. In this section these items are included as separate items, so we decided to split D4 into four separate items as well. The item "diagnosis position (D4)" was excluded from the final list of items as we considered this item not to be relevant for drug exposures. The resulting eleven items are listed in Table 1.

TABLE 1. Items pertaining to the quality of reporting of exposure definition in pharmaco-epidemiological research. These items are selected from the ISPE-ISPOR Joint Task Force guideline[5]

Item	Explanation	ISPE-ISPOR item
1. Type of exposure	The type of exposure that is captured or measured, for example, drug vs procedure, new use, incident, prevalent, cumulative, time-varying.	D1
2. Exposure risk window (ERW)	The ERW is specific to an exposure and the outcome under investigation. For drug exposures, it is equivalent to the time between the minimum and maximum hypothesized induction time following ingestion of the molecule.	D2
3. Induction period	Days on or following study entry date during which an outcome would not be counted as "exposed time" or "comparator time."	D2a
4. Stockpiling	The algorithm applied to handle leftover days' supply if there are early refills.	D2b
5. Bridging exposure episodes	The algorithm applied to handle gaps that are longer than expected if there was perfect adherence (eg, non-overlapping dispensation + day's supply).	D2c
6. Exposure extension	The algorithm applied to extend exposure past the days' supply for the last observed dispensation in a treatment episode.	D2d
7. Switching/add on	The algorithm applied to determine whether exposure should continue if another exposure begins.	D3
8. Codes	The exact drug, diagnosis, procedure, lab or other codes used to define inclusion/ exclusion criteria.	D4
9. Frequency and temporality of codes	The temporal relation of codes in relation to each other as well as the study entry date (SED). When defining temporality, be clear whether or not the SED is included in assessment windows (eg, occurred on the same day, 2 codes for A occurred within 7 d of each other during the 30 d prior to and including the SED).	D4

TABLE 1. Continued.

Item	Explanation	ISPE-ISPOR item
10. Care setting	The restrictions on codes to those identified from certain settings, for example, inpatient, emergency department, nursing home.	D4
11. Exposure assessment window (EAW)	A time window during which the exposure status is assessed. Exposure is defined at the end of the period. If the occurrence of exposure defines cohort entry, for example, new initiator, then the EAW may be a point in time rather than a period. If EAW is after cohort entry, follow-up window must begin after EAW.	D5

To ensure uniform interpretation of the listed items when assessing the articles, eight randomly chosen articles were reviewed independently by two reviewers (M.H. and K.L.) and discrepancies of the scores were discussed. This resulted in a formalized data extraction form that was used for the remaining articles.

Of each reviewed article, we scrutinized the methods sections and, if referenced to, the supplementary materials for the data extraction. Each of the 11 items was scored as "not reported," "reported," or "not applicable." An item could be scored as "not applicable" if this item was not relevant for that specific study. For example, if exposure was defined as receiving 1 or more drug prescriptions, it was not relevant to report how stockpiling and handling gaps were dealt with. The items 3 (induction period, D2a) and 6 (exposure extension, D2d) (see Table 1) could also be mentioned implicitly. For example, if an author stated that "the follow-up started on the day of the first prescription and ended after the duration of the last prescription," it is implicit that there was no induction period and no extension of the exposure risk window. For items 3 and 6 we therefore also scored whether reporting was "explicit" or "implicit."

All articles were reviewed independently by two reviewers (M.H. and K.L.) and discrepancies were discussed until consensus was reached. The interobserver agreement (kappa) was 0.53. The eight studies that were used to refine the assessment tool were excluded from the calculation of the kappa.

In addition to the 11 ISPE-ISPOR items, we assessed whether the exposure definition was accompanied by a figure for graphical representation, since this is recommended in the ISPE-ISPOR guideline for study design in general (item B1).

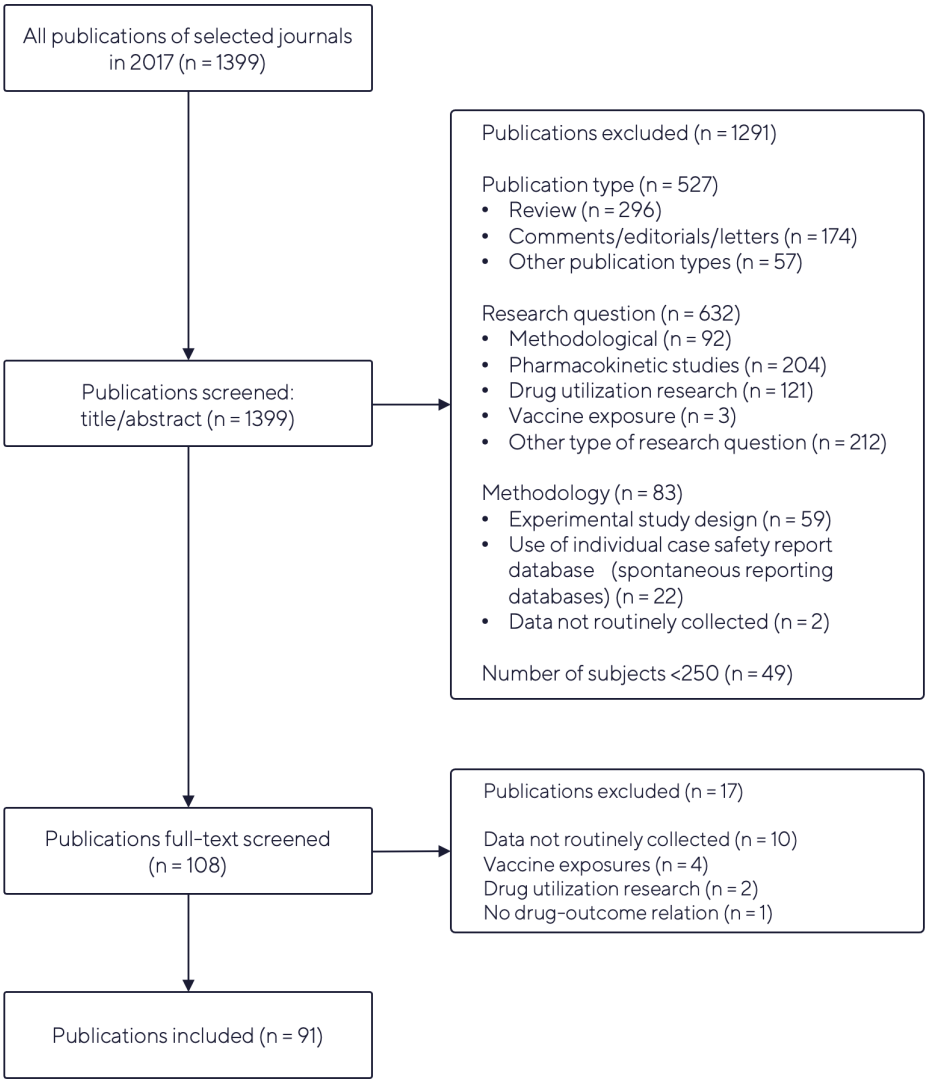


FIGURE 2. Flow chart of the search and screening process to select pharmacoepidemiological studies using routinely collected data.
All articles published in 2017 in the following six journals were included in the first step: Annals of Pharmacotherapy, British Journal of Clinical Pharmacology, Drug Safety, European Journal of Clinical Pharmacology, Pharmacoepidemiology and Drug Safety, and Pharmacotherapy.

2.4. Data analysis

For each of the ISPE-ISPOR reporting items, the primary outcome was the percentage of articles that reports this item, where applicable. The results were stratified by study design, the number of patients included, the type of outcome, the route of administration, the exposure definition, the type of database used, and word limit of the article. The percentage of studies that included a graphical presentation of the exposure definition was considered as a secondary outcome. In

the case of multiple types of exposure, designs or types of databases within one publication, we analyzed them as one unit within the main analysis, including all information that was mentioned in the publication. For stratification purposes, we only used the information provided for that specific design, exposure type or database.

3. RESULTS

3.1. Selection and characteristics of studies

A total of 91 articles were included (Figure 2, see supplementary materials for all references). The characteristics of all 91 articles are summarized in Table 2. Different types of exposure were applied; 24 (26%) studies performed an intention to treat analysis, 43 (47%) studies applied a time-varying exposure assessment; 19 (21%) studies assessed the occurrence of one or more prescriptions during a certain period, 4 (4%) used measures of adherence as exposure, and 3 (3%) investigated the effect of (cumulative) dose. Two studies applied multiple definitions in their study.

TABLE 2. Characteristics of the studies included for evaluation of quality of reporting of pharmacoepidemiological studies (n = 91)

Journal	n (%)
Annals of Pharmacotherapy	7 (8)
British Journal of Clinical Pharmacology	16 (18)
Drug Safety	8 (9)
European Journal of Clinical Pharmacology	17 (19)
Pharmacoepidemiology and Drug Safety	27 (30)
Pharmacotherapy	16 (18)
Design [a]	
Cohort	64 (70)
Case-control	25 (28)
Case-crossover	4 (4)
Type of outcome	
Beneficial effects	18 (20)
Adverse effects	67 (74)
Beneficial and adverse effects	6 (7)
Number of subjects included	
250-1000	13 (14)
1001-10 000	30 (33)
10 001-100 000	24 (26)
>100 000	24 (26)

TABLE 2. Continued.

Type of database [a]	n (%)
Claims database	41 (44)
GP database	17 (19)
Hospital database	18 (20)
Pharmacy database	16 (18)
Unclear	2 (2)
Geographical area [a]	
Europe	39 (43)
Asia	15 (17)
North America	36 (40)
Australia	2 (2)
Route of administration	
Oral and inhaled	80 (88)
Intravenous and subcutaneous	11 (12)
Exposure definition [a]	
Intention to treat	24 (26)
≥1 prescription/dispense during a certain period	19 (21)
Time-varying	43 (47)
Measures of adherence	4 (4)
(Cumulative) dose at index date	3 (3)

[a] Sum of n may exceed 91.

3.2. Reporting quality

An average of 6.6 (SD 1.8) items were reported out of the 11 items pertaining to quality of reporting of exposure definition. The median number of items reported was 7, ranging between 2 and 10 per study. The reporting of each item is presented in Table 3. Most studies reported the type of exposure (eg, current use, cumulative dose) (n = 91, 100%), the exposure risk window in general terms (n = 77, 85%), and the exposure assessment window (n = 89, 98%). The operationalization of the exposure window was infrequently described: of 90 studies that should report on an induction period, 14 (16%) studies explicitly reported the presence or absence of an induction period, and another 67 (74%) reported this implicitly. Among the 49 studies where exposure extension was possible, 17 (35%) studies reported explicitly how long the exposure was extended and 10 (20%) studies mentioned this implicitly. Stockpiling and bridging of exposure episodes was reported in 5 of 47 (11%) and 18 of 44 (41%) studies. How switching between drugs or add-on was dealt with was reported in 50 of 81 (62%) studies where this item was applicable. Specific drug codes and care setting were reported in 24 of 91 (26%) and 67 of 91 (74%) studies. Temporality of codes was reported in 77 of 91 (85%) studies.

Eleven studies (12%) supported the reporting of their exposure definition with a graphical representation, nine of them in the article itself and two in the supplementary materials.

TABLE 3. Quality of Reporting of exposure for the included studies. For each specific item, the number of studies reporting that item is shown. (n = 91)

Item	Studies, n[b]	Reported, n (%)
1. Type of exposure	91	91 (100)
2. Exposure risk window (ERW)	91	77 (85)
3. Induction period[a]	90	81 (90)
Explicit		14 (16)
Implicit		67 (74)
4. Stockpiling	47	5 (11)
5. Bridging exposure episodes	44	18 (41)
6. Exposure extension[a]	49	27 (55)
Explicit		17 (35)
Implicit		10 (20)
7. Switching/ add on	81	50 (62)
8. Codes	91	24 (26)
9. Frequency and temporality of codes	91	77 (85)
10. Care setting	91	67 (74)
11. Exposure Assessment Window (EAW)	91	89 (98)

[a] When explicitly mentioning an induction period, a period after the index date is clearly excluded in the exposure risk window. Stating that follow-up started on the day of the first prescription implies implicitly that there was no induction period. The same reasoning applies to the extension period.
[b] Total number of studies to which this item was applicable.

3.3. Stratification by study characteristics

The exposure definition determined which details needed to be reported regarding the exposure assessment. Stratification by exposure definition showed that studies using time-varying definitions report on average more items compared with all other definitions (7.4 (SD 1.7) vs 6.0 (SD 1.7), Table S1). The items stockpiling (item 4) and handling gaps (item 5) were considered to be relevant only for the time-varying definitions, where they were reported in 7% and 42% of the studies respectively. Exposure extension (item 6) was also reported more often in the studies with a time-varying exposure assessment (42%) vs studies with another exposure definition (6%).

Stratification by route of administration also showed differences, studies on intravenous or subcutaneous administered drugs reported less frequently on nearly all items than studies on oral or inhaled drugs.

Stratification by study design, number of subjects included in the study, type of outcome, type of database used, and word limit of the article, did not reveal major differences. The results of the stratified analyses are available in Tables S2-S7.

4. DISCUSSION

This systematic review of quality of reporting of drug exposure in pharmacoepidemiological studies showed that none of the studies assessed met all requirements of reporting of drug exposure as defined by the ISPE-ISPOR guideline. The number of reported items varied widely between studies, ranging from 2 to 10. In general, the conceptual details about the exposure risk window and the exposure assessment window were reported relatively often (85% and 98%, respectively). However, the operational details concerning the construction of the exposure risk window were reported less often. For example, handling gaps and overlapping episodes were reported in only 11% and 41% of studies, where this type of reporting was applicable, thereby impeding reproducibility.

Our findings on the substandard quality of reporting of pharmacoepidemiological database studies are in line with the results of a study by Wang et al.[14] In their attempt to reproduce 31 pharmacoepidemiological database studies, they noted that code lists for outcomes, covariates and inclusion/exclusion criteria were reported in only 11 of 31 studies (35%). Likewise, we found that code lists for drug exposure were only reported in 24 of 91 studies (26%). Although not all details were reported, Wang et al. were able to reproduce several database studies with high accuracy, but mention that this was partly due to *"the efforts of the reproduction team, a group of pharmacoepidemiologists with decades of experience, to make informed guesses regarding variable definitions or other key decisions when these were not clearly specified in the original articles."*[14] It is debatable whether this level of expertise could be expected from the general reader of pharmacoepidemiological studies. Therefore, it is important that details are reported clearly for correct interpretation and reproducibility of study results.

Besides reporting operational details, it is important to clearly describe the choices regarding the exposure definition, such as the exposure risk window and what type of exposure is examined. Currently there are a number of guidelines to support these methodological choices, like the ISPE Guidelines for good pharmacoepidemiology practices,[15] the ENCePP Guide on Methodological Standards in Pharmacoepidemiology,[16] the FDA guidance for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data,[17] the User's Guide for Developing a Protocol for Observational Comparative Effectiveness Research by the Agency for Healthcare Research and Quality,[18] and the EU Good Pharmacovigilance Practice (Module VIII).[19] Without clear reporting of all key decisions, assessment of the validity of the results will be difficult.

A strength of this study was the independent assessment of all studies by two researchers. This also revealed one of the limitations of this study: substantial interpretation was needed by the researchers to score all studies, which is also reflected in a moderate kappa of 0.53. This could partially be explained by the fact that not all questions of the checklist applied to each study included in our study. The listed guideline items could be scored most easily for a study design with time-varying treatment episodes. The items 4 (stockpiling, D2b), 5 (handling gaps, D2c), 6

(exposure extension, D2d) and 7 (switching, D3) were for other types of exposure, (eg, intention to treat) not relevant and thus scored NA. This is also reflected in the kappa and the percentage agreement of these items (Table S8). When we recalculated kappa, with only a contrast between reporting something (Yes) or not (No or NA), this resulted in a kappa of 0.64.

There was also a difference between the reporting of exposure to drugs that were oral or inhaled administered, compared with the reporting of exposure drugs that were intravenous or subcutaneous administered. This might be explained by the fact that intravenous or subcutaneous drugs are commonly identified by procedural codes instead of drug dispensing information. These data contain other information about the drug exposure, resulting in also another way of reporting of the drug exposure assessment, which might not be captured in the guideline used for this review.

Another possible limitation concerns the inclusion of only publications in six pharmacoepidemiological journals. The results may thus not be generalizable to the quality of reporting of drug exposure in general. Furthermore, we only searched the methods and supplementary materials (if referenced) for exposure assessment information, possibly missing out on information described in other sections of the publication. For transparency reasons, it is however still recommended to describe all methodological choices in the methods section. In addition, it might also be possible that these details are described in other study reports, such as reports provided to the regulator, but are left out of the publication, due to word count limitations. We did, however, not see differences in results between publications in journals with a strict word limit (≤3000 words) compared to publications in journals with less stringent word limits.

Suboptimal reporting is not unique to pharmacoepidemiological research and the effort for more transparent reporting has facilitated the development of various reporting guidelines, such as CONSORT and STROBE. To further stimulate use of these, endorsement by many journals has resulted in improved reporting, but after two decades, adherence to CONSORT is still suboptimal.[20–22] In order to accelerate adherence to RECORD-PE and the ISPE-ISPOR guideline, it might be considered to oblige authors to use one of these two guidelines. Four of the six included journals (*Annals of Pharmacotherapy, British Journal of Clinical Pharmacology, Drug Safety, Pharmacoepidemiology and Drug Safety*) currently recommend authors to adhere to the guidelines available through the EQUATOR network, including the RECORD-PE guideline. One journal (*European Journal of Clinical Pharmacology*) advises to adhere to CONSORT for observational research and one journal (*Pharmacotherapy*) does not recommend a specific reporting guideline. In addition, current good practices can be used as examples. We summarized some good practices of clear reporting of exposure assessment in Textbox 1, which can be helpful for future studies. We also noticed that giving arguments for specific choices was helpful for the interpretation of the conceptual choices, as was the inclusion of a graphical representation for the interpretation of the operational choices.

To conclude, we recommend that publications of pharmacoepidemiological studies should include more details on exposure ascertainment, especially about the operationalization of the exposure risk window (eg, the presence or absence of an induction period or exposure extension, handling of stockpiling and gaps, and specific codes), to allow correct interpretation

of the results and to enable reproducibility, and validity assessment. Authors, reviewers, and editors are encouraged to pay more attention to adhere to relevant reporting guidelines such as the ISPE-ISPOR and RECORD-PE guidelines.

TEXTBOX 1. Examples of good practices for each of the items in the ISPE-ISPOR checklist cited from included articles.

1 Type of exposure	"Those who filled a prescription for an antidepressant during this period <January 1, 2007 and December 31, 2013> with no such fills during the preceding year were considered treatment initiators."[23]
	"In the first model, the mutually exclusive binary indicators of use for each NSAID were (1) current use on the index date, (2) recent use 1 to 30 days ago, (3) past use 31 to 180 days ago, or (4) no use in the last 180 days before the index date."[24]
2 Exposure risk window (ERW)	"For each patient, we defined a period of continuous drug use beginning with the first prescription after their 66th birthday and ending with death, discontinuation of treatment, the end of the study period (March 31, 2014), or 90 days of follow-up, whichever occurred first. <...> We based our selection of a 90-day observation window on existing literature describing heart failure and edema within a few months of pregabalin therapy."[25]
	"The primary outcome was hospitalization for COPD or pneumonia within 30 days after the index date <...>"[26]
3 Induction period	"<...> the effect of insulin on chronic complications may take some time, so we conducted a lag-time analysis, whereby patients with chronic complication events that occurred 3 years after the initiation of insulin were excluded."[27]
	"Outcomes were collected starting 30 days following the index date to ensure that events occurring during the baseline period were not mistakenly captured as study period events."[28]
4 Stockpiling	"For overlapping prescriptions, the individual was assumed to have completed the former one before starting the second."[29]
	"To account for gaps and overlaps in redemptions due to incomplete adherence or lost prescriptions, we presumed that health-insured persons have drug stocks lasting up to 15 days due to incomplete compliance ('15- day rule'), added apparent overlaps up to a maximum overlap duration corresponding to 25% of the quantity of the last overlapping prescription, and applied common recommendations to fill apparent gaps between prescriptions using prospective filling."[30]
5 Bridging exposure episodes	"Discontinuation of use <was> defined as a 60-day gap between the end of one COC prescription and the next COC prescription."[31]
6 Exposure extension	"Observation was extended by half the days supplied from the final prescription to capture outcomes that may have prompted cessation of therapy."[25]

TEXTBOX 1. Continued.

7	Switching/ add on	"If a patient switched from warfarin to rivaroxaban or vice versa during the study period, that was considered discontinuation of the index drug, and they were censored at that time."[32]
		"To assess whether associations varied with different antidepressants, we categorized antidepressants into 3 types (SSRI monotherapy, non-SSRI monotherapy, or both SSRI and non-SSRI antidepressants)."[33]
8	Codes	"Antihypertensive drugs studied were: angiotensin-converting enzyme (ACE) inhibitors: ATC code C09A and C09B, angiotensin receptors blockers (ARBs): ATC code C09C and C09D, calcium channel blockers (CCBs): ATC code C08, β-blockers: ATC code C07, diuretics: ATC code C03 (thiazide or thiazide-like diuretics, loop diuretics and potassium-sparing diuretics) and miscellaneous antihypertensive agents: ATC code C02."[34]
		"We selected all patients who had ever received Fz/Cz <...> according to the Anatomical Therapeutic Chemical (ATC) codes N07CA03 (for Fz) and N07CA02 (for Cz) <...>."[35]
9	Frequency and temporality of codes	"All patients included in the Cohort were followed from the 90th day after the incident ACS occurrence (index date) until the incidence of a major adverse cardiac event (MACE), death, date removed from the database or 31 December 2013, whichever came first."[36]
10	Care setting	"Because of the high patient pharmacy loyalty in the Netherlands, the prescription records for each patient in the database are virtually complete, except for over-the-counter (OTC) drugs and drugs dispensed during hospitalization."[37]
11	Exposure Assessment Window (EAW)	"As diagnosis and treatment start may be registered in different days <...>, we allowed a time interval of ±3 months from diagnosis date and start of treatment."[38]
		"We conducted a <study of residents> prescribed digoxin at any time between January 1, 1994 and December 31, 2012, the last date for which complete data were available."[39]

Author's contribution

MH designed the study, performed the literature search, conducted the evaluation of reporting quality in collaboration with KL, wrote the first draft of the manuscript, and implemented the contribution of the co-authors and external reviewers up to final publication. During the whole process she implemented input and feedback from the other contributors to this study.

REFERENCES

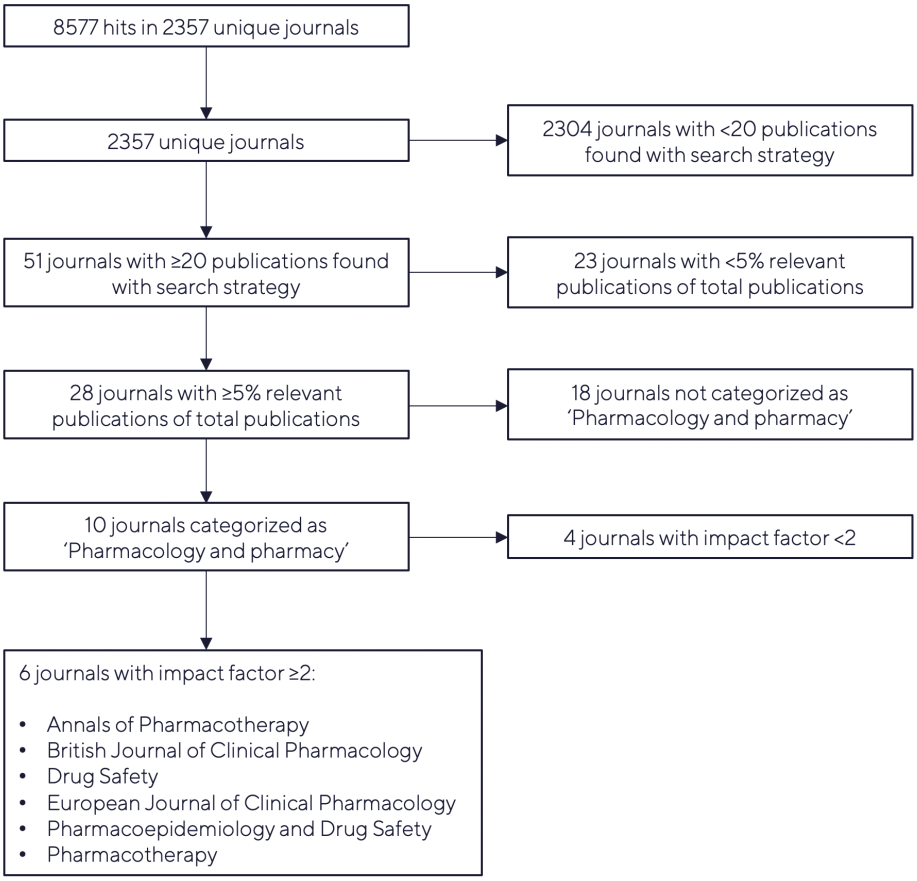
1. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c332. doi:10.1136/bmj.c332
2. von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ*. 2007;335(7624):806-808. doi:10.1136/bmj.39335.541782.AD
3. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLOS Med*. 2015;12(10):e1001885. doi:10.1371/journal.pmed.1001885
4. Langan SM, Schmidt SA, Wing K, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). *BMJ*. 2018;363:k3532. doi:10.1136/bmj.k3532
5. Wang S V., Schneeweiss S, Berger ML, et al. Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies V1.0. *Pharmacoepidemiol Drug Saf*. 2017;26(9):1018-1032. doi:10.1002/pds.4295
6. Gardarsdottir H, Souverein PC, Egberts TCGG, Heerdink ER. Construction of drug treatment episodes from drug-dispensing histories is influenced by the gap length. *J Clin Epidemiol*. 2010;63(4):422-427. doi:10.1016/j.jclinepi.2009.07.001
7. Petrone AB, DuCott A, Gagne JJ, Toh S, Maro JC. The Devil's in the details: Reports on reproducibility in pharmacoepidemiologic studies. *Pharmacoepidemiol Drug Saf*. 2019;28(5):671-679. doi:10.1002/pds.4730
8. de Vries F, de Vries C, Cooper C, Leufkens B, van Staa T-P. Reanalysis of two studies with contrasting results on the association between statin use and fracture risk: the General Practice Research Database. *Int J Epidemiol*. 2006;35(5):1301-1308. doi:10.1093/ije/dyl147
9. Pye SR, Sheppard T, Joseph RM, et al. Assumptions made when preparing drug exposure data for analysis have an impact on results: An unreported step in pharmacoepidemiology studies. *Pharmacoepidemiol Drug Saf*. 2018;27(7):781-788. doi:10.1002/pds.4440
10. Alam M, Rauf M, Ali S, et al. A Systematic Review of Completeness of Reporting in Randomized Controlled Trials in Dermatologic Surgery: Adherence to CONSORT 2010 Recommendations. *Dermatol Surg*. 2016;42(12):1325-1334. doi:10.1097/DSS.0000000000000902
11. Plint AC, Moher D, Morrison A, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust*. 2006;185(5):263-267. doi:10.5694/J.1326-5377.2006.TB00557.X
12. Agha RA, Fowler AJ, Limb C, et al. Impact of the mandatory implementation of reporting guidelines on reporting quality in a surgical journal: A before and after study. *Int J Surg*. 2016;30:169-172. doi:10.1016/J.IJSU.2016.04.032
13. Clarivate. InCites Journal Citation Reports. Published 2018. Accessed December 4, 2018. https://jcr.clarivate.com/
14. Wang S, Verpillat P, Rassen J, Patrick A, Garry E, Bartels D. Transparency and Reproducibility of Observational Cohort Studies Using Large Healthcare Databases. *Clin Pharmacol Ther*. 2016;99(3):325-332. doi:10.1002/cpt.329
15. ISPE. Guidelines for good pharmacoepidemiology practices (GPP). *Pharmacoepidemiol Drug Saf*. 2008;17(2):200-208. doi:10.1002/pds.1471
16. European Medicines Agency. *The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology, 9th Revision.*; 2010. Accessed June 1, 2018. http://www.encepp.eu/standards_and_guidances

17. U.S. Department of Health and Human Services, Administration Food and Drug, Research (CDER) Center for Drug Evaluation and, Research (CBER) Center for Biologics Evaluation and. Guidance for Industry and FDA Staff Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data. *Drug Saf*. Published online 2013. Accessed June 18, 2019. http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm
18. Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. Published online 2013. Accessed May 3, 2019. https://www.ncbi.nlm.nih.gov/books/NBK126190/
19. European Medicines Agency. *Guideline on Good Pharmacovigilance Practices (GVP) - Module VIII – Post-Authorisation Safety Studies (Rev 3).*; 2017. Accessed October 31, 2019. www.ema.europa.eu
20. Huang YQ, Traore K, Ibrahim B, Sewitch MJ, Nguyen LHP. Reporting quality of randomized controlled trials in otolaryngology: review of adherence to the CONSORT statement. *J Otolaryngol - Head Neck Surg*. 2018;47(1):34. doi:10.1186/s40463-018-0277-8
21. Pouwels KB, Widyakusuma NN, Groenwold RHH, Hak E. Quality of reporting of confounding remained suboptimal after the STROBE guideline. *J Clin Epidemiol*. 2016;69:217-224. doi:10.1016/J.JCLINEPI.2015.08.009
22. Adams AD, Benner RS, Riggs TW, Chescheir NC. Use of the STROBE Checklist to Evaluate the Reporting Quality of Observational Research in Obstetrics. *Obstet Gynecol*. 2018;132(2):507-512. doi:10.1097/AOG.0000000000002689
23. Hedna K, Andersson Sundell K, Hamidi A, Skoog I, Gustavsson S, Waern M. Antidepressants and suicidal behaviour in late life: a prospective population-based study of use patterns in new users aged 75 and above. *Eur J Clin Pharmacol*. 2018;74(2):201-208. doi:10.1007/s00228-017-2360-x
24. Bally M, Beauchamp M-E, Abrahamowicz M, Nadeau L, Brophy JM. Risk of acute myocardial infarction with real-world NSAIDs depends on dose and timing of exposure. *Pharmacoepidemiol Drug Saf*. 2018;27(1):69-77. doi:10.1002/pds.4358
25. Ho JM-W, Macdonald EM, Luo J, et al. Pregabalin and heart failure: A population-based study. *Pharmacoepidemiol Drug Saf*. 2017;26(9):1087-1092. doi:10.1002/pds.4239
26. Vozoris NT, Wang X, Austin PC, et al. Incident diuretic drug use and adverse respiratory events among older adults with chronic obstructive pulmonary disease. *Br J Clin Pharmacol*. 2018;84(3):579-589. doi:10.1111/bcp.13465
27. Ou H-T, Lee T-Y, Du Y-F, Li C-Y. Comparative risks of diabetes-related complications of basal insulins: a longitudinal population-based cohort of type 1 diabetes 1999-2013 in Taiwan. *Br J Clin Pharmacol*. 2018;84(2):379-391. doi:10.1111/bcp.13461
28. Samp JC, Joo MJ, Schumock GT, Calip GS, Pickard AS, Lee TA. Comparative Effectiveness of Long-Acting Beta $_2$ -Agonist Combined with a Long-Acting Muscarinic Antagonist or Inhaled Corticosteroid in Chronic Obstructive Pulmonary Disease. *Pharmacother J Hum Pharmacol Drug Ther*. 2017;37(4):447-455. doi:10.1002/phar.1913
29. Rea F, Bonassi S, Vitale C, et al. Exposure to statins is associated to fracture risk reduction in elderly people with cardiovascular disease: evidence from the AIFA-I-GrADE observational project. *Pharmacoepidemiol Drug Saf*. 2017;26(7):775-784. doi:10.1002/pds.4206
30. Meid AD, von Medem A, Heider D, et al. Investigating the Additive Interaction of QT-Prolonging Drugs in Older People Using Claims Data. *Drug Saf*. 2017;40(2):133-144. doi:10.1007/s40264-016-0477-y
31. Larivée N, Suissa S, Coulombe J, Tagalakis V, Filion KB. Drospirenone-Containing Oral Contraceptive Pills and the Risk of Venous Thromboembolism: An Assessment of Risk in First-Time Users and Restarters. *Drug Saf*. 2017;40(7):583-596. doi:10.1007/s40264-017-0525-2
32. Russo-Alvarez G, Martinez KA, Valente M, et al. Thromboembolic and Major Bleeding Events With Rivaroxaban Versus Warfarin Use in a Real-World Setting. *Ann Pharmacother*. 2018;52(1):19-25. doi:10.1177/1060028017727290
33. Momen NC, Munk-Olsen T, Li J, et al. Antidepressant use during pregnancy and childhood cancer in the offspring. *Pharmacoepidemiol Drug Saf*. 2018;27(1):114-118. doi:10.1002/pds.4352

34. Alharbi FF, Souverein PC, de Groot MCH, et al. The impact of serum potassium-influencing antihypertensive drugs on the risk of out-of-hospital cardiac arrest: A case-control study. *Br J Clin Pharmacol*. 2017;83(11):2541-2548. doi:10.1111/bcp.13356
35. Lin H-L, Lin H-C, Tseng Y-F, Chen S-C, Hsu C-Y. Risk of parkinsonism induced by flunarizine or cinnarizine: a population-based study. *Eur J Clin Pharmacol*. 2017;73(3):365-371. doi:10.1007/s00228-016-2181-3
36. Bezin J, Groenwold RHH, Ali MS, et al. Comparative effectiveness of recommended versus less intensive drug combinations in secondary prevention of acute coronary syndrome. *Pharmacoepidemiol Drug Saf*. 2017;26(3):285-293. doi:10.1002/pds.4171
37. Tran Y-H, Groen H, Bergman JEH, Hak E, Wilffert B. Exposure to reactive intermediate-inducing drugs during pregnancy and the incident use of psychotropic medications among children. *Pharmacoepidemiol Drug Saf*. 2017;26(3):265-273. doi:10.1002/pds.4161
38. Giner-Soriano M, Roso-Llorach A, Vedia Urgell C, et al. Effectiveness and safety of drugs used for stroke prevention in a cohort of non-valvular atrial fibrillation patients from a primary care electronic database. *Pharmacoepidemiol Drug Saf*. 2017;26(1):97-107. doi:10.1002/pds.4137
39. Quinn KL, Macdonald EM, Gomes T, et al. Macrolides, Digoxin Toxicity and the Risk of Sudden Death: A Population-Based Study. *Drug Saf*. 2017;40(9):835-840. doi:10.1007/s40264-017-0539-9

SUPPLEMENTARY MATERIALS

FIGURE S1. Flow chart of the search and screening process to select journals



Selection based on the following characteristics:
1. At least 20 hits obtained within 2017 with the next search strategy: *(pharmacoepidemiology OR pharmacy claims OR pharmacy data OR dispensing data OR computerized data OR computerized database OR administrative claims OR prescription claims OR prescription database OR prescription data OR health database OR health care database OR health care claims OR insurance plan OR Medicaid OR managed care Organization OR Veterans Affairs) AND (effectiveness OR efficacy OR side-effects OR safety OR unintended effects) AND English [language]* (based on Andrade, Kahler, Frech, & Chan, 2006)
2. Ratio of the number of publications found with this strategy and the total number of publications > 0.05 *(Note: some journals had> 20 publications with this search strategy, but compared to the total number of publications, this was only a small part. That is why we only searched for journals with a "significant" (> 5%) share of relevant hits)*
3. Categorized in journal category Pharmacology and pharmacy at InCites Journal Citation Reports.[13]
4. Impact Factor in 2017 >2 (Journal Citation Reports)

List of all included studies

1. Admassie E, Chalmers L, Bereznicki LR. Bleeding-related admissions in patients with atrial fibrillation receiving antithrombotic therapy: results from the Tasmanian Atrial Fibrillation (TAF) study. *Eur J Clin Pharmacol*. 2017;73(12):1681-1689. doi:10.1007/s00228-017-2337-9
2. Alharbi FF, Souverein PC, de Groot MCH, et al. The impact of serum potassium-influencing antihypertensive drugs on the risk of out-of-hospital cardiac arrest: A case-control study. *Br J Clin Pharmacol*. 2017;83(11):2541-2548. doi:10.1111/bcp.13356
3. Amer KA, Hurren JR, Edwin SB, Cohen G. Regadenoson versus Dipyridamole: A Comparison of the Frequency of Adverse Events in Patients Undergoing Myocardial Perfusion Imaging. *Pharmacother J Hum Pharmacol Drug Ther*. 2017;37(6):657-661. doi:10.1002/phar.1940
4. Bakhriansyah M, Souverein PC, de Boer A, Klungel OH. Gastrointestinal toxicity among patients taking selective COX-2 inhibitors or conventional NSAIDs, alone or combined with proton pump inhibitors: a case-control study. *Pharmacoepidemiol Drug Saf*. 2017;26(10):1141-1148. doi:10.1002/pds.4183
5. Bali V, Chatterjee S, Johnson ML, Chen H, Carnahan RM, Aparasu RR. Risk of Mortality in Elderly Nursing Home Patients with Depression Using Paroxetine. *Pharmacother J Hum Pharmacol Drug Ther*. 2017;37(3):287-296. doi:10.1002/phar.1898
6. Bally M, Beauchamp M-E, Abrahamowicz M, Nadeau L, Brophy JM. Risk of acute myocardial infarction with real-world NSAIDs depends on dose and timing of exposure. *Pharmacoepidemiol Drug Saf*. 2018;27(1):69-77. doi:10.1002/pds.4358
7. Bérard A, Sheehy O, Zhao J-P, Vinet É, Bernatsky S, Abrahamowicz M. SSRI and SNRI use during pregnancy and the risk of persistent pulmonary hypertension of the newborn. *Br J Clin Pharmacol*. 2017;83(5):1126-1133. doi:10.1111/bcp.13194
8. Bezin J, Groenwold RHH, Ali MS, et al. Comparative effectiveness of recommended versus less intensive drug combinations in secondary prevention of acute coronary syndrome. *Pharmacoepidemiol Drug Saf*. 2017;26(3):285-293. doi:10.1002/pds.4171
9. Biffi A, Rea F, Scotti L, et al. Antidepressants and the risk of arrhythmia in elderly affected by a previous cardiovascular disease: a real-life investigation from Italy. *Eur J Clin Pharmacol*. 2018;74(1):119-129. doi:10.1007/s00228-017-2352-x
10. Brassard P, Wu JW, Ernst P, Dell'Aniello S, Smiechowski B, Suissa S. The effect of statins on influenza-like illness morbidity and mortality. *Pharmacoepidemiol Drug Saf*. 2017;26(1):63-70. doi:10.1002/pds.4112
11. Busby J, Murray L, Mills K, Zhang S-D, Liberante F, Cardwell CR. A combined connectivity mapping and pharmacoepidemiology approach to identify existing medications with breast cancer causing or preventing properties. *Pharmacoepidemiol Drug Saf*. 2018;27(1):78-86. doi:10.1002/pds.4345
12. Cantarutti A, Merlino L, Giaquinto C, Corrao G. Use of antidepressant medication in pregnancy and adverse neonatal outcomes: A population-based investigation. *Pharmacoepidemiol Drug Saf*. 2017;26(9):1100-1108. doi:10.1002/pds.4242
13. Carman WJ, Accortt NA, Anthony MS, Iles J, Enger C. Pregnancy and infant outcomes including major congenital malformations among women with chronic inflammatory arthritis or psoriasis, with and without etanercept use. *Pharmacoepidemiol Drug Saf*. 2017;26(9):1109-1118. doi:10.1002/pds.4261
14. Castelpietra G, Gobbato M, Valent F, De Vido C, Balestrieri M, Isacsson G. Antidepressant use in suicides: a case-control study from the Friuli Venezia Giulia Region, Italy, 2005–2014. *Eur J Clin Pharmacol*. 2017;73(7):883-890. doi:10.1007/s00228-017-2236-0
15. Cieri NE, Kusmierski K, Lackie C, Van Opdorp A, Hassan AK. Retrospective Evaluation of Postoperative Adverse Drug Events in Patients Receiving Rivaroxaban After Major Orthopedic Surgery Compared with Standard Therapy in a Community Hospital. *Pharmacother J Hum Pharmacol Drug Ther*. 2017;37(2):170-176. doi:10.1002/phar.1888
16. Coons JC, Iasella CJ, Chanas T, et al. Comparative Effectiveness and Safety Analysis of Dual Antiplatelet Therapies Within an Integrated Delivery System. *Ann Pharmacother*. 2017;51(8):649-655. doi:10.1177/1060028017706977

17. EuroCoord TEP and PHCC (EPPICC) study group in. Safety of zidovudine/lamivudine scored tablets in children with HIV infection in Europe and Thailand. *Eur J Clin Pharmacol.* 2017;73(4):463-468. doi:10.1007/s00228-016-2182-2
18. Fang G, Annis IE, Farley JF, et al. Incidence of and Risk Factors for Severe Adverse Events in Elderly Patients Taking Angiotensin-Converting Enzyme Inhibitors or Angiotensin II Receptor Blockers after an Acute Myocardial Infarction. *Pharmacother J Hum Pharmacol Drug Ther.* 2018;38(1):29-41. doi:10.1002/phar.2051
19. Ferrajolo C, Verhamme KMC, Trifirò G, et al. Antibiotic-Induced Liver Injury in Paediatric Outpatients: A Case-Control Study in Primary Care Databases. *Drug Saf.* 2017;40(4):305-315. doi:10.1007/s40264-016-0493-y
20. Gagne JJ, Houstoun M, Reichman ME, Hampp C, Marshall JH, Toh S. Safety assessment of niacin in the US Food and Drug Administration's mini-sentinel system. *Pharmacoepidemiol Drug Saf.* 2018;27(1):30-37. doi:10.1002/pds.4343
21. Gieling EM, van den Ham HA, van Onzenoort H, et al. Risk of major bleeding and stroke associated with the use of vitamin K antagonists, nonvitamin K antagonist oral anticoagulants and aspirin in patients with atrial fibrillation: a cohort study. *Br J Clin Pharmacol.* 2017;83(8):1844-1859. doi:10.1111/bcp.13265
22. Giner-Soriano M, Roso-Llorach A, Vedia Urgell C, et al. Effectiveness and safety of drugs used for stroke prevention in a cohort of non-valvular atrial fibrillation patients from a primary care electronic database. *Pharmacoepidemiol Drug Saf.* 2017;26(1):97-107. doi:10.1002/pds.4137
23. González-Pérez A, Sáez ME, Johansson S, Nagy P, García Rodríguez LA. Mortality in patients who discontinue low-dose acetylsalicylic acid therapy after upper gastrointestinal bleeding. *Pharmacoepidemiol Drug Saf.* 2017;26(2):215-222. doi:10.1002/pds.4140
24. Gorricho J, Garjón J, Alonso A, et al. Use of oral antidiabetic agents and risk of community-acquired pneumonia: a nested case-control study. *Br J Clin Pharmacol.* 2017;83(9):2034-2044. doi:10.1111/bcp.13288
25. Grabarczyk TR. Observational Comparative Effectiveness of Pharmaceutical Treatments for Obesity within the Veterans Health Administration. *Pharmacother J Hum Pharmacol Drug Ther.* 2018;38(1):19-28. doi:10.1002/phar.2048
26. Grelle JL, Kutter SN, Giruzzi ME, Tawwater JC. Impact of Insulin Detemir Administration Time on Hypoglycemia Rates in Hospitalized Patients. *Pharmacother J Hum Pharmacol Drug Ther.* 2017;37(12):1523-1529. doi:10.1002/phar.2045
27. Hagberg KW, Divan HA, Nickel JC, Jick SS. Risk of Incident Antidepressant-Treated Depression Associated with Use of 5α-Reductase Inhibitors Compared with Use of α-Blockers in Men with Benign Prostatic Hyperplasia: A Population-Based Study Using the Clinical Practice Research Datalink. *Pharmacother J Hum Pharmacol Drug Ther.* 2017;37(5):517-527. doi:10.1002/phar.1925
28. Ham AC, van Dijk SC, Swart KMA, et al. Beta-blocker use and fall risk in older individuals: Original results from two studies with meta-analysis. *Br J Clin Pharmacol.* 2017;83(10):2292-2302. doi:10.1111/bcp.13328
29. Hedna K, Andersson Sundell K, Hamidi A, Skoog I, Gustavsson S, Waern M. Antidepressants and suicidal behaviour in late life: a prospective population-based study of use patterns in new users aged 75 and above. *Eur J Clin Pharmacol.* 2018;74(2):201-208. doi:10.1007/s00228-017-2360-x
30. Ho JM-W, Macdonald EM, Luo J, et al. Pregabalin and heart failure: A population-based study. *Pharmacoepidemiol Drug Saf.* 2017;26(9):1087-1092. doi:10.1002/pds.4239
31. Hu X, Cheng J, Li C. Effects of rosuvastatin and atorvastatin on nonsustained ventricular tachycardia in patients with ST-elevation myocardial infarction: a retrospective analysis. *Eur J Clin Pharmacol.* 2018;74(1):29-35. doi:10.1007/s00228-017-2338-8
32. Jhang K-M, Huang J-Y, Nfor ON, et al. Extrapyramidal symptoms after exposure to calcium channel blocker-flunarizine or cinnarizine. *Eur J Clin Pharmacol.* 2017;73(7):911-916. doi:10.1007/s00228-017-2247-x

33. Jobski K, Kollhorst B, Garbe E, Schink T. The Risk of Ischemic Cardio- and Cerebrovascular Events Associated with Oxycodone–Naloxone and Other Extended-Release High-Potency Opioids: A Nested Case–Control Study. *Drug Saf.* 2017;40(6):505-515. doi:10.1007/s40264-017-0511-8
34. Jorgenson MR, Descourouez JL, Leverson GE, et al. High-Dose Acyclovir for Cytomegalovirus Prophylaxis in Seropositive Abdominal Transplant Recipients. *Ann Pharmacother.* 2018;52(1):5-10. doi:10.1177/1060028017728296
35. Jorgenson MR, Descourouez JL, Singh T, Astor BC, Panzer SE. Malignancy in Renal Transplant Recipients Exposed to Cyclophosphamide Prior to Transplantation for the Treatment of Native Glomerular Disease. *Pharmacother J Hum Pharmacol Drug Ther.* 2018;38(1):51-57. doi:10.1002/phar.2059
36. Kaye JA, Margulis A V., Fortuny J, et al. Cancer Incidence after Initiation of Antimuscarinic Medications for Overactive Bladder in the United Kingdom: Evidence for Protopathic Bias. *Pharmacother J Hum Pharmacol Drug Ther.* 2017;37(6):673-683. doi:10.1002/phar.1932
37. Kim K, Lee TA, Ardati AK, DiDomenico RJ, Touchette DR, Walton SM. Comparative Effectiveness of Oral Antiplatelet Agents in Patients with Acute Coronary Syndrome. *Pharmacother J Hum Pharmacol Drug Ther.* 2017;37(8):877-887. doi:10.1002/phar.1961
38. Kim TM, Kim H, Jeong YJ, et al. The differences in the incidence of diabetes mellitus and prediabetes according to the type of HMG-CoA reductase inhibitors prescribed in Korean patients. *Pharmacoepidemiol Drug Saf.* 2017;26(10):1156-1163. doi:10.1002/pds.4237
39. King JB, Shah RU, Sainski-Nguyen A, Biskupiak J, Munger MA, Bress AP. Effect of Inpatient Dobutamine versus Milrinone on Out-of-Hospital Mortality in Patients with Acute Decompensated Heart Failure. *Pharmacother J Hum Pharmacol Drug Ther.* 2017;37(6):662-672. doi:10.1002/phar.1939
40. Kreys ED, Frei CR, Villarreal SM, Bollinger MJ, Jones X, Koeller JM. Evaluation of Long-Term Chronic Myeloid Leukemia Treatment Practices with Tyrosine Kinase Inhibitors in a National Cohort of Veterans. *Pharmacother J Hum Pharmacol Drug Ther.* 2017;37(3):278-286. doi:10.1002/phar.1893
41. Kuoppala J, Enlund H, Pulkkinen J, et al. ACE inhibitors and the risk of acute pancreatitis-a population-based case-control study. *Pharmacoepidemiol Drug Saf.* 2017;26(7):853-857. doi:10.1002/pds.4179
42. Lai C-H, Yang Y-H, Chen P-C, King Y-C, Liu C-Y. Retinal vascular complications associated with interferon-ribavirin therapy for chronic hepatitis C: A population-based study. *Pharmacoepidemiol Drug Saf.* 2018;27(2):191-198. doi:10.1002/pds.4363
43. Larivée N, Suissa S, Coulombe J, Tagalakis V, Filion KB. Drospirenone-Containing Oral Contraceptive Pills and the Risk of Venous Thromboembolism: An Assessment of Risk in First-Time Users and Restarters. *Drug Saf.* 2017;40(7):583-596. doi:10.1007/s40264-017-0525-2
44. LaRue HA, Peksa GD, Shah SC. A Comparison of Insulin Doses for the Treatment of Hyperkalemia in Patients with Renal Insufficiency. *Pharmacother J Hum Pharmacol Drug Ther.* 2017;37(12):1516-1522. doi:10.1002/phar.2038
45. Lehto M, Niiranen J, Korhonen P, et al. Quality of warfarin therapy and risk of stroke, bleeding, and mortality among patients with atrial fibrillation: results from the nationwide FinWAF Registry. *Pharmacoepidemiol Drug Saf.* 2017;26(6):657-665. doi:10.1002/pds.4194
46. Leonard CE, Han X, Brensinger CM, et al. Comparative risk of serious hypoglycemia with oral antidiabetic monotherapy: A retrospective cohort study. *Pharmacoepidemiol Drug Saf.* 2018;27(1):9-18. doi:10.1002/pds.4337
47. Lin H-F, Liao K-F, Chang C-M, Lin C-L, Lai S-W. Tamoxifen usage correlates with increased risk of Parkinson's disease in older women with breast cancer: a case–control study in Taiwan. *Eur J Clin Pharmacol.* 2018;74(1):99-107. doi:10.1007/s00228-017-2341-0
48. Lin H-F, Liao K-F, Chang C-M, Lin C-L, Lai S-W. Association of use of selective serotonin reuptake inhibitors with risk of acute pancreatitis: a case-control study in Taiwan. *Eur J Clin Pharmacol.* 2017;73(12):1615-1621. doi:10.1007/s00228-017-2328-x

49. Lin H-F, Liao K-F, Chang C-M, Lin C-L, Lai S-W. Correlation between proton pump inhibitors and risk of pyogenic liver abscess. *Eur J Clin Pharmacol*. 2017;73(8):1019-1025. doi:10.1007/s00228-017-2256-9
50. Lin H-L, Lin H-C, Tseng Y-F, Chen S-C, Hsu C-Y. Risk of parkinsonism induced by flunarizine or cinnarizine: a population-based study. *Eur J Clin Pharmacol*. 2017;73(3):365-371. doi:10.1007/s00228-016-2181-3
51. Lo Re V, Zeldow B, Kallan MJ, et al. Risk of liver decompensation with cumulative use of mitochondrial toxic nucleoside analogues in HIV/hepatitis C virus coinfection. *Pharmacoepidemiol Drug Saf*. 2017;26(10):1172-1181. doi:10.1002/pds.4258
52. Mackenzie IS, Morant S V., Wei L, Thompson AM, MacDonald TM. Spironolactone use and risk of incident cancers: a retrospective, matched cohort study. *Br J Clin Pharmacol*. 2017;83(3):653-663. doi:10.1111/bcp.13152
53. Margulis A V., Hallas J, Pottegård A, et al. Comparison of cardiovascular events among treatments for overactive bladder: a Danish nationwide cohort study. *Eur J Clin Pharmacol*. 2018;74(2):193-199. doi:10.1007/s00228-017-2359-3
54. McGrogan A, Snowball J, Charlton RA. Statins during pregnancy: a cohort study using the General Practice Research Database to investigate pregnancy loss. *Pharmacoepidemiol Drug Saf*. 2017;26(7):843-852. doi:10.1002/pds.4176
55. Meid AD, von Medem A, Heider D, et al. Investigating the Additive Interaction of QT-Prolonging Drugs in Older People Using Claims Data. *Drug Saf*. 2017;40(2):133-144. doi:10.1007/s40264-016-0477-y
56. Momen NC, Munk-Olsen T, Li J, et al. Antidepressant use during pregnancy and childhood cancer in the offspring. *Pharmacoepidemiol Drug Saf*. 2018;27(1):114-118. doi:10.1002/pds.4352
57. Mori N, Kamimura Y, Kimura Y, Hirose S, Aoki Y, Bito S. Comparative analysis of lactic acidosis induced by linezolid and vancomycin therapy using cohort and case–control studies of incidence and associated risk factors. *Eur J Clin Pharmacol*. 2018;74(4):405-411. doi:10.1007/s00228-017-2377-1
58. Mousavi M, Zapolskaya T, Scipione MR, Louie E, Papadopoulos J, Dubrovskaya Y. Comparison of Rates of Nephrotoxicity Associated with Vancomycin in Combination with Piperacillin-Tazobactam Administered as an Extended versus Standard Infusion. *Pharmacother J Hum Pharmacol Drug Ther*. 2017;37(3):379-385. doi:10.1002/phar.1901
59. Muanda FT, Sheehy O, Bérard A. Use of antibiotics during pregnancy and the risk of major congenital malformations: a population based cohort study. *Br J Clin Pharmacol*. 2017;83(11):2557-2571. doi:10.1111/bcp.13364
60. Nardone B, Majewski S, Kim AS, et al. Melanoma and Non-Melanoma Skin Cancer Associated with Angiotensin-Converting-Enzyme Inhibitors, Angiotensin-Receptor Blockers and Thiazides: A Matched Cohort Study. *Drug Saf*. 2017;40(3):249-255. doi:10.1007/s40264-016-0487-9
61. Nishtala PS, Chyou T. Real-world risk of diabetes with antipsychotic use in older New Zealanders: a case-crossover study. *Eur J Clin Pharmacol*. 2017;73(2):233-239. doi:10.1007/s00228-016-2158-2
62. Noize P, Grelaud A, Bay J-O, et al. Real-life patterns of use, safety and effectiveness of sunitinib in first-line therapy of metastatic renal cell carcinoma: the SANTORIN cohort study. *Pharmacoepidemiol Drug Saf*. 2017;26(12):1561-1569. doi:10.1002/pds.4228
63. Ou H-T, Chang K-C, Li C-Y, Wu J-S. Comparative cardiovascular risks of dipeptidyl peptidase 4 inhibitors with other second- and third-line antidiabetic drugs in patients with type 2 diabetes. *Br J Clin Pharmacol*. 2017;83(7):1556-1570. doi:10.1111/bcp.13241
64. Ou H-T, Lee T-Y, Du Y-F, Li C-Y. Comparative risks of diabetes-related complications of basal insulins: a longitudinal population-based cohort of type 1 diabetes 1999-2013 in Taiwan. *Br J Clin Pharmacol*. 2018;84(2):379-391. doi:10.1111/bcp.13461
65. Perreault S, de Denus S, White M, et al. Older adults with heart failure treated with carvedilol, bisoprolol, or metoprolol tartrate: risk of mortality. *Pharmacoepidemiol Drug Saf*. 2017;26(1):81-90. doi:10.1002/pds.4132

66. Poluzzi E, Diemberger I, De Ridder M, et al. Use of antihistamines and risk of ventricular tachyarrhythmia: a nested case-control study in five European countries from the ARITMO project. *Eur J Clin Pharmacol*. 2017;73(11):1499-1510. doi:10.1007/s00228-017-2317-0
67. Qian CJ, Coulombe J, Suissa S, Ernst P. Pneumonia risk in asthma patients using inhaled corticosteroids: a quasi-cohort study. *Br J Clin Pharmacol*. 2017;83(9):2077-2086. doi:10.1111/bcp.13295
68. Quinn KL, Macdonald EM, Gomes T, et al. Macrolides, Digoxin Toxicity and the Risk of Sudden Death: A Population-Based Study. *Drug Saf*. 2017;40(9):835-840. doi:10.1007/s40264-017-0539-9
69. Quinn KL, Macdonald EM, Mamdani MM, Diong C, Juurlink DN, (CDSERN) for the CDS and ERN. Lipophilic Statins and the Risk of Intracranial Hemorrhage Following Ischemic Stroke: A Population-Based Study. *Drug Saf*. 2017;40(10):887-893. doi:10.1007/s40264-017-0552-z
70. Rea F, Bonassi S, Vitale C, et al. Exposure to statins is associated to fracture risk reduction in elderly people with cardiovascular disease: evidence from the AIFA-I-GrADE observational project. *Pharmacoepidemiol Drug Saf*. 2017;26(7):775-784. doi:10.1002/pds.4206
71. Reichman ME, Wernecke M, Graham DJ, et al. Antihypertensive drug associated angioedema: effect modification by race/ethnicity. *Pharmacoepidemiol Drug Saf*. 2017;26(10):1190-1196. doi:10.1002/pds.4260
72. Russo-Alvarez G, Martinez KA, Valente M, et al. Thromboembolic and Major Bleeding Events With Rivaroxaban Versus Warfarin Use in a Real-World Setting. *Ann Pharmacother*. 2018;52(1):19-25. doi:10.1177/1060028017727290
73. Rutter WC, Burgess DS. Acute Kidney Injury in Patients Treated with IV Beta-Lactam/Beta-Lactamase Inhibitor Combinations. *Pharmacother J Hum Pharmacol Drug Ther*. 2017;37(5):593-598. doi:10.1002/phar.1918
74. Samp JC, Joo MJ, Schumock GT, Calip GS, Pickard AS, Lee TA. Risk of Cardiovascular and Cerebrovascular Events in COPD Patients Treated With Long-Acting β_2-Agonist Combined With a Long-Acting Muscarinic or Inhaled Corticosteroid. *Ann Pharmacother*. 2017;51(11):945-953. doi:10.1177/1060028017719716
75. Samp JC, Joo MJ, Schumock GT, Calip GS, Pickard AS, Lee TA. Comparative Effectiveness of Long-Acting Beta$_2$-Agonist Combined with a Long-Acting Muscarinic Antagonist or Inhaled Corticosteroid in Chronic Obstructive Pulmonary Disease. *Pharmacother J Hum Pharmacol Drug Ther*. 2017;37(4):447-455. doi:10.1002/phar.1913
76. Samuel NG, Seifert CF. Risk of Bleeding in Patients on Full-Dose Enoxaparin With Venous Thromboembolism and Selective Serotonin Reuptake Inhibitors. *Ann Pharmacother*. 2017;51(3):226-231. doi:10.1177/1060028016677309
77. Shalev V, Sharman Moser S, Goldshtein I, et al. Adherence With Bisphosphonates and Long-Term Risk of Hip Fractures: A Nested Case-Control Study Using Real-World Data. *Ann Pharmacother*. 2017;51(9):757-767. doi:10.1177/1060028017710482
78. Shin J-Y, Eberg M, Ernst P, Filion KB. Statin potency and the risk of hospitalization for community-acquired pneumonia. *Br J Clin Pharmacol*. 2017;83(6):1319-1327. doi:10.1111/bcp.13208
79. Shlomo M, Gorodischer R, Daniel S, et al. The Fetal Safety of Enoxaparin Use During Pregnancy: A Population-Based Retrospective Cohort Study. *Drug Saf*. 2017;40(11):1147-1155. doi:10.1007/s40264-017-0573-7
80. Simard P, Presse N, Roy L, et al. Association Between Metformin Adherence and All-Cause Mortality Among New Users of Metformin: A Nested Case-Control Study. *Ann Pharmacother*. 2018;52(4):305-313. doi:10.1177/1060028017743517
81. Stolk LM, de Vries F, Ebbelaar C, et al. Risk of myocardial infarction in patients with atrial fibrillation using vitamin K antagonists, aspirin or direct acting oral anticoagulants. *Br J Clin Pharmacol*. 2017;83(8):1835-1843. doi:10.1111/bcp.13264
82. Su X, Yuan W, Chen J, et al. Prenatal exposure to β2-adrenoreceptor agonists and the risk of autism spectrum disorders in offspring. *Pharmacoepidemiol Drug Saf*. 2017;26(7):812-818. doi:10.1002/pds.4214

83. Svendsen T de K, Nørregaard Hansen P, García Rodríguez LA, et al. Statins and polyneuropathy revisited: case-control study in Denmark, 1999-2013. *Br J Clin Pharmacol*. 2017;83(9):2087-2095. doi:10.1111/bcp.13298
84. Thygesen LC, Pottegård A, Ersbøll AK, Friis S, Stürmer T, Hallas J. External adjustment of unmeasured confounders in a case-control study of benzodiazepine use and cancer risk. *Br J Clin Pharmacol*. 2017;83(11):2517-2527. doi:10.1111/bcp.13342
85. Tran Y-H, Groen H, Bergman JEH, Hak E, Wilffert B. Exposure to reactive intermediate-inducing drugs during pregnancy and the incident use of psychotropic medications among children. *Pharmacoepidemiol Drug Saf*. 2017;26(3):265-273. doi:10.1002/pds.4161
86. Valent F. New oral anticoagulant prescription rate and risk of bleeding in an Italian region. *Pharmacoepidemiol Drug Saf*. 2017;26(10):1205-1212. doi:10.1002/pds.4279
87. Vozoris NT, Wang X, Austin PC, et al. Adverse cardiac events associated with incident opioid drug use among older adults with COPD. *Eur J Clin Pharmacol*. 2017;73(10):1287-1295. doi:10.1007/s00228-017-2278-3
88. Vozoris NT, Wang X, Austin PC, et al. Incident diuretic drug use and adverse respiratory events among older adults with chronic obstructive pulmonary disease. *Br J Clin Pharmacol*. 2018;84(3):579-589. doi:10.1111/bcp.13465
89. Wändell P, Carlsson AC, Holzmann M, et al. Association between antithrombotic treatment and hemorrhagic stroke in patients with atrial fibrillation—a cohort study in primary care. *Eur J Clin Pharmacol*. 2017;73(2):215-221. doi:10.1007/s00228-016-2152-8
90. Wang C-Y, Lin Z-F, Lee C-M, et al. Concomitant use of calcium channel blockers with dual antiplatelet therapy and re-hospitalization for acute coronary syndrome. *Pharmacoepidemiol Drug Saf*. 2017;26(3):229-238. doi:10.1002/pds.4147
91. Wei L, Ratnayake L, Phillips G, et al. Acid-suppression medications and bacterial gastroenteritis: a population-based cohort study. *Br J Clin Pharmacol*. 2017;83(6):1298-1308. doi:10.1111/bcp.13205

TABLE S1. Reporting quality of the studies included in this systematic review of the quality of reporting in pharmacoepidemiology stratified per exposure definition. For each specific item, the number of studies reporting that item is shown.

	Intention to treat (n=24)†		**>= 1 prescription during a certain period (n=19)†**		**Time-varying exposure (n=43)†**	
	Studies, n§	Reported, n (%)	Studies, n§	Reported, n (%)	Studies, n§	Reported, n (%)
1 Type of exposure	24	24 (100)	19	19 (100)	43	43 (100)
2 Exposure risk window (ERW)	24	17 (71)	19	16 (84)	43	39 (91)
3 Induction period‡	24	20 (83)	19	17 (89)	43	39 (91)
Explicit		5 (21)		2 (11)		6 (14)
Implicit		15 (63)		15 (79)		33 (77)
4 Stockpiling	0	NA	0	NA	42	3 (7)
5 Bridging exposure episodes	0	NA	0	NA	42	18 (43)
6 Exposure extension‡	0	NA	3	2 (67)	43	24 (56)
Explicit				1 (33)		16 (37)
Implicit				1 (33)		8 (19)
7 Switching/ add on	21	12 (57)	14	7 (50)	42	29 (69)
8 Codes	24	5 (21)	19	8 (42)	43	10 (23)
9 Frequency and temporality of codes	24	18 (75)	19	16 (84)	43	38 (88)
10 Care setting	24	18 (75)	19	13 (68)	43	33 (77)
11 Exposure Assessment Window (EAW)	24	23 (96)	19	18 (95)	43	43 (100)

† The exposure definition was divided in five categories: 1) intention to treat: exposure at baseline is included as a time-fixed variable in the model; 2) the presence of ≥1 prescriptions during a certain time period, for example during pregnancy or during the last 12 months prior to the event; 3) time-varying: episodes of (non) exposure are constructed based on duration of each prescription; and 4) other, including measures of adherence and (cumulative) dose and cumulative dose.
‡ When explicitly mentioning an introduction period, a period after the index date is clearly excluded in the exposure risk window. Stating that the follow-up started on the day of the first prescription implies implicitly that there was no induction period. The same reasoning also applies to the extension period.
§ Total number of studies for which this item was applicable.

TABLE S2. Reporting quality of the studies included in this systematic review of the quality of reporting in pharmacoepidemiology stratified per type of study design. For each specific item, the number of studies reporting that item is shown.

	Cohort (n=64)		**Case-control or case-crossover (n=29)**	
	Studies, n ‡	Reported, n (%)	Studies, n ‡	Reported, n (%)
1 Type of exposure	64	62 (95)	29	29 (100)
2 Exposure risk window (ERW)	64	51 (80)	29	27 (93)
3 Induction period †	63	55 (87)	29	28 (97)
Explicit		9 (14)		5 (17)
Implicit		46 (73)		23 (79)
4 Stockpiling	30	1 (3)	18	4 (22)
5 Bridging exposure episodes	30	12 (40)	15	6 (40)
6 Exposure extension †	34	14 (41)	16	13 (81)
Explicit		6 (18)		11 (69)
Implicit		8 (24)		2 (13)
7 Switching/ add on	55	34 (62)	28	17 (61)
8 Codes	64	12 (19)	29	13 (45)
9 Frequency and temporality of codes	64	51 (80)	29	27 (93)
10 Care setting	64	48 (75)	29	21 (72)
11 Exposure Assessment Window (EAW)	64	63 (98)	29	28 (97)

† When explicitly mentioning an introduction period, a period after the index date is clearly excluded in the exposure risk window. Stating that the follow-up started on the day of the first prescription implies implicitly that there was no induction period. The same reasoning also applies to the extension period.
‡ Total number of studies for which this item was applicable.

TABLE S3. Reporting quality of the studies included in this systematic review of the quality of reporting in pharmacoepidemiology stratified per number of included subjects.
For each specific item, the number of studies reporting that item is shown.

	250 – 1,000 patients (n=13)		**1,001 – 10,000 patients (n=30)**		**10,001 – 100,000 patients (n=24)**		**>100,000 patients (n=24)**	
	Studies, n ‡	Reported, n (%)	Studies, n ‡	Reported, n (%)	Studies, n ‡	Reported, n (%)	Studies, n ‡	Reported, n (%)
1 Type of exposure	13	11 (85)	30	30 (100)	24	24 (100)	24	24 (100)
2 Exposure risk window (ERW)	13	7 (54)	30	27 (90)	24	22 (92)	24	21 (88)
3 Induction period †	12	10 (83)	30	27 (90)	24	22 (92)	23	22 (96)
Explicit		0 (0)		6 (20)		3 (13)		5 (22)
Implicit		10 (83)		21 (70)		19 (79)		17 (74)
4 Stockpiling	6	0 (0)	16	3 (19)	14	1 (7)	11	1 (9)
5 Bridging exposure episodes	6	0 (0)	14	5 (36)	13	4 (31)	11	9 (82)
6 Exposure extension †	8	2 (25)	16	9 (56)	14	10 (71)	11	6 (55)
Explicit		1 (13)		5 (31)		6 (43)		5 (45)
Implicit		1 (13)		4 (25)		4 (29)		1 (9)
7 Switching/ add on	12	5 (42)	25	14 (56)	23	15 (65)	21	16 (76)
8 Codes	13	0 (0)	30	11 (37)	24	6 (25)	24	7 (29)
9 Frequency and temporality of codes	13	6 (46)	30	27 (90)	24	21 (88)	24	23 (96)
10 Care setting	13	12 (92)	30	21 (70)	24	16 (67)	24	18 (75)
11 Exposure Assessment Window (EAW)	13	13 (100)	30	28 (93)	24	24 (100)	24	24 (100)

† When explicitly mentioning an introduction period, a period after the index date is clearly excluded in the exposure risk window. Stating that the follow-up started on the day of the first prescription implies implicitly that there was no induction period. The same reasoning also applies to the extension period.
‡ Total number of studies for which this item was applicable.

TABLE S4. Reporting quality of the studies included in this systematic review of the quality of reporting in pharmacoepidemiology stratified per type of outcome. For each specific item, the number of studies reporting that item is shown.

	Beneficial effects (n=18)		**Adverse effects (n=67)**		**Both (n=6)**	
	Studies, n ‡	Reported, n (%)	Studies, n ‡	Reported, n (%)	Studies, n ‡	Reported, n (%)
1 Type of exposure	18	17 (94)	67	66 (99)	6	6 (100)
2 Exposure risk window (ERW)	18	16 (89)	67	55 (82)	6	6 (100)
3 Induction period †	17	17 (100)	67	58 (87)	6	6 (100)
Explicit		6 (35)		8 (12)		0 (0)
Implicit		11 (65)		50 (75)		6 (100)
4 Stockpiling	10	3 (30)	36	2 (6)	1	0 (0)
5 Bridging exposure episodes	7	4 (57)	36	14 (39)	1	0 (0)
6 Exposure extension †	7	4 (57)	40	22 (55)	2	1 (50)
Explicit		1 (14)		6 (15)		0 (0)
Implicit		3 (43)		16 (40)		1 (50)
7 Switching/ add on	16	12 (75)	60	35 (60)	5	3 (60)
8 Codes	18	5 (28)	67	18 (27)	6	1 (17)
9 Frequency and temporality of codes	18	16 (89)	67	55 (82)	6	6 (100)
10 Care setting	18	13 (72)	67	49 (73)	6	5 (83)
11 Exposure Assessment Window (EAW)	18	18 (100)	67	65 (97)	6	6 (100)

† When explicitly mentioning an introduction period, a period after the index date is clearly excluded in the exposure risk window. Stating that the follow-up started on the day of the first prescription implies implicitly that there was no induction period. The same reasoning also applies to the extension period.
‡ Total number of studies for which this item was applicable.

TABLE S5. Reporting quality of the studies included in this systematic review of the quality of reporting in pharmacoepidemiology stratified per type of database. For each specific item, the number of studies reporting that item is shown.

	Claims (n=41)		**Pharmacy (n=16)**		**GP (n=17)**		**Hospital (n=18)**	
	Studies, n ‡	Reported, n (%)	Studies, n ‡	Reported, n (%)	Studies, n ‡	Reported, n (%)	Studies, n ‡	Reported, n (%)
1 Type of exposure	41	41 (100)	16	16 (100)	17	17 (100)	18	18 (100)
2 Exposure risk window (ERW)	41	39 (95)	16	15 (94)	17	15 (88)	18	9 (50)
3 Induction period †	41	39 (95)	15	14 (93)	17	15 (88)	18	14 (78)
Explicit		6 (15)		5 (33)		1 (6)		2 (11)
Implicit		33 (80)		9 (60)		14 (82)		12 (67)
4 Stockpiling	20	5 (25)	9	0 (0)	14	0 (0)	6	0 (0)
5 Bridging exposure episodes	17	12 (71)	9	3 (33)	14	3 (21)	6	0 (0)
6 Exposure extension †	18	12 (67)	9	5 (56)	15	11 (73)	9	2 (22)
Explicit		5 (28)		0 (0)		4 (27)		2 (22)
Implicit		7 (39)		5 (56)		7 (47)		0 (0)
7 Switching/ add on	37	27 (73)	14	7 (50)	16	10 (63)	15	9 (60)
8 Codes	41	10 (24)	16	11 (69)	17	5 (29)	18	0 (0)
9 Frequency and temporality of codes	41	40 (98)	16	15 (94)	17	15 (88)	18	9 (50)
10 Care setting	41	22 (54)	16	13 (81)	17	17 (100)	18	18 (100)
11 Exposure Assessment Window (EAW)	41	39 (95)	16	16 (100)	17	17 (100)	18	18 (100)

† When explicitly mentioning an introduction period, a period after the index date is clearly excluded in the exposure risk window. Stating that the follow-up started on the day of the first prescription implies implicitly that there was no induction period. The same reasoning also applies to the extension period.
‡ Total number of studies for which this item was applicable.

TABLE S6. Reporting quality of the studies included in this systematic review of the quality of reporting in pharmacoepidemiology stratified per maximum of words allowed.
For each specific item, the number of studies reporting that item is shown.

	Short reports		**Original research articles**					
	Word limit ≤1500 (n=4)		Word limit 1500-3000 (n=32) [§]		Word limit 3000-4000 (n=30) [§]		Word limit ≥4000 (n=25) [§]	
	Studies, n [‡]	Reported, n (%)	Studies, n [‡]	Reported, n (%)	Studies, n [‡]	Reported, n (%)	Studies, n [‡]	Reported, n (%)
1 Type of exposure	4	4 (100)	32	32 (100)	30	30 (100)	25	25 (100)
2 Exposure risk window (ERW)	4	3 (75)	32	30 (94)	30	25 (83)	25	19 (76)
3 Induction period [†]	4	3 (75)	32	30 (94)	30	26 (87)	24	22 (92)
Explicit		1 (25)		6 (19)		6 (20)		20 (83)
Implicit		2 (50)		24 (75)		20 (67)		2 (8)
4 Stockpiling	1	0 (0)	20	4 (20)	13	0 (0)	13	1 (7)
5 Bridging exposure episodes	1	0 (0)	17	8 (47)	13	6 (46)	13	4 (31)
6 Exposure extension [†]	2	1 (50)	18	10 (56)	16	9 (56)	13	7 (54)
Explicit		0 (0)		5 (28)		6 (38)		6 (46)
Implicit		1 (50)		5 (28)		3 (19)		1 (7)
7 Switching/ add on	4	3 (75)	28	16 (57)	27	17 (62)	22	14 (64)
8 Codes	4	1 (25)	32	8 (25)	30	7 (23)	25	8 (32)
9 Frequency and temporality of codes	4	3 (75)	32	28 (88)	30	25 (83)	25	21 (84)
10 Care setting	4	2 (50)	32	24 (75)	30	24 (80)	25	17 (68)
11 Exposure Assessment Window (EAW)	4	4 (100)	32	31 (97)	30	29 (97)	25	24 (96)

[†] When explicitly mentioning an introduction period, a period after the index date is clearly excluded in the exposure risk window. Stating that the follow-up started on the day of the first prescription implies implicitly that there was no induction period. The same reasoning also applies to the extension period.
[‡] Total number of studies for which this item was applicable.
[§] Word limit per journal, according to the author guidelines of the journals: Annals of Pharmacotherapy 3000; British Journal of Clinical Pharmacology 3000-4000; Drug Safety 6000; European Journal of Clinical Pharmacology 8-10 pages of around 450 words (=4500 words maximum); Pharmacoepidemiology and Drug Safety 3500; and Pharmacotherapy 3500.

TABLE S7. Reporting quality of the studies included in this systematic review of the quality of reporting in pharmacoepidemiology stratified per route of administration (oral/inhaled vs intravenous/subcutaneous).
For each specific item, the number of studies reporting that item is shown.

	Oral/inhaled (n=80)		**Intravenous/ subcutaneous (n=11)**	
	Studies, n ‡	Reported, n (%)	Studies, n ‡	Reported, n (%)
1 Type of exposure	80	80 (100)	11	11 (100)
2 Exposure risk window (ERW)	80	71 (89)	11	6 (55)
3 Induction period †	79	73 (92)	11	8 (72)
Explicit		11 (14)		3 (27)
Implicit		62 (78)		5 (45)
4 Stockpiling	44	5 (11)	3	0 (0)
5 Bridging exposure episodes	41	18 (44)	3	0 (0)
6 Exposure extension †	43	26 (60)	6	1 (17)
Explicit		17 (40)		0 (0)
Implicit		9 (21)		1 (17)
7 Switching/ add on	74	46 (62)	7	4 (57)
8 Codes	80	23 (29)	11	1 (9)
9 Frequency and temporality of codes	80	71 (89)	11	6 (55)
10 Care setting	80	58 (73)	11	9 (82)
11 Exposure Assessment Window (EAW)	80	78 (98)	11	11 (100)

† When explicitly mentioning an introduction period, a period after the index date is clearly excluded in the exposure risk window. Stating that the follow-up started on the day of the first prescription implies implicitly that there was no induction period. The same reasoning also applies to the extension period.
‡ Total number of studies for which this item was applicable.

TABLE S8. Interobserver agreement per item

Item	Scores			Kappa	Overall agreement *
	Yes	*No*	*NA*		
1 Type of exposure	91	0	0	0.11	0.88
2 Exposure risk window (ERW)	77	14	0	0.49	0.82
3 Induction period	71	9	1	0.20	0.66
4 Stockpiling	5	42	33	0.33	0.64
5 Bridging exposure episodes	18	26	47	0.35	0.58
6 Exposure extension	27	22	42	0.28	0.52
7 Switching/ add on	50	31	10	0.33	0.61
8 Codes	24	67	0	0.75	0.90
9 Frequency and temporality of codes	77	14	0	0.35	0.76
10 Care setting	67	24	0	0.43	0.78
11 Exposure Assessment Window (EAW)	89	2	0	0.23	0.93

2.2

BIAS IN OBSERVATIONAL STUDIES ON THE EFFECTIVENESS OF HYDROXYCHLOROQUINE IN COVID-19

Mirjam Hempenius, Rolf H. H. Groenwold, Anthonius de Boer, Olaf H. Klungel, Helga Gardarsdottir

Submited

ABSTRACT

Introduction

During the first waves of the coronavirus pandemic, evidence was urgently needed about potential effective treatments. Results from observational studies on the effectiveness of hydroxychloroquine (HCQ) were conflicting, potentially due to biases. We aimed to assess the quality of observational studies on HCQ and its relation to effect sizes.

Methods

Observational studies on the effectiveness of in-hospital use of HCQ in COVID-19 patients were reviewed. Study quality was assessed regarding seven items, based on the ROBINS-I tool: confounding, selection bias, misclassification of interventions and of outcomes, deviation from intended intervention, missing data, and reporting. Effects sizes found in observational studies were compared to those from RCTs, and differences were related to apparent study quality.

Results

None of the 33 included observational studies were free of risk of bias, most commonly related to confounding (n=26, 79%) and misclassification of interventions (n=22, 67%). Observational studies with estimates closer to those of RCTs appeared less often at risk of bias than studies with more diverging estimates (p=0.02).

Discussion

Overall, the quality of observational HCQ studies was poor, and studies reporting more extreme estimates appeared of lower quality than studies reporting estimates closer to those of RCTs. Synthesis of evidence of effectiveness of HCQ in COVID-19 should focus on RCTs and carefully consider the added value and quality of observational evidence.

1. INTRODUCTION

During the first waves of the pandemic with severe acute respiratory syndrome coronavirus (SARS-CoV-2) and its associated 2019 coronavirus disease (COVID-19), effective treatments were urgently needed to reduce mortality, the severity of symptoms, and the need for hospitalization. Doctors were forced to make choices regarding which treatments were likely to save the lives of critically ill patients, without sufficient evidence-based knowledge about effective treatments for this new disease.[1]

Hydroxychloroquine (HCQ) was one of the drugs that caught early attention of researchers, clinicians, and the public during the pandemic. Preclinical studies indicated HCQ as a potentially effective treatment for the symptoms of COVID-19 because of its *in vitro* antiviral effects.[2] Researchers conducted many observational studies while waiting for results from randomized controlled trials (RCTs). As HCQ was already used early during the pandemic in the treatment of COVID-19, there was an opportunity to perform observational studies using routinely collected patient data.[3] Indeed, well-designed observational studies can be helpful in the generation of hypotheses about the potential effects of drugs; however, observational studies can also provide biased results when not properly designed and analyzed.[4]

The observational studies on the effectiveness of HCQ reported divergent results, from a five-fold reduction in mortality risk to an eight-fold increased risk of intensive care unit (ICU) admission,[5,6] leading to a heated debate about the effectiveness of HCQ.[6–10] Today, RCTs have convincingly shown that HCQ has no benefit in the treatment of COVID-19 patients and may even be harmful.[11,12]

Variation in the estimated effects of HCQ treatment and divergence from the RCT results led to discussions on how the lack of a high-quality and proper study design, quick review times, and a possible lack of expertise when reviewing these studies might have led to the unjustified conclusions.[13,14] For example, many studies suffered from immortal time bias, confounding bias, or bias due to inadequately accounting for competing risks.[15,16] The aim of the current study was to provide a comprehensive overview and to assess the overall quality of observational studies on HCQ and to relate this quality to the observed effect sizes.

2. METHODS

2.1. Search strategy and inclusion criteria

On March 15, 2021, we searched PubMed for all observational studies on COVID-19 and the use of HCQ published between 01/01/2020 and 01/03/2021. The included studies were peer-reviewed primary research articles, were published in English, and used an observational design to investigate in-hospital treatment with HCQ. Moreover, studies were included if they measured one of the following clinical outcomes: mortality, duration of hospitalization, need for mechanical ventilation, or time to clinical improvement. We included only studies in which HCQ (with or without azithromycin) as treatment for COVID-19 was compared to standard care. All studies focusing on the use of HCQ as prophylaxes for COVID-19 were excluded.

For each observational study identified, we extracted the journal name, the impact factor of the scientific journal, and the journal ranking according to the InCites Journal Citation Reports (Q1, Q2, Q3, or Q4).[17] We also extracted the date of first submission, acceptance, and first publication; the geographical region of the study population (Africa, Asia, North America, and Europe); the study design (cohort, case-control, or other); the number of included subjects; and the proportion of study subjects treated with HCQ.

2.2. Quality of studies

The primary aim of this research was to assess the quality of observational studies on HCQ for treatment of COVID19 outcomes. Quality was assessed based on the risk of bias according to seven different domains from the ROBINS-I tool.[18] These seven domains are bias due to confounding, selection of participants, classification of interventions, deviations from intended interventions, missing data, measurement of outcomes, and selective reporting.

For each study included in the review, we scored whether there was risk of bias in the primary analysis of the study, for each of the seven domains. Within each domain, we identified issues that could occur specifically in in-hospital studies on the effectiveness of HCQ. These specific issues are further elaborated in the supplementary materials (Section A) for each domain. If insufficient information was included in the publication to assess a domain, this domain was scored as "insufficient information."

We assessed the overall quality per study by counting the number of domains that were sufficiently reported on and were considered not to be at risk of bias. A higher count indicated higher quality.

2.3. Effect size

The effect estimates reported in the observational studies were compared to the effect estimates found in RCTs; the latter served as a reference. Since effect estimates might also differ between RCTs, we selected benchmark estimates from the meta-analysis by Siemieniuk et al. This meta-analysis, published in *the BMJ*, is a living review, with the last update (at the time of study) on the 6th of April 2021.[19] The outcomes included in this meta-analysis were mortality, need for mechanical ventilation, duration of hospital stay, ventilator-free days, and time to clinical improvement. Clinically relevant effects of HCQ were not observed for any of the outcomes (Table 1). For each of the outcomes, we extracted the point estimates and corresponding confidence intervals.

We note that the effect estimates obtained from RCTs are also estimates and hence do not concur with the *true* effect of HCQ. Moreover, this approach cannot distinguish between deviations between effect estimates from observational studies and RCTs that arise due to bias or due to (random) sampling variability. Nevertheless, we hypothesize that in observational studies where effect estimates deviate more from those found in RCTs, the potential for bias is larger.

TABLE 1. Estimates of the effect of hydroxychloroquine (HCQ) in Covid-19 patients from randomized controlled trials (RCTs)

Outcome	Estimate from meta-analysis/RCTs
Mortality	1.09 (0.93–1.27) [19]
Mechanical ventilation	1.15 (0.92–1.46) [19]
Duration of hospital stay	0.1 (-1.8–2.0) [19]
Ventilator-free days	-1.4 (-4.9–2.2) [19]
Time to clinical improvement	-0.9 (-2.9–2.1) [19]
Composite outcome *	1.13 (0.60–2.14) ** [43]
Intensive care unit (ICU) admission	1.13 (0.60–2.14) ***

* Composite outcome indicating disease aggravation, for example ICU admission, need for mechanical ventilation, or mortality.
** The meta-analysis by Siemieniuk did not provide an estimate for a composite outcome of disease aggravation, ICU admission, and death, or for ICU admission alone. The odds ratio of the composite outcome was found to be 1.13 (0.60–2.14) in the RCT by Self et al.
*** In other studies, the occurrence of ICU admission was categorized as disease aggravation. We assume that the effect estimate of ICU admission is in line with the estimates for mortality, ventilation, and the composite outcome, and it was set at 1.13.

For each observational study, we extracted the point estimates for the primary outcome. If a study included multiple primary outcomes, we included the effect estimate for mortality, if present. For all relative measures, we subsequently calculated the extent to which this effect deviated from the benchmark estimates. This deviation was calculated as abs(log(HR_obs) – log(HR_RCT)).

2.4 Data analysis

Publication and study characteristics were described using descriptive statistics. The publication date was dichotomized as before or in June 2020, or after June 2020, which was the month in which the interim results of the RECOVERY trial were published and the FDA decided to revoke the emergency use authorization for HCQ.[20,21] For each of the domains, we described the number of studies that were or were not considered to be at risk of bias or that were scored as having "insufficient information." The relation between publication details (journal ranking, publication date, and time between submission and publication) and the overall quality of the studies as well as the relation between the effect size and the overall quality were assessed using Poisson regression.

3. RESULTS

Our search strategy yielded 2,331 hits in PubMed, 79 of which were selected on the basis of title and abstract. Of those studies, 33 were included in this review. The reasons for inclusion and exclusion are depicted in Figure 1. A list of all included studies is presented in the supplementary materials (Section C).

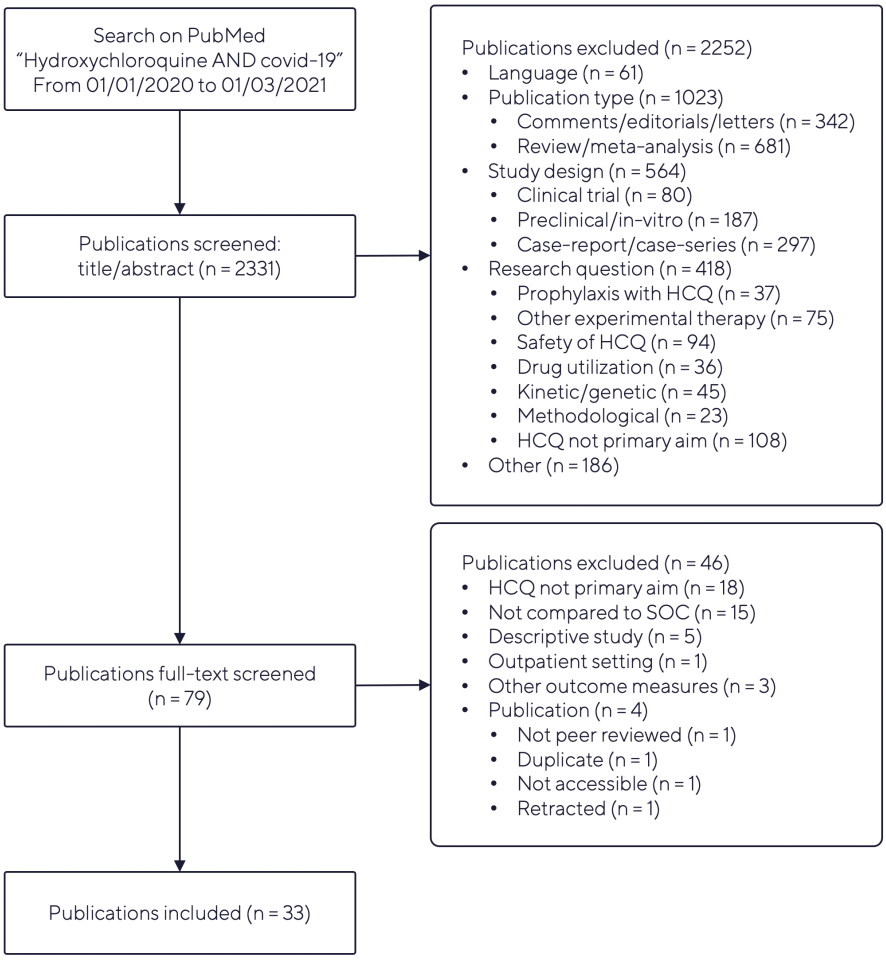


FIGURE 1. Flow chart of the search and screening process according to inclusion and exclusion criteria

3.1. Publication and study characteristics

The characteristics of the included studies are summarized in Table 2. Out of 33 studies, 14 (42%) were published in journals that were ranked in the first quartile (Q1), according to the InCites Journal Citation Reports. Four of those 14 were published in a journal with an impact factor >30. The first observational study was published on the 7th of May 2020, and the last on the 18th of February 2021. Seven studies (21%) were published before or in June 2020. The median time from submission to first publication was 64 days (interquartile range [IQR] 34.5–83.5 days). Studies were most often performed in Europe (45%) and the US (39%), and all studies used a cohort design. The median number of subjects included in the observational studies was 807 (IQR 307–1,949), and the median proportion of the study population treated with HCQ was 54.6% (IQR 39.6%–74.8%).

TABLE 2. Characteristics of the 33 observational studies of HCQ in Covid-19 that were included in the systematic review

Journal ranking*	*n* (%)
Journals ranked as "Q1"	14 (42)
Journals ranked as "Q2"	12 (36)
Journals ranked as "Q3"	4 (12)
Journals ranked as "Q4"	0 (0)
No journal ranking available	3 (9)
Publication date **	
Before or in June 2020	5 (15)
After June 2020	28 (85)
Time from submission to publication	64 (34.5-83.5)
0–45 days	10 (30)
46–90 days	13 (39)
>90 days	4 (12)
Insufficient information for estimation	6 (18)
Geographical area	
Africa	1 (3)
Asia	4 (12)
North America	13 (39)
Europe	15 (45)
Study design	
Cohort	33 (100)
Case-control	0 (0)
Other design	0 (0)
Number of subjects included (median [IQR])	807 (307–1,949)
1–250	8 (24)
251–1,000	10 (30)
1,001–2,500	9 (27)
>2,500	6 (18)
% of cohort treated with HCQ (median [IQR])	54.6 (39.6–74.8)
<20%	1 (3)
20%–80%	25 (76)
>80%	6 (18)
Unclear	1 (3)

* according to the InCites Journal Citation Reports[17]
** In June 2020, the interim results of the RECOVERY trial were published and the FDA decided to revoke the emergency use authorization for HCQ.[20,21]
Abbreviations: HCQ – hydroxychloroquine, IQR – interquartile range

3.2. Quality of studies

The quality of the studies is summarized per domain in Table 3. The identified increased risk of bias most often involved bias due to confounding (Domain 1, *n* = 25, 76%) and bias due to the classification of interventions (Domain 3, *n* = 22, 67%). The fewest number of issues were noted for bias due to measurement of outcomes (Domain 6, *n* = 5, 15%) and bias due to selective reporting (Domain 7, *n* = 8, 24%).

All included studies were considered to be at risk within at least one of the domains. The median number of domains in which there was risk of bias per individual study was three (range: one to six), and 14 studies (42%) were in 4 domains or more considered not to be at risk of bias. In addition, almost two third of all studies (21 of 33) reported insufficient information for one or more domains.

TABLE 3. Risk of bias in observational studies of HCQ in COVID-19, stratified by different bias domains

Domain 1: Bias due to confounding	**n (%)**
Not at risk of bias	**8 (24)**
At risk of bias	**25 (76)**
At least 1 but not all confounders included	22 (67)
No confounders included	3 (9)
Domain 2: Bias due to selection of participants	
Not at risk of bias	**17 (52)**
At risk of bias*	**15 (45)**
Inclusion based on discharge data	9 (27)
Different index date for HCQ users and non-HCQ users	3 (9)
Exclusion of patients who discontinued HCQ early	1 (3)
Exclusion of subjects who received HCQ >x hours after admission	2 (6)
Exclusion of patients who had the outcome prior to receiving their first dose of HCQ	1 (3)
Insufficient information	**1 (3)**
Domain 3: Bias due to classification of interventions	
Not at risk of bias	**4 (12)**
All HCQ use started on day of admission	2 (6)
HCQ use defined in a time-varying manner	1 (3)
Comparison of hospital strategies	1 (3)
At risk of bias	**22 (67)**
HCQ use defined as "ever during hospitalization"	15 (45)
HCQ use defined as "started within 48/72 hours of hospitalization"	5 (15)

TABLE 3. Continued.

HCQ use defined when subjects used HCQ for at least three days	2 (6)
Insufficient information	**7 (21)**
Domain 4: Bias due to deviations from intended interventions	
Not at risk of bias	**9 (27)**
Exclusion of subjects using corticosteroids	0 (0)
No (expected) differences in use of corticosteroids between HCQ users and non-HCQ users	8 (24)
Time-varyingly adjusted for other interventions during hospitalization	0 (0)
Adjusted for other interventions that were determined at baseline	1 (3)
At risk of bias	**10 (30)**
Adjusted for other interventions during hospital stay inappropriately	6 (18)
Significant differences in steroid treatment between HCQ users and non-HCQ users	4 (12)
Insufficient information	**14 (42)**
Domain 5: Bias due to missing data	
Not at risk of bias	**12 (36)**
Data complete for >95%	8 (24)
Multiple imputations used for missing data	4 (12)
At risk of bias	**13 (39)**
Presence of missing data, no report of how these were handled	3 (9)
Use of a biased method without substantiation for appropriateness	10 (30)
Complete case analysis	*7 (21)*
Missing indicator method	*2 (6)*
Single value substitution	*1 (3)*
Insufficient information	**8 (24)**
Domain 6: Bias due to measurements of outcomes	
Not at risk of bias	**26 (79)**
Outcome in-hospital mortality	16 (48)
Fixed follow-up for all subjects	6 (18)
Cox regression with discharge alive as competing risk	1 (3)
Patients discharged alive censored at the end of the study period	2 (6)
Length of stay as outcome, with all patients discharged alive	1 (6)
At risk of bias	**5 (15)**
Differences in outcome measurement between HCQ users and non-HCQ users	0 (0)

TABLE 3. Continued.

Length of stay as outcome, no differentiation between discharge alive or dead	2 (6)
Patients discharged alive censored at discharge in analysis of mortality	1 (3)
Transferred to hospice, not counted under mortality	1 (3)
Censoring of patients who died in length-of-stay analysis	1 (3)
Insufficient information	**2 (6)**
Domain 7: Bias due to selective reporting	
Not at risk of bias	**25 (76)**
At risk of bias	**8 (24)**
Reporting incomplete, unclear, or inconsistent	7 (21)
(Subgroup) analyses were performed in a data-driven manner	1 (3)
Insufficient information	**NA**

* more than one reason could apply
Abbreviations: HCQ – hydroxychloroquine

3.3. Outcome measure and effect size

Twenty-one of 33 studies (64%) measured the effect of HCQ on mortality. Other outcome measures that were used were survival (*n* = 2, 6%); ICU admission (*n* = 2, 6%); hospital length of stay (*n* = 2, 6%); or a composite outcome of mortality, ventilation, and/or ICU admission (*n* = 6, 18%). Since the outcomes for mortality and ICU admission as well as the composite outcome had comparable benchmark estimates for the use of HCQ compared to standard care (HR 1.09, 1.13, and 1.13, respectively), these effect estimates are depicted in Figure 2. The effect estimates for survival were HR 0.96 (95% confidence interval [CI] 0.47–2.02) and HR 0.9 (0.4–2.1), and for hospital length of stay, they were -2.12 days (95% CI 0.47–4.50) and -5.41 days (95% CI -10.49–0.32) when comparing HCQ to standard care.

Almost all included studies (30 of 33) reported relative measures of the effectiveness of HCQ for which the deviation from the RCT estimates could be calculated on a log scale. Of the three remaining studies, two reported differences in length of stay, and one presented only *p*-values. The estimate in the study by Peters et al. was closest to the benchmark estimates (deviation 0.00). The authors found an HR of 1.09 (95% CI 0.81–1.47) for mortality, which is equal to the benchmark estimate.[22] The study with the most deviating estimate (deviation 0.87 on the log scale) was by Su et al., where HR 0.15 (95% CI 0.040–0.575) was reported for disease aggravation.[23]

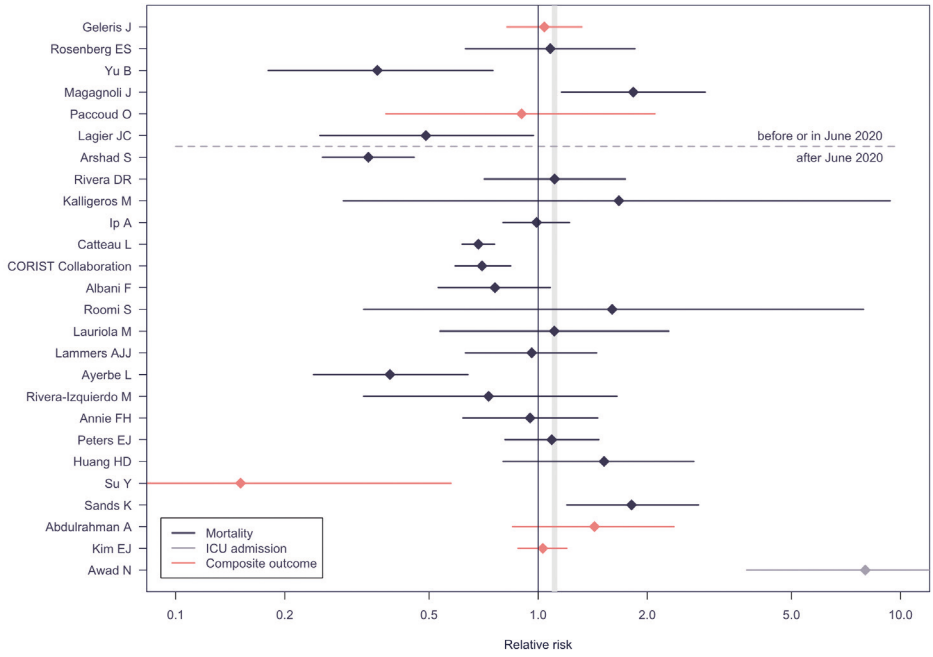


FIGURE 2. Effect estimates and 95% confidence intervals of the effect of HCQ on outcomes mortality, ICU admission, and a composite outcome of disease aggravation.
Estimates are in chronological order. The dashed line indicates June 2020, in which both the interim results of the RECOVERY trial were published and the FDA decided to revoke the emergency use authorization for HCQ.[20,21] Benchmark estimates for mortality, ICU admission, and the composite outcome are HR 1.09, 1.13, and 1.13, respectively. The grey area indicates estimates between 1.09 and 1.13. Two studies were excluded from this plot because no CIs were presented.

3.4. Relation between publication details and overall quality

The results of the Poisson regression estimating the relation between publication details and the overall quality are summarized in Table 4. Studies published in journals ranked as "Q1" had a higher overall quality than studies published in journals ranked as "Q2" or "Q3." Furthermore, we found no relation between the overall quality and the length of the reviewing process, the publication date, or the cohort size.

3.5. Relation between effect size and overall quality

We found a negative relation between the overall quality and the effect size (Table 4). The deviation from the RCT estimates decreased with an increasing number of domains that were considered not to be at risk of bias (p-value 0.02).

TABLE 4. Relation between publication characteristics, effect size, and overall quality of the studies. Results of the univariate Poisson regression analysis of the relationship between publication characteristics and effect size and the number of domains that were considered to be not at risk of bias per study

	No of domains considered to be not at risk of bias Median (IQR)	**Regression coefficient Poisson regression (95%CI)**	***p*-value**
Journal ranking			<0.001
Journals ranked as "Q1"	4.0 (3.0–4.75)	Ref	
Journals ranked as "Q2"	2.0 (1.75–3.0)	-0.58 (-1.05 to -0.11)	
Journals ranked as "Q3"	0.5 (0.0–1.0)	-2.02 (-3.43 to -0.61)	
Time from submission to publication			0.89
0–45 days	3.0 (2.0–4.0)	Ref	
46–90 days	3.0 (2.0–3.0)	-0.11 (-0.59 to 0.37)	
>90 days	2.5 (0.75–4.0)	-0.12 (-0.81 to 0.57)	
Publication date[$]			0.10
Before or in June 2020	4.0 (3.5–4.5)	Ref	
After June 2020	3.0 (1.25–4.0)	-0.37 (-0.81 to 0.07)	
Effect size (deviation from RCTs on the log scale)*	NA	-1.45 (-2.68 to -0.22)	0.02

* Calculated as abs(log(HR_obs) – log(HR_RCT).
[$] In June 2020, both the interim results of the RECOVERY trial were published and the FDA decided to revoke the emergency use authorization for HCQ.[20,21]
Abbreviations: IQR – interquartile range

4. DISCUSSION

4.1. Summary of findings

In this review, we observed that none of the included observational studies were completely free of risk of bias. Studies published in journals with a higher impact factor had a higher overall quality, which was assessed as the number of domains that were considered not to be at risk of bias. In addition, studies with effect estimates that diverged less from the RCT estimates had a higher overall quality than studies that diverged more.

Biases such as immortal time bias and competing risk bias may impact the effect estimates,[24,25] which the sensitivity analyses in some of the included studies also suggested. Studies assessing, for example, the impact of immortal time bias found large differences in the effect estimates, with estimates changing, for instance, from 1.08 to 1.46[26] or from 0.68 to 0.82[27] when HCQ use was time-varyingly defined instead of "any HCQ use during hospitalization." Interestingly, these differences in effect estimates were neither presented nor discussed in the main text, but only presented in the supplementary materials. In contrast, we found that different methods to handle missing data or different approaches for confounder adjustment (e.g., multivariable

analysis and propensity score adjustment) had limited impact on the effect estimates in this selection of studies.[22,27-31]

Furthermore, in two thirds of all included studies, insufficient information was reported in the article to fully comprehend all methodological choices. We observed this most often in the assessment of bias due to classification of interventions, bias due to missing data, and bias due to deviations from intended effects. This poor reporting is not specific to COVID-19 research and has also been observed for pharmacoepidemiologic studies in general.[32,33] Understandably, due to word limits, authors are unable to elaborate on all methodological decisions in their manuscript. However, in situations where methods deviate from generally accepted methods, substantiation of choices that were made is needed for correct interpretation of the results of a study as well as for assessment of its validity.

4.2. Strengths and limitations

The strengths of our study were the systematic assessment and analysis of all biases defined in the ROBINS-I tool and the strict inclusion criteria that were used to define our study sample. As a result, a relatively homogeneous set of included studies was obtained, which enabled us a) to specify issues in each of the domains of the ROBINS-I tool for these particular studies that could potentially lead to bias and b) to compare effect estimates.

The limitations of our study were that we scored all domains as either being at risk of bias or not, without differentiating between the severity of these potential biases. For example, studies corrected for most confounders (but not all), received the same score as studies that did not adjust for confounders at all. In addition, some potential biases might have been missed, as we paid attention to other aspects within a particular domain. For instance, within the domain of bias due to confounding, we assessed whether studies adjusted for a minimal set of confounders, whereas Martinuka et al. examined the method of confounder adjustment. One can also assess the performance of the confounder adjustment or the potential of residual confounding. However, zooming in on all different aspects was beyond the scope of this research project. Therefore, we reported in detail how we assessed the different domains in this review. Another limitation was that the quality assessment was performed on the basis of information that was reported in the publication. This is actually an indirect way of assessing the risk of bias, as the extent to which a correct assessment is possible depends on the quality of reporting.

4.3. Implications and recommendations

During the COVID-19 pandemic, with a high need for evidence-based therapy decisions, the results of observational studies formed the basis for clinical guidelines, at a time when RCT results were not yet available. Although none of the studies were free of potential biases, there was an association between their overall quality and the extent to which the effect estimates deviated from the RCT estimates. However, the promising results of some early observational studies led to strong recommendations to treat with HCQ in some countries or healthcare organizations, instead of waiting for the results of RCTs or high quality observational studies.[6,8-10] Our first recommendation is that, if uncertainty exists about the effectiveness of a potential treatment, these drug candidates should be prescribed and tested in an RCT setting, in order to gain evidence, rather than prescribing them off-label.

Second, when treatment with HCQ is already strongly advised in the clinical guidelines, it is difficult to make a valid comparison of the outcome risk between treated and untreated patients, as there must be underlying reasons why subjects are not treated, for example contra-indications for the use of HCQ.[6,9,10] Since subjects without contra-indications are likely to be healthier than those with contra-indications, the results of these studies are likely to confirm the positive results of previous studies that led to these strong recommendations, due to confounding by indication. In circumstances with such strong beliefs in the effectiveness of a potential treatment, random allocation of treatments is necessary to truly assess the effect of potential drugs. Therefore, the appropriateness of an observational design should always be assessed.

Our third recommendation is that the quality of observational studies must be improved. Observational studies can provide valid effect estimates, if properly designed and analyzed. The publication of results from invalid or flawed observational studies, however, will likely create confusion, as has been seen in the debate on the effectiveness of HCQ. On the one hand, to improve the quality, researchers should work in multidisciplinary teams that include clinicians, methodologists, and database experts, among others, to combine their knowledge and research skills. On the other hand, guidelines for designing observational studies should be used. There are currently a number of guidelines to support the design of a pharmacoepidemiologic study and to avoid potential biases.[34-38] In addition to these general pharmacoepidemiologic guidelines, recommendations specific to pharmacoepidemiologic COVID-19 research were published at the beginning of the pandemic (May 5th, 2020).[4] One can also design an observational study as if it was an RCT.[39] This "emulated trial design" framework may be helpful in avoiding biases that can otherwise easily occur in observational studies.[40]

Fourth, journals and their editors also have a responsibility to guard the quality of the studies that are published, both in the process of peer review and in their final decision regarding whether or not to publish the study results. The fact that studies published in journals with a higher impact factor were of higher quality than studies published in journals with a lower impact factor, may be partially due to the efforts of the editorial teams. To guarantee sufficient quality, journals should encourage or even oblige the use of checklists by authors and reviewers, such as ROBINS-I or RECORD-PE.[18,41] Moreover, reviewers must have sufficient expertise to critically review the quality of submitted studies against the presence of potential biases. As an aid, reviewer guidelines have recently been published on how to assess and interpret real-world evidence from observational studies.[42]

4.4. Conclusions

To conclude, the overall quality of observational studies on the effectiveness of in-hospital use of HCQ for the treatment of COVID-19 symptoms was poor, and studies reporting more extreme estimates appeared to be of lower quality than studies that found estimates closer to those of RCTs. The urgency of situations such as a pandemic should never be an argument for conducting and publishing observational studies that are of low quality, more so because in such situations, results quickly find their way into daily practice, and the results of biased studies can have potentially harmful consequences for patients.

Author's contribution

MH designed the study, performed the literature search, conducted the bias assessment, wrote the first draft of the manuscript, and implemented the contribution of the co-authors. During the whole process she implemented input and feedback from the other contributors to this study.

REFERENCES

1. Rubin EJ, Harrington DP, Hogan JW, Gatsonis C, Baden LR, Hamel MB. The Urgency of Care during the Covid-19 Pandemic — Learning as We Go. *N Engl J Med*. 2020;382(25):2461-2462. doi:10.1056/nejme2015903
2. Fantini J, Di Scala C, Chahinian H, Yahi N. Structural and molecular modelling studies reveal a new mechanism of action of chloroquine and hydroxychloroquine against SARS-CoV-2 infection. *Int J Antimicrob Agents*. 2020;55(5):105960.
3. Gautret P, Lagier J-C, Parola P, et al. Clinical and microbiological effect of a combination of hydroxychloroquine and azithromycin in 80 COVID-19 patients with at least a six-day follow up: A pilot observational study. *Travel Med Infect Dis*. 2020;34:101663. doi:10.1016/j.tmaid.2020.101663
4. Pottegård A, Kurz X, Moore N, Christiansen CF, Klungel O. Considerations for pharmacoepidemiological analyses in the SARS-CoV-2 pandemic. *Pharmacoepidemiol Drug Saf*. 2020;29(8):825-831. doi:10.1002/pds.5029
5. Awad N, Schiller DS, Fulman M, Chak A. Impact of hydroxychloroquine on disease progression and ICU admissions in patients with SARS-CoV-2 infection. *Am J Heal Pharm*. Published online February 2021. doi:10.1093/ajhp/zxab056
6. Lagier J-C, Million M, Gautret P, et al. Outcomes of 3,737 COVID-19 patients treated with hydroxychloroquine/azithromycin and other regimens in Marseille, France: A retrospective analysis. *Travel Med Infect Dis*. 2020;36:101791. doi:10.1016/j.tmaid.2020.101791
7. Alves CP de L, Segundo J de DB, Costa GG da, et al. How a few poorly designed COVID-19 studies may have contributed to misinformation in Brazil: the case for evidence-based communication of science. *BMJ Open Sci*. 2021;5(1):e100202. doi:10.1136/BMJOS-2021-100202
8. Arshad S, Kilgore P, Chaudhry ZS, et al. Treatment with hydroxychloroquine, azithromycin, and combination in patients hospitalized with COVID-19. *Int J Infect Dis*. 2020;97:396-403. doi:10.1016/j.ijid.2020.06.099
9. Lauriola M, Pani A, Ippoliti G, et al. Effect of Combination Therapy of Hydroxychloroquine and Azithromycin on Mortality in Patients With COVID-19. *Clin Transl Sci*. 2020;13(6):1071-1076. doi:10.1111/cts.12860
10. Lee TC, MacKenzie LJ, McDonald EG, Tong SYC. An observational cohort study of hydroxychloroquine and azithromycin for COVID-19: (Can't Get No) Satisfaction. *Int J Infect Dis*. 2020;98:216-217. doi:10.1016/j.ijid.2020.06.095
11. The RECOVERY Collaborative Group. Effect of Hydroxychloroquine in Hospitalized Patients with Covid-19. *N Engl J Med*. 2020;383(21):2030-2040. doi:10.1056/NEJMoa2022926
12. Pan H, Peto R, Karim QA, et al. Repurposed Antiviral Drugs for Covid-19 — Interim WHO Solidarity Trial Results. *N Engl J Med*. 2021;384(6):497-511. doi:10.1056/nejmoa2023184
13. Kim AHJ, Sparks JA, Liew JW, et al. A Rush to Judgment? Rapid Reporting and Dissemination of Results and Its Consequences Regarding the Use of Hydroxychloroquine for COVID-19. *Ann Intern Med*. 2020;172(12):819-821. doi:10.7326/M20-1223
14. Fiolet T, Guihur A, Rebeaud ME, Mulot M, Peiffer-Smadja N, Mahamat-Saleh Y. Effect of hydroxychloroquine with or without azithromycin on the mortality of coronavirus disease 2019 (COVID-19) patients: a systematic review and meta-analysis. *Clin Microbiol Infect*. 2021;27(1):19-27. doi:10.1016/j.cmi.2020.08.022
15. Renoux C, Azoulay L, Suissa S. Biases in evaluating the safety and effectiveness of drugs for COVID-19: designing real-world evidence studies. *Am J Epidemiol*. 2021;190(8):1452-1456. doi:10.1093/aje/kwab028
16. Martinuka O, von Cube M, Wolkewitz M. Methodological evaluation of bias in observational coronavirus disease 2019 studies on drug effectiveness. *Clin Microbiol Infect*. 2021;27(7):949-957. doi:10.1016/j.cmi.2021.03.003
17. Clarivate. InCites Journal Citation Reports. Published 2021. Accessed August 4, 2021. https://jcr.clarivate.com

18. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355. doi:10.1136/bmj.i4919
19. Siemieniuk RAC, Bartoszko JJ, Ge L, et al. Drug treatments for covid-19: Living systematic review and network meta-Analysis. *BMJ*. 2020;370. doi:10.1136/bmj.m2980
20. RECOVERY trial to stop enrolling patients to "ineffective" hydroxychloroquine arm - The Pharmaceutical Journal. Accessed September 28, 2021. https://pharmaceutical-journal.com/article/news/recovery-trial-to-stop-enrolling-patients-to-ineffective-hydroxychloroquine-arm
21. Coronavirus (COVID-19) Update: FDA Revokes Emergency Use Authorization for Chloroquine and Hydroxychloroquine | FDA. Accessed October 6, 2021. https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-revokes-emergency-use-authorization-chloroquine-and
22. Peters EJ, Collard D, Van Assen S, et al. Outcomes of persons with coronavirus disease 2019 in hospitals with and without standard treatment with (hydroxy)chloroquine. *Clin Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis*. 2021;27(2):264-268. doi:10.1016/j.cmi.2020.10.004
23. Su Y, Ling Y, Ma Y, et al. Efficacy of early hydroxychloroquine treatment in preventing COVID-19 pneumonia aggravation, the experience from Shanghai, China. *Biosci Trends*. 2021;14(6):408-414. doi:10.5582/bst.2020.03340
24. Suissa S, Dell'Aniello S. Time-related biases in pharmacoepidemiology. *Pharmacoepidemiol Drug Saf*. 2020;29(9):1101-1110. Accessed June 17, 2021. https://onlinelibrary.wiley.com/doi/full/10.1002/pds.5083
25. Wolkewitz M, Schumacher M. Survival biases lead to flawed conclusions in observational treatment studies of influenza patients. *J Clin Epidemiol*. 2017;84:121-129. doi:10.1016/J.JCLINEPI.2017.01.008
26. Rosenberg ES, Dufort EM, Udo T, et al. Association of Treatment With Hydroxychloroquine or Azithromycin With In-Hospital Mortality in Patients With COVID-19 in New York State. *JAMA*. 2020;323(24):2493-2502. doi:10.1001/jama.2020.8630
27. Catteau L, Dauby N, Montourcy M, et al. Low-dose hydroxychloroquine therapy and mortality in hospitalised patients with COVID-19: a nationwide observational study of 8075 participants. *Int J Antimicrob Agents*. 2020;56(4):106144. doi:10.1016/j.ijantimicag.2020.106144
28. Albani F, Fusina F, Giovannini A, et al. Impact of Azithromycin and/or Hydroxychloroquine on Hospital Mortality in COVID-19. *J Clin Med*. 2020;9(9). doi:10.3390/jcm9092800
29. Ayerbe L, Risco-Risco C, Ayis S. The association of treatment with hydroxychloroquine and hospital mortality in COVID-19 patients. *Intern Emerg Med*. 2020;15(8):1501-1506. doi:10.1007/s11739-020-02505-x
30. Collaboration COVID-19 RISK and Treatments (CORIST). Use of hydroxychloroquine in hospitalised COVID-19 patients is associated with reduced mortality: Findings from the observational multicentre Italian CORIST study. *Eur J Intern Med*. 2020;82:38-47. doi:10.1016/j.ejim.2020.08.019
31. Magagnoli J, Narendran S, Pereira F, et al. Outcomes of Hydroxychloroquine Usage in United States Veterans Hospitalized with COVID-19. *Med (New York, NY)*. 2020;1(1):114-127.e3. doi:10.1016/j.medj.2020.06.001
32. Hempenius M, Luijken K, de Boer A, Klungel O, Groenwold R, Gardarsdottir H. Quality of reporting of drug exposure in pharmacoepidemiological studies. *Pharmacoepidemiol Drug Saf*. 2020;29(9):1141-1150. doi:10.1002/pds.5020
33. Hunt NB, Gardarsdottir H, Bazelier MT, Klungel OH, Pajouheshnia R. A systematic review of how missing data are handled and reported in multi-database pharmacoepidemiologic studies. *Pharmacoepidemiol Drug Saf*. 2021;30(7):819-826. doi:10.1002/pds.5245
34. ISPE. Guidelines for good pharmacoepidemiology practices (GPP). *Pharmacoepidemiol Drug Saf*. 2008;17(2):200-208. doi:10.1002/pds.1471

35. European Medicines Agency. *The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology, 9th Revision*. Accessed August 4, 2021. http://www.encepp.eu/standards_and_guidance
36. U.S. Department of Health and Human Services, Administration Food and Drug, Research (CDER) Center for Drug Evaluation and, Research (CBER) Center for Biologics Evaluation and. Guidance for Industry and FDA Staff Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data. *Drug Saf*. Published online 2013. Accessed June 18, 2018. http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm
37. Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. Published online 2013. Accessed May 3, 2019. https://www.ncbi.nlm.nih.gov/books/NBK126190/
38. European Medicines Agency. *Guideline on Good Pharmacovigilance Practices (GVP) - Module VIII – Post-Authorisation Safety Studies (Rev 3).*; 2017. Accessed October 31, 2019. www.ema.europa.eu
39. Hernán MA, Robin J. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016;183(8):758-764. doi:10.1093/AJE/KWV254
40. Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. 2016;79:70-75. Accessed June 17, 2021. https://pubmed.ncbi.nlm.nih.gov/27237061/
41. Langan SM, Schmidt SA, Wing K, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). *BMJ*. 2018;363:k3532. doi:10.1136/bmj.k3532
42. Wang S V., Schneeweiss S. Assessing and Interpreting Real-World Evidence Studies: Introductory Points for New Reviewers. *Clin Pharmacol Ther*. Published online August 20, 2021. doi:10.1002/CPT.2398

SUPPLEMENTARY MATERIALS

A) Description of specific issues that could occur specifically in in-hospital studies on the effectiveness of HCQ

Domain 1: Bias due to confounding

Bias due to confounding could exist when there are prognostic factors related to both the treatment status and the outcome. A minimal set of potential confounders that play a role in the association between in-hospital treatment with HCQ and clinical outcomes in COVID-19 patients was defined (see supplementary material B). This set was based on baseline characteristics that were presented in articles about RCTs on the effectiveness of HCQ, which were published in high-impact journals (impact factor > 30).[1–5] The minimal set of confounders included the following: age; sex; any measure of COVID-19 severity; body mass index (BMI); and several comorbidities including diabetes, hypertension, heart disease, lung disease, and kidney disease. We considered studies to be at risk of bias if not corrected for all potential confounders.

Domain 2: Bias due to selection of participants

The inclusion and exclusion criteria for patients were assessed with regard to potential introduction of selection bias. We considered studies to be at risk of bias when any of the following inclusion and exclusion criteria were used: 1) inclusion of subjects based on discharge data, since excluding those who were still hospitalized at the end of the study period could lead to a selection bias; 2) exclusion of subjects who prematurely discontinued the use of HCQ; 3) different start of follow-up applied for HCQ users and non-users (e.g., admission date for non-HCQ users and start treatment for HCQ users); or 4) exclusion of subjects experiencing the outcome before the start of HCQ treatment. Studies were considered not to be at risk of selection bias when these criteria were not used.

Domain 3: Bias due to classification of treatments

To assess the risk of bias due to treatment misclassification, we extracted information on how HCQ treatment was defined. Studies were considered to be at risk of bias when the treatment status was not correctly classified during follow-up. This may occur, for example, when the complete study follow-up is classified as "treated with HCQ" even though the treatment is not directly initiated on admission, but is instead perhaps initiated several days thereafter. In this case, misclassification of treatment status can lead to immortal time bias.[6] Studies were considered not to be at risk of bias when treatment was analyzed as a time-varying exposure (i.e., accounting for time on and off treatment), or when all treatments and thus follow-up started on the day of admission.

Domain 4: Bias due to deviations from intended interventions

Bias due to deviations from intended interventions concerns both adherence to the treatment protocol of the study drug and the use of concomitant treatments that could differ between treatment arms. At the time when most observational studies of HCQ were conducted, there was great uncertainty about the potential effectiveness of treatments. Therefore, patients were often treated with multiple (experimental) treatments at the same time. These other treatments could also have impacted the risk of the outcome (i.e., mortality or need for mechanical ventilation). By now, RCTs have shown that corticosteroids are an effective treatment.[7] Therefore, we assessed whether there were differences in cotreatment with corticosteroids, which could have led to

bias. In addition, inadequate adjustment for treatments delivered during hospitalization could also have led to bias, for example if not analyzed in a time-varying manner.[8]

Studies were thus considered to be at risk of bias when 1) there were differences in the use of corticosteroids between HCQ users and non-HCQ users, and these differences were not appropriately corrected for; or 2) when adjustment was made for intermediates, such as ventilator use or use of other drugs, assessed at any time during hospitalization. Furthermore, we considered studies not to be at risk of bias when 1) subjects using corticosteroids were excluded; 2) there were no (expected) differences in treatments with corticosteroids between HCQ users and non-HCQ users; 3) the studies employed time-varying methods to adjust for time-varying cotreatments; or 4) adjustment was only made for other treatments that were started before or at study baseline.

Domain 5: Bias due to missing data
Missing information on treatment, COVID-19 related outcomes, and study subject characteristics may also cause bias. We hence considered studies to be at risk of bias in the following situations: 1) the presence of >5% or an unknown proportion of missing data without any report of how missing values were handled or 2) the use of a method to deal with missing data (>5%) that is often considered inappropriate (e.g., complete case analysis, missing indicator method, or single value substitution), without substantiation of the appropriateness of that method.[9–12] Studies were considered not to be at risk of bias in the following situations: 1) there were no, or less than 5%, missing data per variable; 2) missing data were handled using multiple imputations; or 3) supportive evidence was provided for the use of other methods to deal with the missing data.

Domain 6: Bias due to measurements of outcomes
Within the domain of bias due to outcome measurement, studies were considered to be at risk of bias in the following situations: 1) when there were differences in the measurement of COVID-19 outcomes for those treated with HCQ and those not treated with HCQ; 2) when outcomes were included that only analyzed length of hospitalization, without distinguishing between discharge alive or death; or 3) when discharged patients were censored in time-to-event analyses. Studies were considered not to be at risk of bias when 1) in-hospital mortality was analyzed, 2) there was a fixed duration of follow-up for each subject (e.g., 21 days), or 3) discharge alive was included as a competing risk in the model (e.g., according to the method by Fine and Gray).[13]

Domain 7: Bias due to selection of the reported results
We considered studies to be at risk of bias due to selective reporting of results when 1) not all analyses presented in the methods sections were reported in the results section; 2) decisions regarding subgroup analysis were based on the results of data analysis; 3) inconsistencies between the main text and tables meant that the results could not be interpreted correctly; or 4) only *p*-values or relative numbers were reported without absolute numbers for the outcomes.

B) selection of minimal set of confounders

For the relation between in-hospital treatment with HCQ and clinical outcomes in covid-19 patients we defined a minimal set of confounders that should be accounted for, based on randomized clinical trials published in high impact journals (*the British Medical Journal (The BMJ), the New England Journal of Medicine (NEJM), the Journal of the American Medical Association (JAMA),* and *The Lancet, and their subjournals). These RCTs were the WHO solidarity trial, the RECOVERY trial, the trial by Self et al, Tang et al and Cavalcanti et al.*[1–5] We defined a minimal set of confounders that should be included in the observational studies, which were those included in at least 3 of 5 RCTs.

TABLE S1. Baseline characteristics measured in RCTs (as reported in Table 1)

Parameter	Cavalcanti	Tang	Solidarity	RECOVERY	Self
Age	x	x	x	x	x
Sex	x	x	x	x	x
Race/ethnicity				x	x
Geographical region			x		
Living at home/nursing home					x
Smoking status	x		x		
BMI					x
Location of randomization	x				x
Disease severity					
On ordinal scale	x	x			
Symptoms (shortness of breath, cough, fever)		x			x
Respiratory support at entry			x	x	x
Presence of lesions in both lungs			x		x
Previous days in hospital	x		x	x	x
No of days since symptom onset	x	x		x	x
SOFA score					x
Laboratory measurements					x
White cell count		x			x
Lymphocyte count		x			
Neutrophil count		x			
Platelet count		x			x
Haemoglobin		x			
Aspartate aminotransferase		x			x
Alanine aminotransferase		x			x

TABLE S1. Continued.

Parameter	Cavalcanti	Tang	Solidarity	RECOVERY	Self
γ-glutamyl transpeptidase		x			
Total bilirubin		x			
Albumin		x			
Lactate dehydrogenase		x			
Creatine kinase		x			x
Creatine kinase isoenzyme-MB		x			
Creatinine		x			
Blood urea nitrogen		x			
Urea		x			
International normalized ratio		x			
C reactive protein		x			
Erythrocyte sedimentation rate		x			
Tumour necrosis factor α		x			
Interleukin 6		x			
Coexisting conditions					
Diabetes	x	x	x	x	x
Hypertension	x	x			x
Heart disease	x		x	x	x
Chronic lung disease	x		x	x	x
Asthma	x		x		
Chronic liver disease			x	x	
Tuberculosis				x	
HIV-infection	x			x	
Kidney disease	x			x	x
Cancer	x				
Previous medication use	x	x			
QTc interval					x
BMI		x			
Exposure history		x			

Note Gray shading indicate baseline characteristics that were included in at least 3 of 5 RCTs.

C) List of all included studies in alphabetical order

Abdulrahman A, AlSayed I, AlMadhi M, et al. The Efficacy and Safety of Hydroxychloroquine in Patients with COVID-19: A Multicenter National Retrospective Cohort. Infect Dis Ther. Published online January 2021:1-17. doi:10.1007/s40121-021-00397-8

Albani F, Fusina F, Giovannini A, et al. Impact of Azithromycin and/or Hydroxychloroquine on Hospital Mortality in COVID-19. J Clin Med. 2020;9(9). doi:10.3390/jcm9092800

Almazrou SH, Almalki ZS, Alanazi AS, Alqahtani AM, Alghamd SM. Comparing the impact of Hydroxychloroquine based regimens and standard treatment on COVID-19 patient outcomes: A retrospective cohort study. Saudi Pharm J SPJ Off Publ Saudi Pharm Soc. 2020;28(12):1877-1882. doi:10.1016/j.jsps.2020.09.019

Annie FH, Sirbu C, Frazier KR, Broce M, Lucas BDJ. Hydroxychloroquine in Hospitalized Patients with COVID-19: Real-World Experience Assessing Mortality. Pharmacotherapy. 2020;40(11):1072-1081. doi:10.1002/phar.2467

Arshad S, Kilgore P, Chaudhry ZS, et al. Treatment with hydroxychloroquine, azithromycin, and combination in patients hospitalized with COVID-19. Int J Infect Dis IJID Off Publ Int Soc Infect Dis. 2020;97:396-403. doi:10.1016/j.ijid.2020.06.099

Ashinyo ME, Duti V, Dubik SD, et al. Clinical characteristics, treatment regimen and duration of hospitalization among COVID-19 patients in Ghana: a retrospective cohort study. Pan Afr Med J. 2020;37(Suppl 1):9. doi:10.11604/pamj.supp.2020.37.9.25718

Awad N, Schiller DS, Fulman M, Chak A. Impact of hydroxychloroquine on disease progression and ICU admissions in patients with SARS-CoV-2 infection. Am J Heal Pharm AJHP Off J Am Soc Heal Pharm. Published online February 2021. doi:10.1093/ajhp/zxab056

Ayerbe L, Risco-Risco C, Ayis S. The association of treatment with hydroxychloroquine and hospital mortality in COVID-19 patients. Intern Emerg Med. 2020;15(8):1501-1506. doi:10.1007/s11739-020-02505-x

Catteau L, Dauby N, Montourcy M, et al. Low-dose hydroxychloroquine therapy and mortality in hospitalised patients with COVID-19: a nationwide observational study of 8075 participants. Int J Antimicrob Agents. 2020;56(4):106144. doi:10.1016/j.ijantimicag.2020.106144

Use of hydroxychloroquine in hospitalised COVID-19 patients is associated with reduced mortality: Findings from the observational multicentre Italian CORIST study. Eur J Intern Med. 2020;82:38-47. doi:10.1016/j.ejim.2020.08.019

Dubernet A, Larsen K, Masse L, et al. A comprehensive strategy for the early treatment of COVID-19 with azithromycin/hydroxychloroquine and/or corticosteroids: Results of a retrospective observational study in the French overseas department of Réunion Island. J Glob Antimicrob Resist. 2020;23:1-3. doi:10.1016/j.jgar.2020.08.001

Geleris J, Sun Y, Platt J, et al. Observational Study of Hydroxychloroquine in Hospitalized Patients with Covid-19. N Engl J Med. 2020;382(25):2411-2418. doi:10.1056/NEJMoa2012410

Grimaldi D, Aissaoui N, Blonz G, et al. Characteristics and outcomes of acute respiratory distress syndrome related to COVID-19 in Belgian and French intensive care units according to antiviral strategies: the COVADIS multicentre observational study. Ann Intensive Care. 2020;10(1):131. doi:10.1186/s13613-020-00751-y

Huang HD, Jneid H, Aziz M, et al. Safety and Effectiveness of Hydroxychloroquine and Azithromycin Combination Therapy for Treatment of Hospitalized Patients with COVID-19: A Propensity-Matched Study. Cardiol Ther. 2020;9(2):523-534. doi:10.1007/s40119-020-00201-7

Ip A, Berry DA, Hansen E, et al. Hydroxychloroquine and tocilizumab therapy in COVID-19 patients-An observational study. PLoS One. 2020;15(8):e0237693. doi:10.1371/journal.pone.0237693

Kalligeros M, Shehadeh F, Atalla E, et al. Hydroxychloroquine use in hospitalised patients with COVID-19: An observational matched cohort study. J Glob Antimicrob Resist. 2020;22:842-844. doi:10.1016/j.jgar.2020.07.018

Kim EJ, Coppa K, Hirsch JS, et al. Examination of patient characteristics and hydroxychloroquine use based on the US Food and Drug Administration's recommendation: a cross-sectional analysis in New York. BMJ Open. 2021;11(2):e042965. doi:10.1136/bmjopen-2020-042965

Lagier J-C, Million M, Gautret P, et al. Outcomes of 3,737 COVID-19 patients treated with hydroxychloroquine/azithromycin and other regimens in Marseille, France: A retrospective analysis. Travel Med Infect Dis. 2020;36:101791. doi:10.1016/j.tmaid.2020.101791

Lammers AJJ, Brohet RM, Theunissen REP, et al. Early hydroxychloroquine but not chloroquine use reduces ICU admission in COVID-19 patients. Int J Infect Dis IJID Off Publ Int Soc Infect Dis. 2020;101:283-289. doi:10.1016/j.ijid.2020.09.1460

Lauriola M, Pani A, Ippoliti G, et al. Effect of Combination Therapy of Hydroxychloroquine and Azithromycin on Mortality in Patients With COVID-19. Clin Transl Sci. 2020;13(6):1071-1076. doi:10.1111/cts.12860

Lecronier M, Beurton A, Burrel S, et al. Comparison of hydroxychloroquine, lopinavir/ritonavir, and standard of care in critically ill patients with SARS-CoV-2 pneumonia: an opportunistic retrospective analysis. Crit Care. 2020;24(1):418. doi:10.1186/s13054-020-03117-9

Magagnoli J, Narendran S, Pereira F, et al. Outcomes of Hydroxychloroquine Usage in United States Veterans Hospitalized with COVID-19. Med (New York, NY). 2020;1(1):114-127.e3. doi:10.1016/j.medj.2020.06.001

Mahévas M, Tran V-T, Roumier M, et al. Clinical efficacy of hydroxychloroquine in patients with covid-19 pneumonia who require oxygen: observational comparative study using routine care data. BMJ. 2020;369:m1844. doi:10.1136/bmj.m1844

Paccoud O, Tubach F, Baptiste A, et al. Compassionate use of hydroxychloroquine in clinical practice for patients with mild to severe Covid-19 in a French university hospital. Clin Infect Dis an Off Publ Infect Dis Soc Am. Published online June 2020. doi:10.1093/cid/ciaa791

Peters EJ, Collard D, Van Assen S, et al. Outcomes of persons with coronavirus disease 2019 in hospitals with and without standard treatment with (hydroxy)chloroquine. Clin Microbiol Infect Off Publ Eur Soc Clin Microbiol Infect Dis. 2021;27(2):264-268. doi:10.1016/j.cmi.2020.10.004

Rivera DR, Peters S, Panagiotou OA, et al. Utilization of COVID-19 Treatments and Clinical Outcomes among Patients with Cancer: A COVID-19 and Cancer Consortium (CCC19) Cohort Study. Cancer Discov. 2020;10(10):1514-1527. doi:10.1158/2159-8290.CD-20-0941

Rivera-Izquierdo M, Valero-Ubierna MDC, R-delAmo JL, et al. Therapeutic agents tested in 238 COVID-19 hospitalized patients and their relationship with mortality. Med Clin (English ed). 2020;155(9):375-381. doi:10.1016/j.medcle.2020.06.024

Roomi S, Ullah W, Ahmed F, et al. Efficacy of Hydroxychloroquine and Tocilizumab in Patients With COVID-19: Single-Center Retrospective Chart Review. J Med Internet Res. 2020;22(9):e21758. doi:10.2196/21758

Rosenberg ES, Dufort EM, Udo T, et al. Association of Treatment With Hydroxychloroquine or Azithromycin With In-Hospital Mortality in Patients With COVID-19 in New York State. JAMA. 2020;323(24):2493-2502. doi:10.1001/jama.2020.8630

Sands K, Wenzel R, McLean L, et al. No clinical benefit in mortality associated with hydroxychloroquine treatment in patients with COVID-19. Int J Infect Dis. 2021;104:34-40. doi:10.1016/j.ijid.2020.12.060

Su Y, Ling Y, Ma Y, et al. Efficacy of early hydroxychloroquine treatment in preventing COVID-19 pneumonia aggravation, the experience from Shanghai, China. Biosci Trends. 2021;14(6):408-414. doi:10.5582/bst.2020.03340

Vernaz N, Agoritsas T, Calmy A, et al. Early experimental COVID-19 therapies: associations with length of hospital stay, mortality and related costs. Swiss Med Wkly. 2020;150:w20446. doi:10.4414/smw.2020.20446

Yu B, Li C, Chen P, et al. Low dose of hydroxychloroquine reduces fatality of critically ill patients with COVID-19. Sci China Life Sci. 2020;63(10):1515-1521. doi:10.1007/s11427-020-1732-2

References (supplementary materials)

1. Pan H, Peto R, Karim QA, et al. Repurposed Antiviral Drugs for Covid-19 — Interim WHO Solidarity Trial Results. *N Engl J Med*. 2021;384(6):497-511. doi:10.1056/nejmoa2023184
2. The RECOVERY Collaborative Group. Effect of Hydroxychloroquine in Hospitalized Patients with Covid-19. *N Engl J Med*. 2020;383(21):2030-2040. doi:10.1056/NEJMoa2022926
3. Tang W, Cao Z, Han M, et al. Hydroxychloroquine in patients with mainly mild to moderate coronavirus disease 2019: Open label, randomised controlled trial. *BMJ*. 2020;369. doi:10.1136/bmj.m1849
4. Cavalcanti AB, Zampieri FG, Rosa RG, et al. Hydroxychloroquine with or without Azithromycin in Mild-to-Moderate Covid-19. *N Engl J Med*. 2020;383(21):2041-2052. doi:10.1056/NEJMoa2019014
5. Self WH, Semler MW, Leither LM, et al. Effect of Hydroxychloroquine on Clinical Status at 14 Days in Hospitalized Patients with COVID-19: A Randomized Clinical Trial. *JAMA*. 2020;324(21):2165-2176. doi:10.1001/jama.2020.22240
6. Suissa S, Dell'Aniello S. Time-related biases in pharmacoepidemiology. *Pharmacoepidemiol Drug Saf*. 2020;29(9):1101-1110. Accessed June 17, 2021. https://onlinelibrary.wiley.com/doi/full/10.1002/pds.5083
7. Siemieniuk RAC, Bartoszko JJ, Ge L, et al. Drug treatments for covid-19: Living systematic review and network meta-Analysis. *BMJ*. 2020;370. doi:10.1136/bmj.m2980
8. Schisterman EF, Cole SR, Platf RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology*. 2009;20(4):488-495. doi:10.1097/EDE.0b013e3181a819a1
9. Martín-Merino E, Calderón-Larrañaga A, Hawley S, et al. The impact of different strategies to handle missing data on both precision and bias in a drug safety study: A multidatabase multinational population-based cohort study. *Clin Epidemiol*. 2018;10:643-654. doi:10.2147/CLEP.S154914
10. Knol MJ, Janssen KJM, Donders ART, et al. Unpredictable bias when using the missing indicator method or complete case analysis for missing confounder values: an empirical example. *J Clin Epidemiol*. 2010;63(7):728-736. doi:10.1016/j.jclinepi.2009.08.028
11. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009;338:b2393. doi:10.1136/bmj.b2393
12. Greenland S, Finkle WD. A Critical Look at Methods for Handling Missing Covariates in Epidemiologic Regression Analyses. *Am J Epidemiol*. 1995;142(12):1255-1264. doi:10.1093/oxfordjournals.aje.a117592
13. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *J Am Stat Assoc*. 1999;94(446):496-509. doi:10.1080/01621459.1999.10474144

3.

THE EXTENT AND IMPACT OF EXPOSURE MISCLASSIFICATION

3.1

PRIMARY NONADHERENCE TO DRUGS PRESCRIBED BY GENERAL PRACTITIONERS: A DUTCH DATABASE STUDY

Mirjam Hempenius, Simone Rijken, Rolf H. H. Groenwold, Karin Hek, Anthonius de Boer, Olaf H. Klungel, Helga Gardarsdottir

Submitted

ABSTRACT

Background

Primary nonadherence (PNA) is defined as not filling the first prescription for a drug treatment. PNA can lead not only to poor patient outcomes but also to exposure misclassification in prescription databases and consequently biased estimates. This study aims to estimate PNA in primary care in the Netherlands and to investigate factors associated with PNA.

Methods

Patients from the Nivel Primary Care Database who received a new prescription (>1 year not prescribed) from a general practitioner in 2012 were linked to public pharmacy dispensing information. PNA was defined as receiving a new prescription without a record of dispensing at a pharmacy within 30 days. PNA was assessed overall and per drug class. The associations between PNA and patient- and prescription-related characteristics (sex, age, neighborhood socioeconomic status, comorbidities, prescription date, and reimbursement status) were assessed using mixed effects logistic regression models.

Results

This study included 65,877 subjects who received 181,939 new drug prescriptions, for which PNA was present in 11.5%. PNA was lowest for thyroid hormones (5.5%) and highest for proton pump inhibitors (12.8%). Several factors were associated with PNA. Patients with >3 active diagnoses were more likely to be primary nonadherent (OR 1.46 95% CI [1.37–1.56] compared to no active diagnoses). Patients were more likely to be primary nonadherent to not reimbursed drugs compared to fully reimbursed (2.78 [2.65–2–92]).

Discussion

One out of 10 newly prescribed drugs were not dispensed by a pharmacy. This can lead to overestimation of the actual drug exposure status when using prescription databases.

1. INTRODUCTION

Medication nonadherence is the process of patients not using their medication as prescribed. Nonadherence can occur at several stages during medication use, which are commonly classified as the initiation phase (taking the first dose), the implementation phase (taking the right dose at the right regimen), and the discontinuation phase (discontinuing drug use at the right time).[1,2] Nonadherence in each of these phases may lead to poor patient outcomes, such as risk of (re) hospitalization, morbidity, and mortality, since patients do not receive the treatment they need.[3–7]

Not only is nonadherence a problem from a medical point of view, but it can also impact pharmacoepidemiologic studies. In studying the relation between drug treatments and health outcomes, routinely collected health data are often used for the assessment of drug exposure, including prescription or dispensing information from primary care. Nonadherence can lead to misclassification of exposure status using these databases, which may in turn lead to biased estimates of the exposure–outcome relationship.[8] Particularly when nonadherence is related to factors that are also associated with the outcome risk, the bias can be unpredictable and may lead to attenuated or exaggerated effect estimates.[9] Insight into the expected level of nonadherence during all phases is therefore important when conducting and interpreting pharmacoepidemiologic research.

Most studies on adherence focus on the implementation and discontinuation phases, whereas the initiation phase is less studied.[10] Nonadherence in the initiation phase is also called primary nonadherence (PNA) and is often measured as the proportion of newly prescribed drugs that are not dispensed at the pharmacy within a certain time window.[11] The main challenge in measuring PNA is that information on prescriptions and dispensings, often from different data sources, must be linked at the patient level for the estimation.[12] A few studies have been able to do so and have assessed PNA for specific drug classes, such as antidepressants, statins, and antihypertensives,[13–17] or across all different drug classes – one study from Denmark, one from Canada, and three from the US.[18–22] The reported PNA estimates showed large variation within these studies,[23] which can be partly explained by differences in the methods employed, including the duration of the time window in which PNA is measured.[11] The differences in PNA could also be driven by the drug class and, in relation to that, the beliefs patients may have about the efficacy.[24] For instance, PNA was described as being higher for statins, which are used in the prevention of cardiovascular disease (20.8%), than for drugs that are used for the treatment of depression (10.8%).[23] The population in which PNA is studied could also impact estimates, since PNA has been found to be associated with patient characteristics, such as age, sex, and socioeconomic status (SES).[18–21,23] Lastly, differences in reimbursement systems may also explain differences in PNA. A meta-analysis revealed that PNA was twice as high in North America compared to Europe (17.0% vs 8.5%) due to the presence of universal health coverage in most European countries, but not in the US.[23] Several other studies have also indicated that costs and reimbursement status are important drivers of PNA.[21,25,26]

Since the underlying health system may play a significant role in PNA, it is important to provide insights into PNA in different countries. We aim to assess PNA in the Netherlands, with its own healthcare and reimbursement system. In this country, all citizens are obliged to have health insurance, which reimburses all care provided by general practitioners (GPs). For almost all other

provided care, patients are required to pay a deductible excess of a few hundred euros of the total healthcare costs per year themselves. After this deductible excess is spent, most drugs are fully reimbursed by the health insurance, without copayment, which contrasts with most other European countries.[27] Information about PNA in the Dutch general population is available for a limited number of drugs,[14–16] and there is currently no overview of PNA for all drug classes. The aim of this study is thus to provide an overview of PNA in primary care within the Dutch general practice and to assess the possible factors associated with PNA.

2. METHODS

2.1. Databases and linkage procedure

Data were obtained from the Nivel Primary Care Database (Nivel-PCD),[28] the Foundation for Pharmaceutical Statistics (SFK),[29] Statistics Netherlands,[30] and the G-Standaard of the Z-Index.[31] The Nivel-PCD provides a nationally representative database comprising routine data from the electronic medical records of patients from approximately 10% (*n* = 529) of general practices in the Netherlands. Data include a patient's sex and age, morbidity data coded according to the International Classification of Primary Care (ICPC-1)[32] and any prescribed drugs including date of prescription, and the Anatomical Therapeutic Classification (ATC) code.[33] The SFK databases contain information on pharmacy dispensings, including ATC code, dispensing date, and reimbursement status (yes or no). Neighborhood SES was obtained from Statistics Netherlands, and drug pricing information and maximum reimbursed price per drug were obtained from the G-Standaard.

Subjects in the Nivel-PCD records were linked to individual data from one of the nearby participating SFK pharmacies that consented to linkage. To ensure matching, the sex, year of birth, and 4-digit postal code from the Nivel-PCD records had to fully match the SFK data, and at least 50% of the ATC codes of all drugs prescribed by GPs had to match per patient within a lag period of six days. After matching, each patient was assigned a unique patient identifier indicating the match in the Nivel-PCD and SFK data. Neighborhood SES was linked with patient's 4-digit-postal code, and pricing information was linked with ATC code. For this study, we reused the most recently linked Nivel-PCD and SFK dataset, with data linked for patients registered at NIVEL-PCD from 2011-2013.

2.2. Study population

All successfully matched subjects who received a prescription for a new drug in 2012 – defined as not having a prescription as not being prescribed in the prior 365 days – were included in the study population. Patients could receive a new prescription for multiple drugs. All prescriptions with invalid ATC codes (e.g., "Y" or "Z") were excluded for the analysis, as well as prescriptions that are not dispensed via the outpatient pharmacy in the Netherlands, such as influenza vaccines or expensive drugs (Table S1). To ensure follow-up in the SFK database, new prescriptions were only included when there was at least one drug dispensing (any) registered in the SFK database for a patient after or on the prescription date until 31/12/2013. To ensure the inclusion of newly prescribed drugs only, all prescriptions with a record of dispensing before the first record of prescription were excluded.

2.3. Definition of primary nonadherence

PNA was assessed for all new prescriptions prescribed in 2012. For this assessment, the SFK database was searched for a record of dispensing from a pharmacy within 30 days of the prescription date, matched on ATC code (fifth level). PNA was defined as not having a prescription dispensed within 30 days from the prescribing date.

2.4. Assessment of associated factors

On the patient level, we assessed the following characteristics: sex, age (categorized as 0–20 years, 21–40 years, 41–60 years (ref), 61–80 years, and 80 years and older), neighborhood SES (the highest and lowest quintiles were categorized as high and low SES scores, respectively, while the middle three quintiles were categorized as a medium SES score), the number of active diagnoses on the first day of the prescription month (categorized as 0, 1–3, and >3), the number of GP contact moments in the 12 months preceding the prescription month (categorized as 0, 1–5, and >5), the number of different drugs dispensed in the three months preceding the prescription month (defined on the fourth ATC level; categorized as 0, 1–5, and >5), and the presence of specific comorbidities (cardiovascular diseases, diabetes mellitus, respiratory diseases, psychological disorders, and malignancies; for ICPC codes, see supplementary materials Table S2).

On the prescription level, we assessed the quarter of the year in which the prescription date fell and the reimbursement status. Reimbursement status was categorized as follows: fully reimbursed, partially reimbursed (if the costs are higher than the maximum reimbursed price), conditionally reimbursed (only reimbursed after drug use for more than 6 months; Table S3),[34] or not reimbursed (e.g., vitamins or acetaminophen). The reimbursement status could change every month, thus information about reimbursement status was updated on the first day of every month.

2.5. Data analysis

Baseline characteristics of the included subjects were assessed on 01/01/2012 and described as proportions.

PNA was calculated as the proportion of the total number of new prescriptions that were not dispensed within 30 days of the prescription date. PNA was assessed overall and per ATC class (first level). In addition, PNA was assessed for drug classes that are frequently prescribed in primary care in the Netherlands (Table 1).[35]

The association between PNA and the patient- and prescription-related characteristics was assessed using mixed effects logistic regression, with a random intercept per subject, per general practice, and per pharmacy. The following characteristics were assessed: age, sex, neighborhood SES, the number of active diagnoses, the number of GP contact moments in the preceding 12 months, the number of different drugs dispensed in the preceding three months, the quartile in which the prescription fell, and the reimbursement status. All fixed effects were estimated both separately in a univariable analysis and combined in a multivariable analysis. Multicollinearity was checked, and variables were removed if necessary.

TABLE 1. Anatomical Therapeutic Classification (ATC) codes of frequently prescribed drug (classes)

Drug class	ATC code(s)
Proton pump inhibitors	A02BC
Laxatives	A06
Insulins	A10A
Oral antidiabetics	A10B
Acetylicsalicylic acid	B01AC06
Antihypertensives	C02, C03, C07, C08, C09
Statins	C10AA
Dermal steroids	D07
Hormonal anticonceptives	G03A
Thyroid hormones	H03A
Systemic antibiotics	J01
Non-steroidal anti-inflammatory drugs	M01A
Benzodiazepines	N05AB, N05CD
Selective serotonin inhibitors	N06AB
Inhalation drugs for asthma/COPD	R03A, R03B
Antihistaminics for systemic use	R06

Abbreviations: COPD - chronic obstructive pulmonary disease

As sensitivity analyses, we assessed PNA applying a definition of dispensing within 90 days and 365 days of the prescription date and by matching prescription data and dispensing data at the fourth ATC level.

3. RESULTS

The Nivel-PCD records included 396,251 subjects with at least one prescription during 2011-2013. Of those, 86,361 (21.8%) were matched to subjects in the SFK database. The characteristics of matched and unmatched subjects are presented in the supplementary materials (Table S4). Of the 86,361 matched subjects, 65,877 subjects initiated one or more new drug treatments in 2012 from 119 different GP practices, that were dispensed by 126 different pharmacies. In total, 181,939 new prescriptions were prescribed in 2012. The median number of new prescriptions per subject prescribed during 2012 was two (range: one to four).

TABLE 2. Baseline characteristics of the 65,877 included subjects, assessed at 01/01/2012.

	Number of subjects (%)	Number of new prescriptions (%)
Sex		
male	26974 (40.9)	68999 (37.9)
female	38903 (59.1)	112940 (62.1)
Age		
0-20	9782 (14.8)	20754 (11.4)
21-40	12065 (18.3)	32032 (17.6)
41-60	21964 (33.3)	60548 (33.3)
61-80	19067 (28.9)	58278 (32.0)
80+	2999 (4.6)	10327 (5.7)
Neighborhood socioeconomic status[§]		
Low	13529 (20.5)	37331 (20.5)
Medium	32705 (49.6)	90593 (49.8)
High	19494 (29.6)	53505 (29.4)
missing	149 (0.2)	510 (0.3)
The number of active diagnoses on January 1st, 2012		
0	14663 (22.3)	32089 (17.6)
1-3	37662 (57.2)	100641 (55.3)
>3	13552 (20.6)	49209 (27.0)
The number of GP contact moments in 2011		
0	6964 (10.6)	14805 (8.1)
1-5	36206 (55)	90207 (49.6)
>5	22707 (34.5)	76927 (42.3)
The number of different drugs dispensed last quartile of 2011[$]		
0	21835 (33.1)	51251 (28.2)
1-5	36183 (54.9)	100407 (55.2)
>5	7859 (11.9)	30281 (16.6)
Comorbidities		
Cardiovascular disease	6764 (10.3)	22504 (12.4)
Diabetes mellitus	6052 (9.2)	19941 (11.0)
Respiratory diseases	8877 (13.5)	29757 (16.4)
Psychiatric diseases	4453 (6.8)	14702 (8.1)
Malignancy (excl skin malignancy)	3325 (5.0)	11357 (6.2)

$ Neighborhood socioeconomic status (SES) was divided by quintile (the highest and lowest quintiles were categorized as high and low SES scores, respectively; the middle three quintiles were categorized as a medium SES score). Quintiles were based on the total Nivel-PCD population and not on the matched population.
[$] defined on the fourth ATC level
Abbreviations: GP - general practitioner

The baseline characteristics of the included 65,877 subjects are presented in Table 2. Most patients were aged 41–60 years (33.3%), and 59.1% were women. Approximately half of the study population was classified as living in a neighborhood with a medium SES, 20% in a neighborhood with a low SES, and 30% in a neighborhood with a high SES. The majority had at least one active diagnosis on 01-01-2012 (77.7%), at least one contact moment with the GP in 2011 (89.4%), and one or more drugs dispensed in the last quartile of 2011 (66.9%). Comorbidities that were most present were respiratory diseases (13.5%) and cardiovascular diseases (10.3%)

3.1. Primary nonadherence (PNA) overall and per drug class

The overall PNA was 11.5%, defined as newly prescribed drugs that were not dispensed at the pharmacy within 30 days of the prescription date. PNA varied among ATC classes (Table 3), with lower PNA for drugs prescribed to treat cardiovascular disease (ATC Class C, 8.3%) and genito-urinary system drugs and sex hormones (ATC Class G, 8.5%). In contrast, PNA was highest for antineoplastic and immunomodulating drugs (ATC Class L, 19.5%) and drugs for blood and blood-forming organs (ATC Class B, 16.1%). Within drug classes that are frequently used in primary care, PNA was 9.9%. Furthermore, PNA was highest for proton pump inhibitors (PPIs) and nonsteroidal anti-inflammatory drugs (NSAIDs) (12.8% and 11.8%, respectively) and lowest for thyroid hormones and oral antidiabetics, with 5.5% and 5.6% PNA, respectively (Table 3).

TABLE 3. Primary nonadherence[§] according to different drug classes

Drug class Anatomical chemical therapeutic class (first level)	Not dispensed/ prescribed	Proportion primary nonadherence (95% CI)
Alimentary tract and metabolism (A)	3262/23360	14.0 (13.5 – 14.4)
Blood and blood forming organs (B)	790/4992	15.8 (14.8 – 16.8)
Cardiovascular system (C)	1180/14262	8.3 (7.8 – 8.7)
Dermatologicals (D)	2674/24888	10.7 (10.4 – 11.1)
Genito urinary system and sex hormones (G)	762/8948	8.5 (7.9 – 9.1)
Systemic hormonal preparations (H) [$]	498/4243	11.7 (10.8 – 12.7)
Anti-infective for systemic use (J)	3170/28295	11.2 (10.8 – 11.6)
Antineoplastic and immunomodulating agents (L)	134/730	18.4 (15.5 – 21.2)
Musculo-skeletal system (M)	1828/15213	12 (11.5 – 12.5)
Nervous system (N)	2524/19202	13.1 (12.7 – 13.6)
Antiparasitic products, insecticides, and repellents (P)	161/1578	10.2 (8.7 – 11.7)
Respiratory system (R)	2232/23149	9.6 (9.3 – 10.0)
Sensory organs (S)	1740/13011	13.4 (12.8 – 14.0)
Various (V)	15/68	22.1 (12.2 – 31.9)

TABLE 3. Continued.

Drug class Specific drug classes	Not dispensed/ prescribed	Proportion primary nonadherence (95% CI)
Proton pump inhibitors	890/6965	12.8 (12.0 – 13.6)
Laxatives	820/7110	11.5 (10.8 – 12.3)
Insulins	45/461	9.8 (7.1 – 12.5)
Oral antidiabetics	60/1079	5.6 (4.2 – 6.9)
Acetylicsalicylic acid	78/715	10.9 (8.6 – 13.2)
Antihypertensives	593/8593	6.9 (6.4 – 7.4)
Statins	194/2686	7.2 (6.2 – 8.2)
Dermal steroids	986/13032	7.6 (7.1 – 8.0)
Hormonal anticonceptives	233/2835	8.2 (7.2 – 9.2)
Thyroid hormones	18/325	5.5 (3.1 – 8.0)
Systemic antibiotics	2736/24640	11.1 (10.7 – 11.5)
Non-steroidal anti-inflammatory drugs	1625/13727	11.8 (11.3 – 12.4)
Benzodiazepines	113/1705	6.6 (5.4 – 7.8)
Selective serotonin inhibitors	52/829	6.3 (4.6 – 7.9)
Inhalation drugs for asthma/COPD	544/6554	8.3 (7.6 – 9.0)
Antihistaminics for systemic use	540/5159	10.5 (9.6 – 11.3)

[§]Primary nonadherence was defined as not having a prescription dispensed within 30 days from prescription date.
[§] Excl sex hormones and insulines
Abbreviations: COPD - chronic obstructive pulmonary disease

3.2. Patient-related factors associated with PNA

Different patient characteristics were associated with PNA in both the univariable and multivariable analyses (Table 4). On the patient level, females were less likely to be primary nonadherent than males (OR 0.96 [0.92–0.99]). Moreover, patients aged 0–20 years and 21–40 years were more likely to be primary nonadherent than patients aged 41–60 years (OR 1.13 [1.06–1.20] and 1.19 [1.12–1.25], respectively). Patients living in a neighborhood with a high or medium SES were less likely to be primary nonadherent (OR 0.93 [0.86–0.99] and 0.92 [0.85–0.99], respectively) compared to those in low SES neighborhoods. In addition, having more different diagnoses or GP contact moment increased the likelihood of displaying PNA. Due to collinearity between these two factors, only the number of active diagnoses was included in the multivariate model, resulting in ORs of 1.24 (1.17–1.31) and 1.46 (1.37–1.56) for one to three active diagnosis and more than three active diagnoses, respectively, compared to subjects with no active diagnoses. Prevalent drug users were less likely to be primary nonadherent. The OR for PNA for patients having one to five drugs dispensed in the preceding 90 days was 0.85 [0.82-0.89] compared to naïve drug users, and 0.80 [0.75-0.85] for patients that had more than five drugs dispensed.

3.3. Prescription-related factors associated with PNA

On the prescription level, drugs with partial reimbursement were less likely to not be dispensed compared to fully reimbursed drugs (OR 0.88 [0.81–0.96]), whereas drugs that were not reimbursed and those that were reimbursed conditionally were more likely to not be dispensed (OR 2.78 [2.65–2.92] and 1.09 [1.04–1.15], respectively). Patients receiving prescriptions that should be filled during the first quarter of 2012 were more likely to exhibit PNA when compared to those receiving prescriptions in the other quarters, with PNA being least likely for patients receiving prescriptions that should be filled during the last quarter of 2012 (OR 0.68 [0.65–0.71], compared to the first quarter). This decreasing PNA over time was observed for both fully reimbursed drugs (12.3% to 8.9%) and partially reimbursed drugs (12.5% to 6.7%), but to a lesser extent for drugs that were only reimbursed after use for more than six months (12.7% to 11.4%) or drugs that were not reimbursed (23.5% to 21.4%).

3.4. Sensitivity analyses

The sensitivity analyses using different durations of time for defining PNA showed similar results to when 30 days were applied, namely 10.9% and 9.4% PNA for 90 and 365 days, respectively, compared to 11.5%. The sensitivity analysis assessing PNA at the fourth ATC level resulted in a PNA estimate of 11.2%, similar to the estimate at the fifth ATC level.

TABLE 4. Results of the mixed effects logistic regression model* assessing the association between patient and prescription characteristics and primary nonadherence§

		% Primary nonadherence (not dispensed/ prescribed)	Univariate analysis OR (95% CI)	Multivariable analysis OR (95% CI)
Sex				
	male	11.7 (8092/68999)	ref	ref
	female	11.4 (12878/112940)	0.95 (0.92 – 0.99)	0.96 (0.92 – 0.99)
Age				
	0-20	11.3 (2345/20754)	1.13 (1.07 – 1.20)	1.13 (1.06 – 1.20)
	21-40	12.1 (3889/32032)	1.15 (1.09 – 1.21)	1.19 (1.12 – 1.25)
	41-60	11.0 (6634/60548)	ref	ref
	61-80	11.8 (6870/58278)	1.08 (1.03 – 1.13)	1.04 (0.99 – 1.09)
	80+	11.9 (1232/10327)	1.11 (1.02 – 1.21)	1.05 (0.96 – 1.14)
Socioeconomic status§				
	Low	10.9 (4143/37841)	ref	ref
	Medium	10.9 (9944/91103)	0.92 (0.86 – 0.98)	0.93 (0.86 – 0.99)
	High	12.4 (6679/54015)	0.90 (0.83 – 0.98)	0.92 (0.85 – 0.99)
The number of active diagnoses on the 1st of the prescription month				
	0	9.8 (2689/27352)	ref	ref
	1-3	11.4 (11223/98512)	1.10 (1.05 – 1.16)	1.24 (1.17 – 1.31)

TABLE 4. Continued.

	% Primary nonadherence (not dispensed/ prescribed)	Univariate analysis OR (95% CI)	Multivariable analysis OR (95% CI)
>3	12.6 (7058/56075)	1.18 (1.12 – 1.25)	1.46 (1.37 – 1.56)
The number of GP contact moments in the year before the prescription month			
0	9.3 (1187/12742)	ref	NA[¶]
1-5	11.0 (9534/86372)	1.23 (1.15 – 1.31)	NA
>5	12.4 (10249/82825)	1.39 (1.29 – 1.49)	NA
The number of different drugs dispensed in the 90 days before the prescription month[#]			
0	13.5 (6128/45523)	ref	ref
1-5	11.4 (3815/33349)	0.87 (0.84 – 0.91)	0.85 (0.82 – 0.89)
>5	10.7 (2658/24880)	0.87 (0.82 – 0.92)	0.80 (0.75 – 0.85)
Prescription date			
Q1–2012	13.2 (6677/50505)	ref	ref
Q2–2012	12.3 (5792/47032)	0.90 (0.86 – 0.93)	0.90 (0.86 – 0.93)
Q3–2012	10.2 (4181/41101)	0.69 (0.66 – 0.72)	0.68 (0.65 – 0.71)
Q4–2012	10.0 (4320/43301)	0.67 (0.64 – 0.70)	0.68 (0.65 – 0.71)
Reimbursement status[^]			
Fully reimbursed	10.6 (15244/143607)	ref	ref
Not reimbursed	21.7 (3081/14171)	2.73 (2.60 – 2.86)	2.78 (2.65 – 2.92)
Partially reimbursed	9.1 (677/7413)	0.87 (0.80 – 0.94)	0.88 (0.81 – 0.96)
Conditionally reimbursed	11.8 (1968/16748)	1.08 (1.02 – 1.14)	1.09 (1.04 – 1.15)

[*] Mixed effects logistic regression, with patient, general practice, and pharmacy as random effects.
[§] Primary nonadherence was defined as not having a prescription dispensed within 30 days from prescription date.
[$] Neighborhood socioeconomic status (SES) was divided by quintile (the highest and lowest quintiles were categorized as high and low SES scores, respectively; the middle three quintiles were categorized as a medium SES score).
[#] Defined on the fourth ATC level.
[^] Drugs were categorized as partially reimbursed if the costs were higher than the maximum reimbursed price. Drugs were categorized as conditionally reimbursed if they were only reimbursed after use for more than six months.
[¶] Excluded from the multivariate analysis due to multicollinearity with the number of active diagnoses.

4. DISCUSSION AND CONCLUSION

Overall, 11.5% of all newly prescribed drugs that were included in this study did not have a record of dispensing in the pharmacy database within 30 days of the prescription date. Among specific drug classes that are frequently prescribed in primary care, PNA was found to be 9.9%, with the lowest level of PNA for thyroid hormones (5.5%), and the highest for PPIs (12.8%). Several patient characteristics were associated with PNA – the strongest associations were observed for patients with more comorbidities (OR for more than three active diagnoses 1.46 [1.37–1.56], compared to no active diagnoses) and for patients using more than five drugs (OR 0.80 [0.75–0.85], compared to patients using no drugs). Age, sex, and neighborhood SES were also found to be associated with PNA. On the prescription level, the strongest associations with PNA were seen for drugs without reimbursement (OR 2.78 [2.65–2.92], compared to fully reimbursed), and for the date of prescription (OR for drugs prescribed in the last quarter of 2012 0.68 [0.65–0.71], compared to the first quarter).

The estimate of PNA was in line with results from other European studies on PNA, which obtained estimates around 9%. Moreover, patterns of PNA between drug classes were similar to patterns found in Denmark, such as relatively high levels of PNA for PPIs, salicylic acid, and NSAIDs compared to a lower level of PNA for antidepressants, antihypertensives, and antidiabetic agents.[18]

The association between SES and PNA was also in line with what others have assessed.[18] The relation between age and PNA, however, varied between the different studies. We found a U-shaped association, which was also found by Shin et al., in drugs used for chronic conditions, but not for drugs used for acute conditions. Other studies have also shown an effect of age, but these results are all inconsistent: the level of PNA either increased with age,[20,21] or it decreased.[18,22,36] With regard to sex, some studies found no significant association,[18,21,22] while Shin et al. found, in contrast to our study, that men were less likely to be primarily nonadherent.[19]

The negative relation between the number of drugs in use and the likelihood of a patient being primary nonadherent was also observed in Canada. In Denmark, however, an inverse relation was observed.[18,22] In the Netherlands, this negative relation might be explained by the reimbursement system, where patients are required to pay a deductible excess of a few hundred euros of the total healthcare costs per year themselves (220 euro in 2012). After this deductible excess is spent, most drugs are fully reimbursed by the health insurance, without copayment. The more drugs in use, the higher the chance that the deductible excess is used and new drugs are reimbursed, resulting in lower levels of PNA. Since the deductible excess is reset to zero at the beginning of each year,[37] this system may also explain the association between prescription date and PNA.

The strong association between drugs without reimbursement and PNA was also noted in studies in Canada and the US.[21,22] In addition, we found an association between drugs being conditionally reimbursed and the level of PNA. In the Netherlands, there are specific conditions attached to the reimbursement of laxatives, antihistamines used for allergies, antidiarrheal agents, gastric emptying agents, and agents to protect the eyes against dehydration, which are only reimbursed for chronic use, defined as being used for six months or longer. From January

1st, 2012, PPIs and H2-receptor antagonists were added to this list of conditionally reimbursed drugs.[34] We found higher levels of PNA for drug classes thar are conditionally reimbursed.[34] Other studies investigating the impact of this conditional reimbursement measure for PPIs in the Netherlands found that the number of PPI prescriptions temporarily decreased, but stabilized in the second quarter of 2012 to the same levels as 2011.[37] However, another report found that the proportion of NSAID users who use a PPI decreased in 2012 compared to 2011 (69.0% vs. 73.3%).[38] The stabilized number of prescriptions does not necessarily mean that the number of dispensings also remains stable, due to PNA.

In contrast, drugs that were partially reimbursed showed a lower level of PNA. However, the copayment for these drugs was generally low (e.g., €1.50 per month for digoxin). Moreover, the treating physician may have a reason to initiate specifically the drug with the copayment among other options without copayment, which may explain why PNA is lower for these drugs.

The fact that these prescription-related factors associated with PNA could largely be explained by the Dutch health and reimbursement system highlights the need for transparent reporting on the health system and reimbursement rules that are in place.

4.1. Strengths and limitations

A strength of this study was the fact that prescription data and dispensing data could be linked for a large representative sample from the general population. Furthermore, we provided an overview of PNA for all medication instead of a select set of drug classes.

One of the limitations of this study was the fact that the matching procedure was based, among other things, on a minimum of 50% matching ATC codes in the Nivel-PCD data and the SFK database. Subjects with a higher degree of nonadherence were more likely to be excluded, which may have led to an underestimation of PNA. In addition, subjects that could be matched received more prescriptions than subjects that could not be matched, and had probably also more comorbidities (Table S4). The matching procedure could thus have led to biased estimates of the association between the number of drugs in use and the number of comorbidities on the one hand and PNA on the other hand.

Furthermore, the presence of a dispensing record does not automatically mean that the drug is taken by the patient. For example, for statins, antidepressant agents, and antihypertensive agents, it has been shown that approximately 20%–30% of all new users fill only one prescription, of which a proportion do not initiate at all.[39–41] Moreover, patients do not always collect the drugs that have been dispensed for them. Information on whether or not drugs are being collected was not available in the SFK data and can also lead to an underestimation of PNA.

PNA could also be overestimated for drugs that may be obtained without being recorded in the outpatient pharmacy database. This may be the case for drug prescriptions that can also be obtained over the counter, such as NSAIDs, PPIs, and antihistamines. This may also be the case for drugs that are dispensed in the in-hospital outpatient pharmacy, such as antineoplastic medicines. The PNA estimates found for these types of drugs should therefore be interpreted with caution. Patients visiting multiple pharmacies may be another reason that dispensations

may not be recorded. However, most patients (>80%) receive all their medicines from a single pharmacy,[42] and the matching procedure based on ATC codes also limits the impact of patients visiting multiple pharmacies. In addition, patients can collect their prescription after more than 30 days, although sensitivity analyses with longer windows indicate that this share was limited (10.4% for 90 days instead of 11.5%). Prescription errors can further explain a proportion of PNA. However, we do not expect this to be common for the frequently prescribed drug classes, and sensitivity analysis with fourth ATC-level matching did not lead to significantly different estimates (11.2% instead of 11.5%).

Another limitation was the fact that we used data from 2012. Nevertheless, we expect no major differences to the current situation. The associations between patient characteristics and PNA are assumed to remain the same, and there were also no major changes in the reimbursement system, except for the increase in the deductible excess (385 euro in 2021, compared to 220 euro in 2012).

4.2. Implications

Although the level of PNA differed among drug classes, the amount of PNA was around 10% for most frequently prescribed drug classes. This means that roughly 10% of all drug treatments are not initiated, which has both clinical and methodological implications. For clinical practice, the implication is that nonadherent patients are not being treated as intended by their physician, potentially leading to poor patient outcomes, increased health expenditure, and hence increased costs.[3–7,43,44] From a methodological point of view, the implication is that exposure status estimations based on prescription data may suffer from exposure misclassification due to PNA, potentially leading to biased estimates of the association between drug treatments and health outcomes.[8] Since PNA is found to be associated with patient characteristics that may also be associated with the outcomes being studied, the misclassification can be differential, potentially resulting in unpredictable bias of the effect estimate.[9] Yet, the impact of 10% exposure misclassification due to nonadherence and the difference in the level of PNA between subgroups are likely limited. For example, a simulation study of the impact of exposure misclassification on effect estimates revealed that a 10% nondifferential nonadherence could cause an approximate 10% bias toward the null effect, and the impact of differential misclassification was also limited.[45]

4.3. Conclusions

To conclude, 1 out of 10 prescriptions initiated by a GP is not dispensed from a pharmacy. PNA varies across drug classes, ranging between 5.5% for thyroid hormones and 12.8% for PPIs. PNA was found to be associated with several patient- and prescription-related characteristics, which could to some extent be explained by reimbursement levels. Therefore, in pharmacoepidemiologic studies, we recommend that researchers elaborate on the health and reimbursement system and the potential for exposure misclassification due to PNA. Moreover, when researchers utilize prescribing data, we recommend that they provide estimates of a) the amount of PNA for the specific drugs under investigation and b) the possible impact that PNA might have on effect estimates. In addition, for drugs with high levels of PNA, GP databases are less suitable, and claims or dispensings databases should be chosen instead.

Acknowledgments

We would like to thank the Dutch Foundation for Pharmaceutical Statistics for providing data.

Author's contribution

MH designed the study, conducted the data-analysis, wrote the first draft of the manuscript, and implemented the contribution of the co-authors. During the whole process she implemented input and feedback from the other contributors to this study.

REFERENCES

1. Vrijens B, De Geest S, Hughes DA, et al. A new taxonomy for describing and defining adherence to medications. *Br J Clin Pharmacol*. 2012;73(5):691-705. doi:10.1111/j.1365-2125.2012.04167.x
2. DiMatteo MR, Giordani PJ, Lepper HS, Croghan TW. Patient adherence and medical treatment outcomes: a meta-analysis. *Med Care*. 2002;40(9):794-811. doi:10.1097/01.MLR.0000024612.61915.2D
3. Ho MP, Bryson CL, Rumsfeld JS. Medication adherence: its importance in cardiovascular outcomes. *Circulation*. 2009;119(23):3028-3035.
4. Ozaki A, Choi A, Le Q, et al. Real-World Adherence and Persistence to Direct Oral Anticoagulants in Patients With Atrial Fibrillation: A Systematic Review and Meta-Analysis. *Circ Cardiovasc Qual Outcomes*. 2020;13(3). doi:10.1161/CIRCOUTCOMES.119.005969
5. Lee JS, Joyce G, Mccombs J. Outcomes Associated With Primary and Secondary Nonadherence to Cholesterol Medications. *Am J Pharm Benefits*. 2016;8(2):54-60. Accessed September 24, 2021. www.ajpb.com
6. Ho P, Rumsfeld J, Masoudi F, et al. Effect of medication nonadherence on hospitalization and mortality among patients with diabetes mellitus. *Arch Intern Med*. 2006;166(17):1836-1841. doi:10.1001/ARCHINTE.166.17.1836
7. Jackevicius CA, Li P, Tu J V. Prevalence, predictors, and outcomes of primary nonadherence after acute myocardial infarction. *Circulation*. 2008;117(8):1028-1036. doi:10.1161/CIRCULATIONAHA.107.706820
8. Copeland KT, Checkoway H, Mcmichael AJ, Holbrook RH. Bias due to misclassification in the estimation of relative risk. *Am J Epidemiol*. 1977;105(5):488-495. doi:10.1093/oxfordjournals.aje.a112408
9. Jurek AM, Greenland S, Maldonado G. Brief Report: How far from non-differential does exposure or disease misclassification have to be to bias measures of association away from the null? *Int J Epidemiol*. 2008;37(2):382-385. doi:10.1093/ije/dym291
10. Cahir C. Primary nonadherence: The forgotten component of medication adherence? *Polish Arch Intern Med*. 2020;130(1):1-3. doi:10.20452/pamw.15164
11. Adams AJ, Stolpe SF. Defining and measuring primary medication nonadherence: Development of a quality measure. *J Manag Care Spec Pharm*. 2016;22(5):516-523. doi:10.18553/jmcp.2016.22.5.516
12. Hutchins DS, Zeber JE, Roberts CS, Williams AF, Manias E, Peterson AM. Initial Medication Adherence - Review and Recommendations for Good Practices in Outcomes Research: An ISPOR Medication Adherence and Persistence Special Interest Group Report. *Value Heal*. 2015;18(5):690-699. doi:10.1016/j.jval.2015.02.015
13. Freccero C, Sundquist K, Sundquist J, Ji J. Primary adherence to antidepressant prescriptions in primary health care: a population-based study in Sweden. *Scand J Prim Health Care*. 2016;34(1):83-88. doi:10.3109/02813432.2015.1132884
14. van Geffen EC, Gardarsdottir H, van Hulten R, van Dijk L, Egberts AC, Heerdink ER. Initiation of antidepressant therapy: do patients follow the GP's prescription? *Br J Gen Pract*. 2009;59(559):81-87. doi:10.3399/bjgp09X395067
15. Holvast F, Oude Voshaar RC, Wouters H, et al. Non-adherence to antidepressants among older patients with depression: A longitudinal cohort study in primary care. *Fam Pract*. 2018;36(1):3-11. doi:10.1093/FAMPRA/CMY106
16. Holvast F, Wouters H, Hek K, et al. Non-adherence to cardiovascular drugs in older patients with depression: A population-based cohort study. *Int J Cardiol*. 2019;274:366-371. doi:10.1016/J.IJCARD.2018.08.100
17. Thengilsdóttir G, Pottegård A, Linnet K, Halldórsson M, Almarsdóttir AB, Gardarsdóttir H. Do patients initiate therapy? Primary non-adherence to statins and antidepressants in Iceland. *Int J Clin Pract*. 2015;69(5):597-603. doi:10.1111/ijcp.12558

18. Pottegård A, Christensen R dePont, Houji A, et al. Primary non-adherence in general practice: a Danish register study. *Eur J Clin Pharmacol*. 2014;70(6):757-763. doi:10.1007/s00228-014-1677-y
19. Shin J, McCombs JS, Sanchez RJ, Udall M, Deminski MC, Cheetham TC. Primary nonadherence to medications in an integrated healthcare setting. *Am J Manag Care*. 2012;18(8):426-434.
20. Fischer MA, Stedman MR, Lii J, et al. Primary medication non-adherence: Analysis of 195,930 electronic prescriptions. *J Gen Intern Med*. 2010;25(4):284-290. doi:10.1007/s11606-010-1253-9
21. Fischer MA, Choudhry NK, Brill G, et al. Trouble Getting Started: Predictors of Primary Medication Nonadherence. *Am J Med*. 2011;124(11):1081.e9-1081.e22. doi:10.1016/J.AMJMED.2011.05.028
22. Tamblyn R, Eguale T, Huang A, Winslade N, Doran P. The incidence and determinants of primary nonadherence with prescribed medication in primary care: A cohort study. *Ann Intern Med*. 2014;160(7):441-450. doi:10.7326/M13-1705
23. Lemstra M, Nwankwo C, Bird Y, Moraros J. Primary nonadherence to chronic disease medications: A meta-analysis. *Patient Prefer Adherence*. 2018;12:721-731. doi:10.2147/PPA.S161151
24. Wouters H, Amin DFH, Taxis K, Heerdink ER, Egberts ACG, Gardarsdottir H. Associations Between Personality Traits and Adherence to Antidepressants Assessed Through Self-Report, Electronic Monitoring, and Pharmacy Dispensing Data. *J Clin Psychopharmacol*. 2016;36(5):465-471. doi:10.1097/JCP.0000000000000541
25. Briesacher BA, Gurwitz JH, Soumerai SB. Patients At-Risk for Cost-Related Medication Nonadherence: A Review of the Literature. *J Gen Intern Med 2007 226*. 2007;22(6):864-871. doi:10.1007/S11606-007-0180-X
26. Khera R, Valero-Elizondo J, Das SR, et al. Cost-Related Medication Nonadherence in Adults With Atherosclerotic Cardiovascular Disease in the United States, 2013 to 2017. *Circulation*. 2019;140(25):2067-2075. doi:10.1161/CIRCULATIONAHA.119.041974
27. World Health Organization. *Medicines Reimbursement Policies in Europe*. WHO Regional Office for Europe; 2018. Accessed October 11, 2021. https://www.euro.who.int/__data/assets/pdf_file/0011/376625/pharmaceutical-reimbursement-eng.pdf
28. Nivel Primary Care Database | Nivel. Accessed September 30, 2021. https://www.nivel.nl/en/nivel-zorgregistraties-eerste-lijn/nivel-primary-care-database
29. Stichting Farmaceutische Kengetallen [Pharmaceutical Key Figures Foundation]. Accessed September 30, 2021. https://www.sfk.nl/
30. Centraal Bureau voor de Statistiek (CBS) [Central Bureau of Statistics]. Accessed September 30, 2021. https://www.cbs.nl/
31. Z-Index. Accessed September 30, 2021. https://www.z-index.nl/
32. WONCA Working Party: International Classification (WICC) . Accessed September 30, 2021. https://www.globalfamilydoctor.com/groups/workingparties/wicc.aspx
33. WHO-ATC DDD index. Accessed October 5, 2021. http://www.whocc.no/atc_ddd_index/?code=C01BD01
34. Rijksoverheid. Regeling zorgverzekering - Bijlage 2 [Government. Health Insurance Regulations - Appendix 2]. Published January 1, 2012. Accessed September 6, 2021. https://wetten.overheid.nl/BWBR0018715/2012-01-01/1#Bijlage2
35. Zorginstituut Nederland [Healthcare Institute Netherlands]. GIP Databank. Accessed October 11, 2021. https://gipdatabank.nl
36. Charlton A, Vidal X, Sabate M, Ballarin E, Martinez Leguizamo LM, Ibanez L. Factors associated with primary nonadherence to newly initiated direct oral anticoagulants in patients with nonvalvular atrial fibrillation. *J Manag Care Spec Pharm*. 2021;27(9):1210-1220. doi:10.18553/jmcp.2021.27.9.1210
37. Flinterman LE, Hek K, Korevaar JC, van Dijk L. Impact of a Restriction in Reimbursement on Proton Pump Inhibitors in Patients with an Increased Risk of Gastric Complications. *Front Public Heal*. 2018;6:51. doi:10.3389/fpubh.2018.00051

38. Meulepas M, Lambooij A. *Rapport Maagzuurremmergebruik Als Protectie Bij NSAID En ASA. [Report Gastric Acid Use As Protection With NSAID And ASA]*; 2013. Accessed September 23, 2021. https://www.medicijngebruik.nl
39. van Geffen ECG, van Hulten R, Bouvy ML, Egberts ACG, Heerdink ER. Characteristics and reasons associated with nonacceptance of selective serotonin-reuptake inhibitor treatment. *Ann Pharmacother*. 2008;42(2):218-225. doi:10.1345/aph.1K516
40. Evans CD, Eurich DT, Remillard AJ, Shevchuk YM, Blackburn D. First-fill medication discontinuations and nonadherence to antihypertensive therapy: an observational study. *Am J Hypertens*. 2012;25(2):195-203. doi:10.1038/ajh.2011.198
41. Lemstra M, Blackburn D. Nonadherence to statin therapy: discontinuation after a single fill. *Can J Cardiol*. 2012;28(5):567-573. doi:10.1016/J.CJCA.2012.03.018
42. Stichting Farmaceutische Kengetallen [Pharmaceutical Key Figures Foundation]. Ruim 80% medicijngebruikers bezoekt slechts één apotheek [More than 80% drug users visit only one pharmacy]— PW | Pharmaceutisch Weekblad [Pharmaceutical Weekly Magazine]. *Pharmaeutisch Weekbl*. Published online 2018. Accessed January 16, 2020. https://www.pw.nl/vaste-rubrieken/sfk/2018/ruim-80-medicijngebruikers-bezoekt-slechts-een-apotheek-1
43. Sokol M, McGuigan K, Verbrugge R, Epstein R. Impact of medication adherence on hospitalization risk and healthcare cost. *Med Care*. 2005;43(6):521-530. doi:10.1097/01.MLR.0000163641.86870.AF
44. Roebuck M, Liberman J, Gemmill-Toyama M, Brennan T. Medication adherence leads to lower health care use and costs despite increased drug spending. *Heal Aff*. 2011;30(1):91-99. doi:10.1377/HLTHAFF.2009.1087
45. Hempenius M, Groenwold RHH, de Boer A, Klungel OH, Gardarsdottir H. Drug exposure misclassification in pharmacoepidemiology: Sources and relative impact. *Pharmacoepidemiol Drug Saf*. Published online September 7, 2021. doi:10.1002/PDS.5346

SUPPLEMENTARY MATERIALS

TABLE S1. Excluded ATC-codes, which are not dispensed at the outpatient pharmacy in 2012

ATC-code	Description
J07BB02	Influenza vaccine
V03AN	Medicinal gasses
V04	Diagnostic agents
V06	General nutrients
V07	All other non-therapeutic products
V08	Contrast media
V09	Diagnostic radiopharmaceuticals
V10	Therapeutic radiopharmaceuticals
L01EC01	vemurafenib
L04AB01	Etanercept
L04AB02	Infliximab
L04AB04	Adalimumab
L04AB05	Certolizumab pegol
L04AB06	Golimumab
L04AC03	Anakinra
L04AC05	Ustekinumab
A16AX06	Miglustat
B01AC09	Epoprostenol
B01AC21	Treprostinil
B02BX04	Romiplostim
B03XA01	Epoetine
B03XA02	Darbepoetine alfa
B03XA03	Methoxypolyethyleenglycolepoetine beta
C02KX01	Bosentan
H01AC01	Somatropine
H01AX01	Pegvisomant
H05AA02	Teriparatide
H05AA03	Parathyroidhormoon
J01XB01	Colistine
J06BA02	Immunoglobuline normaal intravasculair
L01XE05	Sorafenib

TABLE S1. Continued.

ATC-code	Description
L02AE02	Leuproreline
L03AA13	Pegfilgrastim
L03AB07	Interferon beta 1a
L03AB08	Interferon beta 1b
L03AB10	Peginterferon alfa 2b
L03AX13	Glatirameer

TABLE S2. International Classification of Primary Care (ICPC) codes used for identification of comorbidities

Comorbidity	Code
cardiovascular diseases	K74, K75, K76, K77, K78, K79, K80, K83, K84, K89, K90, K91
diabetes mellitus	T89, T90
respiratory disorders	R95, R96
psychological disorder	P72, P73, P74, P76, P79, P98
malignancy	A79, B72, B73, B74, D74, D75, D76, D77, L71, N74, R84, R85, T71, U75, U76, U77, W72, X75, X76, X77, Y77, Y78

TABLE S3. Drugs that are only reimbursed for chronic use (more than six months in use) in 2012

ATC-code	Description
A02BA	H2-receptor antagonists
A02BC	Proton pump inhibitors
M01AE52	Naproxen/esomeprazole
A06AD11	Lactulose
A06AD15	Macrogol
A06AC01	Psyllium
A06AC03	Steruliagom
A06AB06	Sennosiden
A06AG11	sodium lauryl sulfoacetate, incl. combinations
A06AG10	docusate sodium, incl. combinations
R06AE07	Cetirizine
R06AX27	Desloratidine

TABLE S3. Continued.

ATC-code	Description
R06AX13	Loratidine
A07DA03	Loperamide
S01XA20	artificial tears and other indifferent preparations
A03FA03	Domperidon

Sources: Dutch government and SFK data[34]

TABLE S4. Characteristics of matched and unmatched subjects in the Nivel Primary Care Database (Nivel-PCD)

	Matched, n (%) N=86 361	Unmatched, n (%) N=309 890
Sex		
Male	36 337 (42.1)	148 132 (47.8)
Female	50 024 (57.9)	161 758 (52.2)
Age*		
0-4	1954 (2.3)	12 637 (4.1)
5-17	8276 (9.6)	46 314 (15.0)
18-44	23 803 (27.6)	106 689 (34.4)
45-64	30 641 (35.5)	92 224 (29.8)
65-74	13 263 (15.4)	29 742 (9.6)
75-84	6886 (8.0)	16 270 (5.3)
85 and older	1538 (1.8)	6014 (1.9)
Number of prescriptions**		
Q1 – 1 to 3	4153 (4.8)	89 769 (29.0)
Q2 – 4 to 11	21 540 (24.9)	84 627 (27.3)
Q3 – 12 to 33	29 713 (34.4)	67 257 (21.7)
Q4 – 34 or more	30 955 (35.8)	68 237 (22.0)

*Age assessed on 1-1-2012
**All prescriptions during 2011–2013, including repeat prescription and invalid ATC codes

3.2

AMIODARONE USE AND THE RISK OF ACUTE PANCREATITIS: INFLUENCE OF DIFFERENT EXPOSURE DEFINITIONS

Mirjam Hempenius, Rolf H. H. Groenwold, Anthonius de Boer, Olaf H. Klungel, Helga Gardarsdottir

Pharmacoepidemiol Drug Saf, 2019 Dec;28(12):1563-1571

ABSTRACT

Purpose

The antiarrhythmic drug amiodarone has a long half-life of 60 days, which is often ignored in observational studies. This study aimed to investigate the impact of different exposure definitions on the association between amiodarone use and the risk of acute pancreatitis.

Method

Using data from the Dutch PHARMO Database Network, incident amiodarone users were compared to incident users of a different type of antiarrhythmic drug. Eighteen different definitions were applied to define amiodarone exposure, including dichotomized, continuous and categorized cumulative definitions with lagged effects to account for the half-life of amiodarone. For each exposure definition, a Cox proportional hazards model was used to estimate the hazard ratio (HR) of hospitalization for acute pancreatitis.

Results

This study included 15,378 starters of amiodarone and 21,394 starters of other antiarrhythmic drugs. Adjusted HRs for acute pancreatitis ranged between 1.21–1.43 for dichotomized definitions of exposure to amiodarone, between 1.13-1.22 for dose definitions (per DDD) and between 0.52-1.72 for cumulative dose definitions, depending on the category. Accounting for lagged effects had little impact on estimated HRs.

Conclusions

This study demonstrates the relative insensitivity of the association between amiodarone and the risk of acute pancreatitis against a broad range of different exposure definitions. Accounting for possible lagged effects had little impact, possibly because treatment switching and discontinuation was uncommon in this population.

1. INTRODUCTION

Amiodarone is a class III antiarrhythmic drug used for rhythm control in patients with atrial fibrillation. In the Netherlands, it is preserved as a second-line treatment because of its various side effects.[1,2] Amiodarone is a highly fat-soluble drug and accumulates in the body after long-term use.[3] This results in a long half-life of about 60 days (range 9-107 d),[1,3-6] which increases with longer exposure to amiodarone.[7] As a consequence, both the positive and the adverse effects of amiodarone mainly occur after prolonged use, when the drug has accumulated in the body.[8,9] Adverse reactions may therefore also occur several weeks to months after discontinuation of the intake of amiodarone.

The long half-life of amiodarone may have consequences for observational studies of the effects of the drug. In such studies, information about exposure to amiodarone is mostly based on prescription or dispensing information. Assuming that patients take their pills as prescribed, the resulting exposure classification may inadequately reflect actual exposure status as the patient might be much longer physically exposed because of the long half-life. However, in observational studies of adverse effects of amiodarone, these pharmacokinetic characteristics are not always considered when defining exposure to amiodarone. In fact, exposure to amiodarone has been defined in different ways, eg, ever vs never use[10]; current/recent/past use vs never use[11]; recent vs nonrecent use[12]; cumulative dose[10,13]; and duration of use.[14] An exception to this is the study by Taylor et al who accounted for the relatively long half-life of amiodarone by extending the exposure period with 2 months after discontinuation of use.[15]

Various exposure definitions were also used in research into the association between amiodarone use and the risk of acute pancreatitis. Whether acute pancreatitis is a direct or cumulative effect is still unclear. Case reports on use of amiodarone and the occurrence of acute pancreatitis suggest either an immediate reaction (3-5 d following initiation)[16,17] or a cumulative effect (1-36 mo following initiation).[18,19] The association between amiodarone and acute pancreatitis has been investigated in two case-control studies, both using different definitions for amiodarone exposure. In the study by Lai et al, a comparison was made between current use (most recent prescription < 3 mo before the event) and never use, which resulted in an odds ratio (OR) of 5.21 (95% confidence interval [CI], 3.22-8.43).[11] Alsonso et al compared ever use of amiodarone before the event date with never use. This resulted in an OR of 1.53 (95% CI, 1.24-1.88).[10] These very different effect estimates are possibly due to the different methods of defining exposure to amiodarone. In addition, both studies did not take into account the pharmacokinetic properties of amiodarone.

The aim of this study is therefore to investigate the impact of different amiodarone exposure definitions on the association between amiodarone and the risk of acute pancreatitis, taking into account the pharmacokinetic properties.

2. METHODS

2.1. Data source

Data were obtained from the PHARMO Database Network in the Netherlands, which includes information about more than four million inhabitants of the Netherlands (approximately 25% of

the Dutch population) with an average follow-up of 10 years.[20] The PHARMO Database Network links data from different health care settings. For this study, the Out-patient Pharmacy Database and the Hospitalization Database were used. The Out-patient Pharmacy Database comprises drug dispensing history prescribed by either general practitioners (GPs) or specialists. The dispensing records include information about type of drug, dispensing date, dosage, quantity, and the dosage regimen. Drug dispensings are coded according to the Anatomical Therapeutic Chemical (ATC) Classification System.[21] The Hospitalization Database comprises information about hospital admissions from the Dutch Hospital Data Foundation. These records include information about discharge diagnoses and hospital admission and discharge dates. Diagnoses are coded according to the International Classification of Diseases (ICD) version 9.[22]

2.2. Study population

All subjects in the Out-patient Pharmacy Database with a first dispensing of the class III antiarrhythmic drug amiodarone (ATC code C01BD01) between 1 January 2005 and 31 December 2013 were included in the study. The comparison group consisted of all subjects with a first dispensing of a class I or III antiarrhythmic drug other than amiodarone during the same period (ATC code C01B, excl C01BD01). The date of the first dispensing was defined as index date for both groups. Inclusion criteria for both groups were an age of ≥18 years at the index date and the presence of at least 6 months of enrolment in the database prior to the index date to ensure the selection of incident users. All subjects with a known history of acute pancreatitis in the 6 months before the index date were excluded.

Each subject was followed up until acute pancreatitis was diagnosed, death, deregistration from the concerning pharmacy, or the end of the study period, whichever came first. Subjects were allowed to switch from amiodarone to another antiarrhythmic drug, to use both types of drugs at the same time, or to stop using any antiarrhythmic medication. This was taken into account in the analysis (see Section 2.6).

2.3. Outcome definition

Occurrence of acute pancreatitis was defined using the hospitalization data. ICD-9 code 577.0 was used for identification of the outcome.[23]

2.4. Exposure definitions

Days exposed was identified on the basis of the theoretical duration of each individual dispensing. The assessment was based on the dispensing date, quantity dispensed, strength, and written dosage instruction of each dispensing. In case of unknown dosage instructions (<1% of all amiodarone dispensings), the maintenance dose was set at 1 daily defined dose (DDD), 200 mg once daily.[21] Treatment with amiodarone usually starts with a loading scheme. Therefore, a standard loading scheme was applied to all dispensings in which a loading scheme was mentioned, on the basis of the most frequently applied scheme in the study population: 7 days 3 × 200 mg followed by 7 days 2 × 200 mg. Several exposure definitions were applied, which we describe below (see Figure 1 for illustration).

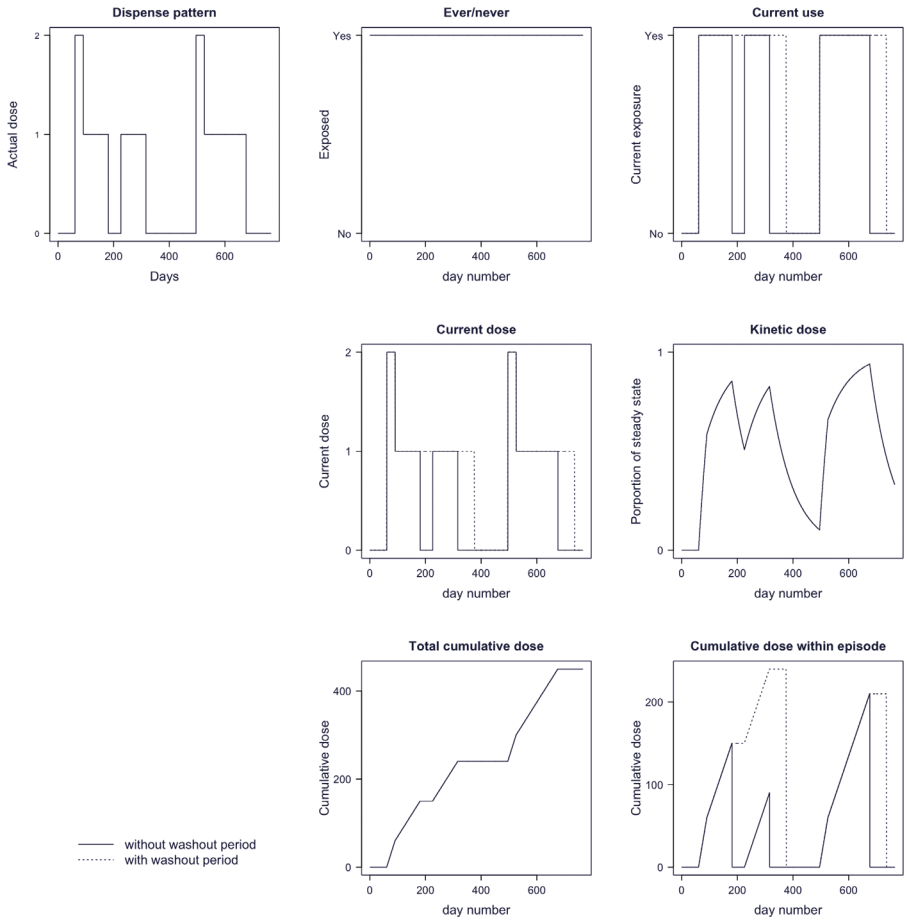


FIGURE 1. Illustration of different exposure definitions in a drug exposure study. Top left panel shows dispensing pattern. Other panels show the result of different exposure definitions. Washout periods are set at 60 d.

2.4.1. Dichotomous exposure definitions

Intention to treat

All subjects with an index dispensing of amiodarone were considered as exposed to amiodarone throughout the whole follow-up. All other subjects were defined as nonexposed. This definition can lead to a biased estimate of the relation between actual amiodarone use and risk of pancreatitis because it ignores the duration of amiodarone treatment; a treatment duration of 1 week could, for example, result in an "exposure episode" of 9 years. Nevertheless, it was used in previous research on the adverse events of amiodarone and therefore included in our analysis to allow for comparison.

Current use vs noncurrent use

All constructed episodes of exposure were considered as "current use." In the first definition of "current use," overlapping periods and gaps between two dispensings were ignored. For the second definition, overlapping periods between two dispensings were added to the end of the concerning exposure episode, with a maximum of 90 days. Gaps were still ignored. For the third definition, these "overlap-adjusted" episodes of current use were additionally prolonged with different washout periods of 30, 60, and 90 days. When these washout periods had overlap with a subsequent exposure episode, this overlap was not added to the end of the next episode.

2.4.2. Continuous exposure definitions

Current dose

"Current dose" was defined as the dose in DDD during the episodes of "current use" after correction for overlapping periods. Different washout periods were applied (30, 60, and 90 d). During these washout periods, the dose was considered to be equal to the latest dispensed dose.

Kinetic dose-model

For the "kinetic dose" model, the cumulative dose present in the body was estimated in DDD. Parameters needed for the estimation of this cumulative dose included the half-life (30, 60, or 90 d), the strength, and the dose regimen. A simplified calculation of the "kinetic dose" was obtained by summing the dispensed dose at each day and the remaining fraction of the "kinetic" dose of the previous day; the latter is calculated as $0.5^{1/\text{half-life}}$.

For example, when a subject receives a dose of 1 DDD on three consecutive days with an assumed half-life of 30 days, the kinetic dose on day 1 is 1 DDD. On the second day, the remaining fraction of day 1 is $1 * 0.5^{1/30} = 0.98$ DDD. Summed with the dose of day 2, the kinetic dose on day 2 is 1.98 DDD. On day 3, the remaining fraction of day 2 is $1.98 * 0.5^{1/30} = 1.93$ DDD, and the kinetic dose will amount 2.93 DDD. When on day 4 no new dose is taken, the kinetic dose on day 4 is $2.93 * 0.5^{1/30} = 2.87$ DDD. This kinetic dose will gradually drop, until the remaining amount can be neglected, or a new dose is taken by the subject. Approximately four times the half-life is needed to reach steady state. When steady state is reached, it takes also about four times the half-life to eliminate the drug from the body.

2.4.3. Categorized exposure definitions

Cumulative exposure

The cumulative exposure was expressed as number of DDDs. The resulting cumulative dose was then divided into three categories—1 to 90 DDD, 91 to 360 DDD, and >360 DDD—to enable a comparison between short-term, medium-term, and long-term users. The cumulative dose was a time-dependent variable. Two different definitions were applied: "overall cumulative exposure" and "cumulative exposure within episode." Overall cumulative exposure was defined as the amount of DDDs dispensed during the whole study period, starting from 0 and accumulating with each day of use. After each exposure episode, the cumulative dose did not change, until a new exposure episode started.

In the second definition, the cumulative dose was calculated for each episode separately, starting each episode from 0 and accumulating with each day of use. When there were gaps between two dispensings, the cumulative dose was set to 0 at the end of a treatment episode and started again from 0 when a new episode started. In addition to this definition, a washout period (30, 60, and 90 d) was added to each exposure period. In this washout period, the cumulative dose was held constant and started from that level when a new dispensing started within the washout period. When no new prescription was dispensed during the washout period, the cumulative dose was set to 0 after the washout period and started again from 0 when a new episode started.

2.5. Potential confounders

Potential confounders that were considered as covariates in the models were age, sex, comorbidities (diabetes mellitus, hypertriglyceridemia, and biliary stones), and (co)medication use (antiarrhythmic drugs, acetaminophen, opiates, simvastatin, atorvastatin, enalapril, estrogens, furosemide, hydrochlorothiazide, doxycycline, and steroids), because these have been reported as possible risk factors for developing acute pancreatitis.[24] ATC and ICD codes for both comedication and comorbidities are included in Table S1.

2.6. Data analysis

The characteristics of subjects included in the study were determined for each group separately. For all exposure definitions, a Cox proportional hazards model was used to estimate the relation between exposure to amiodarone and the risk of acute pancreatitis presented as hazard ratios (HRs) with 95% CIs. The reference category for all analyses was "no exposure to amiodarone." We corrected for current use of other antiarrhythmic drugs, measured per day. Other confounders related to comorbidities and comedication use were also included as time-varying covariates in all analyses, measured per day. Since none of the time-dependent confounders were considered to be affected by previous amiodarone use, we expected no bias by including the time-dependent confounders as time-dependent covariates in the Cox proportional hazards models. The covariates diabetes, simvastatin, enalapril, and estrogens were excluded from the final model, because of a limited number of events in the PHARMO database. These covariates were selected on their low prevalence and/or the strength of their relationship with the outcome. In addition, two sensitivity analyses were performed. The first sensitivity analysis excluded all subjects exposed for more than 95% of their follow-up time, since in these subjects, the different exposure definitions would not result in very different patterns of exposure. In the second sensitivity analysis, all amiodarone users with baseline use of another antiarrhythmic drug were excluded to minimize the risk of confounding by indication. The assumption of proportional hazard functions over time was checked graphically with a "log-log" plot. Data analysis was performed using the statistical software package R.[25]

3. RESULTS

On the basis of cohort entry medication, the study included 15 378 amiodarone starters and 21 394 starters of other antiarrhythmic drugs. The characteristics of the study populations are presented in Table 1. The mean age for the amiodarone starters was 70.7 (+/–11.0) years and for starters of other antiarrhythmic drugs 61.3 (+/–14.4) years. Of all subjects in the amiodarone group, 40.2% were women, whereas this percentage was 53.9% for other antiarrhythmic drug users.

TABLE 1. Baseline characteristics of amiodarone starters and starters of another antiarrhythmic drug

	Amiodarone starters	**Starters of other antiarrhythmic drugs**
No. of subjects	15,378	21,394
Age (mean, SD)	70.7 (+/- 11.0)	61.3 (+/- 14.4)
Comorbidities (n, %)		
Diabetes	2519 (16.4)	1676 (7.8)
Hypertriglyceridemia	104 (0.7)	68 (0.3)
Biliary stones	28 (0.2)	38 (0.2)
Comedication (n, %)		
Sotalol	3485 (22.6)	3862 (18.1)
Other antiarrhythmics†	1867 (12.1)	0 (0.0)
Simvastatin	2898 (18.8)	2536 (11.9)
Atorvastatin	2017 (13.1)	1305 (6.1)
Hydrochlorothiazide	1388 (9.0)	1775 (8.3)
Furosemide	3713 (24.1)	1032 (4.8)
Enalapril	1011 (6.6)	747 (3.5)
Acetaminophen	1298 (8.4)	1392 (6.5)
Opiates	1265 (8.2)	1553 (7.3)
Doxycycline	1257 (8.2)	1153 (5.4)
Oral steroids	1740 (11.3)	1805 (8.4)
Estrogens	209 (1.4)	686 (3.2)

† Includes class I and III antiarrhythmics and excludes amiodarone and sotalol

TABLE 2. Hazard ratios of acute pancreatitis for different amiodarone exposure definitions

Definition[a]	Person-years[b]		No. of events		Hazard ratio (95% confidence interval)	
	Exp	Unexp	Exp	Unexp	Crude	Fully adjusted
Dichotomous[c]						
Intention to treat	53.6	72.4	45	30	2.04 (1.28 – 3.23)	1.43 (0.82 – 2.05)
Current use not adj. for overlaps	21.9	104.1	22	53	1.98 (1.19 – 3.30)	1.36 (0.78 – 2.38)
Overlap-adjusted current use						
No washout period	23.7	102.4	22	53	1.79 (1.08 – 2.97)	1.21 (0.69 – 2.10)
Washout period of 30 days	25.4	100.7	24	51	1.87 (1.13 – 3.07)	1.26 (0.73 – 2.19)
Washout period of 60 days	26.5	99.4	25	50	1.88 (1.15 – 3.08)	1.27 (0.74 – 2.20)
Washout period of 90 days	27.5	98.5	26	49	1.91 (1.17 – 3.11)	1.30 (0.75 – 2.24)
Continuous						
Current dose (DDD)[d]						
No washout period	*NA*	*NA*	*NA*	*NA*	1.57 (1.05 – 2.36)	1.19 (0.74 – 1.91)
Washout period of 30 days	*NA*	*NA*	*NA*	*NA*	1.48 (1.02 – 2.15)	1.13 (0.71 – 1.80)
Washout period of 60 days	*NA*	*NA*	*NA*	*NA*	1.46 (1.02 – 2.09)	1.13 (0.72 – 1.78)
Washout period of 90 days	*NA*	*NA*	*NA*	*NA*	1.44 (1.03 – 2.01)	1.13 (0.73 – 1.77)
Kinetic dose (DDD)[e]						
Half-life of 30 days	*NA*	*NA*	*NA*	*NA*	1.71 (1.09 – 2.70)	1.22 (0.73 – 2.06)
Half-life of 60 days	*NA*	*NA*	*NA*	*NA*	1.74 (1.07 – 2.83)	1.21 (0.69 – 2.10)
Half-life of 90 days	*NA*	*NA*	*NA*	*NA*	1.74 (1.04 – 2.91)	1.17 (0.65 – 2.10)
Categorized[f]						
Cumulative dose of 1–90 DDD						
Reset after 0 days	7.7	102.4	9	53	2.57 (1.21 – 5.47)	1.72 (0.78 – 3.81)
Reset after 30 days	5.2	100.7	3	51	1.19 (0.32 – 4.39)	0.77 (0.20 – 3.01)
Reset after 60 days	4.8	99.5	4	50	1.96 (0.59 – 6.49)	1.30 (0.37 – 4.56)
Reset after 90 days	4.6	98.5	4	49	2.14 (0.62 – 7.32)	1.40 (0.38 – 5.13)
No reset	11.1	72.4	8	30	1.75 (0.78 – 3.93)	1.29 (0.54 – 3.09)
Cumulative dose of 91–360 DDD						
Reset after 0 days	8.5	102.4	10	53	2.26 (1.11 – 4.59)	1.50 (0.71 – 3.16)
Reset after 30 days	8.8	100.7	11	51	2.48 (1.20 – 5.12)	1.66 (0.77 – 3.59)
Reset after 60 days	9.0	99.5	9	50	1.90 (0.86 – 4.21)	1.26 (0.55 – 2.93)
Reset after 90 days	9.1	98.5	9	49	1.91 (0.85 – 4.27)	1.26 (0.54 – 2.97)

TABLE 2. Continued.

Definition[a]	Person-years[b]		No. of events		Hazard ratio (95% confidence interval)	
	Exp	Unexp	Exp	Unexp	Crude	Fully adjusted
No reset	19.4	72.4	18	30	2.16 (1.19 – 3.93)	1.54 (0.78 – 3.03)
Cumulative dose of >360 DDD						
Reset after 0 days	7.5	102.4	3	53	0.74 (0.23 – 2.39)	0.52 (0.16 – 1.72)
Reset after 30 days	11.4	100.7	10	51	1.70 (0.85 – 3.38)	1.17 (0.57 – 2.40)
Reset after 60 days	12.8	99.5	12	50	1.85 (0.97 – 3.52)	1.27 (0.65 – 2.50)
Reset after 90 days	13.8	98.5	13	49	1.87 (1.00 – 3.48)	1.29 (0.67 – 2.50)
No reset	23.1	72.4	19	30	2.07 (1.14 – 3.75)	1.40 (0.72 – 2.74)

Note. A Cox proportional hazards model was used to estimate the relation between exposure to amiodarone and the risk of acute pancreatitis. The reference category for all analyses was "no exposure to amiodarone."

[a] Exposure definitions: (1) Dichotomized definitions: (a) intention to treat: each subject with one or more dispensings of amiodarone was considered as exposed from the index date and throughout the whole study period; (b) current use: episodes of use/nonuse based on the start date of each dispensing and duration; (c) current use + overlap: the overlap between two dispensings was added to the end of the according exposure episode (max 90 d) whether or not prolonged with a washout period. (2) Continuous definitions: (a) current dose: dose during episodes of overlap-adjusted current use, whether or not prolonged with a washout period; (b) kinetic dose, based on the half-life and the dose regimen. (3) Categorized cumulative definitions: (a) cumulative exposure expressed in daily defined dose (DDD) and calculated for each episode (corrected for overlaps), a washout period was added with different lengths where cumulative exposure did not further increase, and the cumulative exposure was not reset to zero either or during the whole study period.
Adjusted for age, sex, diabetes mellitus, hypertriglyceridemia, biliary stones, antiarrhythmic drugs, acetaminophen, opiates, atorvastatin, furosemide, hydrochlorothiazide, doxycycline, and steroids.
[b] Expressed in 1000 person-years
[c] Hazard ratio (HR) expressed for being exposed vs nonexposed.
[d] HR expressed per 1 DDD.
[e] HR expressed for steady-state dose.
[f] HR expressed for this category vs nonuse or past-use (past-use is not applicable when no reset was applied).

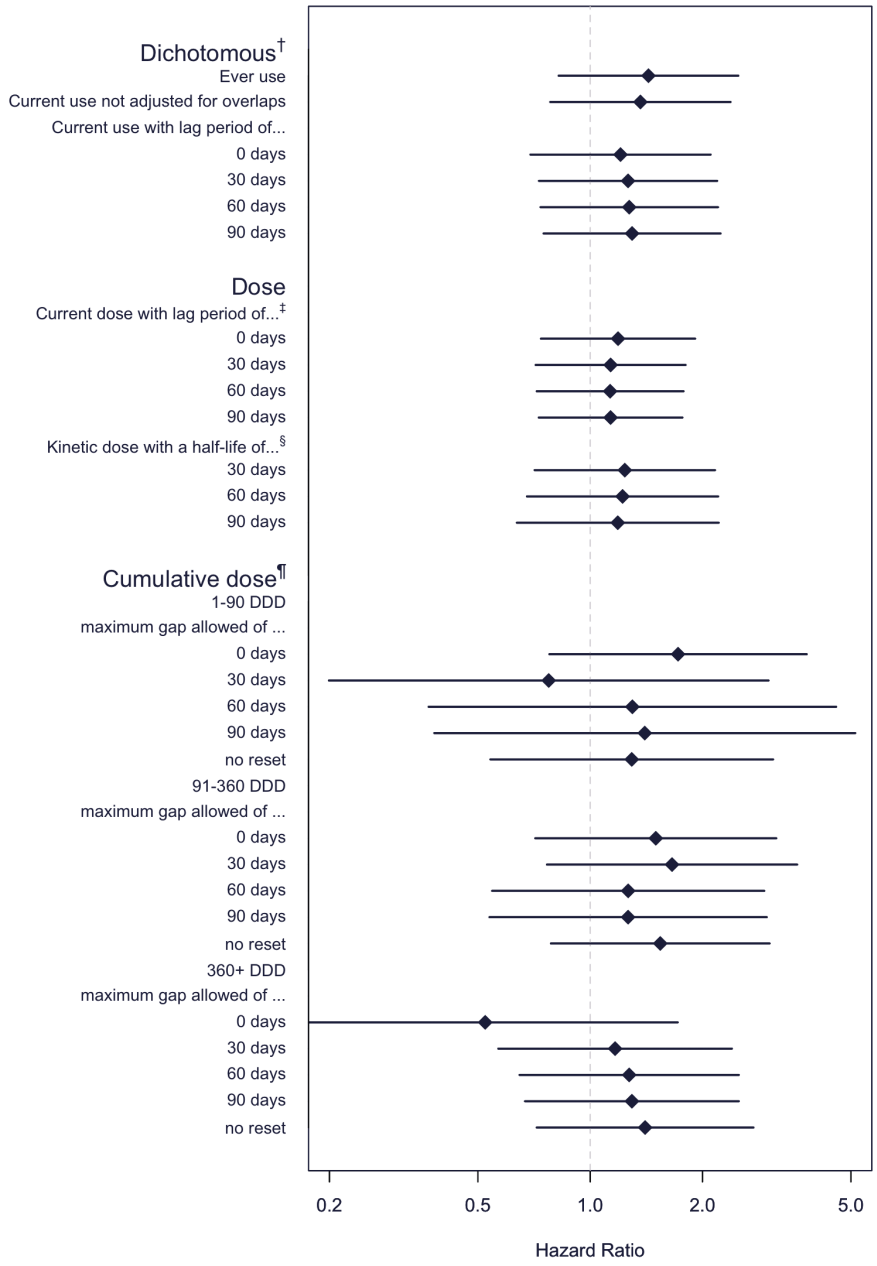


FIGURE 2. Hazard ratios (HRs) of exposure to amiodarone compared with exposure to another antiarrhythmic drug and the risk of acute pancreatitis.
†HR expressed for being exposed vs nonexposed. ‡HR expressed per 1 daily defined dose (DDD). §HR expressed for steady-state dose. ¶HR expressed for this category vs nonuse or past-use (past-use is not applicable when no reset was applied). HRs adjusted for age, sex, diabetes mellitus, hypertriglyceridemia, biliary stones, antiarrhythmic drugs, acetaminophen, opiates, atorvastatin, furosemide, hydrochlorothiazide, doxycycline, and steroids.

Median follow-up time for the amiodarone starters and the starters of other antiarrhythmic drugs was 3.1 and 3.0 years after initial cohort entry, respectively. Mean exposed time in the amiodarone group, measured as proportion of days covered (PDC), was 55.7%, with 28.8% of all subjects exposed for more than 95% of their follow-up time. There were 75 pancreatitis events during the study period, 45 among the amiodarone starters and 30 among the starters of other antiarrhythmic drugs. Median time-to-event was 857 days after starting with amiodarone and 686 days after starting with another antiarrhythmic drug.

The effects of amiodarone exposure on the risk of acute pancreatitis are presented in Table 2 and Figure 2 for different amiodarone exposure definitions. For the dichotomous definitions, the adjusted HR varied between 1.21 (95% CI, 0.69-2.10) for the overlap-adjusted definition of current use without washout period and 1.43 (95% CI, 0.82-2.05) for the intention-to-treat definition. The adjusted HR of continuous definitions (expressed per 1 DDD) varied between 1.13 (95% CI, 0.73-1.77) for the current dose with a washout period of 90 days and 1.22 (95% CI, 0.73-2.06) for the kinetic dose with an assumed half-life of 30 days (HR expressed per steady-state dose). Most of the variation in HR was seen when cumulative exposure was measured in different categories, with subjects switching differently between the categories for each different definition. Depending on the definition used, the adjusted HR varied between 0.52 (95% CI, 0.16-1.72) for the cumulative use of more than 360 DDD and 1.72 (95% CI, 0.78-3.81) for the use of 1 to 90 DDD, both for cumulative dose within an episode and when no washout or gap was allowed.

The results of the two sensitivity analyses in which all amiodarone users exposed for more than 95% of their follow-up time or with a baseline use of another antiarrhythmic drug were excluded did not show any relevant differences in the estimates (Tables S2 and S3).

4. DISCUSSION

In this study, we found no association between exposure to amiodarone and the risk of acute pancreatitis within a cohort of antiarrhythmic drug users. The different ways of defining exposure to amiodarone did not result in materially different effect estimates.

The results of our study differ from those of Alonso et al and Lai et al, which may be explained by the differences in patient characteristics between the studies and the relatively low number of events in our study. In contrast to the studies by Alonso et al and Lai et al, we were able to apply different definitions of exposure to amiodarone. The choice of the most appropriate exposure definition is however dependent on the etiological relation between the drug and outcome under investigation. Some adverse drug reactions require current exposure in order to occur, whereas others depend on cumulative exposure, or may occur years after the drug is discontinued, such as cancer. The effect of the long half-life of amiodarone could thus be different for different outcomes. In addition, amiodarone is a drug with many known drug interactions, mediated by CYP enzymes. It is therefore also in drug-drug interaction studies of importance to take the half-life of amiodarone into account.

Case reports on use of amiodarone and the occurrence of acute pancreatitis suggest either an idiosyncratic reaction (3-5 d following initiation)[16,17] or a cumulative effect (1-36 mo following

initiation).[18,19] In our data, we observed a median time to hospitalization for pancreatitis of 857 days since starting amiodarone with an interquartile range of 341 to 1361 days, suggesting that this effect is a cumulative effect, possibly related to the plasma concentrations of amiodarone as reflected by the kinetic dose model.

One potential limitation of our study was the low predictive value of the outcome, since ICD-9 code 577.0 does not differentiate between different causes of acute pancreatitis, such as drug-induced, alcoholic, biliary, and idiosyncratic pancreatitis. The effect estimates found in our study reflect thus the relation between amiodarone and "all-cause pancreatitis." In general, it is not advisable to exclude cases with a "known" cause.[26] The effect of amiodarone use on drug-induced pancreatitis is likely to be diluted if we assume that the amiodarone use is not associated with the other forms of acute pancreatitis.

A second limitation was the small number of events and consequently limited power to show differences between the effect estimates from the different exposure definitions if such differences exist. Furthermore, discontinuation of amiodarone treatment was uncommon in the population. More than a quarter of all subjects were exposed to amiodarone for more than 95% of the duration of the follow-up, leaving no room for lag periods or changes in the kinetic dose. Sensitivity analyses in which subjects exposed for more than 95% of their follow-up time were excluded did however not result in different estimates, yet the power was even lower in these analyses. Another limitation of our study concerns the kinetic dose definition. In this study, the half-life was assumed to be the same for all individuals, but this is rather a simplification of reality. The half-life varies between subjects, probably caused by a different distribution of fatty tissue,[4] of which no information is available in the used database. There was also no information available on the alcohol consumption, which is a known risk factor for acute pancreatitis, thus potentially resulting in unmeasured confounding.

To conclude, in this study, the association between amiodarone and the risk of acute pancreatitis was insensitive for a broad range of different exposure definitions. Accounting for lagged effects had little impact, possibly because treatment switching was uncommon in this population. To further assess the impact of different exposure definitions in research practice, we recommend replication of this study in larger databases, with more variation in amiodarone use.

Author's contribution

MH designed the study, conducted the data-analysis, wrote the first draft of the manuscript, and implemented the contribution of the co-authors and external reviewers up to final publication. During the whole process she implemented input and feedback from the other contributors to this study.

REFERENCES

1. KNMP. Amiodaron. In: *Informatorium Medicamentorum*. KNMP Medicijn Media; 2017.
2. Kirchhof P, Benussi S, Kotecha D, et al. 2016 ESC Guidelines for the Management of Atrial Fibrillation Developed in Collaboration With EACTS. *Eur Heart J*. 2016;37(38):2893–2962. doi:10.1016/j.rec.2016.11.033
3. Freedman MD, Somberg JC. Pharmacology and pharmacokinetics of amiodarone. *J Clin Pharmacol*. 1991;31(11):1061–1069. Accessed August 8, 2017. http://www.ncbi.nlm.nih.gov/pubmed/1753010
4. Kashima A, Funahashi M, Fukumoto K, et al. Pharmacokinetic characteristics of amiodarone in long-term oral therapy in Japanese population. *Biol Pharm Bull*. 2005;28(10):1934–1938. Accessed August 8, 2017. http://www.ncbi.nlm.nih.gov/pubmed/16204949
5. Pollak P, Bouillon T, Shafer SL. Population pharmacokinetics of long-term oral amiodarone therapy. *Clin Pharmacol Ther*. 2000;67(6):642–652. doi:10.1067/mcp.2000.107047
6. Stäubli M, Troendle A, Schmid B, et al. Pharmacokinetics of amiodarone, desethylamiodarone and other iodine-containing amiodarone metabolites. *Eur J Clin Pharmacol*. 1985;29(4):417–423.
7. Poirier JM, Escoubet B, Jaillon P, Coumel P, Richard MO, Cheymol G. Amiodarone pharmacokinetics in coronary patients: differences between acute and one-month chronic dosing. *Eur J Drug Metab Pharmacokinet*. 13(1):67–72. Accessed August 8, 2017. http://www.ncbi.nlm.nih.gov/pubmed/3396616
8. Shukla R, Jowett NI, Thompson DR, Pohl JE. Side effects with amiodarone therapy. *Postgrad Med J*. 1994;70(825):492–498. doi:10.1136/PGMJ.70.825.492
9. Wilson JS, Podrid PJ. Side effects from amiodarone. *Am Heart J*. 1991;121(1):158–171. doi:10.1016/0002-8703(91)90969-O
10. Alonso A, MacLehose RF, Lutsey PL, Konety S, Chen LY. Association of amiodarone use with acute pancreatitis in patients with atrial fibrillation: a nested case-control study. *JAMA Intern Med*. 2015;175(3):449–450. doi:10.1001/jamainternmed.2014.6927
11. Lai S-W, Lin C-Y, Liao K-F. Amiodarone use and risk of acute pancreatitis: A population-based case-control study. *Hear Rhythm*. 2015;12(1):163–166. doi:10.1016/j.hrthm.2014.08.037
12. Douros A, Bronder E, Andersohn F, et al. Drug-induced liver injury: results from the hospital-based Berlin Case-Control Surveillance Study. *Br J Clin Pharmacol*. 2015;79(6):988–999. doi:10.1111/bcp.12565
13. Bouvy ML, Heerdink ER, Hoes AW, Leufkens HG. Amiodarone-induced thyroid dysfunction associated with cumulative dose. *Pharmacoepidemiol Drug Saf*. 2002;11(7):601–606. doi:10.1002/pds.735
14. Siccama R, Balk AH, de Herder WW, van Domburg R, Vantrimpont P, van Gelder T. Amiodarone therapy before heart transplantation as a predictor of thyroid dysfunction after transplantation. *J Hear Lung Transpl*. 2003;22(8):857–861.
15. Taylor CJ, Hodgkinson J, Hobbs FD. Rhythm control agents and adverse events in patients with atrial fibrillation. *Int J Clin Pr*. 2010;64(8):1069–1075. doi:10.1111/j.1742-1241.2010.02426.x
16. Bosch X, Bernadich O. Acute pancreatitis during treatment with amiodarone. *Lancet*. 1997;350:1300.
17. Famularo G, Minisola G, Nicotra GC, De Simone C. Acute pancreatitis caused by amiodarone. *Eur J Emerg Med*. 2004;11(5):305–306. Accessed October 12, 2017. http://www.ncbi.nlm.nih.gov/pubmed/15359209
18. Chen Y-Y, Chen C-Y, Leung K-K. Acute pancreatitis and amiodarone: a case report. *World J Gastroenterol*. 2007;13(6):975–977. doi:10.3748/WJG.V13.I6.975
19. Sastri S V, Diaz-Arias AA, Marshall JB. Can pancreatitis be associated with amiodarone hepatotoxicity? *J Clin Gastroenterol*. 1990;12(1):70–73. Accessed October 12, 2017. http://www.ncbi.nlm.nih.gov/pubmed/1689330
20. Herings RM, Pedersen L. Pharmacy-based medical record linkage systems. In: Strom B, Kimmel S, Hennessy S, eds. *Pharmacoepidemiology*. 5th ed. Wiley; 2012:270–86. Accessed November 22, 2018. www.wiley.com/

21. WHO-ATC DDD index. Accessed November 29, 2016. http://www.whocc.no/atc_ddd_index/?code=C01BD01
22. Conversion ICD-10 to ICD-9. Accessed November 29, 2016. http://www.icd10codesearch.com/
23. Moores K, Gilchrist B, Carnahan R, Abrams T. A systematic review of validated methods for identifying pancreatitis using administrative data. *Pharmacoepidemiol Drug Saf*. 2012;21:194-202. doi:10.1002/pds.2334
24. Lankisch PG, Apte M, Banks PA. Acute pancreatitis. *Lancet*. 2015;386(9988):85-96. doi:10.1016/s0140-6736(14)60649-8
25. R Core Team. R: A Language and Environment for Statistical Computing. Published online 2015. https://www.r-project.org/
26. Rothman KJ, Ray W. Should cases with a ?known? cause of their disease be excluded from study? (Commentary). *Pharmacoepidemiol Drug Saf*. 2002;11(1):11-14. doi:10.1002/pds.683

SUPPLEMENTARY MATERIALS

TABLE S1. ATC and ICD codes of comorbidities and comedication

Drug/comorbidity	Type	ICD-9	ICD-10	ATC
Biliary stones	comorbidity	574	K80	
Diabetes	comorbidity	250	E08 E09 E10 E11 E13	A10A A10B A10X
Sotalol	comedication			C07AA07 C07FX02 C07BA07
Hypertriglyceridemia	comorbidity	272.1, 272.2	E78.1 E78.2	C10AB C10BA03 C10BA04
Acetaminophen	comedication			N02AJ01 N02AJ06 N02AJ13 N02AJ17 N02BE01 N02BE51 N02BE71
Antiarrhythmic drugs	comedication			C01B
Atorvastatin	comedication			C10AA05 C10BA05 C10BX03 C10BX06 C10BX08 C10BX11 C10BX12
Doxycycline	comedication			J01AA02
Enalapril	comedication			C09AA02 C09BA02 C09BB02 C09BB06
Estrogens	comedication			G03AA G03AB G03C G03F
Furosemide	comedication			C03CA01 C03CB01 C03EB01
Hydrochlorothiazide	comedication			C03AA03 C03AB03 C03AX01 C03EA01 C09DX01 C09DX03 C09XA52 C09XA54
Opiates	comedication			N02A
Simvastatin	comedication			C10AA01 C10BA02 C10BA04 C10BX01 C10BX04
Steroids (oral)	comedication			H02A H02B

TABLE S2. Sensitivity analysis: hazard Ratios of acute pancreatitis for different amiodarone exposure definitions.
Incident amiodarone users with exposed for less than 95% of their follow-up were compared to incident users of other antiarrhythmic drugs.

Definition [a]	Person-years [b]		No. of events		Hazard ratio (95% confidence interval)	
	exp.	uexp.	exp.	unexp.	crude	fully adjusted
Dichotomous [c]						
Ever use	44.6	72.4	34	30	1.85 (1.13 – 3.02)	1.36 (0.77 – 2.42)
Current use not adj. for overlaps	13.5	103.5	11	53	1.66 (0.85 – 3.22)	1.17 (0.58 – 2.36)
Overlap-adjusted current use						
No washout period	14.8	102.3	11	53	1.49 (0.77 – 2.90)	1.05 (0.52 – 2.10)
Washout period of 30 days	16.4	100.6	13	51	1.64 (0.88 – 3.06)	1.15 (0.59 – 2.24)
Washout period of 60 days	17.5	99.5	14	50	1.67 (0.91 – 3.07)	1.18 (0.61 – 2.26)
Washout period of 90 days	18.5	98.5	15	49	1.72 (0.95 – 3.12)	1.22 (0.64 – 2.31)
Continuous						
Current dose (DDD) [d]						
No washout period	*NA*	*NA*	*NA*	*NA*	1.41 (0.82 – 2.40)	1.08 (0.60 – 1.98)
Washout period of 30 days	*NA*	*NA*	*NA*	*NA*	1.34 (0.82 – 2.21)	1.05 (0.59 – 1.85)
Washout period of 60 days	*NA*	*NA*	*NA*	*NA*	1.35 (0.84 – 2.15)	1.06 (0.61 – 1.83)
Washout period of 90 days	*NA*	*NA*	*NA*	*NA*	1.35 (0.87 – 2.11)	1.08 (0.63 – 1.83)
Kinetic dose (DDD) [e]						
Half-life of 30 days	*NA*	*NA*	*NA*	*NA*	1.59 (0.84 – 3.00)	1.12 (0.56 – 2.27)
Half-life of 60 days	*NA*	*NA*	*NA*	*NA*	1.63 (0.83 – 3.20)	1.12 (0.53 – 2.35)
Half-life of 90 days	*NA*	*NA*	*NA*	*NA*	1.63 (0.80 – 3.32)	1.09 (0.50 – 2.38)
Categorized [f]						
Cumulative dose of 1–90 DDD						
Reset after 0 days	5.9	102.3	5	53	1.90 (0.73 – 4.93)	1.30 (0.48 – 3.49)
Reset after 30 days	4.2	100.6	1	51	0.50 (0.06 – 3.94)	0.34 (0.04 – 2.75)
Reset after 60 days	3.9	99.5	2	50	1.22 (0.26 – 5.70)	0.85 (0.17 – 4.15)
Reset after 90 days	3.7	98.5	2	49	1.33 (0.28 – 6.43)	0.91 (0.18 – 4.62)
No reset	10.2	72.4	6	30	1.47 (0.60 – 3.59)	1.13 (0.44 – 2.91)
Cumulative dose of 91–360 DDD						
Reset after 0 days	5.5	102.3	5	53	1.83 (0.71 – 4.68)	1.24 (0.47 – 3.26)
Reset after 30 days	6.3	100.6	7	51	2.39 (1.02 – 5.57)	1.70 (0.70 – 4.12)
Reset after 60 days	6.6	99.5	5	50	1.59 (0.59 – 4.25)	1.12 (0.40 – 3.09)

TABLE S2. Continued.

Definition [a]	Person-years [b]		No. of events		Hazard ratio (95% confidence interval)	
	exp.	uexp.	exp.	unexp.	crude	fully adjusted
Reset after 90 days	6.8	98.5	5	49	1.59 (0.59 – 4.31)	1.12 (0.40 – 3.16)
No reset	17.0	72.4	14	30	1.99 (1.05 – 3.78)	1.49 (0.73 – 3.03)
Cumulative dose of >360 DDD						
Reset after 0 days	3.4	102.3	1	53	0.53 (0.07 – 3.86)	0.40 (0.06 – 2.97)
Reset after 30 days	5.9	100.6	5	51	1.60 (0.63 – 4.05)	1.12 (0.43 – 2.89)
Reset after 60 days	7.1	99.5	7	50	1.89 (0.85 – 4.21)	1.33 (0.58 – 3.04)
Reset after 90 days	8.0	98.5	8	49	1.91 (0.89 – 4.07)	1.35 (0.62 – 2.97)
No reset	17.4	72.4	14	30	1.93 (1.00 – 3.70)	1.37 (0.67 – 2.81)

Note. A Cox proportional hazards model was used to estimate the relation between exposure to amiodarone and the risk of acute pancreatitis. The reference category for all analyses was "no exposure to amiodarone."

[a] Exposure definitions: (1) Dichotomized definitions: (a) intention to treat: each subject with one or more dispensings of amiodarone was considered as exposed from the index date and throughout the whole study period; (b) current use: episodes of use/nonuse based on the start date of each dispensing and duration; (c) current use + overlap: the overlap between two dispensings was added to the end of the according exposure episode (max 90 d) whether or not prolonged with a washout period. (2) Continuous definitions: (a) current dose: dose during episodes of overlap-adjusted current use, whether or not prolonged with a washout period; (b) kinetic dose, based on the half-life and the dose regimen. (3) Categorized cumulative definitions: (a) cumulative exposure expressed in daily defined dose (DDD) and calculated for each episode (corrected for overlaps), a washout period was added with different lengths where cumulative exposure did not further increase, and the cumulative exposure was not reset to zero either or during the whole study period.
Adjusted for age, sex, diabetes mellitus, hypertriglyceridemia, biliary stones, antiarrhythmic drugs, acetaminophen, opiates, atorvastatin, furosemide, hydrochlorothiazide, doxycycline, and steroids.
[b] Expressed in 1000 person-years
[c] Hazard ratio (HR) expressed for being exposed vs nonexposed.
[d] HR expressed per 1 DDD.
[e] HR expressed for steady-state dose.
[f] HR expressed for this category vs nonuse or past-use (past-use is not applicable when no reset was applied).

TABLE S3. Sensitivity analysis: Hazard Ratios of acute pancreatitis for different amiodarone exposure definitions.
Incident amiodarone users without baseline use of another antiarrhythmic drug were compared to incident users of other antiarrhythmic drugs

Definition [a]	**Person-years** [b]		**No. of events**		**Hazard ratio (95% confidence interval)**	
	exp.	unexp.	exp.	unexp.	crude	fully adjusted [b]
Dichotomous [c]						
Ever use	53.6	72.4	45	30	2.20 (1.38 – 3.52)	1.55 (0.87 – 2.66)
Current use not adj. for overlaps	21.9	104.1	22	53	2.14 (1.28 – 3.60)	1.44 (0.81 – 2.56
Overlap-adjusted current use						
No washout period	23.7	102.4	22	53	1.94 (1.16 – 3.26)	1.28 (0.72 – 2.27)
Washout period of 30 days	25.4	100.7	24	51	2.04 (1.12 – 3.39)	1.36 (0.77 – 2.39)
Washout period of 60 days	26.5	99.4	25	50	2.06 (1.25 – 3.42)	1.38 (0.78 – 2.42)
Washout period of 90 days	27.5	98.5	26	49	2.10 (1.28 – 3.46)	1.41 (0.80 – 2.47)

Note. A Cox proportional hazards model was used to estimate the relation between exposure to amiodarone and the risk of acute pancreatitis. The reference category for all analyses was "no exposure to amiodarone."

[a] Exposure definitions: (1) Dichotomized definitions: (a) intention to treat: each subject with one or more dispensings of amiodarone was considered as exposed from the index date and throughout the whole study period; (b) current use: episodes of use/nonuse based on the start date of each dispensing and duration; (c) current use + overlap: the overlap between two dispensings was added to the end of the according exposure episode (max 90 d) whether or not prolonged with a washout period. (2) Continuous definitions: (a) current dose: dose during episodes of overlap-adjusted current use, whether or not prolonged with a washout period; (b) kinetic dose, based on the half-life and the dose regimen. (3) Categorized cumulative definitions: (a) cumulative exposure expressed in daily defined dose (DDD) and calculated for each episode (corrected for overlaps), a washout period was added with different lengths where cumulative exposure did not further increase, and the cumulative exposure was not reset to zero either or during the whole study period.
Adjusted for age, sex, diabetes mellitus, hypertriglyceridemia, biliary stones, antiarrhythmic drugs, acetaminophen, opiates, atorvastatin, furosemide, hydrochlorothiazide, doxycycline, and steroids.
[b] Expressed in 1000 person-years
[c] Hazard ratio (HR) expressed for being exposed vs nonexposed.

3.3

IMPACT OF ANTICOAGULANT EXPOSURE MISCLASSIFICATION ON THE BLEEDING RISK OF DIRECT ORAL ANTICOAGULANTS

Mirjam Hempenius, Rolf H. H. Groenwold, Patrick C. Souverein, Anthonius de Boer, Olaf H. Klungel, Helga Gardarsdottir

Br J Clin Pharmacol 2021 Sep;87(9):3508-3517

ABSTRACT

Aims

Drug exposure status based on routinely collected data might be misclassified when the database contains only prescriptions from 1 type of prescriber (e.g. general practitioner and not specialist). This study aims to quantify the impact of such exposure misclassification on the risk of major bleeding and stroke/transient ischaemic attack (TIA) associated with direct oral anticoagulants (DOACs) *vs.* vitamin K antagonists (VKAs).

Methods

Incident anticoagulant users (>12 mo free of anticoagulation use) in the Dutch PHARMO Database Network between 2008 and 2017 were included. Drug exposure was assessed using pharmacy dispensing information. The risks of hospital admission of major bleeding for DOAC *vs.* VKA users was assessed with Cox regression analysis, where exposure was based on all dispensings, on general practitioner (GP)-prescribed dispensings only or on specialist-prescribed dispensings only. Hazard ratios (HRs) were estimated also for hospitalization for gastrointestinal bleeding, intracranial bleeding and stroke/TIA.

Results

We included 99182 VKA-initiators and 21795 DOAC-initiators. Use of DOAC was associated with a lower risk of major bleeding compared to VKA use; HR 0.79 (95% confidence interval 0.70–0.90), 0.78 (0.68–0.91) and 0.62 (0.50–0.76), for exposure based on complete dispensing information, only GP- and only specialist-prescribed dispensings, respectively. Similar results were found for the other bleeding outcomes. For stroke/TIA the HRs were 0.96 (0.84–1.09), 1.00 (0.84–1.18) and 0.72 (0.58–0.90), respectively.

Conclusion

Including only GP-prescribed anticoagulant dispensings in this case did not materially impact the effect estimates compared to including all anticoagulant dispensings. Including only specialist-prescribed dispensings, however, strengthened the effect estimates.

1. INTRODUCTION

Observational studies on safety and effectiveness of pharmacological agents are often performed using routinely collected data from administrative or health-care databases. Different types of databases are available, such as insurance databases, outpatient pharmacy databases, general practitioner (GP) databases or hospital databases. These data sources differ in the information they contain.[1] For example, pharmacy or insurance databases hold information of all prescriptions collected at the pharmacy, whereas GP-databases contain only information about GP prescriptions and hospital databases usually contain only information about specialist prescriptions. The use of a single prescriber prescription database may lead to a misclassification of drug exposure status when a subject is treated for the same condition by 2 different types of prescribers, or is being treated by a prescriber whose prescribing information is not included in the database that is being used.[2] Moreover, this misclassification can be differential, for example when different prescribers (e.g. specialists and GPs) are treating different types of patients, who have different distributions of (unmeasured) risk factors for developing the outcome, such as frailty[3,4], which may lead to selection and information bias. Although these databases are being used widely in pharmacoepidemiology, the extent and impact of such misclassification in research practice is largely unknown.

To provide insight in the impact of exposure misclassification due to differences in data sources used for pharmaco-epidemiological studies, we used direct oral anticoagulants (DOACs) and the risk of a major bleeding and stroke/transient ischaemic attack (TIA) as an example. In the first years after licensing, DOACs were prescribed predominantly by specialists, such as cardiologists, internists or orthopaedics.[5,6] Furthermore, patient characteristics, including risk of stroke and major bleeding, differ between patients who receive their DOAC prescriptions from a GP and those who receive their prescription from a specialist.[7] Currently, a lot of attention is paid to real-world evidence about the effectiveness and safety of DOACs, as reflected in the number of recent publications and planned studies.[8,9] Different types of databases are used in these studies, including GP databases, hospital databases, health-care insurance databases and pharmacy databases, with different data capture on drug use.[10–39]

The primary aim of this study was to quantify the extent to which the estimated effects of DOACs *vs.* vitamin K antagonists (VKAs) on the risk of major bleeding and stroke are affected by misclassification caused by the use of a database that contains only prescriptions of 1 type of prescriber. Secondary aims were to describe the characteristics of DOAC users treated by the GP only, by the specialist only or by both, and to describe the prescribing patterns over time.

2. METHODS

To investigate the impact of the absence of prescriptions, we used the PHARMO Database Network, containing - among other things - drug dispensing information from community pharmacies in the Netherlands, including information on the type of prescribing physician for most dispensed prescriptions. This enabled us to carry out separate analyses in which we included all anticoagulant dispensings or only a subset of anticoagulant dispensings that were prescribed by either a GP or a specialist.

The study protocol is based on the protocol of an European Medicines Agency-sponsored study, which aimed at characterising the risk of major bleeding in patients with nonvalvular atrial fibrillation (NVAF).[31]

2.1. PHARMO database network

The PHARMO Database Network contains drug dispensing information from a representative sample of Dutch community pharmacies (Outpatient Pharmacy Database) that is linked with the national registry of hospital discharge diagnoses (Hospital Database) and electronic patient records registered by GPs (GP Database). More than 4 million inhabitants of the Netherlands (approximately 25% of the Dutch population) with an average follow-up of 10 years are included in the PHARMO Database Network.

The Outpatient Pharmacy Database comprises information about basic demographic information and about dispensed drugs, including the type of prescriber (i.e. GP, specialist, or other types of prescribers, such as dentists), type of drug, dispensing date, dose, quantity and the dosage regimen. Drug type is coded according to the Anatomical Therapeutic Chemical (ATC) Classification System.[40] The Outpatient Pharmacy Database was used to determine exposure to the anticoagulant drugs and comedication (see section on potential confounders).

The Hospital Database contains information about hospital admissions from the Dutch Hospital Data Foundation. The records include information about discharge diagnoses and hospital admission and discharge dates. Diagnoses are coded according to the International Classification of Diseases (ICD) version 9 (January 2008–September 2015) or version 10 (October 2015–December 2017).[41,42] Procedures are coded according to the Dutch CBV (operations file), the CVV (Classification of Operations) system, or the Dutch ZA (care activities) procedural codes. The Hospital Database was used to determine the outcomes, the potential indication of the anticoagulant drug use, and other comorbidities (see section on potential confounders).

The GP Database comprises information from electronic patient records registered by GPs. The records include, among other things, information on diagnoses and symptoms, coded according to the International Classification of Primary Care.[43] Information from the GP Database was available for approximately 25% of the study population and this information was used to complement the potential indication of the anticoagulant drug use.

2.2. Exposure–outcome example

2.2.1. Study population

A cohort was constructed consisting of all incident anticoagulant users between January 2008 and December 2017. Incident users were defined as patients initiating a DOAC (dabigatran etelixate, rivaroxaban, apixaban, and edoxaban) or VKA (acenocoumarol and phenprocoumon) during the study period without any use of either of the 2 drugs for at least 365 days prior to the index date. The index date was defined as the first dispensing date of an anticoagulant drug. Inclusion criteria were an age at index date of 18 years or older and at least 12 months of enrolment in the database prior to the index date. All subjects with a registered knee or hip replacement, a diagnosis of valvular atrial fibrillation, deep venous thrombosis or pulmonary embolism in the 90 days before or after the index date, and without a diagnosis of NVAF in the 90 days before or after the index date, were excluded from the study population. Each subject was followed until the outcome of interest was diagnosed, death, deregistration from the concerning pharmacy or the end of the study period, whichever came first. Subjects were allowed to switch from DOACs to VKAs or vice versa, or to stop using anticoagulant medication.

2.2.2. Outcome definition

The primary outcome of interest was hospitalization for major bleeding (haemorrhagic stroke/intracranial bleeding, gastrointestinal bleeding or other extracranial or unclassified bleeding and traumatic intracranial bleeding). Secondary outcomes included hospitalization for gastrointestinal bleeding, intracranial bleeding and stroke (haemorrhagic as well as infarction) and TIA. Only the primary hospitalization diagnoses were used for the outcome assessment. ICD codes for the outcomes are given in Table S2.

2.2.3. Exposure definition

The theoretical duration of each DOAC dispensing (ATC codes B01AE and B01AF; Table S1) and VKA dispensing (ATC code B01AA; Table S1) was based on the dispensing date, quantity, strength and the dosage regimen. In case of missing information about dose regimen, which is often the case with VKAs, the theoretical duration of each dispensing was for each individual defined by the median time between the dispensings. When only 1–3 dispensings were available for an individual patient or when the estimated duration exceeded 100 days, the duration was based on the most frequently occurring estimated dispensing duration for the specific drug in the study. For the construction of the treatment episodes, a maximum gap of 30 days was allowed between the theoretical end of a dispensing and the start of a next dispensing. Overlapping episodes were added to the end of the treatment episode with a maximum of 90 days. If the subsequent dispensing was another type of anticoagulant drug, the patient was considered to have switched therapy and the remaining tablet days from the prior dispensing were disregarded.

2.2.4. Potential confounders

The assessment of and adjustment for potential confounders were conducted in line with the European Medicines Agency-sponsored study.[31] As potential confounders of the relation

between DOAC/VKA and the different outcomes, we considered the risk factors for the various outcomes. Important risk factors considered for major bleeding are: thrombocytopenia; hypertension or use of antihypertensive drugs (Table S5); history of stroke/TIA; history of major bleeding event; presence of malignancy; concomitant use of medicines that increase bleeding risk (nonsteroidal anti-inflammatory drugs, corticosteroids, selective serotonin inhibitors and antiplatelet drugs; Table S6); concomitant use of medications that have pharmacokinetic interactions with DOACs (assessed per DOAC separately; see supplementary materials Table S7); history of pulmonary embolism or deep venous thrombosis; peptic ulcer diseases; kidney disease; and hepatic impairment (for ICD codes; see Tables S3 and S4). Important risk factors considered for stroke/TIA were concomitant use of medications that have pharmacokinetic interactions with DOACs (assessed per DOAC separately), prior stroke/TIA, pulmonary embolism/deep venous thrombosis, hypertension, diabetes mellitus, congestive heart failure, other (cardio)vascular disease (angina, myocardial infarction, coronary heart disease, aortic plaque and peripheral arterial disease), kidney disease and hepatic impairment. The use of comedication was assessed using the outpatient pharmacy database. The presence or history of comorbidities was assessed using the hospitalization database, or the outpatient pharmacy database in case of medication use as proxy.

Age, comorbidities (various time intervals prior to index date, Tables S3 and S4), and comedication use (6 months before the index date) were considered as time dependent confounders and their status was updated whenever the exposure status changes, or every 6 months, whichever comes first.

2.3. Data analysis

Cox proportional hazards regression analysis was applied to estimate the effect of current DOAC treatment compared to current VKA use on the risk of major bleeding and stroke/TIA, with and without adjusting for the abovementioned confounders. We assumed no misclassification of the information available in the data systems and we used all dispensed drugs to determine concomitant medication use or the presence of comorbidities, regardless of the prescribing physician.

The abovementioned analysis was repeated 3 times: (i) using all anticoagulant dispensing information; (ii) using only the information about the anticoagulant dispensings prescribed by GPs; and (iii) using only the information about the anticoagulant dispensings prescribed by specialists. The different effect estimates were compared. In these 3 analyses, the size of the study population, the index date per subject and the time on treatment per subject differed, depending on the anticoagulant dispensings that were included in the exposure assessment. These analyses were repeated for the different outcomes.

Differences in patient characteristics (age, sex and the presence of risk factors for bleeding or stroke) and drug dispensing patterns were also assessed. Patient characteristics were summarized as means and standard deviations or proportions where appropriate and presented stratified by prescriber. Patient characteristics were compared between VKA and DOAC users on index date and differences between these groups were quantified by means of standardized differences.[44] For all treatment episodes, the initiating prescriber was determined and whether

subjects received their prescriptions from multiple types of prescribers. In addition, exposure time caused by specialist-prescribed dispensings and GP-prescribed dispensings was determined for VKA and DOAC use.

Several sensitivity analyses were performed. First, we restricted our analysis to subjects with a registered NVAF diagnosis in the 90 days before or after the index date. Second, we stratified on several characteristics: age (<65, 65–85, >85 y), sex, index date (before or after 1 January 2013, which is halfway the study period) and all DOACs individually.

3. RESULTS

Based on cohort entry medication, the study included 99 182 VKA initiators and 21 795 DOAC initiators, when all anticoagulant dispensings were used. When only GP-prescribed dispensings were included, the study included 87 106 VKA initiators and 14 542 DOAC initiators. Including only specialist-prescribed dispensings resulted in 62 566 VKA initiators and 18 809 DOAC initiators. The characteristics of the study populations are presented in Table 1. The mean age for all VKA initiators was 70.1 (± 13.7) years and for all DOAC initiators 69.8 (± 11.8) years. Of both the VKA and DOAC initiators, 52.6% were men. The characteristics of the DOAC and VKA initiators treated only by a GP, only by a specialist, or treated by both a GP and a specialist are also presented in Table 1. Mean standardized differences of these subpopulation are visually depicted in Figure 1. In general, VKA users who receive their VKA prescriptions only by the specialist have more comorbidities and use more comedication than patients who only receive their VKA prescriptions from the GP. For prescribing of DOACs, the opposite was seen: patients who only receive DOACs from the specialist generally had fewer comorbidities and received fewer comedications.

TABLE 1. Baseline characteristics of subjects with index use of VKA or DOAC, stratified per prescribing physician (general practitioner, specialist or both)

	VKA				DOAC			
	All subjects	**Subjects treated by GP only**	**Subjects treated by specialist only**	**Subjects treated by GP & specialist**	**All subjects**	**Subjects treated by GP only**	**Subjects treated by specialist only**	**Subjects treated by GP & specialist**
Number of subjects	99182	35 500	11 046	51 632	21 795	4314	7000	9932
Age (y), mean (standard deviation)	70.1 (13.7)	70.97 (14.1)	69 (14.4)	69.7 (13.2)	69.78 (11.8)	71.98 (12.4)	67.18 (12.1)	70.79 (10.9)
Sex, *n* males (%)	52 168 (52.6)	17 702 (49.9)	6185 (56)	27 794 (53.8)	11 469 (52.6)	2165 (50.2)	3652 (52.2)	5383 (54.2)
Diagnosis of NVAF ± 3 months from index date	32 234 (32.5)	9015 (25.4)	3682 (33.3)	19 301 (37.4)	8415 (38.6)	1686 (39.1)	2096 (29.9)	4502 (45.3)
Use of comedication, *n* (%)								
Glucocorticosteroids	11 152 (11.2)	3533 (10)	1533 (13.9)	5950 (11.5)	2280 (10.5)	464 (10.8)	682 (9.7)	1086 (10.9)
NSAIDs	14 750 (14.9)	4788 (13.5)	1557 (14.1)	8271 (16)	3549 (16.3)	631 (14.6)	1443 (20.6)	1361 (13.7)
Antiplatelet agents	39 178 (39.5)	10 841 (30.5)	4462 (40.4)	23 538 (45.6)	8121 (37.3)	1564 (36.3)	2074 (29.6)	4330 (43.6)
SSRIs	3670 (3.7)	1410 (4)	357 (3.2)	1853 (3.6)	762 (3.5)	183 (4.2)	197 (2.8)	359 (3.6)
Presence or history of comorbidities, *n* (%)								
Malignancy	3259 (3.3)	1014 (2.9)	616 (5.6)	1589 (3.1)	600 (2.8)	130 (3)	210 (3)	238 (2.4)
Thrombocytopenia	174 (0.2)	38 (0.1)	44 (0.4)	90 (0.2)	36 (0.2)	9 (0.2)	11 (0.2)	15 (0.2)
Major bleeding	1134 (1.1)	381 (1.1)	167 (1.5)	572 (1.1)	193 (0.9)	45 (1)	55 (0.8)	89 (0.9)
Alcohol abuse	355 (0.4)	109 (0.3)	55 (0.5)	183 (0.4)	116 (0.5)	36 (0.8)	31 (0.4)	49 (0.5)
Gastrointestinal ulcer	206 (0.2)	59 (0.2)	37 (0.3)	109 (0.2)	17 (0.1)	4 (0.1)	2 (0)	11 (0.1)
Hepatic failure	273 (0.3)	76 (0.2)	52 (0.5)	139 (0.3)	40 (0.2)	2 (0)	14 (0.2)	12 (0.1)

TABLE 1. Continued.

	VKA				DOAC			
	All subjects	**Subjects treated by GP only**	**Subjects treated by specialist only**	**Subjects treated by GP & specialist**	**All subjects**	**Subjects treated by GP only**	**Subjects treated by specialist only**	**Subjects treated by GP & specialist**
Stroke/TIA	3159 (3.2)	1056 (3)	239 (2.2)	1821 (3.5)	735 (3.4)	237 (5.5)	124 (1.8)	357 (3.6)
DVT/PE	466 (0.5)	214 (0.6)	71 (0.6)	173 (0.3)	89 (0.4)	17 (0.4)	32 (0.5)	38 (0.4)
Other cardiovascular disease	10 733 (10.8)	2752 (7.8)	1487 (13.5)	6382 (12.4)	1558 (7.1)	331 (7.7)	412 (5.9)	771 (7.8)
Renal failure	2811 (2.8)	823 (2.3)	572 (5.2)	1378 (2.7)	524 (2.4)	140 (3.2)	126 (1.8)	241 (2.4)
Hypertension	69 019 (69.6)	23 841 (67.2)	7398 (67)	37 149 (71.9)	14 884 (68.3)	2972 (68.9)	4159 (59.4)	7458 (75.1)
Diabetes mellitus	17 670 (17.8)	6217 (17.5)	1751 (15.9)	9518 (18.4)	3563 (16.3)	782 (18.1)	886 (12.7)	1827 (18.4)
Congestive heart failure	27 313 (27.5)	9029 (25.4)	3422 (31)	14 544 (28.2)	4114 (18.9)	950 (22)	972 (13.9)	2107 (21.2)

Abbreviations – VKA: vitamin K antagonist; DOAC: direct oral anticoagulant; GP: general practitioner; NVAF: non-valvular atrial fibrillation; NSAIDs: non-steroidal anti-inflammatory drugs; SSRIs: selective serotonin inhibitors; TIA: transient ischemic attack; DVT: deep venous thrombosis; PE: pulmonary embolis

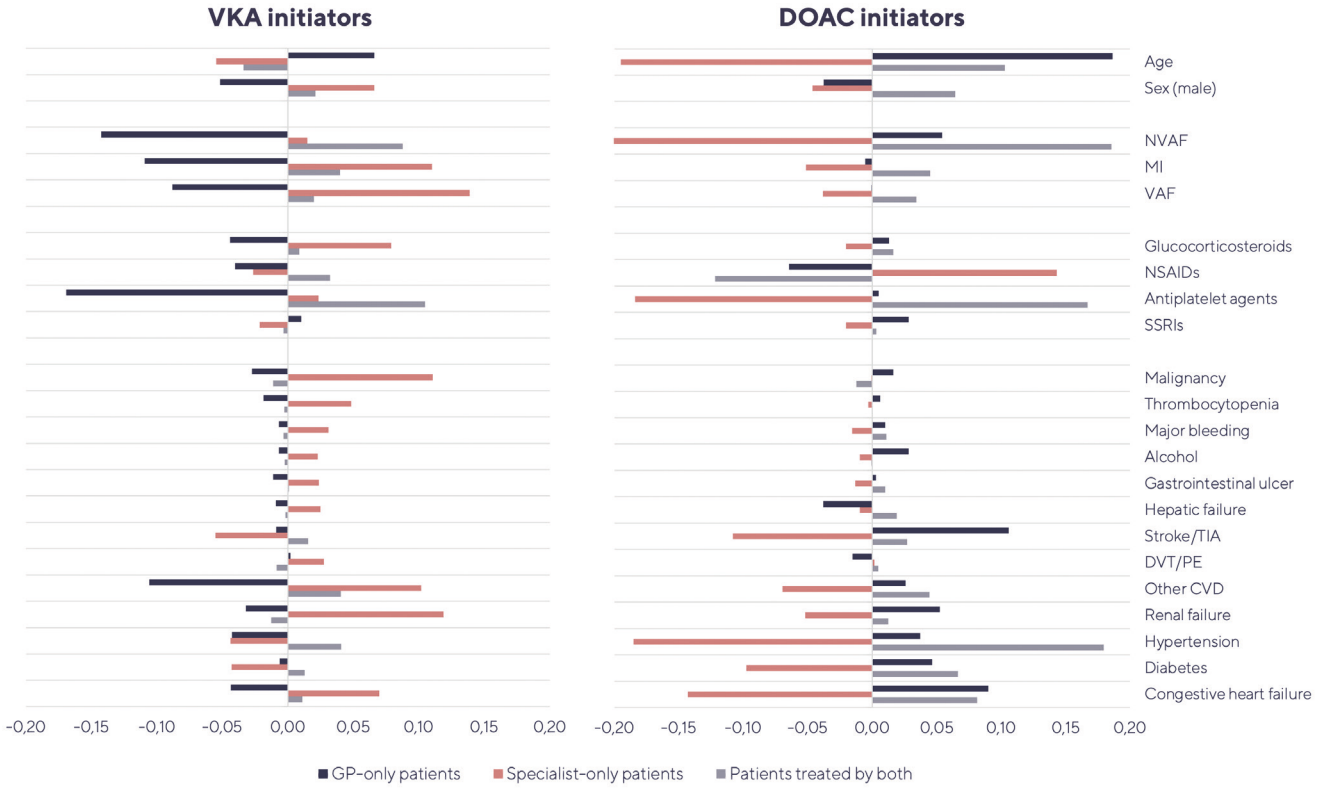


FIGURE 1. Standardized mean differences of the baseline characteristics of VKA or DOAC initiators stratified per prescribing physician (GP, specialist, or both), compared to all VKA or DOAC initiators. VKA: vitamin K antagonist; DOAC: direct oral anticoagulant; GP: general practitioner; NVAF: nonvalvular atrial fibrillation; NSAIDs: nonsteroidal anti-inflammatory drugs; SSRIs: selective serotonin inhibitors; TIA: transient ischaemic attack; DVT: deep venous thrombosis; PE: pulmonary embolism

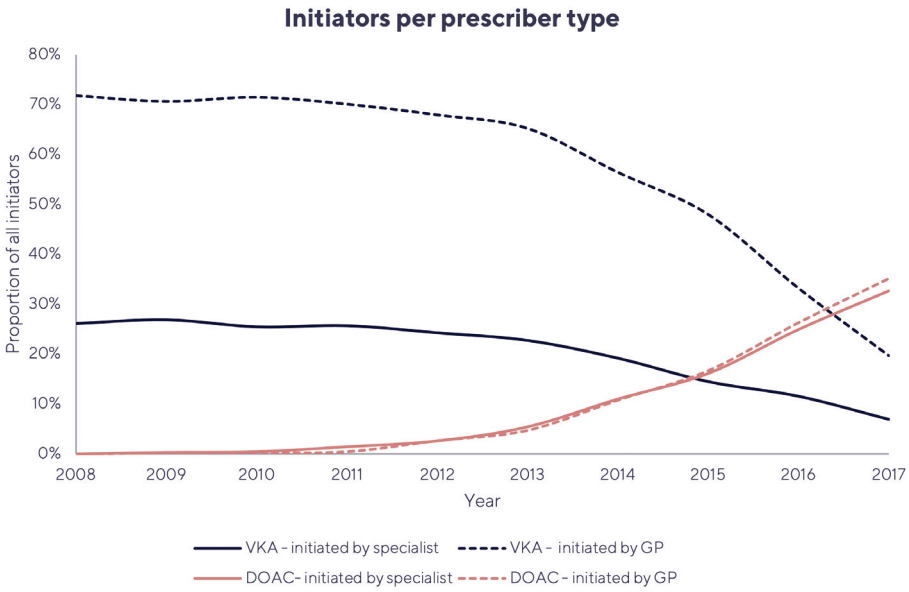


FIGURE 2. Distribution of anticoagulant initiators per study drug and prescriber type, per calendar year. VKA: vitamin K antagonist; DOAC: direct oral anticoagulant; GP: general practitioner

3.1. Prescribing physicians

Figure 2 shows the number of initiators per prescriber type for VKAs and DOACs. From 2012, the number of people who started using VKA decreased rapidly and was replaced by DOAC initiators. For about 25% of all VKA initiators, the first anticoagulant dispensed was prescribed by a specialist, compared to 50% for the DOAC initiators. In total, GP-prescribed dispensings accounted for about 80% of all VKA exposure time and about 65% of all DOAC exposure time.

During the whole study period, about half of the VKA and DOAC users had their prescriptions issued by both a GP and a specialist (Figure 3) There were more VKA users who had the prescriptions only issued by a GP compared to the DOAC users (35.7 *vs.* 18.6%). Consequently, more DOAC users had their prescriptions only issued by the specialist compared to the VKA users (31.1 *vs.* 11.2%).

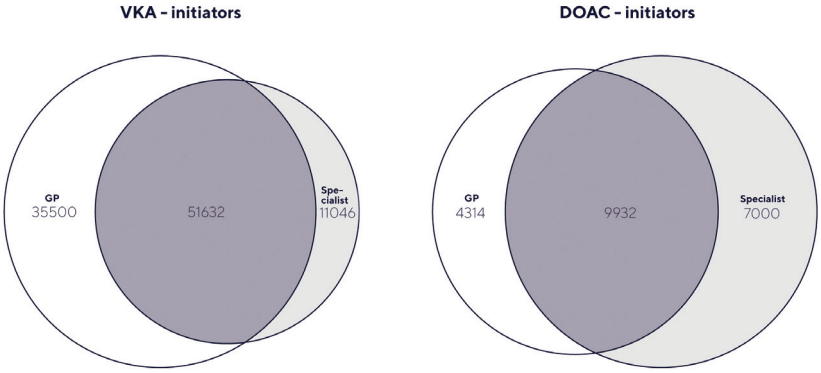


FIGURE 3. Number of VKA and DOAC initiators treated by the general practitioner, by the specialist, or by both. VKA: vitamin K antagonist; DOAC: direct oral anticoagulant; GP: general practitioner

3.2. Primary outcomes

There were 3372 major bleeding events during VKA exposure and 390 during DOAC exposure. For exposure based only on GP-prescribed dispensings, 2706 events occurred during VKA exposure and 248 during DOAC exposure. For exposure based on specialist-prescribed dispensings only, these numbers were 828 and 135 for VKA and DOAC exposure, respectively. Compared to current VKA use, crude hazard ratios of current use of DOACs for major bleeding were 0.77 (95% confidence interval 0.69–0.85), 0.83 (0.73–0.95) and 0.61 (0.51–0.73), for exposure based on complete dispensing information, only GP- and only specialist-prescribed dispensings, respectively. The adjusted hazard ratios of current use of DOACs for major bleeding were 0.79 (0.70–0.90), 0.78 (0.68–0.91) and 0.62 (0.50–0.76), respectively. The effects of DOAC use on gastrointestinal bleeding, intracranial bleeding and stroke/TIA are presented in Table 2, stratified by prescriber.

3.3. Sensitivity analyses

The results of sensitivity analyses are presented in the supplementary materials (Tables S8–S12). Stratification per age category did not result in materially different effect estimates (Table S8). Restriction to only subjects with a registered NVAF indication showed the same patterns as the primary analysis. For the risk of gastrointestinal bleeding, the HRs were 1.07 (0.83–1.37), 1.18 (0.88–1.57) and 0.74 (0.50–1.11) for complete dispensing information, or for only GP- and specialist-prescribed dispensings, respectively. For the risk of intracranial bleeding, the HRs were 0.80 (0.54–1.18), 0.87 (0.55–1.37) and 0.62 (0.33–1.16) and for the risk of stroke, the HRs were 0.89 (0.73–1.08), 1.02 (0.80–1.30) and 0.67 (0.49–0.91) (Table S9).

TABLE 2. Hazard ratios for exposure to DOACs compared to exposure to VKAs for the different outcomes

Included prescriptions	Person-years (×1000)		No of events		Hazard ratios (95% confidence interval)	
	VKA	DOAC	VKA	DOAC	Crude	Adjusted
Any bleeding						
All dispensings	214.71	31.14	3372	390	0.77 (0.69–0.85)	0.79 (0.70–0.90)
GP prescribed dispensings	177.06	18.76	2706	248	0.83 (0.73–0.95)	0.78 (0.68–0.91)
Specialist prescribed dispensings	45.85	12.22	828	135	0.61 (0.51–0.73)	0.62 (0.50–0.76)
Gastrointestinal bleeding						
All dispensings	217.43	31.44	1342	200	0.98 (0.85–1.14)	0.95 (0.80–1.13)
GP prescribed dispensings	179.17	18.95	1056	126	1.07 (0.89–1.29)	0.99 (0.80–1.22)
Specialist prescribed dispensings	46.23	12.30	350	68	0.72 (0.56–0.94)	0.65 (0.48–0.87)
Intracranial bleeding						
All dispensings	218.88	31.62	771	69	0.59 (0.46–0.75)	0.67 (0.51–0.89)
GP prescribed dispensings	180.32	19.06	635	44	0.62 (0.45–0.84)	0.65 (0.46–0.92)
Specialist prescribed dispensings	46.47	12.35	176	24	0.52 (0.34–0.80)	0.58 (0.36–0.93)
Stroke/TIA						
All dispensings	215.46	31.14	2341	329	0.93 (0.83–1.04)	0.96 (0.84–1.09)
GP prescribed dispensings	177.75	18.77	1860	219	1.07 (0.93–1.23)	1.00 (0.84–1.18)
Specialist prescribed dispensings	46.04	46.04	570	113	0.75 (0.62–0.92)	0.72 (0.58–0.90)

Abbreviations: DOAC: direct oral anticoagulant; GP: general practitioner; TIA: transient ischaemic attack; VKA: vitamin K antagonist

Stratification by sex generally showed lower effect estimates for men, except for the effect estimates for stroke/TIA, when only GP-prescribed dispensing information was used, compared to complete dispensing information (Table S10). Analysing the different DOACs separately did show the same patterns as the primary analysis, except for edoxaban (Table S11). There were, however, only a few edoxaban users, resulting in wide confidence intervals. Stratification on index date showed lower effect estimates for the subjects with an index date before 2013 than for subjects with an index date after 2013. The estimates for the analyses with only GP-prescribed dispensing information were however in line with the estimates obtained with the complete dispensing information (Tables S12).

4. DISCUSSION

Compared to using all dispensing information, including only anticoagulant dispensings prescribed by GPs did not materially impact the effect estimates of DOAC use compared to VKA use on the risk of major bleeding and stroke/TIA. However, including only dispensings issued by the specialist, strengthened the effect estimates, compared with the analysis including all dispensing information.

Using only the GP- or specialist-prescribed dispensings lead to misclassification of the exposure status in different ways. Some subjects were later enrolled in the study, some subjects had exposure misclassification during their study follow-up, and other subjects were completely left out of the study population. In total, GP-prescribed dispensings accounted for about 80% of all VKA exposure time and about 65% of all DOAC exposure time.

In addition, subjects treated only by the GP, only by the specialist or treated by both had different characteristics that also differed between initiators of VKAs and DOACs. Also, the proportion of subjects with a registered NVAF diagnosis differed: subjects with only DOAC dispensings prescribed by the GP had more often a registered NVAF diagnosis compared to subjects with only DOAC dispensings prescribed by the specialist (39.1% *vs.* 29.9%), whereas the opposite was true for the VKA users (25.4 *vs.* 33.3%).

Including prescription information from only 1 type of prescriber might have led to different biases in this case study. First, a proportion of subjects was left out of the study sample completely. Because this was unlikely to be a random process, this may have led to a sample that was not representative of the entire treated population (selection bias). The characteristics of DOAC users treated only by the GP, only by the specialist or by both differed, and we also saw these differences in characteristics when comparing initiators of VKAs with initiators of DOACs. Second, because not all information on exposures was recorded correctly, information bias could occur. For some subjects, the extent of misclassified exposure time was larger than for others, and this too was associated with measured patient characteristics that were related to the outcome. Although these measured patient characteristics could be adjusted for in the analysis of the study, misclassification may also depend on unmeasured patient characteristics, suggesting differential misclassification.

We note that the estimates from the analyses with only the GP-prescribed dispensings were in line with the analyses with all dispensings. This can be explained by the fact that the

GP-prescribed dispensings accounted for the majority of all dispensings, resulting in only limited exposure misclassification. However, analyses based on only specialist prescriptions dispensed showed more extreme effect estimates of the risk of bleeding and the risk of stroke/TIA, when compared to using all dispensing information. This was confirmed in the various sensitivity analyses conducted. Since the specialist prescriptions accounted for only 20 and 35% of all VKA and DOAC prescriptions dispensed respectively, selection and information bias might have caused these deviating results and the study population in this analysis might not be representative of the overall population treated. In addition, the involvement of a cardiologist is associated with lower risk of bleeding and stroke in patients treated with anticoagulant drugs,[7,45] therefore these results should be interpreted within the context of secondary care.

The strength of this study was that the utilization and prescribing patterns found in this study were in line with previously found results, such as the sharp increase in DOAC use from 2012 and the percentage of subjects that has only DOAC prescriptions issued by the GP or specialist over time.[7,46,47] Also the estimates for the bleeding risk of DOAC use compared to VKA use were comparable with estimates found in other observational studies.[32,48,49]

One limitation of this study is that the exposure information from pharmacies is still a proxy for actual use of the drug and could also be prone to misclassification. This can happen for example when subjects do pick up the drug, but do not start actually using it.[4] This is, however, not very likely when subjects repeatedly pick up the prescriptions. Moreover, we do not expect that this possible misclassification would have influenced our conclusions, since this could occur for both the GP and the specialist-prescribed dispensings. Inpatient dispensing information was also lacking in this study, which could occur either at the start of the anticoagulant treatment, or during the treatment. Most anticoagulant treatments are, however, initiated in outpatient care and the allowance for a 30-day gap between the theoretical end of a dispensing and the start of a new dispensing would have covered the gaps during treatment hospitalizations. We therefore expect no material effect of these missing dispensings either. Allowing this 30-day gap could also have filled the gaps caused by subjects switching prescriber type, which could have hindered our primary question. However, in the Netherlands, drugs for chronic diseases are most often prescribed for 90 days, so these gaps would not have been filled.

We also did not have complete information about the indication for the anticoagulant treatment. The inclusion criterion in this study was incident anticoagulant use, rather than a diagnosis of NVAF, which is more commonly used. This resulted in a heterogeneous study population. Since DOACs were approved for the indication NVAF between April 2011 and September 2012, we performed a sensitivity analysis in which we stratified between subjects included before and after January 2013 (Table S12). This analysis confirmed our findings from the main analysis**.** We also performed a sensitivity analysis with only subjects with a registered diagnosis of NVAF (Table S9). This analysis, however, excluded the subjects who did have a diagnosis of NVAF that was not recorded in our databases, which could have led to selection bias.[50] The results of this analysis again did not lead to any other conclusions.

In addition, there was limited information about the dose regimen of the VKA treatment, since these dose regimens are highly flexible. Therefore, a proxy was used to estimate the time on treatment. This exposure misclassification was expected to be nondifferential, and to have no

relation with the prescribing physician. Therefore, we expect this not to affect our conclusions. Last, misclassification of the outcome could have occurred. Again, we do not expect that this misclassification would differ between the GP-prescribed and the specialist-prescribed subjects, and thus affecting the conclusions.

To conclude, including only GP-prescribed anticoagulant dispensings did not materially impact the effect estimates of bleeding risk of the use DOACs compared to the use of VKAs in this study. However, including only specialist-prescribed dispensings did have impact on the effect estimates. Specifying the setting in which the study was performed (primary or secondary care) is thus of importance when reporting on the safety and effectiveness of anticoagulant drugs. Whether the same results would be obtained if other databases had been used, or with other drug–outcome relationships, remains unknown, and is highly dependent on the specific characteristics of the database that is being used: which patients, prescriptions or dispensings are included in the database and which are not? Since misclassification in a particular database that contains only prescriptions of 1 type of prescriber is likely to be drug and context-specific, we recommend further research, for example with other drug exposure–outcomes relations or other databases. For now, we recommend using databases that are as complete as possible in terms of prescriptions history for patients without regards to type of prescriber to avoid exposure misclassification and, as a result, biased results.

Author's contribution

MH designed the study, conducted the data-analysis, wrote the first draft of the manuscript, and implemented the contribution of the co-authors and external reviewers up to final publication. During the whole process she implemented input and feedback from the other contributors to this study.

REFERENCES

1. Cox E, Martin BC, Staa T Van, Garbe E, Siebert U, Johnson ML. Good Research Practices for Comparative Effectiveness Research: Approaches to Mitigate Bias and Confounding in the Design of Nonrandomized Studies of Treatment Effects Using Secondary Data Sources: The International Society for Pharmacoeconomics and Outco. *Value Heal.* 2009;12(8):1053-1061. doi:10.1111/j.1524-4733.2009.00601.x
2. U.S. Department of Health and Human Services, Administration Food and Drug, Research (CDER) Center for Drug Evaluation and, Research (CBER) Center for Biologics Evaluation and. Guidance for Industry and FDA Staff Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data. *Drug Saf.* Published online 2013. Accessed April 18, 2018. http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm
3. Rothman K, Greenland S, Lash T. *Modern Epidemiology.* 3rd ed. Lippincott Williams & Wilkins (LWW); 2008.
4. Jonsson Funk M, Landi SN. Misclassification in Administrative Claims Data: Quantifying the Impact on Treatment Effect Estimates. *Curr Epidemiol Reports.* 2014;1(4):175-185. doi:10.1007/s40471-014-0027-z
5. National Institute for Health and Care Excellence. Atrial fibrillation: management. NICE guideline CG180. 2014;(June 2014). Accessed April 24, 2018. https://www.nice.org.uk/guidance/cg180
6. Pharmaceutisch Weekblad [Pharmaceutical Weekly Magazine]. NHG: DOAC en cumarine zijn gelijkwaardig. Accessed April 24, 2018. https://www.pw.nl//nieuws/2016/nhg-noac-en-cumarine-zijn-gelijkwaardig
7. O'Neal WT, Sandesara PB, Claxton JS, et al. Provider Specialty, Anticoagulation Prescription Patterns, and Stroke Risk in Atrial Fibrillation. *J Am Heart Assoc.* 2018;7(6):e007943. doi:10.1161/JAHA.117.007943
8. PubMed. https://www.ncbi.nlm.nih.gov/
9. ENCePP. http://www.encepp.eu/index.shtml
10. Coleman CI, Turpie AGG, Bunz TJ, Beyer-Westendorf J. Effectiveness and safety of rivaroxaban versus warfarin in patients with provoked venous thromboembolism. *J Thromb Thrombolysis.* Published online June 7, 2018. doi:10.1007/s11239-018-1695-1
11. Margolis JM, Deitelzweig S, Kline J, et al. Shorter Hospital Stays and Lower Costs for Rivaroxaban Compared With Warfarin for Venous Thrombosis Admissions. *J Am Heart Assoc.* 2016;5(10):e003788. doi:10.1161/JAHA.116.003788
12. Shiga T, Naganuma M, Nagao T, et al. Persistence of non-vitamin K antagonist oral anticoagulant use in Japanese patients with atrial fibrillation: A single-center observational study. *J Arrhythmia.* 2015;31(6):339-344. doi:10.1016/j.joa.2015.04.004
13. Al-Khalili F, Lindström C, Benson L. The safety and persistence of non-vitamin-K-antagonist oral anticoagulants in atrial fibrillation patients treated in a well structured atrial fibrillation clinic. *Curr Med Res Opin.* 2016;32(4):779-785. doi:10.1185/03007995.2016.1142432
14. Coleman CI, Antz M, Bowrin K, et al. Real-world evidence of stroke prevention in patients with nonvalvular atrial fibrillation in the United States: the REVISIT-US study. *Curr Med Res Opin.* 2016;32(12):2047-2053. doi:10.1080/03007995.2016.1237937
15. Larsen TB, Skjøth F, Nielsen PB, Kjældgaard JN, Lip GYH. Comparative effectiveness and safety of non-vitamin K antagonist oral anticoagulants and warfarin in patients with atrial fibrillation: propensity weighted nationwide cohort study. *BMJ.* 2016;353:i3189. doi:10.1136/BMJ.I3189
16. Halvorsen S, Ghanima W, Fride Tvete I, et al. A nationwide registry study to compare bleeding rates in patients with atrial fibrillation being prescribed oral anticoagulants. *Eur Hear J - Cardiovasc Pharmacother.* 2017;3(1):28-36. doi:10.1093/ehjcvp/pvw031
17. Lamberts M, Staerk L, Olesen JB, et al. Major Bleeding Complications and Persistence With Oral Anticoagulation in Non-Valvular Atrial Fibrillation: Contemporary Findings in Real-Life Danish Patients. *J Am Heart Assoc.* 2017;6(2):e004517. doi:10.1161/JAHA.116.004517

18. Li XS, Deitelzweig S, Keshishian A, et al. Effectiveness and safety of apixaban versus warfarin in non-valvular atrial fibrillation patients in "real-world" clinical practice. A propensity-matched analysis of 76,940 patients. *Thromb Haemost*. 2017;117(6):1072-1082. doi:10.1160/TH17-01-0068
19. Staerk L, Fosbøl EL, Lip GYH, et al. Ischaemic and haemorrhagic stroke associated with non-vitamin K antagonist oral anticoagulants and warfarin use in patients with atrial fibrillation: a nationwide cohort study. *Eur Heart J*. 2016;38(12):ehw496. doi:10.1093/eurheartj/ehw496
20. Noseworthy PA, Yao X, Abraham NS, Sangaralingham LR, McBane RD, Shah ND. No Title. *Chest*. 2016;150(6):1302-1312. doi:10.1016/j.chest.2016.07.013
21. Butt JH, Li A, Xian Y, et al. Direct oral anticoagulant- versus vitamin K antagonist–related gastrointestinal bleeding: Insights from a nationwide cohort. *Am Heart J*. 2019;216:117-124. doi:10.1016/j.ahj.2019.07.012
22. Jun M, Lix LM, Durand M, et al. Comparative safety of direct oral anticoagulants and warfarin in venous thromboembolism: multicentre, population based, observational study. *BMJ*. 2017;359:j4323. Accessed June 12, 2018. http://www.ncbi.nlm.nih.gov/pubmed/29042362
23. Villines TC, Ahmad A, Petrini M, et al. Comparative safety and effectiveness of dabigatran vs. rivaroxaban and apixaban in patients with non-valvular atrial fibrillation: A retrospective study from a large healthcare system. *Eur Hear J - Cardiovasc Pharmacother*. 2019;5(2):80-90. doi:10.1093/ehjcvp/pvy044
24. Deitelzweig S, Keshishian A, Li X, et al. Comparisons between Oral Anticoagulants among Older Nonvalvular Atrial Fibrillation Patients. *J Am Geriatr Soc*. 2019;67(8):1662-1671. doi:10.1111/jgs.15956
25. Amin A, Garcia Reeves AB, Li X, et al. Effectiveness and safety of oral anticoagulants in older adults with non-valvular atrial fibrillation and heart failure. *PLoS One*. 2019;14(3). doi:10.1371/journal.pone.0213614
26. Gupta K, Trocio J, Keshishian A, et al. Effectiveness and safety of direct oral anticoagulants compared to warfarin in treatment naïve non-valvular atrial fibrillation patients in the US Department of defense population. *BMC Cardiovasc Disord*. 2019;19(1). doi:10.1186/s12872-019-1116-1
27. Ohshima A, Koyama T, Ogawa A, et al. Oral anticoagulants usage in Japanese patients aged 18-74 years with non-valvular atrial fibrillation: A retrospective analysis based on insurance claims data. *Fam Pract*. 2019;36(6):685-692. doi:10.1093/fampra/cmz016
28. Datar M, Crivera C, Rozjabek H, et al. Comparison of real-world outcomes in patients with nonvalvular atrial fibrillation treated with direct oral anticoagulant agents or warfarin. *Am J Heal Pharm*. 2019;76(5):275-285. doi:10.1093/ajhp/zxy032
29. Haastrup SB, Hellfritzsch M, Rasmussen L, Pottegård A, Grove EL. Use of Non-Vitamin K Antagonist Oral Anticoagulants 2008-2016: A Danish Nationwide Cohort Study. *Basic Clin Pharmacol Toxicol*. 2018;123(4):452-463. doi:10.1111/bcpt.13024
30. Cerdá M, Cerezo-Manchado JJ, Johansson E, et al. Facing real-life with direct oral anticoagulants in patients with nonvalvular atrial fibrillation: Outcomes from the first observational and prospective study in a Spanish population. *J Comp Eff Res*. 2019;8(3):165-178. doi:10.2217/cer-2018-0134
31. Souverein PC, van den Ham HA, Huerta C, et al. Comparing risk of major bleeding between users of different oral anticoagulants in patients with nonvalvular atrial fibrillation. *Br J Clin Pharmacol*. 2020;87(3):988-1000. doi:10.1111/bcp.14450
32. Lip GYH, Keshishian A, Li X, et al. Effectiveness and safety of oral anticoagulants among nonvalvular atrial fibrillation patients: The ARISTOPHANES study. *Stroke*. 2018;49(12):2933-2944. doi:10.1161/STROKEAHA.118.020232
33. Chan Y-H, See L-C, Tu H-T, et al. Efficacy and Safety of Apixaban, Dabigatran, Rivaroxaban, and Warfarin in Asians With Nonvalvular Atrial Fibrillation. *J Am Heart Assoc*. 2018;7(8):e008150. doi:10.1161/JAHA.117.008150

34. van den Ham HA, Klungel OH, Singer DE, Leufkens HGM, van Staa TP. Comparative Performance of ATRIA, CHADS2, and CHA2DS2-VASc Risk Scores Predicting Stroke in Patients With Atrial Fibrillation. *J Am Coll Cardiol*. 2015;66(17):1851-1859. doi:10.1016/j.jacc.2015.08.033
35. Deitelzweig S, Bruno A, Trocio J, et al. An early evaluation of bleeding-related hospital readmissions among hospitalized patients with nonvalvular atrial fibrillation treated with direct oral anticoagulants. *Curr Med Res Opin*. 2016;32(3):573-582. doi:10.1185/03007995.2015.1131676
36. Noseworthy PA, Yao X, Abraham NS, Sangaralingham LR, McBane RD, Shah ND. Direct Comparison of Dabigatran, Rivaroxaban, and Apixaban for Effectiveness and Safety in Nonvalvular Atrial Fibrillation. *Chest*. 2016;150(6):1302-1312. doi:10.1016/j.chest.2016.07.013
37. Lee SI, Sayers M, Lip GYH, Lane DA. Use of non-vitamin K antagonist oral anticoagulants in atrial fibrillation patients: insights from a specialist atrial fibrillation clinic. *Int J Clin Pract*. 2015;69(11):1341-1348. doi:10.1111/ijcp.12712
38. Nielsen PB, Skjøth F, Søgaard M, Kjældgaard JN, Lip GYH, Larsen TB. Effectiveness and safety of reduced dose non-vitamin K antagonist oral anticoagulants and warfarin in patients with atrial fibrillation: propensity weighted nationwide cohort study. *BMJ*. 2017;356:j510. http://www.ncbi.nlm.nih.gov/pubmed/28188243
39. Forslund T, Wettermark B, Andersen M, Hjemdahl P. Stroke and bleeding with non-vitamin K antagonist oral anticoagulant or warfarin treatment in patients with non-valvular atrial fibrillation: a population-based cohort study. *EP Eur*. 2018;20(3):420-428. doi:10.1093/europace/euw416
40. WHO-ATC DDD index. Accessed June 12, 2018. http://www.whocc.no/atc_ddd_index/?code=C01BD01
41. Conversion ICD-10 to ICD-9. Accessed June 12, 2018. http://www.icd10codesearch.com/
42. WHO ICD-10 classification. Accessed June 12, 2018. http://www.who.int/classifications/icd/en/
43. WHO ICPC Classification. Accessed June 12, 2018. https://www.who.int/classifications/icd/adaptations/icpc2/en/
44. Austin PC. Using the Standardized Difference to Compare the Prevalence of a Binary Variable Between Two Groups in Observational Research. *Commun Stat - Simul Comput*. 2009;38(6):1228-1234. doi:10.1080/03610910902859574
45. Perino AC, Fan J, Schmitt SK, et al. Treating Specialty and Outcomes in Newly Diagnosed Atrial Fibrillation: From the TREAT-AF Study. *J Am Coll Cardiol*. 2017;70(1):78-86. doi:http://dx.doi.org/10.1016/j.jacc.2017.04.054
46. Ibáñez L, Sabaté M, Vidal X, et al. Incidence of direct oral anticoagulant use in patients with nonvalvular atrial fibrillation and characteristics of users in 6 European countries (2008–2015): A cross-national drug utilization study. *Br J Clin Pharmacol*. 2019;85(11):2524-2539. doi:10.1111/bcp.14071
47. Huiart L, Ferdynus C, Renoux C, et al. Trends in initiation of direct oral anticoagulant therapies for atrial fibrillation in a national population-based cross-sectional study in the French health insurance databases. *BMJ Open*. 2018;8(3):e018180. doi:10.1136/bmjopen-2017-018180
48. Yao X, Abraham NS, Sangaralingham LR, et al. Effectiveness and Safety of Dabigatran, Rivaroxaban, and Apixaban Versus Warfarin in Nonvalvular Atrial Fibrillation. *J Am Heart Assoc*. 2016;5(6):e003725. doi:10.1161/JAHA.116.003725
49. Coleman CI, Peacock WF, Bunz TJ, Alberts MJ. Effectiveness and Safety of Apixaban, Dabigatran, and Rivaroxaban Versus Warfarin in Patients With Nonvalvular Atrial Fibrillation and Previous Stroke or Transient Ischemic Attack. *Stroke*. 2017;48(8):2142-2149. doi:10.1161/STROKEAHA.117.017474
50. Hellfritzsch M, Pottegård A, Haastrup SB, Rasmussen L, Grove EL. Cohort selection in register-based studies of direct oral anticoagulant users with atrial fibrillation: An inevitable trade-off between selection bias and misclassification. *Basic Clin Pharmacol Toxicol*. Published online May 4, 2020. doi:10.1111/bcpt.13423

SUPPLEMENTARY MATERIALS

TABLE S1. ATC codes of study drugs

Medication class	Name	ATC
OAC	Phenprocoumon	B01AA04
	Acenocoumarol	B01AA07
DOAC	Dabigatran	B01AE07
	Rivaroxaban	B01AF01
	Apixaban	B01AF02
	Edoxaban	B01AF04

TABLE S2. ICD codes for major bleeding and stroke/TIA

Outcome	ICD-10	ICD-9-CM
Major bleeding		
Haemorrhagic stroke/ intracranial Bleeding	I60 I61 I62	430. 431. 432
Extracranial or unclassified major bleeding	D62, J94.2, H11.3, H31.3, H35.6, H43.1. N95.0. R04, R31. R58, M25.0	285.1. 511.1. 511.89, 372.72, 363.6, 362.81. 379.23, 459.0. 596.7,599.7, 627.1.719.1.784.7, 784.8, 786.3
Gastrointestinal bleeding	I85, K22.6, K22.8, K25.0. K25.2, K25.4, K25.6, K26.0. K26.2, K26.4, K26.6, K27.0. K27.2, K27.4, K27.6,K28.0. K28.2, K28.4, K28.6, K29.01. K29.21. K29.31. K29.41. K29.51. K29.61. K.29.71. K29.81. K29.91. K31.8, K62.5, K66.01. K92.0. K92.1. K92.2, K66.1 K57.01. K57.11. K57.03, K57.13, K57.21. K57.31. K57.23, K57.33, K55.2, K57.41. K57.51. K57.43, K57.53, K57.81. K57.91. K57.83, K57.93	456.0. 530.7, 530.82;531.0. 531.2, 531.4, 531.6, 532.0. 532.2, 532.4, 532.6, 533.0. 533.2, 533.4, 533.6, 534.0, 534.2, 534.4, 534.6, 535.01. 535.11. 535.21. 535.31. 535.41. 535.51. 535.61. 537.83562.02, 562.03, 562.12, 562.13, 568.81. 569.3, 569.85, 578.0. 578.1. 578.9
Traumatic intracranial bleeding	S06.3, S06.4, S06.5, S06.6, S06.8	852.0. 852.1. 852.2, 852.3, 852.4, 852.5, 853.0. 853.1800.1-8, 801.1-8,
Stroke/TIA		
Cerebral infarction	I63	433.x1. 434.01. 434.11. 434.91
Stroke (not specified as haemorrhage or infarction) and TIA	I64, G45	430. 431. 432, 433, 434, 435

TABLE S3. ICD codes of co-morbidities and risk factors for major bleeding

	ICD-9-CM (2)	ICD-10	Time window for identification
Major Bleeding	See table A3.1	See table A3.1	6 months prior to index date and during follow-up
Hypertension	401-405	I10, I11, I12, I13, I15	6 months prior to index date and during follow-up
Stroke/TIA	430, 431, 432, 433, 434, 435	I63, I64, G45	6 months prior to index date and during follow-up
DVT, PE	451.1x, 451.81, 415.1x, 453.4, 453.5	I80.1, I80.2, I80.4, I82.4, I82.5, I26	6 months prior to index date and during follow-up
Alcohol	265, 303.00, 303.01, 303.02, 303.90, 303.91, 303.92, 305.00, 305.01, 305.02, 357.5, 425.5, 535.3, 571.1, 571.2, 571.3, 655.4, 980	E52, F10. G31.2, G62.1. G72.1 I42.6, K29.2, K70. K86.0. O35.4, T51 Z71.4, Z72.1	6 months prior to index date and during follow-up
Any malignancy, except malignant neoplasm of the skin	140.x–209.x, Excluding: 172, 173	C00-C97 excluding C43, C44	6 months prior to index date and during follow-up
Gastrointestinal ulcer	530.2, 531.x–534.x	K22.1. K25-K28	6 months prior to index date and during follow-up
Thrombocytopenia	287, 287.1, 287.3, 287.30, 287.31, 287.32, 287.33, 287.39, 287.4, 287.41. 287.49, 287.5, 279.12	D69.1. D69.3, D69.4, D69.41. D69.42, D69.49, D69.5, D69.51. D69.59, D69.6 D82.0	6 months prior to index date and during follow-up
Hepatic impairment	070.0, 070.01, 070.20, 070.21, 070.30, 070.31, 070.6, 456.0, 456.1, 456.21, 570. 572.2, 572.3, 572.8, 070.22, 070.23, 070.32, 070.33, 070.44, 070.54, 07059, 571.0-6, 571.8, 571.9, 573.3	B15.0. B16.0. B16.2, B19.0. I85, K70.4, K72, K76.6, B18, K70.0-K70.3, K70.9, K71. K73, K74, K76.0	6 months prior to index date and during follow-up
Renal failure, chronic	403.01, 403.11, 403.91, 404.02, 404.03, 404.12, 404.13, 404.92, 404.93, 580.8, 581. 582, 585.3, 584, 585.5, 585.5, 585.6, 585.9, 586, 588, 590.1, 590.20, 590.8, 753.1, V42.0	I12, I13, N00. N01. N03, N04, N05, N07, N08, N10. N11. N12, N14, N17, N18.4, N18.5, N18.9, N19, Q61. Z94.0	6 months prior to index date and during follow-up

TABLE S4. ICD codes of co-morbidities and risk factors for stroke/TIA

	ICD-9-CM (2)	ICD-10	Time window for identification
Stroke/TIA	430, 431, 432, 433, 434, 435	I63, I64, G45	6 months prior to index date and during follow-up
DVT, PE	451.1x, 451.81, 415.1x, 453.4, 453.5	I80.1. I80.2, I80.4, I82.4, I82.5, I26	6 months prior to index date and during follow-up
Hypertension	401-405, medication in Table S5	I10, I11, I12, I13, I15, medication in Table S5	6 months prior to index date and during follow-up
Congestive heart failure	402.01. 402.11. 402.91. 428.x, 404.01. 404.03, 404.11. 404.13, 404.91. 404.93, ATC: C03C	I11.0, I13.0, I13.2, I42.0, I50. ATC: C03C	6 months prior to index date and during follow-up
Diabetes	250, 357.2, 362.0, ATC: A10	E10- E14, ATC: A10	6 months prior to index date and during follow-up
Other cardiovascular disease			
Coronary heart disease (MI/Angina),	410, 411, 413, 414, 429.7	I20, I21, I22, I23, I24, I25	6 months prior to index date and during follow-up
Peripheral vascular disease	433, 434, 440.0, 440.2	I65, I66, I70.0, I70.2,	6 months prior to index date and during follow-up
Atherosclerosis/ aortic plaque	443.9	I73.9	6 months prior to index date and during follow-up

TABLE S5. ATC codes antihypertensive drugs

ATC	Drug
C02AC05	Moxonidin
C02CA04	Doxazosin
C03	Diuretics
C07	Beta blocking agents
C08	Calcium channel blockers
C09A	ACE inhibitors
C09C	Angiotensin II antagonists

TABLE S6. ATC codes for risk factors for major bleeding – Medication

ATC	Drug
M01A	NSAIDs
N06AB	SSRIs
H02	Corticosteroids
B01AC	Antiplatelet drugs
A10	Antidiabetic drugs

TABLE S7. ATC codes for pharmacokinetic interacting drugs.

Drug	ATC	Dabigatran	Apixaban	Edoxaban	Rivaroxaban	CYP-inhibitor	CYP-inductor	Strong	Moderate/weak
Amiodarone	C01BD01	x	x	x	x	x			x
Diltiazem	C08DB01		x			x			x
Quinidine	C01BA01	x	x	x	x	x			x
Verapamil	C08DA01	x				x		x	
Verapamil	C08DA01			x		x			x
Clarithromycine	J01FA09	x	x		x	x			x
Clarithromycine	J01FA09			x		x		x	
Erythromycine	J01FA01	x	x		x	x			x
Erythromycine	J01FA01			x		x		x	
Rifampicin	J04AB02	x	x		x		x	x	
HIV protease inhibitors	J05AE	x	x	x	x	x		x	
Fluconazole	J02AC01				x	x			x
Voriconazole	J02AC03	x	x	x	x	x		x	
Itraconazole	J02AC02	x	x	x	x	x		x	
Posaconazole	J02AC04	x	x	x	x	x		x	
Bicalutamide	L02BB03		x		x	x			x
Tamoxifen	L02BA01	x	x	x	x	x			x
Cyclosporine	L04AD01	x		x		x		x	
Cyclosporine	L04AD01		x		x	x			x
Dexamethasone	H02AB02	x	x	x	x		x	x	

TABLE S7. Continued.

Drug	ATC	Dabi-gatran	Apixa-ban	Edoxa-ban	Rivaro-xaban	CYP-inhibitor	CYP-inductor	Strong	Moderate/ weak
Tacrolimus	L04AD02	x	x	x	x	x		x	
Prednison	H02AB07		x		x		x		x
Carbamazepine	N03AF01	x			x		x	x	
Carbamazepine	N03AF01		x	x			x		x
Levetiracetam	N03AX14	x	x	x	x		x	x	
Oxcabarzine	N03AF02		x		x		x		x
Phenobarbitol	N03AA02	x			x		x	x	
Phenobarbitol	N03AA02		x	x			x		x
Phenytoin	N03AB02	x			x		x	x	
Phenytoin	N03AB02		x	x			x		x
Topiramate	N03AX11		x		x		x		x
Valproic acid	N03AG01	x	x	x	x		x	x	
Naproxen	M01AE02		x			x			x

Based on: Steffel J, Verhamme P, Potpara TS, et al. 2018 European Heart Rhythm Association Practical Guide on the use of non-vitamin K antagonist oral anticoagulants in patients with atrial fibrillation | European Heart Journal | Oxford Academic. *Europace*. 2018;20(8):1231-1

TABLE S8. Sensitivity analysis 1: Hazard Ratios for exposure to DOACs compared to exposure to VKAs for the different outcomes – stratified per age category

Included prescriptions	18-65 years		65-85 years		>85 years	
	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted
Any bleeding						
All dispensings	0.73 (0.55-0.97)	0.62 (0.45-0.86)	0.79 (0.69-0.91)	0.76 (0.65-0.89)	0.85 (0.70-1.03)	0.92 (0.73-1.14)
GP prescribed dispensings	0.78 (0.56-1.10)	0.64 (0.44-0.94)	0.83 (0.69-0.99)	0.78 (0.65-0.95)	0.89 (0.71-1.12)	0.94 (0.73-1.22)
Specialist prescribed dispensings	0.68 (0.46-1.00)	0.64 (0.41-0.98)	0.59 (0.47-0.74)	0.57 (0.44-0.75)	0.85 (0.61-1.18)	1.04 (0.72-1.49)
Gastrointestinal bleeding						
All dispensings	0.87 (0.59-1.28)	0.62 (0.39-0.99)	0.98 (0.80-1.20)	0.89 (0.71-1.12)	1.19 (0.91-1.55)	1.28 (0.94-1.76)
GP prescribed dispensings	0.91 (0.56-1.47)	0.63 (0.36-1.09)	1.08 (0.83-1.39)	1.00 (0.76-1.33)	1.23 (0.89-1.70)	1.35 (0.93-1.96)
Specialist prescribed dispensings	0.74 (0.41-1.31)	0.61 (0.31-1.19)	0.62 (0.44-0.89)	0.52 (0.34-0.77)	1.09 (0.66-1.80)	1.16 (0.65-2.07)
Intracranial bleeding						
All dispensings	0.53 (0.26-1.09)	0.51 (0.25-1.05)	0.54 (0.38-0.76)	0.52 (0.35-0.78)	0.80 (0.53-1.20)	1.06 (0.68-1.67)
GP prescribed dispensings	0.69 (0.30-1.59)	0.69 (0.28-1.73)	0.52 (0.34-0.82)	0.53 (0.32-0.87)	0.81 (0.49-1.32)	1.06 (0.61-1.84)
Specialist prescribed dispensings	0.33 (0.08-1.39)	0.34 (0.08-1.41)	0.5 (0.28-0.87)	0.55 (0.28-1.05)	0.80 (0.38-1.67)	0.97 (0.43-2.16)
Stroke/TIA						
All dispensings	0.90 (0.66-1.23)	0.86 (0.59-1.26)	0.92 (0.78-1.08)	1.00 (0.83-1.20)	1.1 (0.90-1.35)	0.97 (0.77-1.22)
GP prescribed dispensings	1.03 (0.70-1.52)	0.78 (0.48-1.26)	1.06 (0.87-1.28)	1.08 (0.86-1.34)	1.23 (0.97-1.56)	1.01 (0.77-1.33)
Specialist prescribed dispensings	0.75 (0.46-1.21)	0.79 (0.46-1.35)	0.75 (0.57-0.99)	0.77 (0.57-1.04)	0.91 (0.61-1.33)	0.66 (0.43-1.01)

TABLE S9. Sensitivity analysis 2: Hazard Ratios for exposure to DOACs compared to exposure to VKAs for the different outcomes – including only subjects with a registered NVAF indication

Included prescriptions	Person-years (x 1000)		No of events		Hazard Ratios (95% confidence interval)	
	VKA	DOAC	VKA	DOAC	Crude	Adjusted
Any bleeding						
All dispensings	77.00	13.92	1347	208	0.80 (0.69-0.92)	0.88 (0.74-1.05)
GP prescribed dispensings	63.78	8.52	1045	133	0.89 (0.74-1.06)	0.91 (0.74-1.12)
Specialist prescribed dispensings	16.54	5.37	358	75	0.66 (0.52-0.85)	0.66 (0.50-0.88)
Gastrointestinal bleeding						
All dispensings	78.10	14.07	547	111	1.04 (0.84-1.27)	1.07 (0.83-1.37)
GP prescribed dispensings	64.59	8.61	408	71	1.19 (0.93-1.54)	1.18 (0.88-1.57)
Specialist prescribed dispensings	16.71	5.42	153	39	0.81 (0.57-1.15)	0.74 (0.50-1.11)
Intracranial bleeding						
All dispensings	78.71	14.16	341	43	0.70 (0.51-0.96)	0.80 (0.54-1.18)
GP prescribed dispensings	65.06	8.66	260	30	0.78 (0.53-1.14)	0.87 (0.55-1.37)
Specialist prescribed dispensings	16.81	5.44	81	14	0.57 (0.32-1.00)	0.62 (0.33-1.16)
Stroke/TIA						
All dispensings	77.20	13.95	986	171	0.88 (0.74-1.03)	0.89 (0.73-1.08)
GP prescribed dispensings	64.04	8.54	713	111	1.07 (0.87-1.30)	1.02 (0.80-1.30)
Specialist prescribed dispensings	16.59	5.38	300	62	0.67 (0.51-0.88)	0.67 (0.49-0.91)

TABLE S10. Sensitivity analysis 3: Hazard Ratios for exposure to DOACs compared to exposure to VKAs for the different outcomes – stratified per sex

Included prescriptions	**Male**		**Female**	
	Crude	Adjusted	Crude	Adjusted
Any bleeding				
All dispensings	0.71 (0.61-0.81)	0.76 (0.64-0.90)	0.85 (0.73-1.00)	0.84 (0.70-1.00)
GP prescribed dispensings	0.72 (0.61-0.87)	0.72 (0.59-0.88)	0.96 (0.80-1.15)	0.93 (0.76-1.14)
Specialist prescribed dispensings	0.68 (0.54-0.84)	0.75 (0.59-0.96)	0.61 (0.47-0.79)	0.62 (0.46-0.83)
Gastrointestinal bleeding				
All dispensings	0.85 (0.69-1.05)	0.80 (0.62-1.03)	1.16 (0.94-1.43)	1.14 (0.90-1.45)
GP prescribed dispensings	0.87 (0.66-1.14)	0.81 (0.60-1.10)	1.33 (1.03-1.71)	1.30 (0.97-1.73)
Specialist prescribed dispensings	0.69 (0.48-0.98)	0.65 (0.43-0.98)	0.78 (0.53-1.14)	0.70 (0.46-1.07)
Intracranial bleeding				
All dispensings	0.61 (0.45-0.84)	0.72 (0.49-1.06)	0.55 (0.37-0.81)	0.62 (0.41-0.94)
GP prescribed dispensings	0.53 (0.34-0.82)	0.56 (0.33-0.96)	0.73 (0.48-1.11)	0.88 (0.57-1.36)
Specialist prescribed dispensings	0.63 (0.39-1.03)	0.81 (0.46-1.41)	0.31 (0.13-0.77)	0.34 (0.13-0.89)
Stroke/TIA				
All dispensings	0.96 (0.82-1.11)	1.05 (0.88-1.25)	0.89 (0.75-1.07)	0.86 (0.70-1.06)
GP prescribed dispensings	1.11 (0.92-1.34)	1.14 (0.92-1.42)	1.02 (0.82-1.27)	0.86 (0.67-1.09)
Specialist prescribed dispensings	0.78 (0.60-1.02)	0.77 (0.58-1.03)	0.71 (0.51-0.98)	0.65 (0.46-0.93)

TABLE S11. Sensitivity analysis 4: Hazard Ratios for exposure to DOACs compared to exposure to VKAs for the different outcomes – stratified per DOAC separately

Included prescriptions	**Dabigatran**		**Rivaroxaban**		**Apixaban**		**Edoxaban**	
	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted	Crude	Adjusted
Any bleeding								
All dispensings	0.59 (0.50-0.70)	0.69 (0.58-0.82)	0.76 (0.64-0.90)	0.93 (0.78-1.11)	0.67 (0.53-0.84)	0.75 (0.60-0.94)	1.09 (0.64-1.87)	1.22 (0.71-2.10)
GP prescribed dispensings	0.64 (0.52-0.80)	0.71 (0.57-0.88)	0.78 (0.64-0.96)	0.91 (0.74-1.12)	0.79 (0.61-1.02)	0.83 (0.63-1.08)	1.01 (0.45-2.28)	1.12 (0.49-2.53)
Specialist prescribed dispensings	0.52 (0.40-0.67)	0.60 (0.46-0.78)	0.65 (0.50-0.85)	0.85 (0.65-1.12)	0.52 (0.35-0.77)	0.59 (0.39-0.89)	0.98 (0.45-2.12)	1.12 (0.52-2.41)
Gastrointestinal bleeding								
All dispensings	0.90 (0.73-1.13)	1.06 (0.85-1.33)	0.82 (0.64-1.06)	1.00 (0.78-1.30)	0.52 (0.36-0.76)	0.59 (0.40-0.86)	1.40 (0.67-2.90)	1.60 (0.77-3.33)
GP prescribed dispensings	0.98 (0.75-1.30)	1.15 (0.87-1.53)	0.86 (0.63-1.17)	1.03 (0.75-1.41)	0.66 (0.42-1.02)	0.73 (0.47-1.13)	1.62 (0.59-4.44)	1.93 (0.70-5.35)
Specialist prescribed dispensings	0.65 (0.45-0.94)	0.71 (0.49-1.03)	0.58 (0.37-0.92)	0.71 (0.45-1.12)	0.48 (0.26-0.89)	0.52 (0.27-0.99)	0.64 (0.15-2.78)	0.69 (0.16-2.90)
Intracranial bleeding								
All dispensings	0.34 (0.21-0.55)	0.39 (0.23-0.65)	0.70 (0.47-1.05)	0.85 (0.56-1.27)	0.88 (0.56-1.38)	0.97 (0.61-1.53)	1.24 (0.41-3.81)	1.43 (0.47-4.33)
GP prescribed dispensings	0.36 (0.19-0.68)	0.39 (0.20-0.77)	0.71 (0.43-1.16)	0.78 (0.47-1.30)	1.17 (0.70-1.95)	1.20 (0.71-2.03)	0.97 (0.14-6.89)	1.03 (0.14-7.43)
Specialist prescribed dispensings	0.34 (0.16-0.73)	0.39 (0.18-0.86)	0.76 (0.40-1.45)	0.93 (0.48-1.81)	0.46 (0.17-1.27)	0.52 (0.18-1.47)	1.67 (0.41-6.78)	1.89 (0.49-7.28)
Stroke/TIA								
All dispensings	0.87 (0.73-1.04)	0.94 (0.78-1.12)	0.73 (0.59-0.90)	0.89 (0.72-1.11)	1.01 (0.81-1.27)	1.11 (0.88-1.39)	0.77 (0.34-1.72)	0.91 (0.41-2.05)
GP prescribed dispensings	1.01 (0.82-1.25)	1.02 (0.82-1.28)	0.76 (0.58-0.99)	0.85 (0.65-1.11)	1.21 (0.93-1.57)	1.23 (0.94-1.61)	0.54 (0.13-2.14)	0.55 (0.14-2.20)
Specialist prescribed dispensings	0.66 (0.49-0.89)	0.70 (0.52-0.94)	0.71 (0.51-0.98)	0.82 (0.59-1.15)	0.58 (0.37-0.92)	0.61 (0.38-0.97)	0.64 (0.20-2.06)	0.68 (0.21-2.20)

TABLE S12. Sensitivity analysis 5: Hazard Ratios for exposure to DOACs compared to exposure to VKAs for the different outcomes – stratified per index date

Included prescriptions	Index date before 01-01-2013		Index date after 01-01-2013	
	Crude	Adjusted	Crude	Adjusted
Any bleeding				
All dispensings	0.77 (0.69-0.85)	0.71 (0.55-0.93)	0.77 (0.69-0.85)	0.81 (0.70-0.93)
GP prescribed dispensings	0.69 (0.50-0.93)	0.71 (0.52-0.99)	0.83 (0.71-0.96)	0.81 (0.68-0.96)
Specialist prescribed dispensings	0.55 (0.37-0.83)	0.52 (0.33-0.82)	0.59 (0.47-0.75)	0.65 (0.51-0.84)
Gastrointestinal bleeding				
All dispensings	0.98 (0.85-1.14)	0.89 (0.61-1.30)	0.98 (0.85-1.14)	0.96 (0.78-1.17)
GP prescribed dispensings	0.84 (0.54-1.32)	0.91 (0.57-1.45)	0.99 (0.8-1.24)	1.03 (0.81-1.31)
Specialist prescribed dispensings	0.77 (0.45-1.32)	0.58 (0.30-1.12)	0.63 (0.45-0.89)	0.71 (0.49-1.02)
Intracranial bleeding				
All dispensings	0.59 (0.46-0.75)	0.66 (0.36-1.21)	0.59 (0.46-0.75)	0.72 (0.52-1.00)
GP prescribed dispensings	0.51 (0.24-1.08)	0.61 (0.29-1.30)	0.73 (0.51-1.05)	0.68 (0.45-1.00)
Specialist prescribed dispensings	0.46 (0.17-1.26)	0.57 (0.21-1.56)	0.59 (0.35-1.02)	0.58 (0.32-1.06)
Stroke/TIA				
All dispensings	0.77 (0.59-1.01)	0.75 (0.55-1.02)	0.93 (0.81-1.07)	0.99 (0.85-1.16)
GP prescribed dispensings	0.92 (0.67-1.26)	0.80 (0.55-0.12)	1.09 (0.92-1.29)	1.04 (0.87-1.26)
Specialist prescribed dispensings	0.56 (0.34-0.93)	0.49 (0.27-0.88)	0.76 (0.59-0.98)	0.75 (0.57-0.99)

3.4

DRUG EXPOSURE MISCLASSIFICATION IN PHARMACOEPIDEMIOLOGY: SOURCES AND RELATIVE IMPACT

Mirjam Hempenius, Rolf H. H. Groenwold, Anthonius de Boer, Olaf H. Klungel, Helga Gardarsdottir

Pharmacoepidemiol Drug Saf. 2021 Dec;30(12):1703-1715

ABSTRACT

Background

Drug exposure assessment based on dispensing data can be misclassified when patients do not adhere to their therapy or when information about over-the-counter drugs is not captured in the study database. Previous research has considered hypothetical sensitivity and specificity values, whereas this study aims to assess the impact of literature-based real values of exposure misclassification.

Methods

A synthetic cohort study was constructed based on the proportion of exposure theoretically captured in a database (range 0.5–1.0) and the level of adherence (0.5–1.0). Three scenarios were explored: nondifferential misclassification, differential misclassification (misclassifications dependent on an unmeasured risk factor doubling the outcome risk), and nondifferential misclassification in a comparative effectiveness study (RR_A and RR_B both 2.0 compared to nonuse, RR_{A-B} 1.0).

Results

For the scenarios with nondifferential misclassification, 25% nonadherence or 25% uncaptured exposure changed the RR from 2.0 to 1.75, and 1.95, respectively. Applying different proportions of nonadherence or uncaptured use (20% vs. 40%) for subgroups with and without the risk factor, an RR of 0.95 was observed in the absence of a true effect (i.e., true RR = 1). In the comparative effectiveness study, no effect on RR was seen for different proportions of uncaptured exposure; however, different levels of nonadherence for the drugs (20% vs. 40%) led to an underestimation of RR_{A-B} (0.89).

Discussion

All scenarios led to biased estimates, but the magnitude of the bias differed across scenarios. When testing the robustness of findings of pharmacoepidemiologic studies, we recommend using realistic values of nonadherence and uncaptured exposure based on real-world data.

1. INTRODUCTION

Observational studies on the safety and effectiveness of pharmacological agents are commonly performed using routinely collected data from administrative or healthcare databases. Examples include healthcare insurance databases, out-patient pharmacy databases, and general practitioner (GP) databases. Information about drug exposure retrieved from these databases can usually only serve as a proxy for actual use (i.e., the patient ingesting the drug). Therefore, pharmacoepidemiologic research conducted using these databases is prone to exposure misclassification.

The extent and nature of exposure misclassification differs per drug and per type of database that is used (see Table 1 and Figure 1). On the one hand, subjects may be misclassified as exposed to a specific drug based on a prescription or dispensing record in the database, when in fact they do not collect or administer the drug (nonadherence).[1] For example, nonadherence to antidepressants is estimated between 10% and 35%.[2–7] On the other hand, subjects can be misclassified as nonexposed when information about the exposure is not captured in the database.[1,8,9] This type of misclassification can occur for over-the-counter (OTC) drugs, drug samples, drugs with a restrictive drug coverage policy, use of drugs that were originally prescribed to someone else, or use of drugs that are prescribed in a clinical setting that is not captured in the database being used. The sources and extent of uncaptured exposure depend on the drug being studied and the database that is being used for a study, as described in Table 1 and Figure 1.

TABLE 1. Sources of exposure misclassification in the different databases

	Single prescriber database	**Pharmacy dispensing database**	**Claims database**
Exposed misclassified as unexposed	Drug bought as OTC, without prescription	Drug bought as OTC	Drug bought as OTC
	Drug sample	Drug sample	Drug sample
	Use of drugs that were originally prescribed to someone else	Use of drugs that were originally prescribed to someone else	Use of drugs that were originally prescribed to someone else
	Drug prescribed by another prescriber		Drug not reimbursed
Unexposed misclassified as exposed	Drug not collected at pharmacy	Drug collected, but not ingested	Drug collected, but not ingested
	Drug collected, but not ingested		

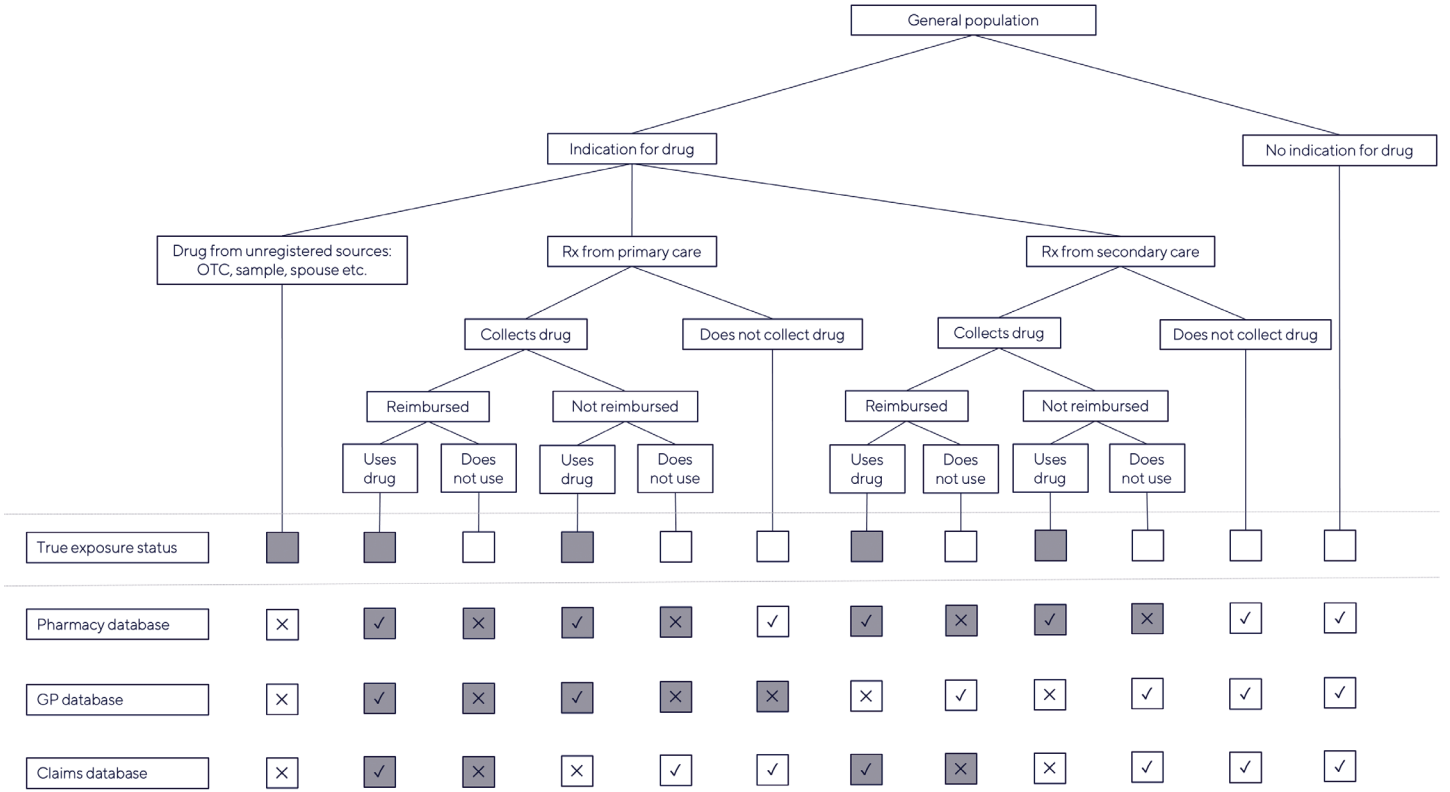


FIGURE 1. Sources of exposure misclassification in pharmacoepidemiology and the effects in different types of databases. Gray squares indicate "exposed"; white squares indicate "unexposed". Rx, prescription; X, misclassified; √, correctly classified.

Reporting guidelines for pharmacoepidemiologic studies indicate that exposure misclassification should always be discussed.[10,11] Although this is generally the case, the quantification of the potential impact of exposure misclassification is uncommon.[12] This is problematic, since nondifferential misclassification of binary exposure variables leads to bias toward the null and may lead to associations not being detected, especially if the effect under study is small. In addition, misclassification can be associated with patient characteristics, such as age,[2,4,5,13–18] sex,[4,13,19] socio-economic status,[3,5,13,15,16] and medical burden[16,19]—characteristics that are often also related to the risk of the outcome. Since this could lead to differential exposure misclassification, thus causing bias toward or away from the null, the potential impact of such misclassification is not trivial.

Key measures to quantify misclassification are sensitivity and specificity.[20] Sensitivity is calculated as the proportion of exposed subjects who are classified as being exposed: True positive/(True positive + False Negative). Specificity is defined as the proportion of unexposed subjects who are classified as being unexposed: True Negative/(True Negative + False Positive). The effect of uncaptured exposure and nonadherence on sensitivity and specificity is illustrated with a numerical example in Figure 2.

Sensitivity is directly related to the proportion of exposure that is captured; an 80% captured exposure equals a sensitivity of 0.8. The value of specificity is affected by both nonadherence and exposure prevalence. A lower exposure prevalence will result in a higher proportion of truly unexposed subjects and thus a higher specificity. For example, 20% nonadherence to a drug with 10% prevalence results in a specificity of 0.987, while the specificity decreases to 0.867 when the exposure prevalence is 50%. On the other hand, in a situation of 10% exposure prevalence, 40% nonadherence results in a specificity of 0.966, compared to 0.987 for 20% nonadherence—both specificity values are high but relate to large differences in adherence rates (Figure 2).

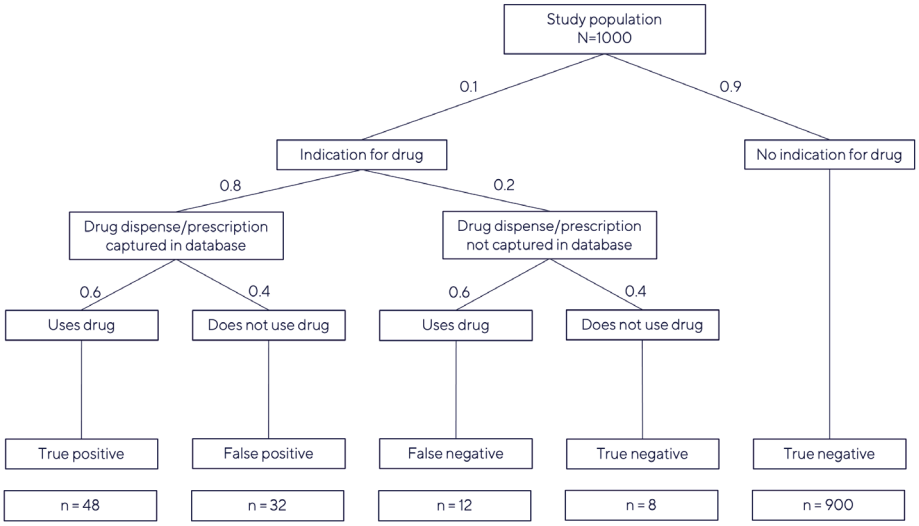


FIGURE 2. Model used for simulation analysis, with values for exposure prevalence (10%), nonadherence (40%) and uncaptured data (20%) and the corresponding exposure status. Sensitivity is in this example: True Positive/(True Positive + False Negative) = 48/(48 + 12) = 0.80; Specificity is True Negative/(True Negative + False Positive) = 908/(908 + 32) = 0.966

It is therefore important to substantiate the values for sensitivity and specificity with known values of exposure prevalence, nonadherence, and uncaptured exposure to apply realistic scenarios in assessing the impact of exposure misclassification. Small deviations in specificity can imply large differences in adherence when the exposure prevalence is low.

In pharmacoepidemiology, to date, research into the impact of exposure misclassification on effect estimates has focused on individual sources of misclassification, such as nonadherence or reimbursement status,[21] or applied hypothetical values to sensitivity and specificity that are not always supported by real data regarding adherence and the proportion of exposure that is captured in the study database.[20,22–25] This study therefore aims to assess the impact of literature-based realistic values of nonadherence and uncaptured use in simulated data, to investigate the relative impact of both sources of exposure misclassification.

2. METHODS

We constructed synthetic datasets of patient cohorts based on predefined exposure prevalence, the proportion of exposure that is captured in the database, and the level of adherence. Patients were divided into four different groups based on their exposure classification: true positive (observed definition as "exposed" is correct), true negative (observed definition as "nonexposed" is correct), false positive (observed definition as "exposed" is incorrect due to nonadherence), and false negative (observed definition as "nonexposed" is incorrect when information about exposure is not captured in database). Outcomes were subsequently assigned as a function of the baseline risk and the relative risk of exposure based on the actual exposure status. Observed relative risks were calculated based on the observed exposure status.

We explored the impact of nonadherence and uncaptured data in three scenarios: nondifferential exposure misclassification, differential exposure misclassification, and nondifferential exposure misclassification when comparing two drugs. We then applied this to two real-world examples to further understand the impact of the different sources of exposure misclassification. Details of these scenarios are described below.

3. CONCEPTUAL SCENARIOS

3.1. Nondifferential exposure misclassification

In the first scenario, we investigated the extent to which nondifferential exposure misclassification could cause bias toward the null. In this scenario, exposure to Drug A was compared with nonexposure. Different levels of nonadherence (0.10, 0.25, and 0.50) and uncaptured exposure (0.10, 0.25, and 0.50) were applied, both separately and in combination. These values were chosen based on the range of values for nonadherence and uncaptured data found in the literature (Table 2). Different levels of true exposure prevalence were used (pr_{true} 0.01, 0.10, and 0.25), again based on the values described in the literature. The observed exposure prevalence (pr_{obs}) was calculated to achieve this true exposure prevalence, accounting for the level of adherence ($pr_{obs} = pr_{true}$/adherence). A baseline risk of 0.1 of the outcome and relative risks of 1.25, 2.0, and 5.0 of the exposure effect were investigated, and observed relative risks were calculated. The percentage bias was calculated as follows: $\%bias = [\log(RR_{obs}) - \log(RR_{true})]/\log(RR_{true}) \times 100\%$. In addition, the sensitivity and specificity of the exposure assessment were also calculated based on both the true and the observed exposure statuses.

TABLE 2. Basic parameters for the two scenarios

	NSAIDs values (literature reference)		**Antidepressant agent values (literature reference)**	
Proportion of general population receiving prescription	0.25 (0.04-0.58) [26-32]		0.10 (0.05-0.20) [26,33]	
Of which from GP	0.85 [34]		0.85 (0.75-0.90) [18,35-38]	
Proportion filling prescription	0.95 (0.91-0.95) [7,39,40]		0.80 (0.65-0.95) [2-7]	
Proportion actual starts using drug[a]	0.95 [40]		0.80 (0.60-0.80) [6,16,18,35,41]	
Proportion users that buy drug OTC	0.5 (0.5-0.9) [29,30,32,42]		NA	
Baseline risk on (gastrointestinal) bleeding	0.01 [43] (10-year risk)		0.025 [44]	
Observed relative risk	3.5 (2.5-4.5) [43]		1.4 [44]	
Reimbursement	Only on prescription [45,46]		Full	
Differential misclassification (old vs. young)	Old: Baseline risk: 0.02 [43] Captured: 0.75 [42]	Young: Baseline risk: 0.01 [43] Captured: 0.25 [42]	Old: Baseline risk: 0.05 Adherence: 0.80 [4,5]	Young: Baseline risk: 0.025 [44] Adherence 0.70 [4,5]
Comparative effectiveness (drug A vs drug B)	Meloxicam: Relative risk: 4.0 [47] Captured: 0.85 [34]	Diclofenac: Relative risk: 4.0 [47] Captured: 0.50 [29,30,32,42]	Escitalopram: Relative risk: 1.5 [44] Adherence: 0.80 [19]	Paroxetine: Relative risk: 1.5 [44] Adherence: 0.60 [19]

[a] The percentages found in these studies are predominantly defined as having only one prescription dispensed. The numbers from these studies comprise thus both patients that do not initiate the use and these that discontinue the use early.

3.2. Differential exposure misclassification

In the second scenario, we investigated the extent to which differential exposure misclassification could cause bias away from the null. For this scenario, it was assumed that the exposure did not influence the risk of the outcome (RR_{true} 1.0), but that the presence of a binary risk factor had an impact on both the amount of exposure misclassification (i.e., the level of nonadherence and uncaptured data) and the risk of the outcome (RR 1.5 and 2.0). This binary risk factor was present in 50% of all subjects.

Exposure to Drug A was compared with nonexposure, the exposure prevalence (pr_{true}) was 0.1, and the baseline risk of the outcome was 0.1. Differences in the level of nonadherence and the proportion of uncaptured prescriptions between subjects with and without the risk factor that would result in an observed relative risk of 0.80, 0.90, 1.10, or 1.25 were plotted.

3.3. Comparative effectiveness research (CER): Drug A versus Drug B

In the third scenario, we examined the extent to which differences in the degree of nondifferential exposure misclassification between two study drugs could cause bias away from the null. In this scenario, exposure to Drug A was compared with exposure to Drug B. Both drugs were considered to increase the risk of the outcome compared to nonuse (either both RR 1.5 and both RR 2.0, with a baseline risk of 0.1), resulting in an RR_{A-B} of 1.0. The exposure misclassification was considered nondifferential, but different levels of adherence and the proportion of prescriptions that were captured were applied for Drugs A and B. Nonadherence to Drugs A or B would place individuals in the nonuser category, not in the other category of exposure. The exposure prevalence (pr_{true}) was 0.1 for both drugs.

Differences in the levels of nonadherence between Drug A and Drug B that would result in an observed relative risk of 0.80, 0.90, 1.10, or 1.25 were plotted. This was also done for differences in the proportion of uncaptured prescriptions.

4. APPLICATION IN TWO CASE STUDIES

In addition to the conceptual scenarios, two real-life examples were investigated (Table 2).

4.1. Nonsteroidal anti-inflammatory drugs (NSAIDs) and the risk of gastrointestinal bleeding

The first example focused on the relation between exposure to NSAIDs and the risk of gastrointestinal bleeding. The baseline risk of gastrointestinal bleeding is 0.01 per 10 person-years.[43] NSAIDs can, however, damage the protective gastric mucus layer via different mechanisms, thereby increasing the risk of gastrointestinal bleeding,[48] which occurs most often immediately after administration.[49] Adherence to NSAIDs is usually quite high (~95%), since patients take it for symptom relief.[7,39,40]

In most countries, some NSAIDs are only accessible through a prescription, while other NSAIDs are available OTC. In the Netherlands, for example, meloxicam is only available through a prescription, whereas diclofenac is available OTC. In the case of OTC NSAIDs, approximately 50% of their use

is without a prescription.[29,30,32,42] OTC use of NSAIDs varies for different age categories: 75% of younger subjects (18–20 years) obtain their NSAIDs without a prescription (i.e., OTC), compared to 25% in those aged 65 years.[42] In addition, the risk of a gastrointestinal ulcer increases with age and is twice as high for subjects aged 75 years or older, compared to younger subjects.[50] The relative risk of gastrointestinal bleeding from meloxicam and diclofenac is comparable (RR ~4.0).[47]

4.2 Selective serotonin reuptake inhibitors (SSRIs) and the risk of bleeding

The second example concerned the relation between exposure to SSRIs and the risk of severe bleeding. The baseline risk of severe bleeding is about 0.025.[44] SSRIs inhibit the platelet serotonin transporter, causing platelets to release less serotonin and hindering the vasoconstriction and aggregation of platelets,[51] resulting in approximately a 1.5 times higher risk of bleeding.[44] SSRIs and other antidepressant drugs are prescription-only drugs, predominantly prescribed by GPs, although they can be prescribed by specialists as well.[18,35–38] Nonadherence is known to be quite high for antidepressant drugs, with ~20% not filling in the first prescription.[2–7] In addition, even when patients do fill their prescription, a large proportion of them do not initiate treatment.[6,16,18,35,41] The level of nonadherence can differ between the individual SSRIs. For this case study, we assumed nonadherence to be twice as high for paroxetine as compared to escitalopram.[19] The level of nonadherence also differs between different age categories and is roughly 1.5 times higher in younger subjects (<=65 years) than those >65 years.[4,5] As mentioned before, the risk of bleeding is increased in older subjects (RR 2.0).[50]

For both examples, we calculated the underlying relative risk that would generate the observed relative risk in case of nondifferential misclassification, given the known numbers for uncaptured exposure and nonadherence (Table 2). Then, we compared meloxicam and diclofenac with a different proportion of captured exposure, and we compared escitalopram with paroxetine with different levels of adherence. Finally, we divided the cohort into two groups, namely, "old" and "young," with different levels of uncaptured exposure and nonadherence and different risks of the outcome, and we calculated crude relative risks with and without correcting for the age effect.

5. RESULTS

5.1. Nondifferential exposure misclassification

The results of the analysis with nondifferential exposure misclassification are presented in Table 3. Nonadherence generally had a greater impact on RR than uncaptured exposure. For example, for a drug with a prevalence of 0.1, applying 25% nondifferential nonadherence to our model changed the RR from 2.0 to 1.75 (% deviation –19.3% on log[RR] scale) while applying 25% nondifferential uncaptured exposure changed the RR from 2.0 to 1.95 (–3.9%). With increasing prevalence of exposure, the effect of uncaptured exposure did, however, increase, while the effect of nonadherence did not. For exposure with a prevalence greater than 40%–50%, the effect of uncaptured data was greater than the effect of nonadherence (Figure 3). The largest effect was observed for the scenario with an exposure prevalence of 25%, 50% nonadherence, 50% uncaptured exposure, and a relative risk of 5.0. In this scenario, the relative risk changed to 1.8—a decrease of 65.5%. In the Supplementary materials, more extensive tables are presented (Tables S1a-d), detailing the impact of different values for nonadherence and uncaptured data for different exposure prevalences.

TABLE 3. The effect of nondifferential exposure misclassification due to data that is uncaptured or nonadherence on the effect estimates

					RR 1.25		RR 2.0		RR 5.0	
True exp prevalence	Uncaptured	Nonadherence	Sens	Spec	RR obs	% dev	RR obs	% dev	RR obs	% dev
0.01	0.1	1	0.9	1.00	1.25	-0.13	2.00	-0.20	4.98	-0.50
0.1	0.1	1	0.9	1.00	1.25	-1.37	1.98	-2.20	4.79	-5.28
0.25	0.1	1	0.9	1.00	1.24	-4.00	1.94	-6.25	4.43	-14.29
0.01	0.25	1	0.75	1.00	1.25	-0.40	2.00	-0.50	4.95	-1.25
0.1	0.25	1	0.75	1.00	1.24	-3.36	1.95	-5.30	4.51	-12.20
0.25	0.25	1	0.75	1.00	1.23	-9.43	1.86	-14.29	3.82	-29.41
0.01	0.5	1	0.5	1.00	1.25	-0.80	1.99	-1.00	4.90	-2.50
0.1	0.5	1	0.5	1.00	1.23	-6.40	1.90	-10.00	4.13	-21.75
0.25	0.5	1	0.5	1.00	1.21	-17.24	1.75	-25.00	3.18	-45.45
0.01	1	0.1	1	1.00	1.23	-10.00	1.90	-10.00	4.60	-10.00
0.1	1	0.1	1	0.99	1.23	-10.00	1.90	-10.00	4.60	-10.00
0.25	1	0.1	1	0.97	1.23	-10.00	1.90	-10.00	4.60	-10.00
0.01	1	0.25	1	1.00	1.19	-25.00	1.75	-25.00	4.00	-25.00
0.1	1	0.25	1	0.97	1.19	-25.00	1.75	-25.00	4.00	-25.00
0.25	1	0.25	1	0.92	1.19	-25.00	1.75	-25.00	4.00	-25.00

TABLE 3. Continued.

					RR 1.25		RR 2.0		RR 5.0	
True exp prevalence	Uncaptured	Nonadherence	Sens	Spec	RR obs	% dev	RR obs	% dev	RR obs	% dev
0.01	1	0.5	1	0.99	1.13	-50.00	1.50	-50.00	3.00	-50.00
0.1	1	0.5	1	0.89	1.13	-50.00	1.50	-50.00	3.00	-50.00
0.25	1	0.5	1	0.67	1.13	-50.00	1.50	-50.00	3.00	-50.00
0.01	0.1	0.1	0.1	1.00	1.23	-10.00	1.90	-10.19	4.58	-10.46
0.1	0.1	0.1	0.1	0.99	1.22	-11.20	1.88	-12.09	4.40	-14.89
0.25	0.1	0.1	0.1	0.97	1.21	-14.05	1.84	-16.13	4.06	-23.53
0.01	0.25	0.25	0.25	1.00	1.19	-25.20	1.75	-25.44	3.96	-26.00
0.1	0.25	0.25	0.25	0.97	1.18	-28.28	1.70	-29.73	3.60	-35.00
0.25	0.25	0.25	0.25	0.92	1.16	-34.69	1.62	-38.46	3.00	-50.00
0.01	0.5	0.5	0.5	0.99	1.12	-50.57	1.49	-50.75	2.94	-51.49
0.1	0.5	0.5	0.5	0.94	1.11	-56.16	1.42	-57.89	2.45	-63.64
0.25	0.5	0.5	0.5	0.83	1.08	-68.00	1.29	-71.40	1.80	-80.00

Applying the scenario of nondifferential exposure misclassification to the example of NSAIDs and the risk of bleeding, assuming 50% uncaptured exposure and 5% nonadherence, we found that when an RR of 3.5 was observed, the true RR was 6.3 (–32.0%). For the case study of SSRIs, we found that when an RR of 1.40 was observed, the true RR was 1.51 (–18.4%), assuming 20% uncaptured exposure, and 20% nonadherence.

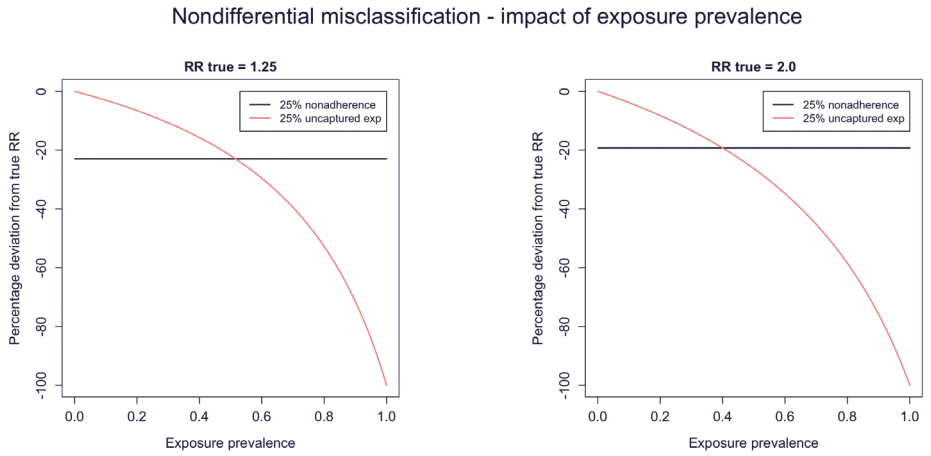


FIGURE 3. Percentage deviation from the true RR with 25% nonadherence or 25% uncaptured data against different exposure prevalences

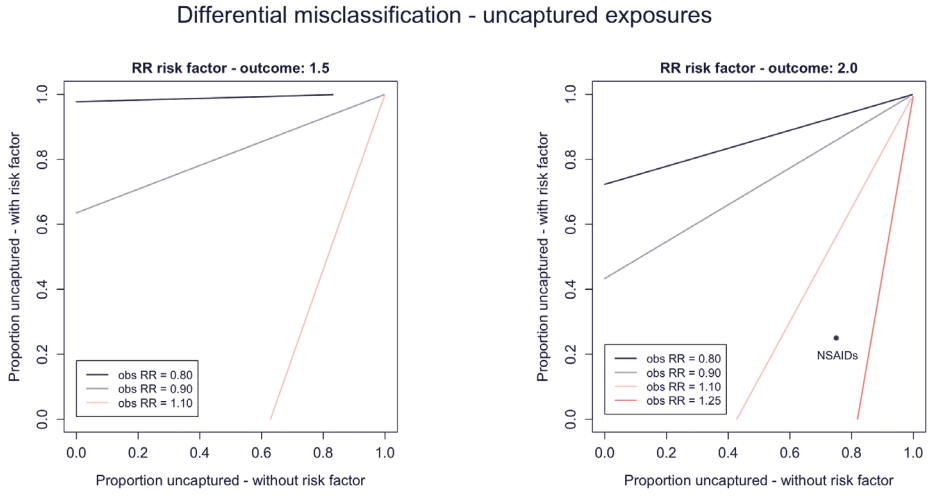


FIGURE 4. Observed relative risks obtained with different proportion of exposure captured for subjects with and without risk factor, in absence of a true effect ($RR_{true} = 1.0$). Observed relative risk for differential misclassification to NSAIDs caused by age was 1.18, assuming a relation of RR 2.0 for older subjects compared to younger subjects, 25% captured exposures for younger subjects, and 75% captured exposures for older subjects

5.2. Differential exposure misclassification

Figure 4 illustrates the different proportions of uncaptured exposure for subjects with or without the risk factor required to observe an RR of 0.80, 0.90, 1.10, or 1.25 in the absence of a true relationship between exposure and outcome (RR_{true} = 1.0). For example, if a risk factor increases the risk of the outcome by a factor of 2, and if 90% and 50% of the exposure for subjects with and without this risk factor, respectively, were recorded in the database, then the resulting observed RR was 1.10. Moreover, with 80% and 60% captured exposure, respectively, an RR of 1.05 would have been observed. If the risk factor instead increased the risk of the outcome by a factor of 1.5, then an RR of 1.03 would have been observed.

The results for the different levels of adherence are depicted in Figure 5. Approximately the same patterns were found for different levels of adherence for subjects with and without the risk factor: with 50% adherence for subjects with the risk factor and 90% adherence for subjects without this risk factor, the resulting observed RR was found to be 1.12. If the risk factor had a stronger effect on the outcome, then the effects were more pronounced. Stratification on the risk factor removed the effect of the differential misclassification in both situations. In the Supplementary materials, more figures are presented, illustrating the impact of differential exposure misclassification with different proportions of subjects with the risk factor and different relative risks between the risk factor and the outcome.

In case of a relation between NSAIDs and gastrointestinal bleeding, the risk factor "age" was considered to increase the risk of the outcome by a factor of 2, and the proportion of captured exposure was 75% and 25% for the "old" and "young" subjects. In this case, a relative risk of 1.18 could have been observed instead of 1.0, when no correction for this risk factor would have been applied (Figure 4). Stratification on age resulted in a relative risk of 1.0 in both subgroups.

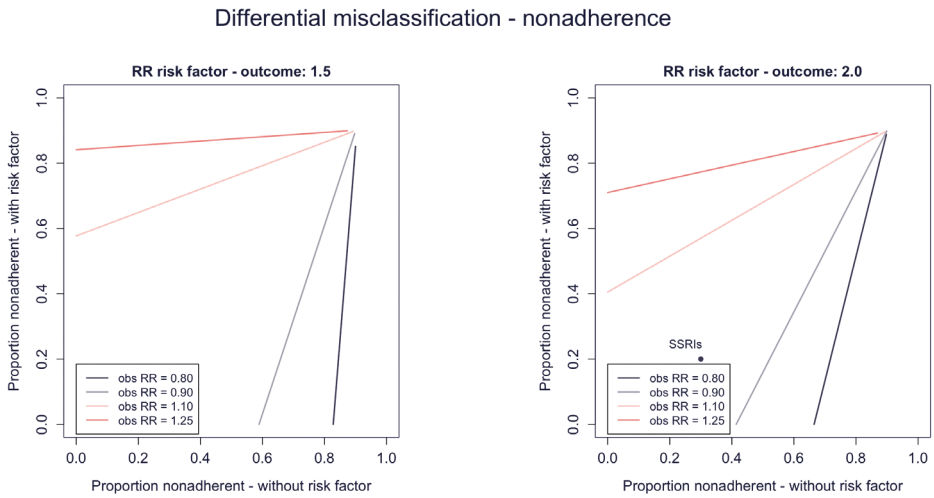


FIGURE 5. Observed relative risks obtained with different levels of nonadherence for subjects with and without confounding risk factor, in absence of a true effect (RR_{true} = 1.0). Observed relative risk for differential misclassification to SSRIs caused by age was 1.05, assuming a relation of RR 2.0 for older subjects compared to younger subjects, 30% nonadherence for younger subjects, and 20% nonadherence for older subjects

For the case study of SSRIs, different levels of adherence were applied for the "old" and "young" subjects (80% vs. 70% respectively). When the risk factor "age" was again considered to increase the risk of the outcome by a factor of 2, there was only a small deviation from the true effect (RR 0.97 instead of 1.0; Figure 5). Again, stratification on age resulted in a relative risk of 1.0 in both subgroups.

5.3 CER

When two drugs were compared with each other, no effect of different levels of captured exposure was seen, as this resulted in sampling of all exposed subjects. As long as this occurred randomly, the risks remained the same, as did the risk ratio. This is illustrated in Table 4, with the case study of meloxicam and diclofenac.

Differences in levels of adherence, however, did generate RRs deviating from 1.0, in the absence of a difference between Drug A and Drug B (Table 5). The different adherence rates required to observe an RR of 0.80, 0.90, 1.10, or 1.25 are shown in Figure 6. For example, 80% and 64% adherence for Drugs A and B, both with an RR of 2.0 with the outcome, resulted in an observed RR_{A-B} of 1.10. Applying this to the comparison between escitalopram (80% adherence) and paroxetine (60% adherence), both with an RR of 1.5 with the outcome, an RR of 1.08 could have been observed when comparing escitalopram to paroxetine. Additional figures are presented in Figure S3 for scenarios where both drugs had a stronger relation with the outcome (RR 5.0 and 10.0).

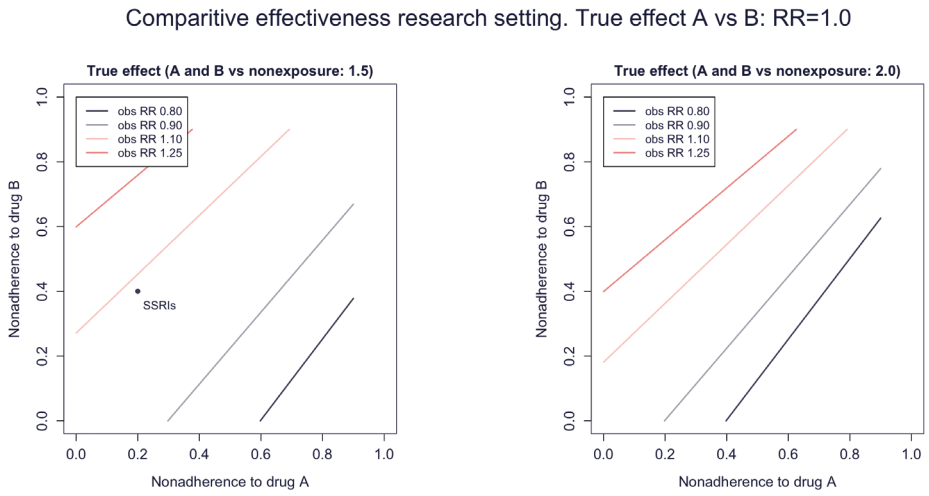


FIGURE 6. Observed relative risks obtained with different adherences rates of drug A and B, in absence of a true effect (RR_{true} = 1.0). Observed relative risk for different levels of adherence to escitalopram (80%) compared paroxetine (60%) was 1.08, assuming a relation of RR 1.5 of both antidepressant agents with the outcome

TABLE 4. Impact of different values of uncaptured data in comparative effectiveness research

	Truth		Observed data	
Diclofenac use versus nonusers	**Drug D1**	**Drug D0**	**Drug D1**	**Drug D0**
Y = 1	6000	8500	3000	11500
Y = 0	9000	76500	4500	81000
Total	15000	85000	7500	92500
Risk	0.4	0.1	0.4	0.124
RR	4.0		3.2	
Meloxicam versus nonusers	**Drug M1**	**Drug M0**	**Drug M1**	**Drug M0**
Y = 1	200	9950	200	9950
Y = 0	300	89550	300	89550
Total	500	99500	500	99500
Risk	0.4	0.1	0.4	0.1
RR	4.0		4.0	
Diclofenac versus meloxicam	**Drug D1**	**DrugM1**	**Drug D1**	**DrugM1**
Y = 1	6000	200	3000	200
Y = 0	9000	300	4500	300
Total	15000	500	7500	500
Risk	0.4	0.4	0.4	0.4
RR_{D-M}	1.0		1.0	

Note: Values used: baseline risk: 0.1; RR: diclofenac 4.0, meloxicam 4.0; exposure prevalence: diclofenac 0.10, meloxicam 0.005; data capture: diclofenac 0.5, meloxicam 1.0; adherence: diclofenac 1.0, meloxicam 1.0.
Abbreviations: D1, exposed to diclofenac; M1, exposed to meloxicam; P0 and E0, nonexposed to diclofenac or meloxicam; Y, outcome.

TABLE 5. Impact of different values of nonadherence in comparative effectiveness research

	Truth		Observed data	
Paroxetine use versus nonusers	**Drug P1**	**Drug P0**	**Drug P1**	**Drug P0**
Y = 1	150	9900	217	9833
Y = 0	850	89100	1450	88500
Total	1000	99000	1667	98333
Risk	0.15	0.10	0.13	0.10
RR	1.5		1.4	
Escitalopram versus nonusers	**Drug E1**	**Drug E0**	**Drug E1**	**Drug E0**
Y = 1	75	9950	87.5	9937.5
Y = 0	425	89550	537.5	89437.5
Total	500	99500	625	99375
Risk	0.15	0.10	0.14	0.10
RR	1.5		1.8	
Escitalopram versus paroxetine	**Drug P1**	**Drug E1**	**Drug P1**	**Drug E1**
Y = 1	150	75	217	87.5
Y = 0	850	425	145	537.5
Total	1000	500	1667	625
Risk	0.15	0.15	0.13	0.14
RR_{E-P}	1.0		1.08	

Note: Values used: baseline risk: 0.1; RR: paroxetine 1.5, escitalopram 1.5; exposure prevalence: paroxetine 0.01, escitalopram 0.005; data capture: paroxetine 1.0, escitalopram 1.0; adherence: paroxetine 0.6, escitalopram 0.8.
Abbreviations: P1: exposed to paroxetine; E1 exposed to escitalopram; P0 and E0 nonexposed to paroxetine or escitalopram; Y = outcome.

6. DISCUSSION AND CONCLUSION

We studied the impact of a range of different values for nonadherence and uncaptured exposure to understand the relative impact of those two sources of exposure misclassification. Among the scenarios considered, we found that for exposure with a prevalence of less than 40%–50%, nonadherence had a greater impact on the RR than uncaptured exposure. To put this in context, in pharmacoepidemiology, the exposure prevalence for most drugs is <10%, unless studies are restricted to those with an indication for the drug, such as exposure to antidepressants within patients diagnosed with depression.

For an exposure with a prevalence of 10%, 25% nondifferential nonadherence changed the RR from 2.0 to 1.75, while applying 25% nondifferential uncaptured exposure changed the RR from 2.0 to 1.95. A substantial degree of nonadherence can therefore lead to associations

being missed, especially if the effect under study is small. Applying nondifferential exposure misclassification to the examples of NSAIDs and SSRIs and the risk of bleeding, we demonstrated that an attenuation of ±20%–30% of the true relative risk can be expected using the values for nonadherence and uncaptured data of antidepressant drugs and NSAIDs, as shown in Table 2. However, these percentages of attenuation are not fixed values, but an example of the degree of bias that can be expected. A range of values has been described in the literature for the degree of nonadherence and uncaptured data, and we used one of many possible combinations. In addition, these scenarios may turn out differently for different databases, as there are varying reasons per database why exposure status can be misclassified (Table 1).

The impact of uncaptured exposure was dependent on exposure prevalence, since uncaptured exposure changed the observed risk of the unexposed group without changing the observed risk among the exposed. The larger the group of truly unexposed was, the smaller the effect of uncaptured exposure was. This was not seen for the effect on nonadherence: in this case, the observed risk of the exposed subjects was changed by misclassifying unexposed subjects as exposed, but nonadherence had no impact on the observed risk of the unexposed. Therefore, the effect of nonadherence was not impacted by exposure prevalence.

Studying the effect of differential misclassification, we found in this simulation RRs deviating away from the null. However, the differences in captured data or adherence between drug users with and without a risk factor with a relative risk of 2.0 with the outcome needed to be large (e.g., 50% vs. 90%) to result in a clinically relevant deviation from the null (arbitrarily set at RR_{obs} 1.10). This has also been demonstrated in the NSAID case and differential misclassification caused by age. In this specific example, the bias can be removed because age is often corrected for in the analysis. However, there are also examples of unmeasured risk factors, such as smoking status, which can lead to biased results if this risk factor is related to both the outcome and the risk of exposure misclassification. In addition, in the studied scenarios, the risk factor was present in 50% of all subjects. However, with a different distribution of the risk factor (e.g., 10% or 90%), the effect of the differential exposure misclassification was even smaller, and the differences between subjects with and without the risk factor needed to be larger to result in a clinically relevant deviation (Figures S1 and S2).

In a comparative study of Drug A versus Drug B, the proportion of uncaptured drug exposure (nondifferential) had no impact on the effect estimates, since including only the captured exposure involved the same process as random sampling, as long as the misclassification due to uncaptured data was nondifferential. Different levels of adherence between Drugs A and B could lead to the estimates of Drug A versus Drug B deviating from the null in the absence of a true different effect. In this case, however, the exposure definition is not dichotomous, but polytomous (exposed to A, B, or none), and it has already been shown that nondifferential misclassification of a polytomous exposure can cause bias away or toward the null.[52] In addition, the differences in adherence needed to be large (e.g., 80% and 64%) to result in a clinically relevant deviation from the null effect (RR_{obs} 1.10), when both drugs had an RR of 2.0 on the outcome, or 80% and 55% when both drugs had an RR of 1.5 on the outcome.

These conclusions are in contrast to prior literature, which has demonstrated that small violations of the assumption of misclassification being nondifferential or differences in misclassification

between Drugs A and B could already result in clinically relevant deviations from the null effect.[22,24,53] In these previous studies, misclassification was introduced by choosing different values for specificity and sensitivity, while we focused on values for nonadherence and uncaptured data. For example, Brenner used a sensitivity of 0.9 and a specificity of 0.9 for the exposure measurement, with an exposure prevalence of 0.01, 0.1, and 0.5. However, the degree of nonadherence required to result in these values of specificity is 91.5%, 50%, and 10% for the different exposure prevalences, respectively. Since most drugs have an exposure prevalence of up to 10%, and nonadherence is often <40%, we considered a specificity of 0.9 to be unlikely for current pharmacoepidemiologic database studies. In the study by Jonsson-Funk and Landi, chosen values for the misclassification of Drug A usage versus nonuse were a sensitivity of 0.85 and a specificity 0.95 (exposure prevalence 0.17), and for Drug B usage versus nonuse, the values were a sensitivity of 0.90 and a specificity of 0.98 (exposure prevalence 0.02). The RR observed for this scenario was 1.20 instead of a true effect of 1.0. In our study, these values translate to 23% and 53% nonadherence, respectively; hence, nonadherence to Drug B was 2.3 times higher compared to Drug A. The strength of our study was that we used literature-based values for nonadherence and uncaptured data, which helps to contextualize the results and enables other researchers to apply these values in their own research.

Another strength of this study was that we were able to examine the effects of both uncaptured exposure and nonadherence in one model. This provided insight into which source of misclassification of exposure has the greatest impact. A limitation of this study was that the model we used was a simplification of the true mechanisms causing exposure misclassification. For example, in the simple 2×2 tables, time effects were ignored in the analysis, such as the fact that subjects prone to a negative outcome quit using the drug earlier than subjects who tolerated the drug better. We also ignored the fact that dosages and associated risks could differ between captured and uncaptured exposure, which is the case, for example, for prescription NSAID use versus OTC use and the risk of bleeding.[54,55]

To conclude, in all scenarios studied, the values for nonadherence and uncaptured data or the differences in these values between subgroups needed to be relatively large to lead to clinically relevant bias. With estimates of the degree of misclassification, for example from pilot studies or published results of drug utilization research, a simple bias analysis can provide insight into the impact of exposure misclassification on the effect estimates. Therefore, we provide additional tables and figures in the Supplementary materials, which can be used to assess the impact of the different sources of misclassification, using values for exposure prevalence, the proportion of nonadherence, and uncaptured data.

It should be kept in mind that scenarios may turn out differently for different databases, as depicted in Figure 1. A prescription-only drug that is fully reimbursed is likely to have nearly 100% captured exposure in a claims database, but a lower percentage in a single prescriber database, when there are multiple prescribers. Nonadherence can also have a different effect in prescribing and dispensing databases, depending on whether a subject decides not to collect the prescribed drug or decides not to ingest the drug after collecting it. We therefore recommend that authors provide estimates of the degree of exposure misclassification, instead of only a vague statement about the possibility of such misclassification, and to use the values

of nonadherence and uncaptured exposure in sensitivity analyses to test the robustness of findings.

Author's contribution

MH designed the study, conducted the simulation, wrote the first draft of the manuscript, and implemented the contribution of the co-authors and external reviewers up to final publication. During the whole process she implemented input and feedback from the other contributors to this study.

REFERENCES

1. Strom BL, Kimmel SE, Hennessy S. *Pharmacoepidemiology*. 5th ed. Wiley-Blackwell; 2011.
2. Xing S, Dipaula BA, Lee HY, Cooke CE. Failure to fill electronically prescribed antidepressant medications: a retrospective study. *Prim care companion CNS Disord*. 2011;13(1). doi:10.4088/PCC.10m00998blu
3. Hansen DG, Vach W, Rosholm J-U, Søndergaard J, Gram LF, Kragstrup J. Early discontinuation of antidepressants in general practice: association with patient and prescriber characteristics. *Fam Pract*. 2004;21(6):623-629. doi:10.1093/fampra/cmh608
4. Thengilsdóttir G, Pottegård A, Linnet K, Halldórsson M, Almarsdóttir AB, Gardarsdóttir H. Do patients initiate therapy? Primary non-adherence to statins and antidepressants in Iceland. *Int J Clin Pract*. 2015;69(5):597-603. doi:10.1111/ijcp.12558
5. Freccero C, Sundquist K, Sundquist J, Ji J. Primary adherence to antidepressant prescriptions in primary health care: a population-based study in Sweden. *Scand J Prim Health Care*. 2016;34(1):83-88. doi:10.3109/02813432.2015.1132884
6. van Geffen EC, Gardarsdottir H, van Hulten R, van Dijk L, Egberts AC, Heerdink ER. Initiation of antidepressant therapy: do patients follow the GP's prescription? *Br J Gen Pract*. 2009;59(559):81-87. doi:10.3399/bjgp09X395067
7. Pottegård A, Christensen R dePont, Houji A, et al. Primary non-adherence in general practice: a Danish register study. *Eur J Clin Pharmacol*. 2014;70(6):757-763. doi:10.1007/s00228-014-1677-y
8. Cepeda MS, Fife D, Denarié M, Bradford D, Roy S, Yuan Y. Quantification of missing prescriptions in commercial claims databases: results of a cohort study. *Pharmacoepidemiol Drug Saf*. 2017;26(4):386-392. doi:10.1002/pds.4165
9. Lauffenburger JC, Balasubramanian A, Farley JF, et al. Completeness of prescription information in US commercial claims databases. *Pharmacoepidemiol Drug Saf*. 2013;22(8):899-906. doi:10.1002/pds.3458
10. Wang S V., Schneeweiss S, Berger ML, et al. Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies V1.0. *Pharmacoepidemiol Drug Saf*. 2017;26(9):1018-1032. doi:10.1002/pds.4295
11. Langan SM, Schmidt SA, Wing K, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). *BMJ*. 2018;363:k3532. doi:10.1136/bmj.k3532
12. Jurek AM, Maldonado G, Greenland S, Church TR. Exposure-measurement error is frequently ignored when interpreting epidemiologic study results. *Eur J Epidemiol*. 2006;21(12):871-876. doi:10.1007/s10654-006-9083-0
13. Burton C, Anderson N, Wilde K, Simpson CR. Factors associated with duration of new antidepressant treatment: analysis of a large primary care database. *Br J Gen Pract*. 2012;62(595):104-112. doi:10.3399/bjgp12X625166
14. Bushnell GA, Stürmer T, White A, et al. Predicting persistence to antidepressant treatment in administrative claims data: Considering the influence of refill delays and prior persistence on other medications. *J Affect Disord*. 2016;196:138-147. doi:10.1016/j.jad.2016.02.012
15. Yau W-Y, Chan M-C, Wing Y-K, et al. Noncontinuous use of antidepressant in adults with major depressive disorders - a retrospective cohort study. *Brain Behav*. 2014;4(3):390-397. doi:10.1002/brb3.224
16. Rossom RC, Shortreed S, Coleman KJ, et al. Antidepressant adherence across diverse populations and healthcare settings. *Depress Anxiety*. 2016;33(8):765-774. doi:10.1002/da.22532
17. Krivoy A, Balicer RD, Feldman B, et al. The impact of age and gender on adherence to antidepressants: a 4-year population-based cohort study. *Psychopharmacology (Berl)*. 2015;232(18):3385-3390. doi:10.1007/s00213-015-3988-9

18. Bambauer KZ, Soumerai SB, Adams AS, Zhang F, Ross-Degnan D. Provider and patient characteristics associated with antidepressant nonadherence: the impact of provider specialty. *J Clin Psychiatry*. 2007;68(6):867-873.
19. Degli Esposti L, Piccinni C, Sangiorgi D, Fagiolini A, Buda S. Patterns of Antidepressant Use in Italy: Therapy Duration, Adherence and Switching. *Clin Drug Investig*. 2015;35(11):735-742. doi:10.1007/s40261-015-0332-4
20. Copeland KT, Checkoway H, Mcmichael AJ, Holbrook RH. Bias due to misclassification in the estimation of relative risk. *Am J Epidemiol*. 1977;105(5):488-495. doi:10.1093/oxfordjournals.aje.a112408
21. Gamble JM, McAlister FA, Johnson JA, Eurich DT. Quantifying the impact of drug exposure misclassification due to restrictive drug coverage in administrative databases: A simulation cohort study. *Value Heal*. 2012;15(1):191-197. doi:10.1016/j.jval.2011.08.005
22. Brenner H. Inferences on the potential effects of presumed nondifferential exposure misclassification. *Ann Epidemiol*. 1993;3(3):289-294. doi:10.1016/1047-2797(93)90032-Y
23. Flegal KM, Brownie C, Haas J. The effects of exposure misclassification on estimates of relatwe risk. *Am J Epidemiol*. 1986;123(4):736-751. doi:10.1093/oxfordjournals.aje.a114294
24. Jonsson Funk M, Landi SN. Misclassification in Administrative Claims Data: Quantifying the Impact on Treatment Effect Estimates. *Curr Epidemiol Reports*. 2014;1(4):175-185. doi:10.1007/s40471-014-0027-z
25. Höfler M. The effect of misclassification on the estimation of association: A review. *Int J Methods Psychiatr Res*. 2005;14(2):92-101. doi:10.1002/mpr.20
26. Zorginstituut Nederland [Healthcare Institute Netherlands]. GIP Databank. Accessed October 11, 2021. https://gipdatabank.nl
27. Fosbøl EL, Gislason GH, Jacobsen S, et al. The pattern of use of non-steroidal anti-inflammatory drugs (NSAIDs) from 1997 to 2005: A nationwide study on 4.6 million people. *Pharmacoepidemiol Drug Saf*. 2008;17(8):822-833. doi:10.1002/pds.1592
28. Zhou Y, Boudreau DM, Freedman AN. Trends in the use of aspirin and nonsteroidal anti-inflammatory drugs in the general U.S. population. *Pharmacoepidemiol Drug Saf*. 2014;23(1):43-50. doi:10.1002/pds.3463
29. Tielemans MM, van Rossum LGM, Eikendal T, et al. Gastrointestinal symptoms in NSAID users in an "average risk population": results of a large population-based study in randomly selected Dutch inhabitants. *Int J Clin Pract*. 2014;68(4):512-519. doi:10.1111/ijcp.12346
30. Wilcox CM, Cryer B, Triadafilopoulos G. Patterns of use and public perception of over-the-counter pain relievers: Focus on nonsteroidal antiinflammatory drugs. *J Rheumatol*. 2005;32(11):2218-2224.
31. Roberto G, Bartolini C, Rea F, et al. NSAIDs utilization for musculoskeletal indications in elderly patients with cerebro/cardiovascular disease. *Eur J Clin Pharmacol*. 2018;74(5):637-643. doi:10.1007/s00228-018-2411-y
32. Castelli G, Petrone A, Xiang J, Shrader C, King D. Rates of Nonsteroidal Anti-Inflammatory Drug Use in Patients with Established Cardiovascular Disease: A Retrospective, Cross-Sectional Study from NHANES 2009-2010. *Am J Cardiovasc Drugs*. 2017;17(3):243-249. doi:10.1007/s40256-016-0212-1
33. Forns J, Pottegård A, Reinders T, et al. Antidepressant use in Denmark, Germany, Spain, and Sweden between 2009 and 2014: Incidence and comorbidities of antidepressant initiators. *J Affect Disord*. 2019;249:242-252. doi:10.1016/j.jad.2019.02.010
34. Visser LE, Graatsma HH, Stricker BHC. Contraindicated NSAIDs are frequently prescribed to elderly patients with peptic ulcer disease. *Br J Clin Pharmacol*. 2002;53(2):183-188. doi:10.1046/j.0306-5251.2001.01190.x
35. Fagot J-P, Cuerq A, Samson S, Fagot-Campagna A. Cohort of one million patients initiating antidepressant treatment in France: 12-month follow-up. *Int J Clin Pract*. 2016;70(9):744-751. doi:10.1111/ijcp.12850

36. Gardarsdottir H, Egberts ACG, Heerdink ER. Transitions from General Practitioner to Psychiatrist Care (or vice versa) during a first antidepressant treatment episode. *Pharmacopsychiatry.* 2010;43(5):179-183. doi:10.1055/s-0030-1249096
37. Etchepare F, Sanglier T, André M, Verdoux H, Tournier M. Antidepressant treatment patterns in younger and older adults from the general population in a real-life setting. *Int J Geriatr Psychiatry.* 2014;29(9):928-935. doi:10.1002/gps.4081
38. Kjosavik SR, Hunskaar S, Aarsland D, Ruths S. Initial prescription of antipsychotics and antidepressants in general practice and specialist care in Norway. *Acta Psychiatr Scand.* 2011;123(6):459-465. doi:10.1111/j.1600-0447.2011.01697.x
39. Linnet K, Halldórsson M, Thengilsdóttir G, Einarsson ÓB, Jónsson K, Almarsdóttir AB. Primary non-adherence to prescribed medication in general practice: Lack of influence of moderate increases in patient copayment. *Fam Pract.* 2013;30(1):69-75. doi:10.1093/fampra/cms049
40. Lanas A, Polo-Tomás M, Roncales P, Gonzalez MA, Zapardiel J. Prescription of and adherence to non-steroidal anti-inflammatory drugs and gastroprotective agents in at-risk gastrointestinal patients. *Am J Gastroenterol.* 2012;107(5):707-714. doi:10.1038/ajg.2012.13
41. Lu CY, Roughead E. New users of antidepressant medications: first episode duration and predictors of discontinuation. *Eur J Clin Pharmacol.* 2012;68(1):65-71. doi:10.1007/s00228-011-1087-3
42. Sarganas G, Buttery AK, Zhuang W, et al. Prevalence, trends, patterns and associations of analgesic use in Germany. *BMC Pharmacol Toxicol.* 2015;16(28). doi:10.1186/s40360-015-0028-7
43. Dubois RW, Melmed GY, Henning JM, Bernal M. Risk of Upper Gastrointestinal Injury and Events in Patients Treated With Cyclooxygenase (COX)-1/COX-2 Nonsteroidal Antiinflammatory Drugs (NSAIDs), COX-2 Selective NSAIDs, and Gastroprotective Cotherapy. *JCR J Clin Rheumatol.* 2004;10(4):178-189. doi:10.1097/01.rhu.0000128851.12010.46
44. Laporte S, Chapelle C, Caillet P, et al. Bleeding risk under selective serotonin reuptake inhibitor (SSRI) antidepressants: A meta-analysis of observational studies. *Pharmacol Res.* 2017;118:19-32. doi:10.1016/j.phrs.2016.08.017
45. FDA. FDA-Approved drugs. Accessed March 12, 2021. https://www.accessdata.fda.gov/
46. CBG-MEB. Drug information bank. Drug information bank. Accessed March 12, 2021. https://db.cbg-meb.nl/
47. Massó González EL, Patrignani P, Tacconelli S, García Rodríguez LA. Variability among nonsteroidal antiinflammatory drugs in risk of upper gastrointestinal bleeding. *Arthritis Rheum.* 2010;62(6):1592-1601. doi:10.1002/art.27412
48. Jones R, Rubin G, Berenbaum F, Scheiman J. Gastrointestinal and Cardiovascular Risks of Nonsteroidal Anti-inflammatory Drugs. *Am J Med.* 2008;121(6):464-474. doi:10.1016/j.amjmed.2008.01.045
49. Graham DY, Smith JL. Aspirin and the stomach. *Ann Intern Med.* 1986;104(3):390-398. doi:10.7326/0003-4819-104-3-390
50. Hallas J, Lauritsen J, Villadsen HD, Gram LF. Nonsteroidal anti-inflammatory drugs and upper gastrointestinal bleeding, identifying high-risk groups by excess risk estimates. *Scand J Gastroenterol.* 1995;30(5):438-444. doi:10.3109/00365529509093304
51. Paton C, Ferrier IN. SSRIs and gastrointestinal bleeding. *Br Med J.* 2005;331(7516):529-530. doi:10.1136/bmj.331.7516.529
52. Brenner H. Bias due to non-differential misclassification of polytomous confounders. *J Clin Epidemiol.* 1993;46(1):57-63. doi:10.1016/0895-4356(93)90009-P
53. Jurek AM, Greenland S, Maldonado G. Brief Report: How far from non-differential does exposure or disease misclassification have to be to bias measures of association away from the null? *Int J Epidemiol.* 2008;37(2):382-385. doi:10.1093/ije/dym291
54. Baigent C, Bhala N, Emberson J, et al. Vascular and upper gastrointestinal effects of non-steroidal anti-inflammatory drugs: Meta-analyses of individual participant data from randomised trials. *Lancet.* 2013;382(9894):769-779. doi:10.1016/S0140-6736(13)60900-9
55. Cohen JS. Why aren't lower, effective, OTC doses available earlier by prescription? *Ann Pharmacother.* 2003;37(1):136-142. doi:10.1345/aph.1A487

SUPPLEMENTARY MATERIALS

TABLE S1a. Observed RR 1.10
True relative risks that would lead to an observed relative risk of 1.10 in presence of different levels of uncaptured exposure or nonadherence

Amount of misclassification		**True RR for different exposure prevalences**		
Prop nonadherence	Prop uncaptured data	p=0.01	p=0.10	p=0.25
0.0	0.0	1.10	1.10	1.10
0.0	0.1	1.10	1.10	1.10
0.0	0.2	1.10	1.10	1.11
0.0	0.3	1.10	1.10	1.11
0.0	0.4	1.10	1.10	1.11
0.0	0.5	1.10	1.11	1.12
0.1	0.0	1.11	1.11	1.11
0.1	0.1	1.11	1.11	1.12
0.1	0.2	1.11	1.11	1.12
0.1	0.3	1.11	1.12	1.13
0.1	0.4	1.11	1.12	1.13
0.1	0.5	1.11	1.12	1.13
0.2	0.0	1.12	1.12	1.12
0.2	0.1	1.12	1.13	1.13
0.2	0.2	1.13	1.13	1.14
0.2	0.3	1.13	1.13	1.14
0.2	0.4	1.13	1.13	1.15
0.2	0.5	1.13	1.13	1.16
0.3	0.0	1.14	1.14	1.14
0.3	0.1	1.14	1.15	1.15
0.3	0.2	1.14	1.15	1.16
0.3	0.3	1.14	1.15	1.17
0.3	0.4	1.14	1.15	1.18
0.3	0.5	1.14	1.16	1.19
0.4	0.0	1.17	1.17	1.17
0.4	0.1	1.17	1.17	1.18
0.4	0.2	1.17	1.17	1.19
0.4	0.3	1.17	1.18	1.21
0.4	0.4	1.17	1.18	1.22
0.4	0.5	1.17	1.18	1.23
0.5	0.0	1.20	1.20	1.20
0.5	0.1	1.20	1.20	1.22
0.5	0.2	1.20	1.21	1.24
0.5	0.3	1.20	1.22	1.27
0.5	0.4	1.20	1.22	1.29
0.5	0.5	1.20	1.23	1.31

TABLE S1b. Observed RR 1.25

True relative risks that would lead to an observed relative risk of 1.25 in presence of different levels of uncaptured exposure or nonadherence

Amount of misclassification		True RR for different exposure prevalences		
Prop nonadherence	**Prop uncaptured data**	**p=0.01**	**p=0.10**	**p=0.25**
0.0	0.0	1.25	1.25	1.25
0.0	0.1	1.25	1.25	1.26
0.0	0.2	1.25	1.26	1.27
0.0	0.3	1.25	1.26	1.28
0.0	0.4	1.25	1.26	1.29
0.0	0.5	1.25	1.27	1.30
0.1	0.0	1.28	1.28	1.28
0.1	0.1	1.28	1.28	1.29
0.1	0.2	1.28	1.29	1.30
0.1	0.3	1.28	1.29	1.32
0.1	0.4	1.28	1.30	1.33
0.1	0.5	1.28	1.30	1.35
0.2	0.0	1.31	1.31	1.31
0.2	0.1	1.31	1.32	1.33
0.2	0.2	1.31	1.32	1.35
0.2	0.3	1.31	1.33	1.37
0.2	0.4	1.31	1.33	1.39
0.2	0.5	1.31	1.34	1.41
0.3	0.0	1.36	1.36	1.36
0.3	0.1	1.36	1.36	1.38
0.3	0.2	1.36	1.37	1.41
0.3	0.3	1.36	1.38	1.43
0.3	0.4	1.36	1.39	1.46
0.3	0.5	1.36	1.39	1.49
0.4	0.0	1.42	1.42	1.42
0.4	0.1	1.42	1.43	1.45
0.4	0.2	1.42	1.44	1.49
0.4	0.3	1.42	1.45	1.53
0.4	0.4	1.42	1.46	1.58
0.4	0.5	1.42	1.47	1.62
0.5	0.0	1.50	1.50	1.50
0.5	0.1	1.50	1.51	1.56
0.5	0.2	1.50	1.53	1.63
0.5	0.3	1.50	1.55	1.70
0.5	0.4	1.50	1.56	1.78
0.5	0.5	1.51	1.58	1.85

TABLE S1c. Observed RR 1.50

True relative risks that would lead to an observed relative risk of 1.50 in presence of different levels of uncaptured exposure or nonadherence

Amount of misclassification		True RR for different exposure prevalences		
Prop nonadherence	**Prop uncaptured data**	**p=0.01**	**p=0.10**	**p=0.25**
0.0	0.0	1.50	1.50	1.50
0.0	0.1	1.50	1.51	1.53
0.0	0.2	1.50	1.52	1.55
0.0	0.3	1.50	1.53	1.58
0.0	0.4	1.50	1.53	1.61
0.0	0.5	1.50	1.54	1.64
0.1	0.0	1.56	1.56	1.56
0.1	0.1	1.56	1.57	1.59
0.1	0.2	1.56	1.58	1.62
0.1	0.3	1.56	1.59	1.66
0.1	0.4	1.56	1.60	1.69
0.1	0.5	1.56	1.61	1.73
0.2	0.0	1.62	1.62	1.62
0.2	0.1	1.63	1.64	1.67
0.2	0.2	1.63	1.65	1.71
0.2	0.3	1.63	1.67	1.76
0.2	0.4	1.63	1.68	1.81
0.2	0.5	1.63	1.69	1.86
0.3	0.0	1.71	1.71	1.71
0.3	0.1	1.72	1.73	1.77
0.3	0.2	1.72	1.75	1.84
0.3	0.3	1.72	1.77	1.91
0.3	0.4	1.72	1.79	1.98
0.3	0.5	1.72	1.81	2.06
0.4	0.0	1.83	1.83	1.83
0.4	0.1	1.83	1.86	1.93
0.4	0.2	1.84	1.88	2.02
0.4	0.3	1.84	1.91	2.13
0.4	0.4	1.84	1.94	2.25
0.4	0.5	1.84	1.96	2.37
0.5	0.0	2.00	2.00	2.00
0.5	0.1	2.00	2.04	2.16
0.5	0.2	2.01	2.08	2.33
0.5	0.3	2.01	2.12	2.53
0.5	0.4	2.01	2.16	2.75
0.5	0.5	2.01	2.20	3.00

TABLE S1d. Observed RR 2.0
True relative risks that would lead to an observed relative risk of 2.0 in presence of different levels of uncaptured exposure or non-adherence

Amount of misclassification		True RR for different exposure prevalences		
Prop non-adherence	**Prop uncaptured data**	**p=0.01**	**p=0.10**	**p=0.25**
0.0	0.0	2.00	2.00	2.00
0.0	0.1	2.00	2.02	2.07
0.0	0.2	2.00	2.05	2.14
0.0	0.3	2.01	2.07	2.22
0.0	0.4	2.01	2.09	2.31
0.0	0.5	2.01	2.12	2.40
0.1	0.0	2.11	2.11	2.11
0.1	0.1	2.11	2.14	2.20
0.1	0.2	2.12	2.17	2.30
0.1	0.3	2.12	2.20	2.40
0.1	0.4	2.12	2.23	2.51
0.1	0.5	2.12	2.26	2.64
0.2	0.0	2.25	2.25	2.25
0.2	0.1	2.25	2.29	2.37
0.2	0.2	2.26	2.32	2.50
0.2	0.3	2.26	2.36	2.64
0.2	0.4	2.26	2.40	2.80
0.2	0.5	2.27	2.44	2.98
0.3	0.0	2.43	2.43	2.43
0.3	0.1	2.43	2.48	2.60
0.3	0.2	2.44	2.53	2.78
0.3	0.3	2.44	2.58	3.00
0.3	0.4	2.44	2.63	3.24
0.3	0.5	2.45	2.69	3.52
0.4	0.0	2.67	2.67	2.67
0.4	0.1	2.67	2.73	2.92
0.4	0.2	2.68	2.80	3.22
0.4	0.3	2.68	2.88	3.57
0.4	0.4	2.69	2.96	4.00
0.4	0.5	2.69	3.04	4.51
0.5	0.0	3.00	3.00	3.00
0.5	0.1	3.01	3.10	3.44
0.5	0.2	3.02	3.21	4.00
0.5	0.3	3.02	3.32	4.71
0.5	0.4	3.03	3.44	5.66
0.5	0.5	3.04	3.57	6.99

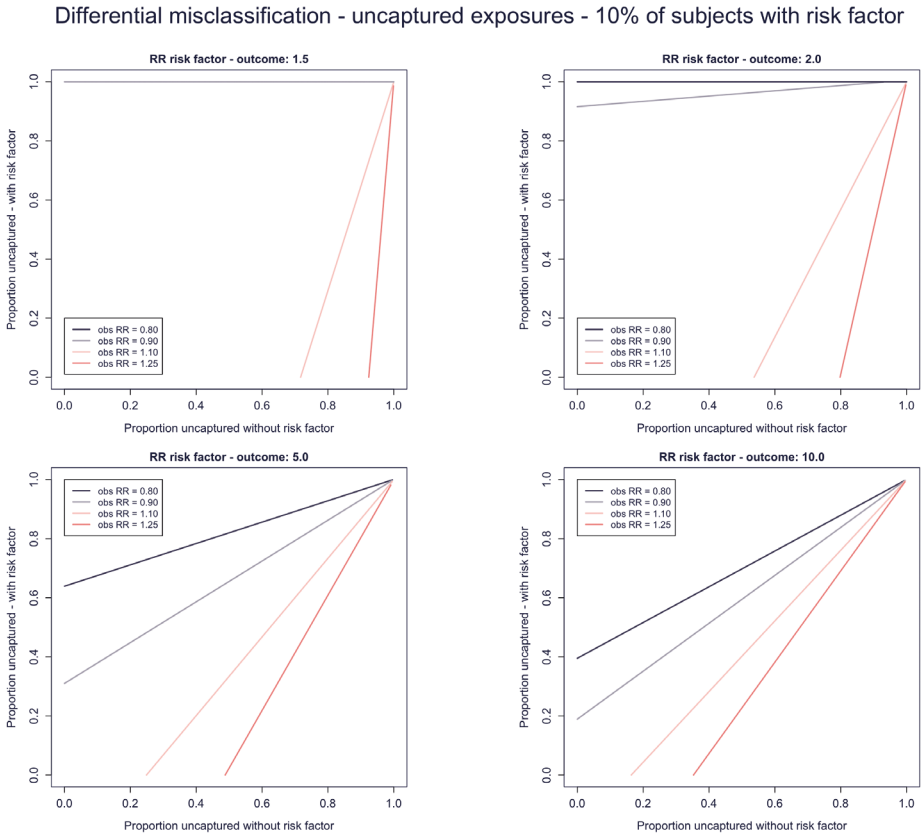


FIGURE S1a. Effect of differential uncaptured exposures - risk factor present in 10% of subjects Observed relative risks obtained with different proportion of exposure captured for subjects with and without risk factor, in absence of a true effect ($RR_{true} = 1.0$).

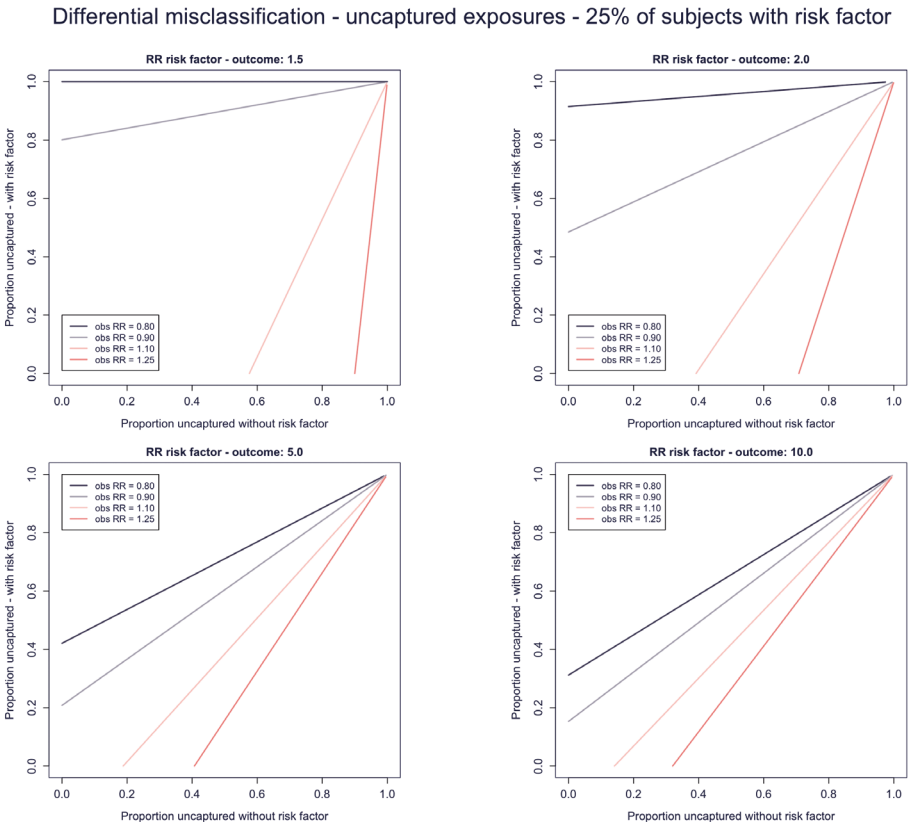


FIGURE S1b. Effect of differential uncaptured exposures – risk factor present in 25% of subjects Observed relative risks obtained with different proportion of exposure captured for subjects with and without risk factor, in absence of a true effect ($RR_{true} = 1.0$).

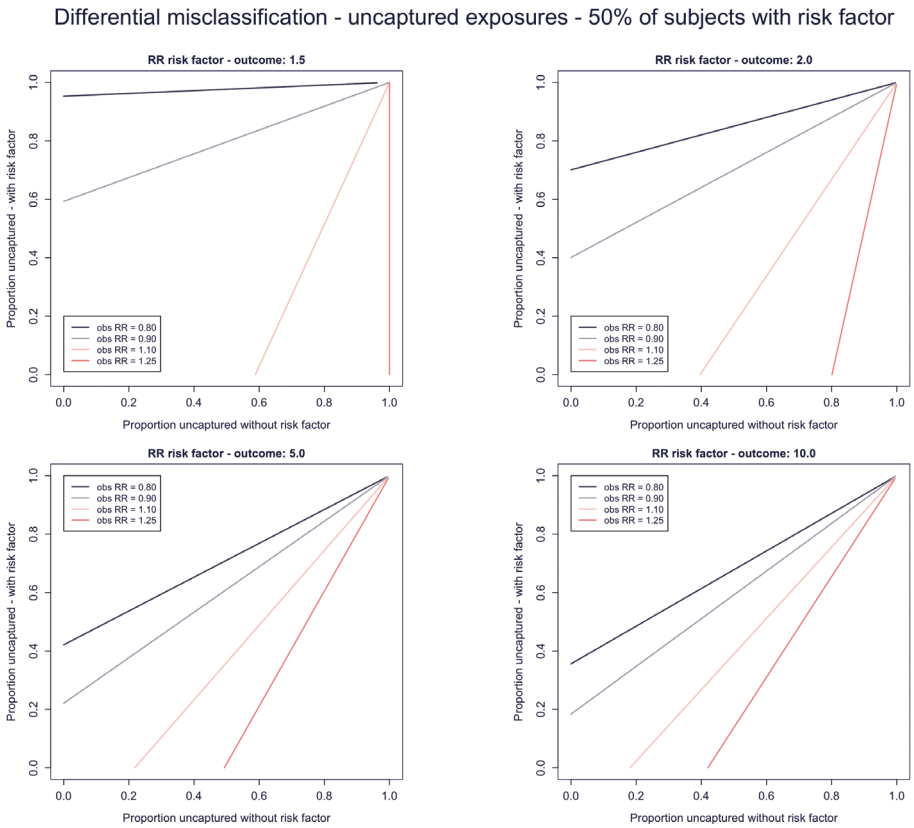


FIGURE S1c. Effect of differential uncaptured exposures - risk factor present in 50% of subjects Observed relative risks obtained with different proportion of exposure captured for subjects with and without risk factor, in absence of a true effect (RR_{true} = 1.0).

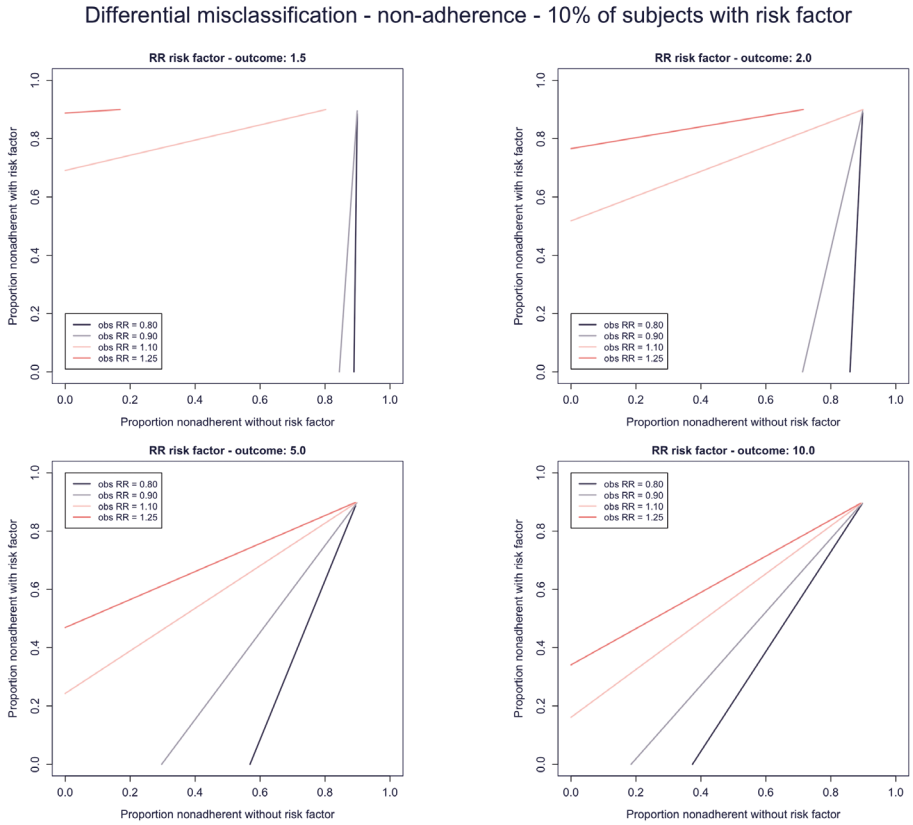


FIGURE S2a. Effect of differential non-adherence - risk factor present in 10% of subjects
Observed relative risks obtained with different proportion of exposure captured for subjects with and without risk factor, in absence of a true effect ($RR_{true} = 1.0$).

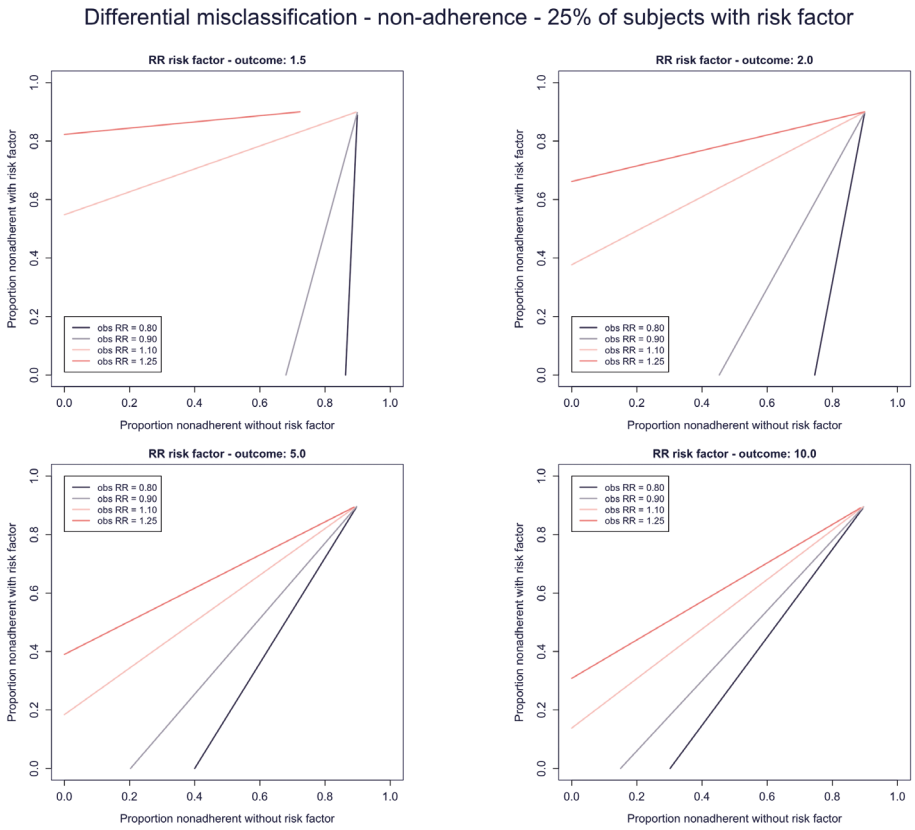


FIGURE S2b. Effect of differential non-adherence - risk factor present in 25% of subjects
Observed relative risks obtained with different proportion of exposure captured for subjects with and without risk factor, in absence of a true effect (RR_{true} = 1.0).

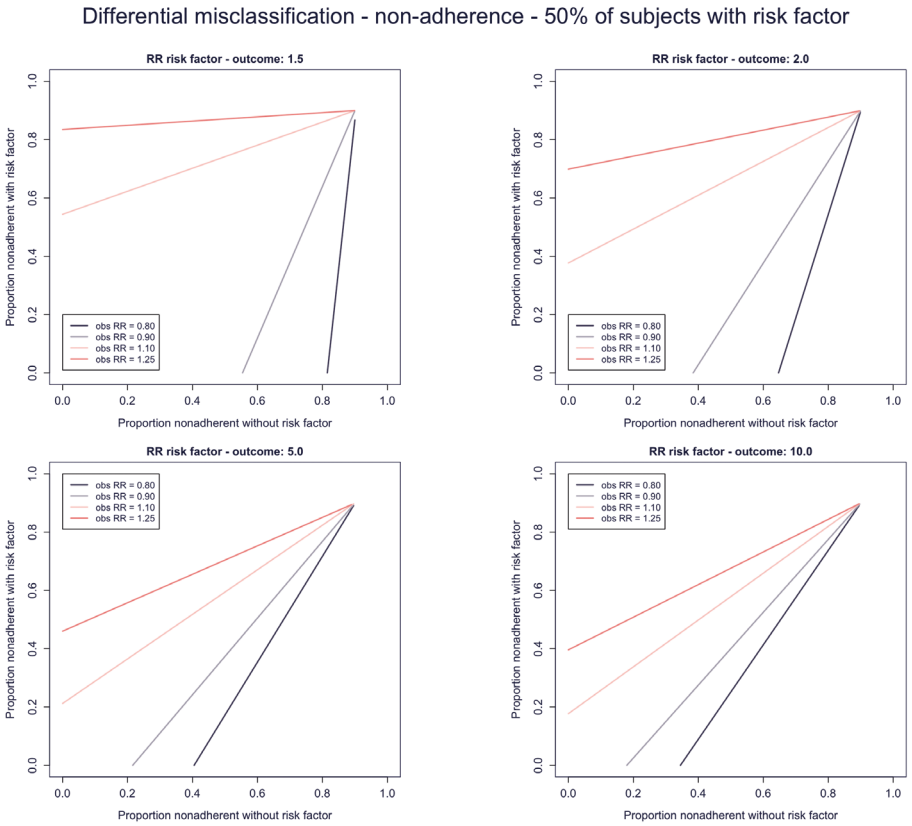


FIGURE S2c. Effect of differential non-adherence - risk factor present in 50% of subjects
Observed relative risks obtained with different proportion of exposure captured for subjects with and without risk factor, in absence of a true effect ($RR_{true} = 1.0$).

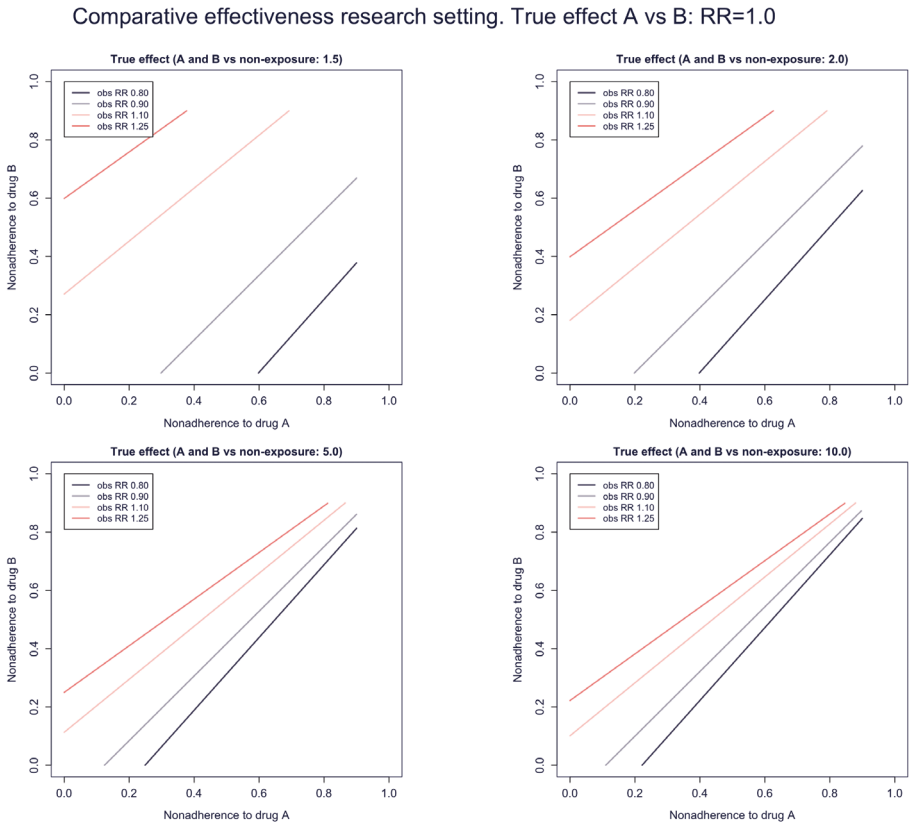
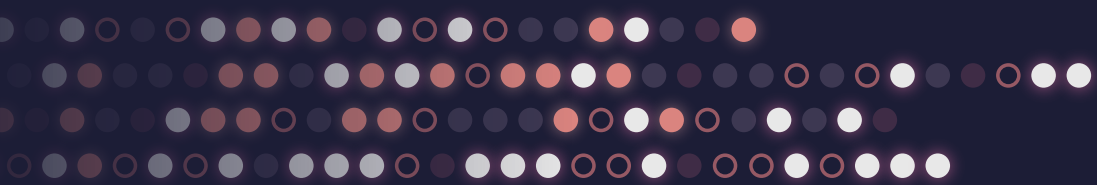


FIGURE S3. Comparative effectiveness research (drug A vs B)



4.

GENERAL DISCUSSION

DRUG EXPOSURE ASSESSMENT IN PHARMACOEPIDEMIOLOGICAL DATABASE STUDIES

Randomized controlled trials (RCTs) are conducted to demonstrate the efficacy and safety of drugs, and they are used to support market approval by regulatory agencies. However, after market approval, unanswered questions often remain regarding the safety and effectiveness of the drug. To complement evidence from pre-approval RCTs, post-approval RCTs and observational studies (also called post authorization studies) can be conducted. These studies take advantage of real-world data (RWD) that represent data on drug use and clinical outcomes in daily clinical practice. The FDA defines RWD as *"data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources."*[1] These sources include pharmacy records, general practitioner (GP) records, medical claims records, and disease registries.[2] However, the data are not primarily collected for research purposes. Information about outcomes, treatment, and other characteristics needed for analysis may consequently be inaccurate or missing, which can impact the estimation of the relation between the drug and the outcome.[3] Since benefit and risk assessments are conducted on the basis of the results of these observational studies, proper study design and measurement of exposures, outcomes, and other characteristics are important to obtain valid estimates of the benefits and risks of a drug treatment.

In this thesis, we focused on how drug exposure is reported and measured in observational studies. In addition, we assessed the impact of exposure measurement on the effect estimate of the relationship between drug exposure and clinical outcomes. Designing an observational study starts with the hypothesized mechanism and hazard function of the relationship between the drug treatment and the outcome. For example, the outcome may occur in a direct reaction to drug exposure, such as anaphylactic reactions, or it may be the result of long-term exposures, such as cancer outcomes.[4,5] The period in which the outcome is expected to occur is also called the risk window.[6] In addition, the period in which subjects are considered to be exposed to a drug must be defined. This period can comprise only the days when a subject uses the drug, but it may also be a longer period for drugs with a longer time of elimination from the body. Both the definition of the exposure window and the risk window are needed to determine whether outcomes are related to the drug exposure.

After defining the exposure and risk window in general, time on and off treatment must be determined for each of the included subjects. Each patient has their own usage patterns, which can consist of complex patterns of use and nonuse over time. To account for these time-variable usage patterns, drug exposure must often be defined in a time-varying manner. In the assessment of time on and off treatment, assumptions are needed, for example how gaps or overlapping periods between two prescriptions should be handled.[7] Moreover, patients rarely take a drug exactly as prescribed by their physician. For example, they can decide not to initiate use, implement the prescribed regimen incorrectly (e.g., taking the wrong dose or skipping days), or prematurely discontinue using the drug.[8] Moreover, for some drugs, it is challenging to perform observational studies, as their use is not captured in the resources employed for observational research, such as in case of the over the counter (OTC) use of drugs.[9] Given this complexity in ascertaining drug use (exposure) in populations, the way in which drug use is defined and measured in observational data can impact the effect estimates of the relation between drug exposure and outcomes.

In this thesis, we assessed the extent of exposure misclassification when using routinely collected health data and the impact thereof on effect estimates. In addition, we investigated how exposure is defined and reported in published pharmacoepidemiologic studies. In this chapter, we discuss our findings in a broader context.

THE ROLE OF PHARMACOKINETICS WHEN DEFINING EXPOSURE

Defining exposure starts with considering the biological mechanism by which the drug causes the outcome. Intended effects are often dose dependent. Adverse events can additionally also be non-dose related, related to cumulative dose, or related to withdrawal.[10,11]

Based on the assumed mechanism, the exposure window and the risk window can be defined. Different guidelines for pharmacoepidemiologic research encourage researchers to provide a rationale for the chosen exposure definition and risk window. For example, the FDA's Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data states that *"The investigator should define the exposure risk window for the outcome of interest and describe in detail the measurement of the window in the selected data source(s). FDA recommends that the investigators obtain information about the postulated exposure risk window from other sources, such as spontaneous report data, to increase the likelihood that only relevant periods of exposure are examined."*[12] Furthermore, the ENCePP Guide on Methodological Standards in Pharmacoepidemiology recommends that researchers consider the biological mechanism of action as well as the pharmacokinetics and pharmacodynamics of the drug.[13] The ISPE-ISPOR Task Force guideline even provides detailed information on how exposure should be defined and reported.[14]

One specific factor when defining the exposure window concerns the half-life of the drug, which is the time needed for the plasma concentration to decrease to half of its starting concentration.[15] After roughly four to five times the half-life, the drug is fully eliminated, and it is assumed that it can no longer exert a direct effect.[15] For most orally taken drugs, the half-life ranges between four and 24 hours[16], but there are exceptions. Examples of drugs with much longer half-lives are the antiarrhythmic drug amiodarone (21–78 days); the antituberculosis drug bedaquiline (more than five months); the antimalarial mefloquine (14–41 days); and dutasteride, which is used for benign prostate hyperplasia (three to five weeks).[17–20] Moreover, newer drug classes, such as monoclonal antibodies, generally have a half-life of several weeks.[16] The consequence of these long half-lives is that both intended and adverse events can occur long after administration, which must be considered when designing observational studies. However, in such studies, the long half-life of medicines is often not accounted for.

Studies that have assessed the association between the use of amiodarone and thyroid disorders, hepatotoxicity, and acute pancreatitis are prime examples of this heterogeneity in how exposure to amiodarone has been defined.[21–25] Only one of these studies extended the defined duration of use beyond ingestion of the last dose with two months for amiodarone to account for its half-life[26]. Since the impact of different exposure definitions in relation to the half-life of amiodarone is unknown, we studied this in relation to the occurrence of acute pancreatitis in *Chapter 3.2* Lai et al. compared current use of amiodarone (most recent prescription <3 months before the event)

and never use of amiodarone in terms of the associated risk of acute pancreatitis; their study revealed a five times higher odds ratio (OR) of current use versus no use (OR = 5.21 [3.22–8.43]).[23] Furthermore, Alsonso et al. compared ever use of amiodarone with never use and reported an OR of 1.53 (1.24–1.88) for the risk of acute pancreatitis.[22] The difference in effect estimates is possibly caused by the different methods of defining exposure to amiodarone. More importantly, both studies did not consider the pharmacokinetic properties of the drug.

To assess the impact of the half-life of amiodarone on effect estimates, we used pharmacy dispensing information and applied various exposure definitions for amiodarone, with and without accounting for the drug half-life. The relation of amiodarone and acute pancreatitis in this study was found to be relatively insensitive to all these different exposure definitions. For example, the hazard ratio (HR) for current compared to noncurrent use was 1.36 (0.78–2.38), while the HR for the model that accounted for the half-life – by gradually reducing the assumed concentration present in the body – was 1.21 (0.69–2.10). It should be noted that within the study population, 28.8% of all subjects were exposed to amiodarone during their complete (>95%) follow-up. The way in which amiodarone was measured in these subjects had little impact, as there was no variation in their exposure status.

Although we found no material impact of the different exposure definitions of amiodarone in this study, the fact that a long exposure period may have an impact was demonstrated for the anti-osteoporotic drug zoledronic acid. This drug has a long residence time in the body due to irreversible binding to bone tissue.[27] The use of zoledronic acid was associated with the occurrence of arrhythmia in a clinical study. However, most of these events occurred more than 30 days after infusion, when the drug is no longer detectable in the circulation.[27] It is therefore recommended that the pharmacokinetics of the study drug should be considered when determining the exposure window. If the exposure window is not identified correctly, associations may be missed, since outcomes are not attributed to the true exposure status.

THE EXTENT OF EXPOSURE MISCLASSIFICATION

After determining the exposure window and the risk window, another element of the study design is to determine which subjects are exposed to the drug and which are not. This is typically based on information captured in routinely collected health data, such as GP databases containing information about prescribed drugs, pharmacy databases containing information about dispensed drugs, or health insurance databases containing administrative claims. Yet, the information about exposure contained in these databases can only serve as a proxy, and exposure status retrieved from these databases may not always reflect the actual exposure status. There are several reasons for these potential mismatches (Figure 1). On the one hand, subjects can be misclassified as "nonexposed" when they are using drugs for which information about the exposure is not captured in the database. On the other hand, subjects may be misclassified as being "exposed" to a specific drug when, in real life, they are nonadherent to the drug prescription.[28] Since exposure misclassification may impact associations with outcomes of interest, It is important to obtain insight into the extent of potential exposure misclassification when conducting a pharmacoepidemiologic study.

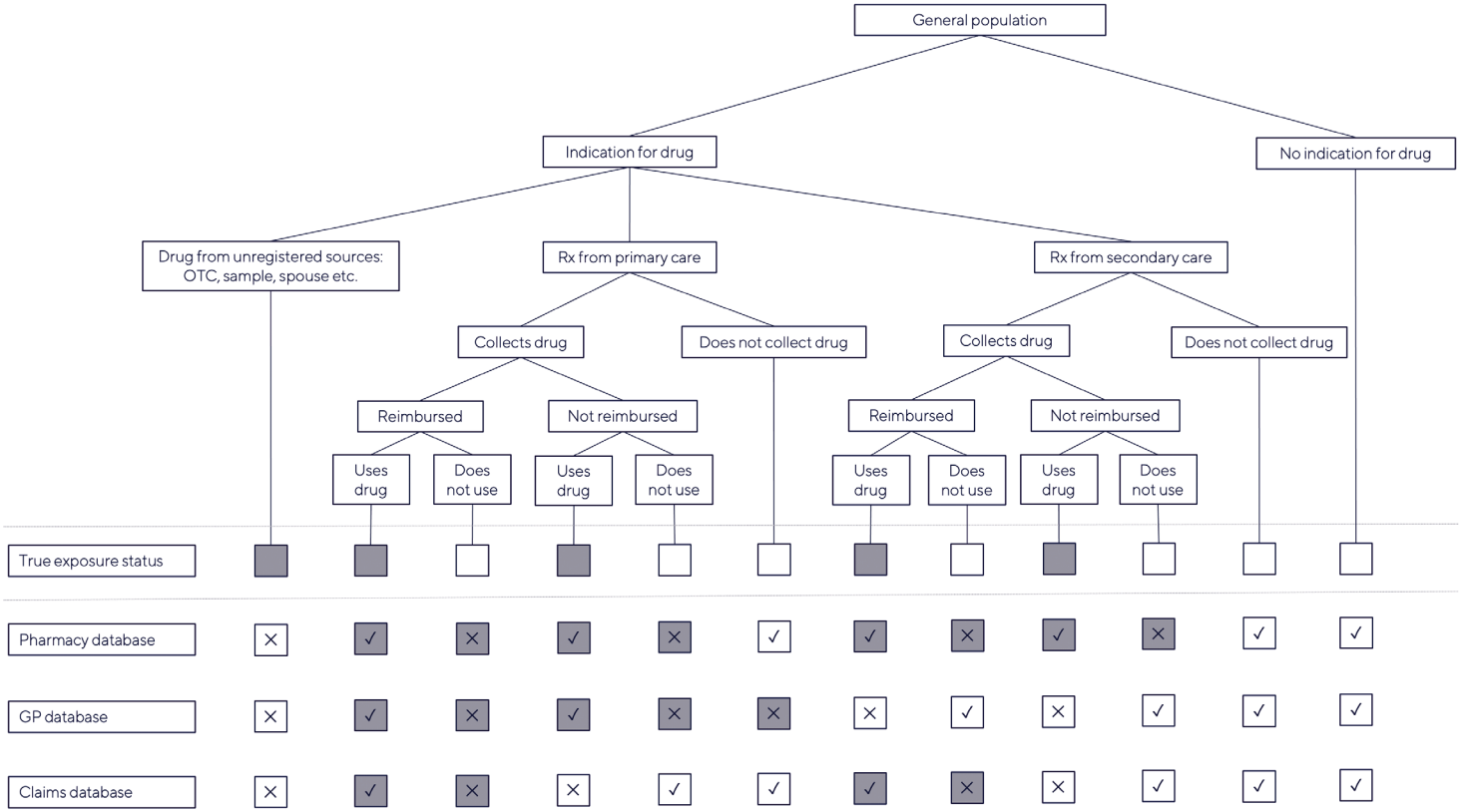


FIGURE 1. Sources of exposure misclassification in pharmacoepidemiology and the effects in different types of databases. Gray squares indicate "exposed." White squares indicate "unexposed." Abbreviations: Rx = prescription, √ = correctly classified, X = misclassified.

Exposure misclassification due to uncaptured exposures

The reason that some drugs are not captured in observational data can be related to the drug class, the database type, or the health care system.[29] For instance, uncaptured use can occur for the use of drugs prescribed in a clinical setting that is not captured in the database being used, for drugs with a restrictive drug coverage policy, for OTC drugs, or for the use of drugs that were originally prescribed to someone else.[28,30,31]

In a GP prescription database, drug exposure may be misclassified when a subject is treated with a drug that is also prescribed by a non-GP prescriber, as these prescriptions are not included in a GP database.[12,32] In *Chapter 3.3,* we showed that in the Netherlands, 33.1% of all DOAC users received their prescriptions only by a hospital specialist, while 45.6% received DOAC prescriptions by both a GP and a hospital specialist. Using only the GP prescription for pharmacoepidemiologic studies may therefore lead to misclassification of the exposure status in different ways. For subjects who are being treated by both a GP and a hospital specialist, exposure time is partially missing, while patients who are only being treated by the hospital specialist will be completely misclassified as unexposed during their follow-up. The extent to which this occurs differs between drug classes, patient groups, and health care systems. A Danish study found that 88% of all dispensings were prescribed by GPs, while only a smaller proportion of prescriptions were issued by hospital physicians (7.4%) and private practicing specialists (4.2%). However, this distribution varied according to age class, with GP prescriptions accounting for 48% of all prescriptions in patients aged 0–17 years vs 90% in patients aged ≥70 years. There were also differences between drug classes, with GP prescriptions accounting for 95% of all cardiovascular drug prescriptions vs 48% for drugs related to the sensory organs. GPs more often prescribed maintenance treatment compared to the initiation of new treatments (90% vs 84%), a pattern which was also seen for the prescribing of antidepressants in the Netherlands.[33] In addition, newly marketed drugs (e.g., DOACs, marketed in 2008, or the antidiabetic sodium-glucose co-transporter 2 [SLGT2] inhibitors, which were marketed in 2012) are more often prescribed by hospital specialists. In *Chapter 3.3,* we also demonstrated that in the Netherlands in 2013, the majority (73.6%) of all DOAC treatments were initiated by a hospital specialist, compared to only half (49.9%) in 2017. The same trends were shown for SLGT-2 inhibitors in Denmark, where the proportion of specialist initiations decreased from 41.% to 23.4% between 2016 and 2018.[34] GP prescription databases may therefore not always be a valid source for drug exposure assessment depending on the specific research question.

In claims databases, drug exposure may not be captured when the drug use is not reimbursed. Whether drugs are reimbursed differs between drug classes and between health care systems. In some countries, drugs are fully reimbursed, while in other countries, there is a co-payment by patients.[29] In addition, some drugs are only reimbursed under specific criteria, such as treatment failure or intolerance with first-line therapies, also called restrictive coverage policies. Patients who do not meet these criteria must pay for the drug themselves.[29,35] This policy affects the chance that these exposures are captured in a claims database. In Canada, for example, it is estimated that 16% of all drug exposures are not captured in a claims database, due to restrictive coverage policies.[36]

Drugs that are available OTC are also often not reimbursed. OTC use of drugs can lead to misclassification, since this type of usage is often not systematically captured in prescription, dispensing, or claims databases. Moreover, the medicines that are available OTC differ per

country. Most countries have NSAIDs available OTC, whereas statins, proton pump inhibitors, and inhaled beta-agonists are available in only a limited number of countries as OTC drugs.[37] The extent of OTC use also differs between different drugs. For NSAIDs, OTC use is estimated to be 10%–50% in the US, Germany, and the Netherlands, while OTC statins are used by less than 1% of the UK population with a potential indication for statin use.[38–42] The difference between NSAIDs and statins is that patients can judge for themselves whether the use of NSAIDs as painkillers is necessary, but the necessity of using statins for cardiovascular prevention is often determined by a GP. Moreover, statins are often prescribed at doses higher than the 10 mg available OTC.

Since the extent and nature of uncaptured exposures differ per drug and per type of database that is used, sufficient knowledge about the database and the underlying health care system is required to understand which exposures are captured and which are not. This knowledge is necessary in choosing the right database for the exposure assessment. Moreover, a description and quantification of potential uncaptured exposures is important in publications of database studies for correct interpretation of the study results.

Exposure misclassification due to nonadherence

Another source of exposure misclassification is misclassification due to nonadherence; that is, patients not adhering to the drug prescription. Nonadherence can occur at several stages during medication use.[8] First, patients can decide not to initiate the use, which is also called primary nonadherence (PNA).[43] Second, the prescribed regimen can be implemented incorrectly, for example by taking the wrong dose or by skipping days, which is also called secondary nonadherence.[43] Third, patients can decide to prematurely discontinue their use of the drug, also called non-persistence. Most studies focus on the assessment of secondary nonadherence and non-persistence, whereas the extent of PNA is less studied.[44] The main challenge in measuring PNA is that for the assessment, multiple data sources are needed, representing both what is prescribed and what is dispensed for the individual patient, to determine the proportion of newly prescribed drugs that is never dispensed at the pharmacy.[45]

In *Chapter 3.1,* we assessed PNA in the Netherlands, which was estimated to be 11.5%. The level of PNA varied across drug classes, ranging between 5.5% for thyroid hormones and 12.8% for proton pump inhibitors. In addition, PNA was found to be associated with several patient- and prescription-related characteristics. A higher likelihood of PNA was observed, for example, among patients with more comorbidities (OR for >3 active diagnoses 1.46 [1.36–1.56] compared to no active diagnoses), for drugs that were not reimbursed (OR 2.74 [2.61–2.88] compared to those that were fully reimbursed), and for drugs that are prescribed in the last quarter of the year compared to the first quarter (OR 0.68 [0.65–0.71]).

Our estimate of PNA was in line with other European estimates (8.5%–9.3%),[46,47] and we found similar patterns in the level of PNA between drug classes compared to those observed in Denmark.[47] In contrast, estimates of PNA in the US are twice as high as estimates from Europe (17.0% vs 8.5%), possibly caused by the presence of universal health coverage in most European countries, but not in the US.[48] Several other studies also indicate that costs and reimbursement status are important drivers for PNA.[49–51] In our study, some of the associations found could be specifically explained by the Dutch reimbursement system. Thus, the degree of nonadherence

and consequently the degree of exposure misclassification differ between drugs and between health systems. Therefore, it is important to provide estimates of the level of misclassification expected in a pharmacoepidemiologic database study due to nonadherence. If these estimates are not available for a specific drug within a specific health system, they can be obtained by using values of comparable drugs in comparable health and reimbursement systems.

THE IMPACT OF EXPOSURE MISCLASSIFICATION

Exposure misclassification can occur when subjects are nonadherent or when they are using drugs for which information about the exposure is not captured in the database. Exposure misclassification may lead to biased estimates of the relationship between drug exposure and clinical outcomes. There are two key types of exposure misclassification: 1) the level of misclassification can be unrelated to the outcome risk (nondifferential), or 2) it can relate to the outcome risk (differential). Both types can be problematic. Nondifferential exposure misclassification generally results in bias toward the null and may lead to weakly associations remain undetected. Differential exposure misclassification can lead to bias in either direction – either masking associations or leading to spurious results. It is therefore important to gain insight into the potential impact of exposure misclassification on effect estimates.

In the 1970s, Copeland et al. provided an extensive overview of the impact of different values of sensitivity and specificity of the exposure assessment on effect estimates (see Textbox 1 and Figure 2 for the relation between nonadherence/uncaptured data and sensitivity/specificity).[52] They showed, for example, that an exposure assessment with 90% sensitivity and 96% specificity changes the true relative risk of an exposure–outcome relationship of 2 to an observed relative risk of 1.5.

TEXTBOX 1. Non-adherence and uncaptured exposures in relation to sensitivity and specificity

Sensitivity and specificity are key measures to quantify misclassification. Sensitivity is calculated as the proportion of exposed subjects who are classified as exposed: $^{\text{True Positive}}/_{\text{(True Positive + False Negative)}}$. Specificity is defined as the proportion of unexposed subjects who are classified as unexposed: $^{\text{True Negative}}/_{\text{(True Negative + False Positive)}}$. The effect of uncaptured exposure and nonadherence on sensitivity and specificity is illustrated with a numerical example in Figure 2.

Sensitivity is directly related to the proportion of exposure that is captured; an 80% captured exposure equals a sensitivity of 0.8. The value of specificity is affected by both nonadherence and exposure prevalence. A lower exposure prevalence will result in a higher proportion of truly unexposed subjects and thus a higher specificity. For example, on the one hand, 20% nonadherence to a drug with a 10% prevalence results in a specificity of 0.987, while the specificity decreases to 0.867 when the exposure prevalence is 50%. On the other hand, in a situation of 10% exposure prevalence, 40% nonadherence results in a specificity of 0.966, compared to 0.987 for 20% nonadherence – both specificity values are high but relate to large differences in adherence rates (Figure 2).

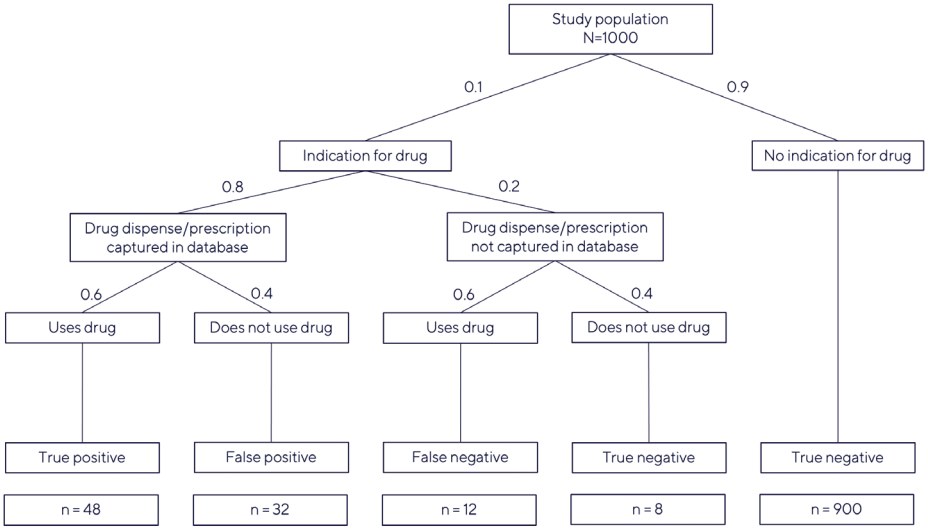


FIGURE 2. Model used for simulation analysis, with values for exposure prevalence (10%), nonadherence (40%), and uncaptured data (20%), and the corresponding exposure status. In this example, sensitivity is True Positive/(True Positive + False Negative) = 48/(48 + 12) = 0.80, and specificity is True Negative/(True Negative + False Positive) = 908/(908 + 32) = 0.966

In addition, other studies have demonstrated that small violations of the assumption of nondifferential misclassification (e.g., specificity 0.90 vs 0.92) or differences in misclassification between Drugs A and B (e.g., specificity 0.95 vs 0.98) could result in clinically relevant deviations from the true effect.[53–55] However, these studies were conducted in an era in which drug exposure was assessed by means of interviews or questionnaires, rather than computerized health care data. The current available databases allow for longitudinal measurement of drug exposures with high specificity, especially for drugs that are taken chronically.[56] For example, 20% nonadherence to a drug with a 10% prevalence results in a specificity of 0.987, which is a high specificity value compared to values used in previous studies.

When studying the impact of exposure misclassification, it is therefore important to substantiate the values for sensitivity and specificity with known values of exposure prevalence, nonadherence, and uncaptured exposure. This will result in more realistic scenarios when assessing the impact of exposure misclassification on study outcomes. Small deviations in specificity can imply large differences in adherence when the exposure prevalence is low. For example, Brenner utilized a sensitivity of 0.9 and a specificity of 0.9 for the exposure measurement, with exposure prevalence values of 0.01, 0.1, and 0.5. However, the degree of nonadherence required to result in these values of specificity is 91.5%, 50%, and 10% for the three different exposure prevalence values, respectively. Since most drugs have an exposure prevalence of up to 10%, and nonadherence is often <40%, we considered a specificity of 0.9 to be unlikely for current pharmacoepidemiologic database studies.

The impact of uncaptured data on effect estimates

In *Chapter 3.4,* we studied the impact of uncaptured exposure on effect estimates in a simulation study using literature-based values for the amount of uncaptured exposure. For nondifferential misclassification, we found that the impact of uncaptured data was generally limited. For example, for exposures with a prevalence of 10%, 25% nondifferential uncaptured exposures changed the relative risk (RR) from 2.0 to 1.95 (-5%). The impact of differential misclassification due to uncaptured exposures was also limited. Differences in the level of uncaptured exposures between subjects with and without a higher outcome risk needed to be relatively large (e.g., 50% vs 90%) to potentially lead to a clinically relevant bias (e.g., RR 1.1 instead of 1.0).

The limited impact of uncaptured exposures on effect estimates was also shown in two other simulation studies. Gaster et al. examined the effect of uncaptured exposure due to OTC use.[9] They found that, based on information about OTC purchases of NSAIDs in Denmark, the maximum attenuation of the effect estimates was 5%. This was comparable to the impact of uncaptured exposures found in our simulation. In addition, Gamble et al. studied the impact of exposure misclassification due to a restrictive coverage policy. This policy means that only subjects meeting specific criteria will be reimbursed for the drug use. Patients who do not meet the criteria will be misclassified as unexposed in a claims database. In this simulation, it was found that randomly misclassifying up to 50% of all exposed subjects as unexposed had a limited impact on the effect estimates (e.g., HR 0.88 for 0% uncaptured data compared to HR 0.82 for 50% uncaptured data).[57] In contrast, when only the first period of use – not the total subject exposure time – was misclassified, the effects were much more pronounced, and the observed relative risk could even be reversed compared to the true effect (e.g., HR 1.34 instead of 0.88 when the first 50% exposure time was defined as unexposed).

The impact of uncaptured data has also been assessed in case studies, using real data. In *Chapter 3.3,* we assessed the impact of exposure misclassification caused by using a GP prescription database in a case study on DOACs and bleeding risk. DOACs were chosen as a case study because in the first years after licensing, DOACs were prescribed predominantly by specialists, such as cardiologists, internists, and orthopedics.[58,59] For this study, we used dispensing information from community pharmacy data in the Netherlands. For each of the included subjects, information on the type of prescribing physician was available for their dispensed DOAC prescriptions. We found that subjects treated with DOACs by only a GP, only a specialist, or both had different characteristics that also differed between initiators of vitamin K antagonists (VKAs) and DOACs, and there were different underlying diagnoses for the anticoagulant use. VKA users who received their VKA prescriptions only by a specialist had more comorbidities and used more comedication than patients who only received their VKA prescriptions from a GP. For the prescribing of DOACs, the opposite was seen: patients who only received DOACs from a specialist generally had fewer comorbidities and received fewer comedications. Since these characteristics were also associated with the bleeding risk, using only GP prescription information could lead to differential exposure misclassification. Interestingly, compared to using all dispensing information, including only anticoagulant dispensings prescribed by GPs did not materially impact the effect estimates of DOAC use compared to VKA use on the risk of major bleeding and stroke/TIA. Another case study assessed the impact of exposure misclassification due to restrictive coverage policies of effect estimates. For this study, a database was used that included all thiazolidinediones dispensings in Saskatchewan, Canada. In addition, information

on their reimbursement status was known for all subjects and was used to determine which subjects were classified as exposed and which were not. This study revealed that although patient characteristics differed between users who were eligible for reimbursement and those who were not, no significant bias of the effect estimate of all-cause hospitalization or death was found when only the subjects with reimbursement were included.[60]

Hence, both the simulation studies and the case studies showed limited impact of uncaptured data on the effect estimates. Whether the same results would be obtained if other databases had been used, or with other drug–outcome relationships, remains unknown and is highly dependent on the specific characteristics of the database that is used. In addition, uncaptured data can have implications on the generalizability of the effect estimates that are found. For example, when a GP prescription database is used, results may only apply to patients treated in primary care, who may be healthier than patients treated by a hospital specialist. When a claims database is used, results may only apply to patients who meet the reimbursement criteria, and they are likely to differ from patients who do not. It is therefore important to report which exposures are included in the study database and which may have been missed, for correct interpretation of the study results.

The impact of nonadherence on effect estimates

In the simulation study in *Chapter 3.4,* we also assessed the impact of literature-based values for nonadherence on effect estimates. We found that nonadherence had a larger impact on effect estimates than uncaptured exposures. For exposures with 10% prevalence, 25% nondifferential nonadherence changed the RR from 2.0 to 1.75 (-25%), instead of the 5% found for uncaptured data. The fact that false-positive exposures (due to nonadherence) have a larger impact than false-negative exposures was also shown by Hernandez-Diaz.[61]

The method by which the impact of misclassification due to nonadherence was measured in both studies also had limitations. In both studies, 2x2 tables were used to simplify the actual mechanisms causing exposure misclassification. Time effects were ignored in the analyses, such as the fact that subjects who were prone to a negative outcome stopped taking the drug earlier than subjects who tolerated the drug better. The level of nonadherence may also vary over time within subjects. To our knowledge, the impact of these complex patterns has not yet been investigated.

Moreover, an additional challenge for studying the impact of misclassification on drug-outcome associations due to nonadherence is that being nonadherent has an impact on the risk itself. Patients who do not use their drugs as prescribed are not optimally treated, as shown, for example, in the relation between the use of drugs for secondary prevention on mortality and cardiovascular morbidity.[62] The impact of nonadherence might also differ across drug-outcome pairs. For outcomes that are related to current use (e.g., the relation between NSAIDs and gastro-intestinal bleeding[63]), it is more important to correctly identify current use than it is for outcomes that are related to cumulative or long-term use (e.g., the relation between statins and cardiovascular events[64]). In addition, the impact of nonadherence also depends on the research question. When the research interest is in the effects of drugs in everyday use, including patients who are nonadherent, this misclassification is not a problem but rather the

subject of study. Conversely, when the focus is on the effects of actual use, nonadherence can lead to biased estimates.[65]

Future research is thus needed a) to comprehend the impact of complex patterns of misclassification due to nonadherence and b) to differentiate between the impact nonadherence itself has on effect estimates and the extent to which the estimate is affected by exposure misclassification. This impact should be studied for different types of exposure–outcome relations, such as outcomes related to current use or related to cumulative use. For now, to allow for correct interpretation of the study results, it is recommended that authors provide estimates of the level of nonadherence and how nonadherence might have impacted the effect estimates.

IMMORTAL TIME BIAS DUE TO EXPOSURE MISCLASSIFICATION

Aside from exposure misclassification due to nonadherence and uncaptured exposures, which can often not fully be avoided, there is another source of exposure misclassification, namely due to incorrect classification of exposure in cohort studies, which can result in immortal time bias. This type of bias occurs when subjects are defined as being exposed during the whole follow-up time, when exposure starts some time during follow-up.[66] The time before treatment starts is misclassified as exposed; however, because the event (e.g., death or the adverse event) cannot occur before the start of treatment, the subjects are "immortal" during this time window. Since this "immortal time" is assigned to the "exposed" category, the risk among the "exposed" will be artificially decreased, resulting in a too optimistic effect estimate for the "exposed" compared to the "nonexposed."[66,67]

This type of bias was already described in 1969, and since then, there have been numerous articles describing and quantifying this problem.[66–69] The impact of immortal time on effect estimates can be large and should always be avoided, either by defining exposure dynamically (i.e., accounting for time on and off treatment) or by starting follow-up when treatment is initiated. Although studies explaining immortal bias and its impact have been extensively published, studies on the effectiveness of hydroxychloroquine (HCQ) in treating patients with COVID-19, providing necessary guidance to clinicians, were not free of immortal time bias.[70] In *Chapter 2.2*, we showed that two thirds of studies that investigated the association between the use of HCQ and clinical outcomes were susceptible to immortal bias. In some of these studies, the potential impact of immortal time bias was assessed as a sensitivity analysis, revealing large differences between estimates with and without immortal time bias. For example, the estimates changed from 1.08 to 1.46[71] or from 0.68 to 0.82[72] when exposure was defined in a time-varying manner instead of "any exposure during hospitalization." However, the conclusions about the effectiveness of HCQ in these studies did not change, because the effect estimates remained on the same side of the null effect (i.e., above or below 1.0). Yet, these large deviations indicate that immortal time bias might have such a large impact on effect estimates that it can alter study conclusions.

It is remarkable that immortal time bias still frequently occurs, since many guidelines on pharmacoepidemiologic studies recommend that exposure should be defined dynamically.[12,14,73–75] Thus, understanding the proper design of pharmacoepidemiologic studies

and how exposure should be defined is vital, and the existing guidelines should therefore always be followed.

REPORTING OF EXPOSURE ASSESSMENT

Since small variations in study design and exposure assessment can impact effect estimates, transparent reporting is important not only for the interpretation of published study results but also for reproducibility and validity assessment.[7,76–79] Reporting guidelines (e.g., CONSORT,[81] for reporting on clinical trials; STROBE,[82] for observational studies; and STROBE-RECORD,[80] for reporting on observational studies using routinely collected health data) support researchers to describe their research in a transparent and complete manner. Specifically for pharmacoepidemiology, there are currently two reporting guidelines: 1) guidelines from the joint ISPE-ISPOR Task Force and 2) an extension of the RECORD statement, namely the RECORD-PE.[14,75]

Many journals endorse the use of CONSORT and STROBE. The publication of these guidelines led to an improved reporting of clinical trials and observational studies, especially for publications in journals that endorsed these guidelines.[81–83] However, it takes a substantial amount of time to see the effects of published guidelines on transparent reporting in practice. In the case of CONSORT, reporting has improved in the 20 years after the first version was published but remains suboptimal, with on average 18 of 37 items being reported in each study over the period 2010–2014. [82,84] In the case of STROBE, reporting has also improved after publishing of the guidelines, but there is still room for improvement, as the median compliance with the 22 items was 77% in 2016, nine years after STROBE was published.[81]

To assess the quality of reporting in pharmacoepidemiological studies, we provided in *Chapter 2.1* a baseline assessment of exposure assessment reporting according to the ISPE-ISPOR Task Force guidelines (published in 2017). This baseline assessment is useful to determine where improvement may be needed and can also be used as a benchmark to assess the effect of the pharmacoepidemiologic publication guidelines over time. We assessed the quality of reporting according to the ISPE-ISPOR Task Force Guidelines, in studies published in 2017 in six epidemiological journals. We included all studies that used routinely collected health data for exposure assessment, such as prescription data, dispensing data, or claims data. We found that none of the 91 included and assessed studies met all requirements for the reporting of drug exposure as defined by the ISPE-ISPOR guidelines. In general, conceptual details about the exposure risk window and the exposure assessment window were reported relatively often (85% and 98%, respectively), whereas details regarding the construction of the exposure risk window were reported poorly. For example, details regarding the presence or absence of an induction period were reported explicitly in only 14 (16%) of all included studies. This suboptimal reporting was also observed by Weisman et al., who studied the reporting of exposure to DPP-4 inhibitors in relation to cardiovascular outcomes.[85] They showed that the lag period was reported in only 5 of 14 studies (38.6%).[85] We also observed this in studies on the effectiveness of HCQ in clinical outcomes in *Chapter 2.2*. Two thirds of all studies reported insufficient information in the article to fully comprehend all methodological choices, with insufficient information provided on the definition of exposure to HCQ in one out of every five studies. This incomplete reporting hinders the correct interpretation, reproducibility, and assessment of validity of the study results.

Authors, reviewers, and editors are therefore encouraged to adhere to relevant reporting guidelines, such as the ISPE-ISPOR and RECORD-PE guidelines. To accelerate adherence, journals could consider obliging authors to use one of these two guidelines. This can be done, for example, by making it compulsory to complete checklists based on these guidelines when submitting a manuscript, which is one step further than the general recommendation that is now being made to adhere to these guidelines.[86–89] For reviewers, the use of these guidelines can also aid in determining which items have not been reported sufficiently to properly assess the quality of the submitted manuscript. In this way, better reporting can also lead to more valid studies.

FUTURE CHALLENGES

Methodological research on drug exposure dates back to 1977, and much has been achieved; however, there are still unanswered questions, and new questions may arise because of the changing landscape of pharmacoepidemiology. Challenges exist regarding drug exposure measurement in multi-database studies when the individual databases differ in the information they contain about drug exposures. There are also challenges regarding the assessment of exposure to newer-generation drugs, which are often not dispensed in primary care. In addition, challenges exist regarding the impact that pandemics such as the COVID-19 pandemic may have on how medicines are prescribed and dispensed. These challenges are discussed in the next section.

Multi-database studies

In the current regulatory and clinical landscape, multi-database studies are the norm.[90] The DARWIN EU initiative of the EMA is expecting to unlock more than 40 data sources over the next five years that will allow them to assess the safety and effectiveness of drug use in Europe.[91] Multi-database studies can include a high number of patients, which is especially advantageous for rare exposures and outcomes, resulting in more precise estimates.[73,90,92] In addition, these databases originate from different populations, health care systems, impacted by differences in the organization of care, reimbursement, and technological aspects. These differences open the opportunity for an understanding of how differences in population characteristics, and health systems might impact drug utilization and hence the estimated benefits and risks.[92,93]

There are different strategies to execute multi-database studies, from local analysis of the data -based on a joint study protocol - to conversion of all data to general or study-specific common data models.[93] In the joint study protocol approach, researchers have the opportunity to apply exposure definitions locally that make optimal use of the data that is available in that database.[94,95] However, the disadvantage is that deviations in the effect estimate could occur due to small differences in the implementation of the study design.[13,79] Nowadays, common data models are becoming the standard approach when conducting multi-database studies. In a common data model, all data are converted into a standardized data structure with common table formats, meanings, and variable names across data partners.[96] While the advantages of common data models are the ease and speed of use,[97] the disadvantage is that information may be lost when data granularity differs between databases. For example, a database may contain detailed information about drug prescriptions, such as dose and duration, while another database may not. In order to harmonize the information in these databases, exposure definitions

can be chosen, which are executable in all participating databases, which may lead to loss of information. In addition, different vocabularies and coding systems may be used across the different databases, potentially leading to information loss due to incomplete conversions of these codes.[99] To offer guidance on how exposure data should be handled in multi-database studies using common data models, future studies should explore the impact of potential information loss due to varying definitions for exposures when using common data models.

In addition, there may be differences between these databases in the potential of exposure misclassification. For example, prescription data suffer from PNA more than dispensing databases do, whereas claims databases do not contain information on drugs that are not reimbursed.[97] Therefore, knowledge of the individual databases and the underlying differences in health care systems is needed to fully comprehend heterogeneity in results that may be obtained in multi-database studies.

Newer-generation drugs

Approximately a quarter of medicines that have entered the market in the EU and the US in the past years are biologicals.[100,101] In addition, until October 2020, 15 advanced-therapy medicinal products (ATMPs) were approved in the EU,[102] including several gene and cell therapies and issue-engineered products.[103] Especially for these types of drugs, the safety profile is often not fully known at the time of market authorization, for example because the safety cannot sufficiently be evaluated in animal models and RCTs. Therefore, PASs are even more important and even obligatory for these therapies compared to traditional small molecules. Biologicals and ATMPs differ from small molecules in various aspects. The half-life is often much longer, and the route of administration differs. For gene therapies, the exposure window may even be lifelong. These aspects should be considered in the definitions of the exposure window of biologicals and ATMPs.

However, there are challenges regarding the assessment of exposure to biologicals and ATMPs. In the Netherlands, for example, there is a regulation that these drugs are not dispensed by a first-line pharmacy, but exclusively by the out-patient pharmacy located in the hospital due to the high costs of biologicals.[104] In addition, these drugs are often prescribed by hospital specialists.[105] For instance, research in the UK shows that the current use of anti-TNF agent was documented in less than 10% in the GP system.[106] This makes databases such as GP prescription databases or pharmacy dispensing databases less suitable for answering questions regarding the effectiveness or safety of biologicals. Instead, exposure to biologicals and ATMPs is often recorded in specific registries, which are often disease-specific registries and do not capture the whole exposed population.[107] To enable PASs of biologicals and ATMPs and to properly measure both exposures and outcomes, new database linkages are needed between the hospital dispensing facilities and medical records from primary or secondary care or between registers and medical records.

The impact of the corona-virus pandemic on drug exposure assessment

The COVID-19 pandemic has had an impact on drug utilization patterns in different ways.[108] In Germany and the US, an overall increase in dispensings was observed just before and at

the start of the lockdown (i.e., the first weeks of March).[109,110] Drugs that were possibly related to COVID-19 outcomes (positively or negatively) showed large differences in the number of dispensings: statins, ACE-inhibitors, HCQ, and the combination of ritonavir/lopinavir showed increases in dispensings of up to 100% compared to the same period in 2019.[109] In the US, there was an increase in the number of prescriptions filled per week for antidepressant, anti-anxiety, and anti-insomnia medications during the first weeks of the pandemic.[111] This increase was also observed in the Netherlands, yet without an increase in new users, suggesting that the existing users stockpiled the drug.[112] The pandemic also had an impact on medication adherence. On the one hand, adherence among patients with asthma and chronic obstructive pulmonary disease increased by 14.3% for instance.[113] On the other hand, patients were more likely to discontinue the use of norgestrel-ethinylestradiol, dexmethylphenidate, and escitalopram.[114]

However, there is currently no comprehensive overview of drug utilization patterns during the COVID-19 pandemic for all frequently used drug classes to fully comprehend the changes in drug utilization that occurred. Creating such an overview is recommended to offer guidance on how data from this unique period should be handled in terms of exposure definition. For example, the disruptions in drug dispensing patterns can be handled by allowing for larger gaps or overlapping periods between dispensings, or by modeling exposure based on exposure patterns that can be expected from clinical practice. In addition, patients being more adherent or less adherent can also impact the effect estimates. These effects can be studied in sensitivity analyses, for instance by estimating drug–outcome relationship before, during, and after the pandemic.

CONCLUSIONS AND RECOMMENDATIONS

To conclude, there are different steps in the assessment of exposure, which can lead to exposure misclassification. Nevertheless, the impact thereof can be kept to a minimum. In this chapter, we present five recommendations for researchers, journal editors, and the readers of pharmacoepidemiologic studies.

First, knowledge of proper pharmacoepidemiologic methods is needed within a team of researchers to correctly design the study and to classify exposure. Biases such as immortal time bias can be avoided by adequate training and by using the existing guidelines. Journal editors and reviewers should have sufficient expertise to critically review the quality of submitted studies against the presence of potential biases that may arise in observational studies. In addition, journals should encourage or even oblige authors and reviewers to use of checklists, such as ROBINS-I or RECORD-PE, for pharmacoepidemiologic studies.[75,115]

Second, researchers must have knowledge of the clinical pharmacological aspects of the drug and the outcome to correctly define the exposure window and the risk window. This knowledge can come, for example, from clinical practice or case reports. For correct interpretation of this information, pharmacological expertise is needed within a research team. If the definition of the exposure window and risk window is uncertain, multiple exposure definitions can be used as a sensitivity analysis.[79]

Third, to understand which exposures are captured and which are not, researchers should have sufficient knowledge about the database that is used and the health system. For correct interpretation, these aspects should be clearly reported on, especially since many readers may not be aware of the differences between their "own" health system and the health system in the study setting. Limitations regarding the data availability should be reported on, and the potential impact of misclassification should be quantified, for example with sensitivity analyses.

Fourth, researchers should provide estimates of the extent of misclassification due to nonadherence and assess the possible impact on effect estimates in sensitivity analyses.

Fifth, readers of pharmacoepidemiologic studies, such as regulators, clinicians, and pharmacists, are encouraged to assess how the exposure is defined and the limitations thereof, for correct interpretation of the study results. This enables them to make the right decisions and inform patients properly about the effectiveness and safety of drugs, in relation to specific exposure patterns.

With these recommendations, exposure assessment in pharmacoepidemiologic studies and the interpretation thereof can be improved, contributing to increased reliability of information on the effective and safe use of drugs.

Author's contribution

The idea and set-up of the general discussion are by MH; she conducted the literature search and wrote the general discussion. During the whole process she implemented input and feedback from her PhD supervisors.

REFERENCES

1. Real-World Evidence | FDA. Accessed September 25, 2021. https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence
2. Pacurariu A, Plueschke K, McGettigan P, et al. Electronic healthcare databases in Europe: Descriptive analysis of characteristics and potential for use in medicines regulation. *BMJ Open*. 2018;8(9). doi:10.1136/bmjopen-2018-023090
3. Izem R, Huang T-YY, Hou L, Pestine E, Nguyen M, Maro JC. Quantifying how small variations in design elements affect risk in an incident cohort study in claims. *Pharmacoepidemiol Drug Saf*. 2020;29(1):84-93. doi:10.1002/PDS.4892
4. van Gaalen RD, Abrahamowicz M, Buckeridge DL. The impact of exposure model misspecification on signal detection in prospective pharmacovigilance. *Pharmacoepidemiol Drug Saf*. 2015;24(5):456-467. doi:10.1002/pds.3700
5. Rivera DR, McGlynn KA, Freedman AN. Connections between pharmacoepidemiology and cancer biology: designing biologically relevant studies of cancer outcomes. *Ann Epidemiol*. 2016;26(11):741-745. doi:10.1016/j.annepidem.2016.10.001
6. McMahon AD, Evans JMM, McGilchrist MM, McDevitt DG, MacDonald TM. Drug exposure risk windows and unexposed comparator groups for cohort studies in pharmacoepidemiology. *Pharmacoepidemiol Drug Saf*. 1998;7(4):275-280. doi:10.1002/(SICI)1099-1557(199807/08)7:4<275::AID-PDS363>3.0.CO;2-N
7. Gardarsdottir H, Souverein PC, Egberts TCGG, Heerdink ER. Construction of drug treatment episodes from drug-dispensing histories is influenced by the gap length. *J Clin Epidemiol*. 2010;63(4):422-427. doi:10.1016/j.jclinepi.2009.07.001
8. Vrijens B, De Geest S, Hughes DA, et al. A new taxonomy for describing and defining adherence to medications. *Br J Clin Pharmacol*. 2012;73(5):691-705. doi:10.1111/j.1365-2125.2012.04167.x
9. Gaster N, Hallas J, Pottegård A, Friis S, Schmidt M. The Validity of Danish Prescription Data to Measure Use of Aspirin and Other Non-Steroidal Anti-Inflammatory Drugs and Quantification of Bias Due to Non-Prescription Drug Use. 2021;2021(13):569-579. doi:10.2147/CLEP.S311450
10. Edwards IR, Aronson JK. Adverse drug reactions: Definitions, diagnosis, and management. *Lancet*. 2000;356(9237):1255-1259. doi:10.1016/S0140-6736(00)02799-9
11. Davies D. *Textbook of Adverse Drug Reactions*. 4th ed. (Davies D, ed.). Oxford University Press; 1991.
12. U.S. Department of Health and Human Services, Administration Food and Drug, Research (CDER) Center for Drug Evaluation and, Research (CBER) Center for Biologics Evaluation and. Guidance for Industry and FDA Staff Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data. *Drug Saf*. Published online 2013. Accessed June 18, 2021. http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm
13. European Medicines Agency. *The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology, 9th Revision*. Accessed August 4, 2021. http://www.encepp.eu/standards_and_guidance
14. Wang SV, Schneeweiss S, Berger ML, et al. Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies V1.0. *Pharmacoepidemiol Drug Saf*. 2017;26(9):1018-1032. doi:10.1002/pds.4295
15. Hallare J, Gerriets V. Half life. *Economist*. 2006;381(8503):81-82. doi:10.3366/edinburgh/9780748639854.003.0009
16. Castelli MS, McGonigle P, Hornby PJ. The pharmacology and therapeutic applications of monoclonal antibodies. *Pharmacol Res Perspect*. 2019;7(6):e00535. doi:10.1002/prp2.535
17. Stäubli M, Troendle A, Schmid B, et al. Pharmacokinetics of amiodarone, desethylamiodarone and other iodine-containing amiodarone metabolites. *Eur J Clin Pharmacol*. 1985;29(4):417-423.

18. van Heeswijk RPG, Dannemann B, Hoetelmans RMW. Bedaquiline: A review of human pharmacokinetics and drug-drug interactions. *J Antimicrob Chemother*. 2014;69(9):2310-2318. doi:10.1093/jac/dku171
19. Karbwang J, White NJ. Clinical Pharmacokinetics of Mefloquine. *Clin Pharmacokinet*. 1990;19(4):264-279. doi:10.2165/00003088-199019040-00002
20. Fookes C. Drug Half-life Explained. Published 2021. Accessed August 9, 2021. https://www.drugs.com/article/drug-half-life.html
21. Siccama R, Balk AH, de Herder WW, van Domburg R, Vantrimpont P, van Gelder T. Amiodarone therapy before heart transplantation as a predictor of thyroid dysfunction after transplantation. *J Hear Lung Transpl*. 2003;22(8):857-861.
22. Alonso A, MacLehose RF, Lutsey PL, Konety S, Chen LY. Association of amiodarone use with acute pancreatitis in patients with atrial fibrillation: a nested case-control study. *JAMA Intern Med*. 2015;175(3):449-450. doi:10.1001/jamainternmed.2014.6927
23. Lai S-W, Lin C-Y, Liao K-F. Amiodarone use and risk of acute pancreatitis: A population-based case-control study. *Hear Rhythm*. 2015;12(1):163-166. doi:10.1016/j.hrthm.2014.08.037
24. Douros A, Bronder E, Andersohn F, et al. Drug-induced liver injury: results from the hospital-based Berlin Case-Control Surveillance Study. *Br J Clin Pharmacol*. 2015;79(6):988-999. doi:10.1111/bcp.12565
25. Bouvy ML, Heerdink ER, Hoes AW, Leufkens HG. Amiodarone-induced thyroid dysfunction associated with cumulative dose. *Pharmacoepidemiol Drug Saf*. 2002;11(7):601-606. doi:10.1002/pds.735
26. Taylor CJ, Hodgkinson J, Hobbs FD. Rhythm control agents and adverse events in patients with atrial fibrillation. *Int J Clin Pr*. 2010;64(8):1069-1075. doi:10.1111/j.1742-1241.2010.02426.x
27. Black DM, Delmas PD, Eastell R, et al. Once-Yearly Zoledronic Acid for Treatment of Postmenopausal Osteoporosis. *N Engl J Med*. 2007;356(18):1809-1822. doi:10.1056/nejmoa067312
28. Strom BL, Kimmel SE, Hennessy S. *Pharmacoepidemiology*. 5th ed. Wiley-Blackwell; 2011.
29. World Health Organization. *Medicines Reimbursement Policies in Europe*. WHO Regional Office for Europe; 2018. Accessed October 11, 2021. https://www.euro.who.int/__data/assets/pdf_file/0011/376625/pharmaceutical-reimbursement-eng.pdf
30. Cepeda MS, Fife D, Denarié M, Bradford D, Roy S, Yuan Y. Quantification of missing prescriptions in commercial claims databases: results of a cohort study. *Pharmacoepidemiol Drug Saf*. 2017;26(4):386-392. doi:10.1002/pds.4165
31. Lauffenburger JC, Balasubramanian A, Farley JF, et al. Completeness of prescription information in US commercial claims databases. *Pharmacoepidemiol Drug Saf*. 2013;22(8):899-906. doi:10.1002/pds.3458
32. Okoli GN, Myles P, Murray-Thomas T, Shepherd H, Wong ICK, Edwards D. Use of Primary Care Data in Research and Pharmacovigilance: Eight Scenarios Where Prescription Data are Absent. *Drug Saf*. Published online July 22, 2021:1-8. doi:10.1007/s40264-021-01093-9
33. Gardarsdottir H, Egberts ACG, Heerdink ER. Transitions from General Practitioner to Psychiatrist Care (or vice versa) during a first antidepressant treatment episode. *Pharmacopsychiatry*. 2010;43(5):179-183. doi:10.1055/s-0030-1249096
34. Pottegård A, Olesen M, Christensen B, Christensen MB, Hallas J, Rasmussen L. Who prescribes drugs to patients: A Danish register-based study. *Br J Clin Pharmacol*. 2021;87(7):2982-2987. doi:10.1111/bcp.14691
35. Gagne JJ. Restrictive Reimbursement Policies: Bias Implications for Claims-Based Drug Safety Studies. *Drug Saf*. 2014;37(10):771-776. doi:10.1007/s40264-014-0220-5
36. Gamble JM, McAlister FA, Johnson JA, Eurich DT. Restrictive Drug Coverage Policies can Induce Substantial Drug Exposure Misclassification in Pharmacoepidemiologic Studies. *Clin Ther*. 2012;34(6):1379-1386. doi:10.1016/j.clinthera.2012.04.009
37. British Medical Association. *Over-the-Counter Medication*. Accessed October 11, 2021. https://www.researchgate.net/publication/281282792_Over-the-counter_medication_BMA_May_2005; 2005.

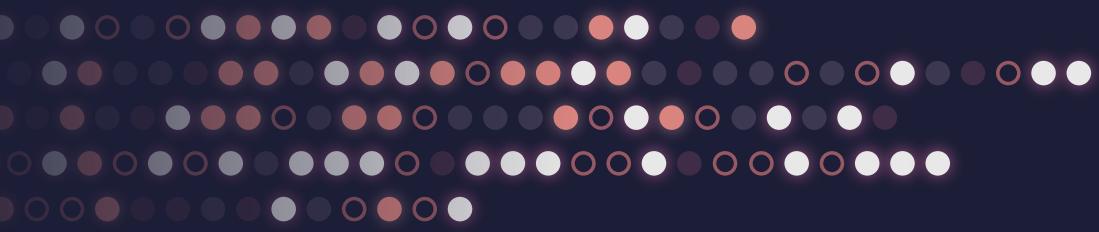
38. Wilcox CM, Cryer B, Triadafilopoulos G. Patterns of use and public perception of over-the-counter pain relievers: Focus on nonsteroidal antiinflammatory drugs. *J Rheumatol*. 2005;32(11):2218-2224.
39. Castelli G, Petrone A, Xiang J, Shrader C, King D. Rates of Nonsteroidal Anti-Inflammatory Drug Use in Patients with Established Cardiovascular Disease: A Retrospective, Cross-Sectional Study from NHANES 2009-2010. *Am J Cardiovasc Drugs*. 2017;17(3):243-249. doi:10.1007/s40256-016-0212-1
40. Sarganas G, Buttery AK, Zhuang W, et al. Prevalence, trends, patterns and associations of analgesic use in Germany. *BMC Pharmacol Toxicol*. 2015;16(28). doi:10.1186/s40360-015-0028-7
41. Tielemans MM, van Rossum LGM, Eikendal T, et al. Gastrointestinal symptoms in NSAID users in an "average risk population": results of a large population-based study in randomly selected Dutch inhabitants. *Int J Clin Pract*. 2014;68(4):512-519. doi:10.1111/ijcp.12346
42. Mainous AG, Baker R, Everett CJ, King DE. Impact of a policy allowing for over-the-counter statins. *Qual Prim Care*. 2010;18(5):301-306.
43. Solomon MD, Majumdar SR. Primary non-adherence of medications: Lifting the veil on prescription-filling behaviors. *J Gen Intern Med*. 2010;25(4):280-281. doi:10.1007/s11606-010-1286-0
44. Cahir C. Primary nonadherence: The forgotten component of medication adherence? *Polish Arch Intern Med*. 2020;130(1):1-3. doi:10.20452/pamw.15164
45. Hutchins DS, Zeber JE, Roberts CS, Williams AF, Manias E, Peterson AM. Initial Medication Adherence - Review and Recommendations for Good Practices in Outcomes Research: An ISPOR Medication Adherence and Persistence Special Interest Group Report. *Value Heal*. 2015;18(5):690-699. doi:10.1016/j.jval.2015.02.015
46. Lemstra M, Nwankwo C, Bird Y, Moraros J. Primary nonadherence to chronic disease medications: A meta-analysis. *Patient Prefer Adherence*. 2018;12:721-731. doi:10.2147/PPA.S161151
47. Pottegård A, Christensen R dePont, Houji A, et al. Primary non-adherence in general practice: a Danish register study. *Eur J Clin Pharmacol*. 2014;70(6):757-763. doi:10.1007/s00228-014-1677-y
48. Lemstra M, Blackburn D, Crawley A, Fung R. Proportion and Risk Indicators of Nonadherence to Statin Therapy: A Meta-analysis. *Can J Cardiol*. 2012;28(5):574-580. doi:10.1016/J.CJCA.2012.05.007
49. Briesacher BA, Gurwitz JH, Soumerai SB. Patients At-Risk for Cost-Related Medication Nonadherence: A Review of the Literature. *J Gen Intern Med 2007 226*. 2007;22(6):864-871. doi:10.1007/S11606-007-0180-X
50. Khera R, Valero-Elizondo J, Das SR, et al. Cost-Related Medication Nonadherence in Adults With Atherosclerotic Cardiovascular Disease in the United States, 2013 to 2017. *Circulation*. 2019;140(25):2067-2075. doi:10.1161/CIRCULATIONAHA.119.041974
51. Fischer MA, Choudhry NK, Brill G, et al. Trouble Getting Started: Predictors of Primary Medication Nonadherence. *Am J Med*. 2011;124(11):1081.e9-1081.e22. doi:10.1016/J.AMJMED.2011.05.028
52. Copeland KT, Checkoway H, Mcmichael AJ, Holbrook RH. Bias due to misclassification in the estimation of relative risk. *Am J Epidemiol*. 1977;105(5):488-495. doi:10.1093/oxfordjournals.aje.a112408
53. Brenner H. Inferences on the potential effects of presumed nondifferential exposure misclassification. *Ann Epidemiol*. 1993;3(3):289-294. doi:10.1016/1047-2797(93)90032-Y
54. Flegal KM, Brownie C, Haas J. The effects of exposure misclassification on estimates of relatwe risk. *Am J Epidemiol*. 1986;123(4):736-751. doi:10.1093/oxfordjournals.aje.a114294
55. Jurek AM, Greenland S, Maldonado G. Brief Report: How far from non-differential does exposure or disease misclassification have to be to bias measures of association away from the null? *Int J Epidemiol*. 2008;37(2):382-385. doi:10.1093/ije/dym291

56. Lau HS, De Boer A, Beuning KS, Porsius A. Validation of pharmacy records in drug exposure assessment. *J Clin Epidemiol*. 1997;50(5):619-625. doi:10.1016/S0895-4356(97)00040-1
57. Gamble JM, McAlister FA, Johnson JA, Eurich DT. Quantifying the impact of drug exposure misclassification due to restrictive drug coverage in administrative databases: A simulation cohort study. *Value Heal*. 2012;15(1):191-197. doi:10.1016/j.jval.2011.08.005
58. National Institute for Health and Care Excellence. Atrial fibrillation: management. NICE guideline CG180. 2014;(June 2014). Accessed April 24, 2018. https://www.nice.org.uk/guidance/cg180
59. Pharmaceutisch Weekblad [Pharmaceutical Weekly Magazine]. NHG: DOAC en cumarine zijn gelijkwaardig [NHG: DOAC and coumarin are equivalent]. Accessed April 24, 2018. https://www.pw.nl//nieuws/2016/nhg-noac-en-cumarine-zijn-gelijkwaardig
60. Gamble JM, Johnson JA, McAlister FA, Majumdar SR, Simpson SH, Eurich DT. Limited Impact of Drug Exposure Misclassification from Non-Benefit Thiazolidinedione Drug Use on Mortality and Hospitalizations from Saskatchewan, Canada: A Cohort Study. *Clin Ther*. 2015;37(3):629-642. doi:10.1016/j.clinthera.2014.12.014
61. Hernández-Díaz S. Estimation of Drug Exposure: Ascertainment and definition [PowerPoint slides]. Department of Epidemiology, Harvard School of Public Health, 2012.
62. Sotorra-Figuerola G, Ouchi D, Giner-Soriano M, Morros R. Impact of adherence to drugs for secondary prevention on mortality and cardiovascular morbidity: A population-based cohort study. IMPACT study. *Pharmacoepidemiol Drug Saf*. 2021;30(9):1250-1257. doi:10.1002/pds.5261
63. Graham DY, Smith JL. Aspirin and the stomach. *Ann Intern Med*. 1986;104(3):390-398. doi:10.7326/0003-4819-104-3-390
64. Maitland-Van Der Zee AH, Klungel OH, Stricker BHC, et al. Comparison of two methodologies to analyze exposure to statins in an observational study on effectiveness. *J Clin Epidemiol*. 2004;57(3):237-242. doi:10.1016/j.jclinepi.2003.07.011
65. Hernán MA, Hernández-Díaz S. Beyond the intention-to-treat in comparative effectiveness research. In: *Clinical Trials*. Vol 9. NIH Public Access; 2012:48-55. doi:10.1177/1740774511420743
66. Suissa S. Immortal time bias in pharmacoepidemiology. *Am J Epidemiol*. 2008;167(4):492-499. doi:10.1093/aje/kwm324
67. Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiol Drug Saf*. 2007;16(3):241-249. doi:10.1002/pds.1357
68. Messmer BJ, Nora JJ, Leachman RD, Cooley DA. Survival-times after cardiac allografts. *Lancet*. 1969;1(7602):954-956. doi:10.1016/s0140-6736(69)91857-1
69. Levesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ*. 2010;340(mar121):b5087-b5087. doi:10.1136/bmj.b5087
70. Martinuka O, von Cube M, Wolkewitz M. Methodological evaluation of bias in observational coronavirus disease 2019 studies on drug effectiveness. *Clin Microbiol Infect*. 2021;27(7):949-957. doi:10.1016/j.cmi.2021.03.003
71. Rosenberg ES, Dufort EM, Udo T, et al. Association of Treatment With Hydroxychloroquine or Azithromycin With In-Hospital Mortality in Patients With COVID-19 in New York State. *JAMA*. 2020;323(24):2493-2502. doi:10.1001/jama.2020.8630
72. Catteau L, Dauby N, Montourcy M, et al. Low-dose hydroxychloroquine therapy and mortality in hospitalised patients with COVID-19: a nationwide observational study of 8075 participants. *Int J Antimicrob Agents*. 2020;56(4):106144. doi:10.1016/j.ijantimicag.2020.106144
73. European Medicines Agency. *The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology, 9th Revision.*; 2010. Accessed June 10, 2021. http://www.encepp.eu/standards_and_guidances
74. Hemkens LG, Benchimol EI, Langan SM, et al. The reporting of studies using routinely collected health data was often insufficient. *J Clin Epidemiol*. 2016;79:104-111. doi:10.1016/J.JCLINEPI.2016.06.005

75. Langan SM, Schmidt SA, Wing K, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). *BMJ*. 2018;363:k3532. doi:10.1136/bmj.k3532
76. Petrone AB, DuCott A, Gagne JJ, Toh S, Maro JC. The Devil's in the details: Reports on reproducibility in pharmacoepidemiologic studies. *Pharmacoepidemiol Drug Saf*. 2019;28(5):671-679. doi:10.1002/pds.4730
77. de Vries F, de Vries C, Cooper C, Leufkens B, van Staa T-P. Reanalysis of two studies with contrasting results on the association between statin use and fracture risk: the General Practice Research Database. *Int J Epidemiol*. 2006;35(5):1301-1308. doi:10.1093/ije/dyl147
78. Pye SR, Sheppard T, Joseph RM, et al. Assumptions made when preparing drug exposure data for analysis have an impact on results: An unreported step in pharmacoepidemiology studies. *Pharmacoepidemiol Drug Saf*. 2018;27(7):781-788. doi:10.1002/pds.4440
79. Eskin M, Simpson SH, Eurich DT. Impact of drug exposure definitions on observed associations in pharmacoepidemiology research. *J Popul Ther Clin Pharmacol*. 2018;25(1):e39-e52. doi:10.22374/1710-6222.25.1.4
80. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLOS Med*. 2015;12(10):e1001885. doi:10.1371/journal.pmed.1001885
81. Agha RA, Fowler AJ, Limb C, et al. Impact of the mandatory implementation of reporting guidelines on reporting quality in a surgical journal: A before and after study. *Int J Surg*. 2016;30:169-172. doi:10.1016/J.IJSU.2016.04.032
82. Plint AC, Moher D, Morrison A, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust*. 2006;185(5):263-267. doi:10.5694/J.1326-5377.2006.TB00557.X
83. Moher D, Jones A, Lepage L. Use of the CONSORT statement and quality of reports of randomized trials. *J Am Med Assoc*. 2001;285(15):1992-1995. doi:10.1001/jama.285.15.1992
84. Alam M, Rauf M, Ali S, et al. A Systematic Review of Completeness of Reporting in Randomized Controlled Trials in Dermatologic Surgery: Adherence to CONSORT 2010 Recommendations. *Dermatol Surg*. 2016;42(12):1325-1334. doi:10.1097/DSS.0000000000000902
85. Weisman A, King LK, Mamdani M. Reporting and variability of constructing medication treatment episodes in pharmacoepidemiology studies: A methodologic systematic review using the case study of DPP-4 inhibitors and cardiovascular outcomes. *Pharmacoepidemiol Drug Saf*. 2020;29(8):939-950. doi:10.1002/PDS.5071
86. British Journal of Clinical Pharmacology - Author Guidelines. Accessed October 1, 2021. https://bpspubs.onlinelibrary.wiley.com/hub/journal/13652125/author-guidelines.html#PREPARING_YOUR
87. Pharmacoepidemiology and Drug Safety - Author Guidelines. Accessed October 1, 2021. https://onlinelibrary.wiley.com/page/journal/10991557/homepage/forauthors.html
88. Annals of Pharmacotherapy - Author Guidelines. Accessed October 1, 2021. doi:10.7326/0003-4819 https://journals.sagepub.com/pb-assets/cmscontent/AOP/AOP%20Submission%20Guidelines.pdf
89. Drug Safety | Submission guidelines. Accessed October 1, 2021. https://www.springer.com/journal/40264/submission-guidelines
90. Bazelier MT, Eriksson I, de Vries F, et al. Data management and data analysis techniques in pharmacoepidemiological studies using a pre-planned multi-database approach: A systematic literature review. *Pharmacoepidemiol Drug Saf*. 2015;24(9):897-905. doi:10.1002/pds.3828
91. Data Analysis and Real World Interrogation Network (DARWIN EU) | European Medicines Agency. Accessed September 27, 2021. https://www.ema.europa.eu/en/about-us/how-we-work/big-data/data-analysis-real-world-interrogation-network-darwin-eu
92. Lai EC-C, Stang P, Yang Y-HK, Kubota K, Wong ICK, Setoguchi S. International Multi-database Pharmacoepidemiology: Potentials and Pitfalls. *Curr Epidemiol Reports*. 2015;2(4):229-238. doi:10.1007/s40471-015-0059-z

93. Gini R, Sturkenboom MCJ, Sultana J, et al. Different Strategies to Execute Multi-Database Studies for Medicines Surveillance in Real-World Setting: A Reflection on the European Model. *Clin Pharmacol Ther*. 2020;108(2):228-235. doi:10.1002/cpt.1833
94. Souverein PC, van den Ham HA, Huerta C, et al. Comparing risk of major bleeding between users of different oral anticoagulants in patients with nonvalvular atrial fibrillation. *Br J Clin Pharmacol*. 2020;87(3):988-1000. doi:10.1111/bcp.14450
95. Margulis A V., Linder M, Arana A, et al. Patterns of use of antimuscarinic drugs to treat overactive bladder in Denmark, Sweden, and the United Kingdom. *PLoS One*. 2018;13(9). doi:10.1371/journal.pone.0204456
96. Sentinel Common Data Model | Sentinel Initiative. Accessed August 10, 2021. https://www.sentinelinitiative.org/methods-data-tools/sentinel-common-data-model
97. Schneeweiss S, Brown JS, Bate A, Trifirò G, Bartels DB. Choosing Among Common Data Models for Real-World Data Analyses Fit for Making Decisions About the Effectiveness of Medical Products. *Clin Pharmacol Ther*. 2020;107(4):827-833. doi:10.1002/cpt.1577
98. Zhou X, Murugesan S, Bhullar H, et al. An Evaluation of the THIN Database in the OMOP Common Data Model for Active Drug Safety Surveillance. *Drug Saf 2012 362*. 2013;36(2):119-134. doi:10.1007/S40264-012-0009-3
99. Zhou X, Murugesan S, Bhullar H, et al. An evaluation of the THIN database in the OMOP common data model for active drug safety surveillance. *Drug Saf*. 2013;36(2):119-134. doi:10.1007/s40264-012-0009-3
100. Mullard A. 2018 FDA drug approvals. *Nat Rev Drug Discov*. 2019;18(2):85-89. doi:10.1038/d41573-019-00014-x
101. Will Biologics Surpass Small Molecules In The Pharma Race? | BioPharmaTrend. Accessed September 27, 2021. https://www.biopharmatrend.com/post/67-will-small-molecules-sustain-pharmaceutical-race-with-biologics/
102. Warreth S, Harris E. The Regulatory Landscape for ATMPs in the EU and US: a Comparison. *Level*. 2020;15(2):Article 5. doi:10.21427/PK3V-G445
103. European Parliament and the Council of the European Union. *Regulation (EC) No 1394/2007 of the European Parliament and of the Council of 13 November 2007 on Advanced Therapy Medicinal Products and Amending Directive 2001/83/EC and Regulation (EC) No 726/2004. Official Journal of the European Union, L 324/.*; 2007. Accessed September 28, 2021. https://eur-lex-europa-eu.proxy.library.uu.nl/LexUriServ/LexUriServ.do?uri=OJ:L:2007:324:0121:0137:en:PDF.
104. Regelgeving dure geneesmiddelen | Medisch-specialistische zorg | Nederlandse Zorgautoriteit [Regulations for expensive medicines | Medical specialist care | Dutch Healthcare Authority]. Accessed August 10, 2021. https://www.nza.nl/zorgsectoren/medisch-specialistische-zorg/registreren-en-declareren-van-geneesmiddelen/regelgeving-dure-geneesmiddelen
105. Gerriets V, Bansal P, Goyal A, Khaddour K. Tumor Necrosis Factor Inhibitors. *StatPearls*. Published online 2021. Accessed August 10, 2021. http://www.ncbi.nlm.nih.gov/pubmed/29494032
106. McCalla P. Rapid response to: Biologicals for rheumatoid arthritis. *BMJ*. 2011;343(d4027). doi:10.1136/bmj.d4027
107. Elseviers M, Poluzzi E, Wettermark B, et al. *Drug Utilization Research: Methods and Applications*. 1st ed.; 2016.
108. Webster-Clark M. Ways COVID-19 may impact unrelated pharmacoepidemiologic research using routinely collected data. *Pharmacoepidemiol Drug Saf*. 2021;30(3):400-401. doi:10.1002/PDS.5182
109. Enners S, Gradl G, Kieble M, Böhm M, Laufs U, Schulz M. Utilization of drugs with reports on potential efficacy or harm on COVID-19 before, during, and after the first pandemic wave. *Pharmacoepidemiol Drug Saf*. Published online July 21, 2021:pds.5324. doi:10.1002/pds.5324
110. Vaduganathan M, Van Meijgaard J, Mehra MR, Joseph J, O'Donnell CJ, Warraich HJ. Prescription Fill Patterns for Commonly Used Drugs during the COVID-19 Pandemic in the United States. *JAMA*. 2020;323(24):2524-2526. doi:10.1001/jama.2020.9184

111. America's State of Mind Report | Express Scripts. Accessed August 9, 2021. https://www.express-scripts.com/corporate/americas-state-of-mind-report
112. Hek K, Weesie Y, Lambooij A, et al. *Voorschrijven van Psychofarmaca Door de Huisarts Tijdens de Coronapandemie. Geneesmiddelenmonitor in Coronatijd. Week 2-24 (6 Januari - 12 Juni 2020).*; 2020.
113. Kaye JA, Margulis A V., Fortuny J, et al. Cancer Incidence after Initiation of Antimuscarinic Medications for Overactive Bladder in the United Kingdom: Evidence for Protopathic Bias. *Pharmacother J Hum Pharmacol Drug Ther*. 2017;37(6):673-683. doi:10.1002/phar.1932
114. Clement J, Jacobi M, Greenwood BN. Patient access to chronic medications during the Covid-19 pandemic: Evidence from a comprehensive dataset of US insurance claims. Leong C, ed. *PLoS One*. 2021;16(4):e0249453. doi:10.1371/journal.pone.0249453
115. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355. doi:10.1136/bmj.i4919



5.

SUMMARIES

5.1
ENGLISH SUMMARY

Randomized controlled trials (RCTs) are conducted to demonstrate the efficacy and safety of drugs, and they are used to support market approval of drugs. However, after market approval, unanswered questions often remain regarding the safety and effectiveness of the drug, for example regarding the effectiveness in subjects who do not meet the RCT eligibility criteria, or regarding rare adverse events and delayed effects. Therefore, post-authorization observational studies can be conducted to augment evidence generated by RCTs. These studies take advantage of real-world data (RWD) that represent data on drug use and clinical outcomes in daily clinical practice. The FDA defines RWD as *"data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources."* These RWD are, for example, obtained from the electronically registered health care usage of patients, including pharmacy records, general practitioner (GP) records, medical claims records, and disease registries. The advantage of these data is that they has already been collected and are therefore relatively quick and easy to use. Another advantage is that, compared to RCTs, observational studies can apply less strict selection criteria. However, the disadvantage is that these data are not primarily collected for research purposes. Information needed when generating evidence on the use of medicines, such as clinical outcomes, treatment, and other characteristics, may consequently be inaccurate or missing. This can impact the estimation of the relation between the drug and the outcome in studies that assess the safety and effectiveness of drugs.

In this thesis, we focus on how drug exposure, or the use of a drug by a patient, is measured in observational studies. Methodological choices and assumptions when defining drug exposure may have an impact on the estimated effect associated with drug use. When assessing the association between exposure and clinical outcomes, certain aspects of their possible relation should be considered, including the pharmacodynamic and pharmacokinetic properties of the drug under study, as well as the timing of the outcome, as the outcome can be a direct reaction to the drug (e.g., anaphylactic reactions) or may be the result of long-term use (e.g., cancer outcomes). In addition, the translation of exposure information contained in the databases to day-to-day patient drug use patterns should be done with care, as exposure status retrieved from these databases may not always reflect the actual exposure status. Mismatches may exist between actual exposure status and assumed exposure status when subjects are using drugs for which information about the exposure is not captured in the database. Moreover, subjects may be misclassified as being "exposed" to a specific drug when, in reality, they are nonadherent to the drug prescription.

Since methodological choices underlying the exposure assessment impact the estimate of the exposure–outcome relationship, transparent reporting is important for the interpretation of published study results and for reproducibility and validity assessment.

The aims of this thesis are hence a) to investigate the current reporting of drug exposure assessment and the risk of bias in pharmacoepidemiologic studies and b) to explore the extent and impact of misclassification of exposure on risk estimates in pharmacoepidemiology. **Chapter 2** provides a review of the reporting of exposure assessment and biases in published pharmacoepidemiologic studies, and in **Chapter 3**, different sources of exposure misclassification are studied, along with the impact thereof. **Chapter 4** then provides a general discussion of the findings within a broader perspective.

QUALITY OF DRUG EXPOSURE REPORTING AND STUDY DESIGN

Chapter 2.1 contains our evaluation of the quality of reporting of drug exposure assessment in pharmacoepidemiologic studies. We systematically reviewed observational pharmacoepidemiologic studies that used routinely collected health data, published in 2017, in six pharmacoepidemiologic journals. The quality of reporting was assessed within 11 items regarding exposure assessment as defined by the ISPE-ISPOR guideline. We found that none of the 91 assessed studies reported on all 11 items, with great variation between studies (range two to 10). Studies more often reported conceptual details about the exposure risk window and the exposure assessment window (85% and 98%, respectively) than on the operational details concerning the construction of the exposure risk window. For example, the handling of gaps and overlapping treatment episodes were reported in only 11% and 41% of studies in which this type of reporting was applicable. Better reporting of exposure ascertainment in pharmacoepidemiologic studies is needed to allow for correct interpretation of the results and to enable both reproducibility and validity assessment.

Chapter 2.2 presents our assessment of the quality of observational studies that measured the effectiveness of in-hospital use of hydroxychloroquine in COVID-19 clinical outcomes. The relation between quality and the reported effect sizes was also assessed. The quality of the 33 included studies was assessed for seven items, based on the ROBINS-I tool: confounding, selection bias, classification of interventions and outcomes, deviation from intended intervention, missing data, and reporting. Effects sizes reported in the observational studies were compared with effect sizes from RCTs on the same outcomes, and differences in these effect sizes were related to apparent study quality. None of the included observational studies were found to be completely free of risk of bias. The median number of domains in which there was risk of bias per individual study was three (range: one to six). The identified increased risk of bias most often involved bias due to confounding (76%) and bias due to the classification of interventions, such as immortal time bias (67%). We also observed that studies with effect estimates that diverged less from the RCT estimates had a higher overall quality than studies that diverged more. As the results of observational studies can quickly find their way into daily practice, the results of biased studies can have potentially harmful consequences for patients. In this chapter, we provide recommendations regarding the appropriateness, quality and reporting of observational studies to improve the validity of published study results.

THE EXTENT AND IMPACT OF EXPOSURE MISCLASSIFICATION

The extent of exposure misclassification and the impact thereof on effect estimates is explored in **Chapter 3**. *Chapter 3.1* presents our assessment of the extent of exposure misclassification due to primary nonadherence (PNA), which is defined as not filling the first prescription for a drug treatment. PNA can lead to exposure misclassification when databases of prescribing physicians are used, as information about dispensing from pharmacies is not known, which consequently may lead to biased estimates. For this study, we included patients from the Nivel Primary Care Database who received a new prescription (>1 year not prescribed) from a GP in 2012. These new prescriptions were linked to public pharmacy dispensing information. We found that 11.5% of

all newly prescribed drugs did not have a record of dispensing in the pharmacy database within 30 days of the prescription date. Among specific drug classes that are frequently prescribed in primary care, PNA was found to be 9.9%, with the lowest level of PNA for thyroid hormones (5.5%), and the highest for PPIs (12.8%). In addition, several patient- and prescription-related characteristics were associated with PNA, such as having more than three active diagnoses registered in the GP database (odds ratio [OR] 1.46, [95% confidence interval (CI) 1.37–1.56]), compared to no active diagnoses), receiving a prescription for drugs that are not reimbursed (OR 2.78 [95% CI 2.65–2.92], compared to fully reimbursed), and the timing of prescribing (OR for drugs prescribed in the last quarter of 2012 0.68 [95% CI 0.65–0.71], compared to the first quarter).

Chapter 3.2 then assesses the impact of accounting for drug half-life on exposure misclassification. For this study, amiodarone, which has a half-life of 60 days, was used as a case study. When exposure to amiodarone is only defined based on day-to-day drug intake, the exposure classification may inadequately reflect actual exposure status, as patients are physically exposed over a longer period of time due to the long half-life. In this study, we applied 18 different exposure definitions to define amiodarone exposure, with and without accounting for the half-life, and we estimated the risk of acute pancreatitis. The hazard ratio (HR) for comparing current to noncurrent use was 1.36 (95% CI 0.78–2.38), while the HR for the model that accounted for the half-life – by gradually reducing the assumed concentration present in the body – was 1.21 (95% CI 0.69–2.10), showing limited difference. This may partly be explained by the fact that almost a third of the population (28.8%) was exposed to amiodarone during their complete (>95%) follow-up, leaving no room for lag periods or changes in the kinetic dose. Sensitivity analyses in which subjects exposed for more than 95% of their follow-up time were excluded, however, did not result in different estimates. To further investigate the impact of different exposure definitions of amiodarone, we recommend replication of this study in larger databases, with more variation in amiodarone use.

In *Chapter 3.3*, we assess the impact of exposure misclassification that occurs when information about prescribing originates from a single care setting (e.g., primary care), while the treatment under investigation can be prescribed by prescribers from various settings. In this study, we assessed exposure misclassification of direct oral anticoagulants (DOACs), which are prescribed by GPs and specialists (cardiologists, internists, or orthopedics), and we estimated the impact on the risk of major bleeding and stroke or transient ischemic attack (TIA). The PHARMO Database Network, including information on the type of prescriber for dispensed drugs, was used for this study. Separate analyses were carried out, including either all anticoagulant dispensings or only a subset of anticoagulant dispensings that were prescribed by either a GP or a specialist. We found that patient characteristics differed between subjects treated only by the GP, only by the specialist or treated by both, with different patterns of use observed between initiators of vitamin K antagonists (VKAs) and DOACs. In general, VKA users who received their VKA prescriptions only by a specialist had more comorbidities and used more comedication than patients who only received their VKA prescriptions from a GP. For the prescribing of DOACs, the opposite was seen: patients who only received DOACs from a specialist generally had fewer comorbidities and received fewer comedications. Yet, compared to the use of all dispensing information, the inclusion of only anticoagulant dispensings prescribed by GPs did not significantly influence the effect estimates of DOAC use compared to VKA use on the risk of major bleeding and stroke/

TIA (HR 0.79 [95% CI 0.70–0.90] for inclusion of all dispensing information vs 0.78 (95% CI 0.68–0.91) for inclusion of only GP prescribed dispensings. The inclusion of only anticoagulant dispensings prescribed by specialists did, however, change the HR to 0.62 (95% CI 0.50–0.76). Therefore, we recommend using databases that are as complete as possible in terms of prescriptions history for patients without regards to type of prescriber to avoid exposure misclassification and, as a result, biased results. Since misclassification in a particular database that contains only prescriptions of one health care setting is likely to be drug- and context-specific, we also recommend further research, for example with other drug exposure–outcome relations or other databases.

While *Chapters 3.2 and 3.3* evaluate the impact of exposure misclassification due to nonadherence and uncaptured exposures in two case studies, using RWD, *Chapter 3.4* investigates the relative impact of misclassification caused either by nonadherence or by uncaptured exposures in a simulation study. To this end, we constructed a hypothetical cohort based on assumed exposure prevalence and (relative) risks on the outcomes. In this cohort, we simulated exposure misclassification either by misclassifying exposed subjects as unexposed (mimicking uncaptured exposures) or by misclassifying unexposed subjects as exposed (mimicking nonadherence). All utilized values were based on values found in the literature on exposure misclassification. Three scenarios were explored to assess the impact of exposure misclassification: nondifferential misclassification, differential misclassification (misclassifications dependent on an unmeasured risk factor doubling the outcome risk), and nondifferential misclassification in a comparative effectiveness study. In all scenarios studied, the values for nonadherence and uncaptured data or the differences in these values between subgroups needed to be relatively large to lead to clinically relevant bias. With estimates of the degree of misclassification, for example from pilot studies or published results of drug utilization research, a simple bias analysis can provide insight into the impact of exposure misclassification on the effect estimates. To this end, we provide tables and figures, which can be used to assess the impact of the different sources of misclassification, using values for exposure prevalence, the proportion of nonadherence, and uncaptured data.

DISCUSSION

Chapter 4 contextualizes our findings within a broader perspective. First, we describe the role of pharmacokinetics in defining exposure. Although we did not observe an impact of a long half-life on effect estimates, the potential impact thereof was demonstrated for other drugs. Knowledge of the clinical pharmacological aspects of the drug and the outcome is therefore needed to correctly define the exposure window and the risk window in order to correctly estimate the relationship between drug exposure and clinical outcomes.

Second, we discuss the extent of exposure misclassification due to nonadherence or uncaptured exposures. For both sources, the extent of exposure misclassification differs between drugs, healthcare systems, and databases. Therefore, sufficient knowledge about the origin of the database and the underlying health care system is required to understand which exposures are captured, which are not captured, and what can be expected when it concerns patient nonadherence to treatment.

Third, we explain the impact of these different sources of exposure misclassification on the effect estimates. We conclude that the impact of exposure misclassification due to uncaptured data on effect estimates is limited. Future research should focus on the impact of misclassification due to nonadherence, to better comprehend the impact thereof. Researchers should provide estimates of the extent of misclassification due to nonadherence or uncaptured data and assess the possible impact on effect estimates in sensitivity analyses.

Fourth, we discuss the impact of immortal time bias due to exposure misclassification. Although many guidelines on pharmacoepidemiologic studies offer recommendations on how to avoid immortal time bias, we conclude that this type of bias still frequently occurs. To avoid biases such as immortal time bias, proper knowledge of pharmacoepidemiologic methods is necessary within a team of researchers to correctly design the study and to classify exposure correctly.

Fifth, we discuss the reporting of exposure in pharmacoepidemiologic studies. The incomplete reporting in these studies hinders the correct interpretation, reproducibility, and assessment of validity of the study results. Authors, reviewers, and editors are therefore encouraged to adhere to relevant reporting guidelines, such as the ISPE-ISPOR and RECORD-PE guidelines. Journals could, for example, consider obliging authors to use one of these two guidelines. In addition, for correct interpretation of the study results, readers of pharmacoepidemiologic studies, such as regulators, clinicians, and pharmacists, are encouraged to assess how the exposure is defined and the limitations thereof. This enables them to make the right decisions and properly inform patients about the efficacy and safety of drugs in relation to specific exposure patterns.

Last, we identify areas within exposure assessment that will remain challenging. These include drug exposure measurement in multi-database studies when the individual databases contain different types and granularity of information on drug exposures. Evaluation of exposure to newer-generation drugs, which are often not dispensed in primary care and therefore are obtained in regional or local databases, will pose challenges in gaining both access and the means to combine results from these databases with results from other (inter)national or regional databases. In addition, pandemics such as the COVID-19 pandemic may have impact on how medicines are prescribed and dispensed. Further research is needed to assess how data generated during such circumstances should be used for pharmacoepidemiologic research.

In conclusion, this thesis demonstrates that exposure assessment in pharmacoepidemiologic studies and the interpretation thereof can be improved, thereby contributing to increased reliability of information on the effective and safe use of drugs.

5.2

NEDERLANDSE SAMENVATTING

Voordat een geneesmiddel wordt toegelaten op de markt, moet eerst de werkzaamheid en veiligheid aangetoond worden. Dit wordt meestal onderzocht in gerandomiseerde gecontroleerde klinische trials (RCT's). Toch zijn er na toelating op de markt vaak nog onbeantwoorde vragen over de veiligheid en effectiviteit. Zo is niet altijd zeker of het geneesmiddel ook effectief is bij patiënten die buiten de studiepopulatie vallen, of zijn zeldzame of bijwerkingen die pas na langere tijd optreden niet uitgebreid onderzocht. Deze aanvullende vragen kunnen worden beantwoord met behulp van observationele onderzoeken. Deze studies kunnen gebruik maken van gegevens uit de dagelijkse klinische praktijk, ook wel *'real world data'* genoemd. *Real world data* zijn gegevens over patiënten die in de dagelijkse klinische praktijk behandeld worden, zonder een strikt gedefinieerd protocol en zonder strikte in- en exclusiecriteria zoals in RCT's. Deze gegevens worden bijvoorbeeld verkregen uit dossiers van apothekers of huisartsen, declaratiegegevens van zorgverzekeraars en specifieke ziekte- of geneesmiddelregistraties. Het voordeel van deze gegevens is dat ze al verzameld zijn, en daarmee dus relatief snel en eenvoudig te gebruiken zijn. Een ander voordeel is dat in, vergelijking met RCT's, in observationeel onderzoek minder strikte selectiecriteria toegepast kunnen worden. Het nadeel is echter dat deze gegevens niet primair verzameld zijn voor onderzoeksdoeleinden, maar voor registratie van zorgactiviteiten. Daarom kan de informatie die nodig is voor het beantwoorden van de onderzoeksvraag onnauwkeurig zijn of ontbreken, bijvoorbeeld informatie over uitkomsten waarvoor geen behandeling is gezocht, de (daadwerkelijk) gekregen behandeling of patiëntkarakteristieken die niet zijn geregistreerd. Dit kan vervolgens weer invloed hebben op de schatting van de relatie tussen het geneesmiddelgebruik en de uitkomst.

In dit proefschrift richten we ons op de bepaling van blootstelling aan geneesmiddelen in observationele studies. We hebben gekeken naar hoe onderzoekers geneesmiddelgebruik definiëren en bepalen in observationele studies. Daarnaast hebben we bestudeerd wat de invloed is van methodologische keuzes en veronderstellingen bij het definiëren van blootstelling aan geneesmiddelen op het geschatte effect dat samenhangt met het gebruik van dat geneesmiddel. Er zijn verschillende aspecten waar rekening mee moet worden gehouden bij het definiëren van blootstelling aan geneesmiddelen. Een van die aspecten is de tijdsrelatie tussen het geneesmiddelgebruik en de uitkomst, waarbij farmacodynamische en farmacokinetische eigenschappen van het geneesmiddel een rol spelen. De uitkomst kan bijvoorbeeld een directe reactie op het geneesmiddel zijn, zoals een anafylactische reactie, of kan het resultaat zijn van langdurig gebruik, zoals carcinogene effecten. Een tweede aspect betreft de vertaling van de informatie uit databases met voorschrijf- of afleverinformatie naar verwachte gebruiksduur en -patronen bij individuele patiënten. Daarnaast kan misclassificatie optreden wanneer geneesmiddelgebruik niet in de gebruikte database is vastgelegd, of wanneer patiënten wel een recept ontvangen, maar de medicijnen niet innemen.

Omdat de methodologische keuzes rondom de blootstellingsbepaling impact hebben op de effectschatting tussen geneesmiddelblootstelling en uitkomsten, is het van belang hier transparant over te rapporteren. Alleen zo is een juiste interpretatie van de gepubliceerde onderzoeksresultaten maar ook reproduceerbaarheid en validiteitsbeoordeling mogelijk.

De doelstellingen van dit proefschrift zijn daarom als volgt: 1) onderzoeken hoe de bepaling van blootstelling wordt gerapporteerd in farmacoepidemiologische studies en wat het risico

op bias is in die studies (**Hoofstuk 2**) en 2) onderzoeken wat de mate van misclassificatie van geneesmiddelblootstelling is en de impact hiervan op effectschattingen (**Hoofdstuk 3**). In **Hoofdstuk 4** worden de bevindingen uit de Hoofdstukken 2 en 3 in een breder perspectief besproken.

KWALITEIT VAN RAPPORTAGE VAN BLOOTSTELLING AAN GENEESMIDDELEN EN ONDERZOEKSDESIGN

Hoofdstuk 2.1 omvat ons onderzoek naar de kwaliteit van het rapporteren van de blootstellingsbepaling in farmacoepidemiologische onderzoeken. Op systematische wijze hebben we observationele farmacoepidemiologische onderzoeken beoordeeld die routinematig verzamelde gezondheidsgegevens gebruikten. Hiervoor keken we naar alle observationele onderzoeken die werden gepubliceerd in 2017 in zes farmacoepidemiologische tijdschriften. De kwaliteit van het rapporteren werd beoordeeld op 11 verschillende items zoals gedefinieerd door de IPSE-ISPOR richtlijn. Geen van de 91 geïncludeerde studies rapporteerden over alle 11 items, en er was grote variatie tussen deze onderzoeken (spreiding: 2– 10 items per studie). Studies rapporteerden vaker over conceptuele keuzes, bijvoorbeeld over de tijdvensters waarin de blootstelling en de uitkomsten werden bepaald (respectievelijk 98% en 85%) dan over hoe dit blootstellings-risicovenster precies werd geconstrueerd. Als voorbeeld, hoe omgegaan is met gaten tussen twee recepten of overlappende periodes werd slechts gerapporteerd in 11% en 41% van de geïncludeerde studies. Beter rapporteren van de blootstellingsbepaling in farmacoepidemiologische onderzoeken is dus nodig voor correcte interpretatie van de resultaten, de mogelijkheid om de studie te reproduceren en de validiteit te kunnen beoordelen.

Hoofdstuk 2.2 presenteert onze beoordeling van de kwaliteit van observationele onderzoeken die de effectiviteit van hydroxychloroquine bestudeerden op klinische uitkomsten van COVID-19 tijdens ziekenhuisopname. Daarnaast hebben we gekeken naar de relatie tussen kwaliteit van deze studies en de uiteindelijke effectschatting. De kwaliteit van de 33 geïncludeerde onderzoeken werd beoordeeld op 7 items, gebaseerd op de ROBINS-I-tool: confounding, selectiebias, classificatie van interventies en van uitkomsten, afwijking van de beoogde interventie, ontbrekende gegevens en het rapporteren van de studieopzet en de resultaten. De effectschattingen die gevonden werden in deze observationele studies hebben we vergeleken met schattingen gevonden in RCT's voor dezelfde uitkomsten. Geen van de geïncludeerde observationele studies bleek volledig vrij te zijn van risico op bias en liep dus risico op een systematische vertekening van de relatie tussen hydroxychloroquine en klinische uitkomsten. Het mediane aantal domeinen met risico op bias per individuele studie was 3 (spreiding: 1–6 domeinen). Bias door confounding (76%) en bias door de misclassificatie van interventies, zoals *immortal time bias* (67%), werden het meest geobserveerd. We zagen ook een relatie tussen de kwaliteit van de studies en de effectschattingen: studies die effectschattingen vonden die minder afweken van de RCT-schattingen hadden een hogere kwaliteit dan studies die meer afweken. Omdat resultaten van observationeel onderzoek snel vertaald kunnen worden naar klinische behandeladviezen, kunnen resultaten van kwalitatief minder onderzoek mogelijk schadelijke gevolgen hebben voor patiënten. In dit hoofdstuk hebben we daarom aanbevelingen gedaan rondom de geschiktheid, de kwaliteit en het rapporteren van observationeel onderzoek, om zo de validiteit van gepubliceerde onderzoeksresultaten te verbeteren.

DE MATE EN IMPACT VAN MISCLASSIFICATIE VAN GENEESMIDDELBLOOTSTELLING

De mate van blootstellingsmisclassificatie en de impact daarvan op effectschattingen zijn bestudeerd in **Hoofdstuk 3**. *Hoofdstuk 3.1* presenteert onze studie naar de mate van misclassificatie die optreedt door primaire therapieontrouw. Primaire therapieontrouw wordt vaak gedefinieerd als het niet ophalen van een nieuw voorgeschreven recept in de apotheek. Dit kan leiden tot misclassificatie van blootstelling wanneer voorschrijfdatabases gebruikt worden. Voor dit onderzoek hebben we patiënten uit de Nivel Eerstelijns Zorgdatabase geïncludeerd die in 2012 een nieuw recept van een huisarts hebben gekregen, wat gedurende minimaal 1 jaar niet voorgeschreven was. Deze voorschrijfinformatie hebben we gekoppeld aan afleverinformatie uit de apotheek. We ontdekten dat 11,5% van alle nieuw voorgeschreven geneesmiddelen niet binnen 30 dagen door de apotheek was verstrekt. Wanneer we keken naar specifieke geneesmiddelklassen die veel worden voorgeschreven in de huisartsenpraktijk, bleek 9,9% van alle voorschriften niet binnen 30 dagen te worden opgehaald. Dit was het laagste voor schildklierhormonen (5,5%) en het hoogste voor protonpompremmers (12,8%). Daarnaast waren verschillende patiënt- en recept-gerelateerde kenmerken geassocieerd met primaire therapieontrouw: patiënten met drie of meer actieve diagnoses in de huisartsendatabase haalden hun recepten vaker niet op dan patiënten zonder actieve diagnoses. Daarnaast haalden patiënten die niet-vergoede geneesmiddelen kregen voorgeschreven of die in het eerste kwartaal van 2012 werden voorgeschreven hun recepten ook minder vaak op. De impact van de misclassificatie die hierdoor kan ontstaan in het gebruik van voorschrijfdatabases is nader onderzocht in *Hoofdstuk 3.4*.

Hoofdstuk 3.2 bespreekt vervolgens de impact van het betrekken van de halfwaardetijd van geneesmiddelen op effectschattingen. Als casestudie is hiervoor amiodaron gebruikt dat een halfwaardetijd heeft van 60 dagen. De daadwerkelijke blootstelling aan amiodaron kan daarom niet alleen worden bepaald op basis van dagelijkse inname, aangezien patiënten ook nadat ze zijn gestopt met amiodaron nog lange tijd worden blootgesteld aan dit middel vanwege de lange halfwaardetijd. In deze studie hebben we 18 verschillende blootstellingsdefinities toegepast om de blootstelling aan amiodaron te definiëren, en het risico op acute pancreatitis geschat. De hazard ratio (HR) voor het vergelijken van huidig (inname vandaag ja/nee) met niet-huidig gebruik was 1,36 (95% betrouwbaarheidsinterval [BI] 0,78-2,38), terwijl de HR voor het model dat wel rekening hield met de halfwaardetijd door de veronderstelde dosering in het lichaam geleidelijk te verminderen 1,21 was (95% BI 0,69– 2,10), een beperkt verschil dus. Dit beperkte verschil kan gedeeltelijk worden verklaard door het feit dat bijna een derde van de studiepopulatie (28,8%) werd blootgesteld aan amiodaron tijdens hun volledige (>95%) follow-up, waardoor de verschillende definities bij deze patiënten weinig verschil maakten. We hebben daarom gevoeligheidsanalyses uitgevoerd waarbij proefpersonen die meer dan 95% van hun follow-uptijd waren blootgesteld werden uitgesloten, wat overigens niet tot andere schattingen leidde. Om te beoordelen of verschillende blootstellingsdefinities van amiodaron daadwerkelijk geen impact hebben op de effectschatting, is replicatie van dit onderzoek in een grotere database nodig, met meer amiodaron gebruikers en vooral meer variatie in de gebruikspatronen.

In *Hoofdstuk 3.3* beoordelen we de impact van misclassificatie van blootstelling die optreedt wanneer voorschrijfinformatie afkomstig is vanuit één type zorginstelling (bijvoorbeeld de eerstelijns zorg), terwijl de onderzochte behandeling door voorschrijvers uit verschillende zorginstellingen (eerste of tweedelijns zorg) kan worden voorgeschreven. Als casestudie gebruikten we hiervoor het voorschrijven van directe orale anticoagulantia (DOAC's), en het risico op een ernstige bloeding en beroerte of *transient ischemic attack* (TIA), aangezien DOAC's worden voorgeschreven door zowel huisartsen als specialisten (cardiologen, internisten of orthopeden). Voor dit onderzoek is gebruik gemaakt van het PHARMO Database Network, waarin onder andere informatie is opgenomen over afgeleverde geneesmiddelen, inclusief het type voorschrijver. Aparte analyses zijn uitgevoerd waarin óf alle afgeleverde recepten voor antistollingsmiddelen meegenomen worden óf alleen een subset van antistollingsmiddelen die door een huisarts of specialist waren voorgeschreven. We vonden dat patiëntkenmerken verschilden tussen patiënten die alleen door de huisarts, alleen door de specialist, of door beide werden behandeld en dat de coumarine- en DOAC-gebruikers hierin van elkaar verschilden. Over het algemeen hadden coumarinegebruikers die onder behandeling bij de specialist zijn meer comorbiditeiten en gebruikten zij meer comedicatie dan coumarinegebruikers die onder behandeling zijn van de huisarts. Bij de DOAC-gebruikers werd het tegenovergestelde gezien: patiënten die hun DOAC's voorgeschreven kregen door de specialist, hadden over het algemeen minder comorbiditeiten en kregen minder comedicatie dan DOAC-gebruikers die onder behandeling waren bij de huisarts. Dit had echter geen groot effect op de effectschattingen van DOAC-gebruik in vergelijking met coumarine-gebruik op het risico op ernstige bloedingen en beroerte of TIA. De HR wanneer alleen huisartsenrecepten werden gebruikt was 0,79 (95% BI: 0,70–0 ,90), vergelijkbaar met de HR wanneer alle aflevergegevens werden meegenomen (HR 0,78 [95% BI: 0,68–0,91]). Alleen recepten meenemen die door specialisten voorgeschreven werden, veranderde de effectschatting wel (HR 0,62 [95% BI: 0,50–0,76]). Het is dus aan te raden om bij blootstelling aan DOACs voor een database te kiezen die zo volledig mogelijk is wat betreft informatie over voorschriften, ongeacht het type voorschrijver. Daarnaast raden we verder onderzoek aan, bijvoorbeeld met andere blootstelling–uitkomst relaties of met gebruik van andere databases, aangezien misclassificatie in een zorginstelling-specifieke database waarschijnlijk specifiek is voor een bepaald geneesmiddel en bepaalde context.

In tegenstelling tot *Hoofdstuk 3.2 en 3.3,* waarin de impact van misclassificatie beoordeeld is in twee casestudies met gegevens uit de klinische praktijk, bestuderen we de impact van misclassificatie in *Hoofdstuk 3.4* met een simulatiestudie. Hiervoor hebben we een hypothetisch cohort geconstrueerd, gebaseerd op een veronderstelde blootstellingsprevalentie en veronderstelde (relatieve) risico's op de uitkomsten. In dit cohort hebben we misclassificatie van blootstelling gesimuleerd door daadwerkelijk blootgestelde patiënten verkeerd te classificeren als "niet blootgesteld" (om zo misclassificatie door niet-vastgelegde blootstellingen na te bootsen) of door daadwerkelijk niet blootgestelde patiënten verkeerd te classificeren als "blootgesteld" (om zo misclassificatie door therapieontrouw na te bootsen). Alle gebruikte waarden voor de mate van misclassificatie zijn gebaseerd op de literatuur. We hebben drie verschillende scenario's onderzocht: niet-differentiële misclassificatie, differentiële misclassificatie (hierbij is de kans op misclassificatie afhankelijk van een niet-gemeten risicofactor die het risico op de uitkomst verdubbelt) en niet-differentiële misclassificatie in een studie waarin twee behandelingen met elkaar vergeleken worden. In alle bestudeerde scenario's moest de mate van therapieontrouw, niet-vastgelegde gegevens of de verschillen tussen

subgroepen relatief groot zijn om tot een klinisch relevante vertekening van de effectschatting te leiden. Pilotstudies of gepubliceerde resultaten rondom geneesmiddelgebruik kunnen inzicht geven in de mate van misclassificatie. Deze waarden kunnen vervolgens in een simpele bias-analyse worden gebruikt om de impact van misclassificatie op de effectschattingen te bepalen. Om andere onderzoekers daarin te faciliteren hebben we hiervoor verschillende tabellen en figuren gemaakt, waarmee de impact van misclassificatie geschat kan worden op basis van waarden voor blootstellingsprevalentie, het percentage therapieontrouw en het percentage niet-vastgelegde gegevens.

DISCUSSIE

In **Hoofdstuk 4** bespreken we onze bevindingen in een breder perspectief. Allereerst gaan we dieper in op de rol van farmacokinetiek bij het definiëren van blootstelling. Hoewel in onze studie over amiodaron geen effect van een lange halfwaardetijd op effectschattingen werd waargenomen, werd het potentiële effect hiervan wel gezien bij andere geneesmiddelen. Het is daarom van belang kennis te hebben van de klinische en farmacologische aspecten van zowel het geneesmiddel als de uitkomst. Deze kennis is nodig om het blootstellingsvenster en het risicovenster juist te kunnen definiëren om zo de relatie tussen geneesmiddelblootstelling en klinische uitkomsten juist te kunnen schatten.

Ten tweede bespreken wat de gevolgen zijn van misclassificatie van blootstelling door therapieontrouw of niet-vastgelegde geneesmiddelblootstellingen. Voor beide bronnen verschilt de mate van misclassificatie van blootstelling tussen geneesmiddelen, gezondheidszorgsystemen en databases. Er is daarom voldoende kennis nodig over de herkomst van de gebruikte database en het onderliggende zorgsysteem om te begrijpen welke blootstellingen wel en niet worden vastgelegd en welke mate van therapietrouw verwacht kan worden.

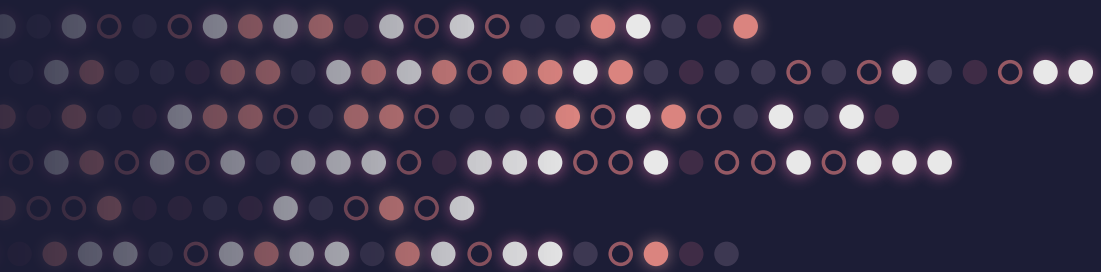
Als derde bespreken we de impact van deze verschillende bronnen van blootstellings-misclassificatie op de effectschattingen. We concludeerden dat blootstellings-misclassificatie door niet-vastgelegde gegevens slechts beperkte impact op effectschattingen heeft. Over de impact van misclassificatie door therapieontrouw is minder bekend en toekomstig onderzoek zou zich hierop moeten concentreren. Daarnaast adviseren we om schattingen te geven van de verwachte mate van misclassificatie door therapieontrouw en niet-vastgelegde geneesmiddelblootstellingen. Deze waardes kunnen vervolgens gebruikt worden in gevoeligheidsanalyses om de mogelijke impact van misclassificatie op de gevonden effectschatting te beoordelen.

Als vierde bespreken we de impact van *immortal time bias* als gevolg van blootstellingsmisclassificatie. Hoewel veel richtlijnen voor farmacoepidemiologische studies aanbevelingen geven hoe dit type bias te vermijden, concludeerden wij dat deze bias nog steeds vaak voorkomt. Om bias zoals *immortal time bias* te voorkomen, is een goede kennis van farmacoepidemiologische methoden nodig binnen een onderzoeksteam voor een juiste studieopzet en bepaling van blootstelling.

Ten vijfde bespreken we het rapporteren van de bepaling van blootstelling in farmacoepidemiologische onderzoeken. Dit rapporteren is vaak onvolledig en belemmert een juiste interpretatie, reproduceerbaarheid en validiteitsbeoordeling van de onderzoeksresultaten. Auteurs, reviewers en redacteuren worden daarom aangemoedigd om zich te houden aan relevante richtlijnen rondom het rapporteren van de bepaling van geneesmiddelblootstelling, zoals de ISPE-ISPOR- en RECORD-PE-richtlijnen. Tijdschriften zouden auteurs bijvoorbeeld kunnen verplichten een van deze twee richtlijnen te hanteren en het invullen van checklists op basis van deze richtlijnen bij het inleveren van een manuscript verplicht te maken. Daarnaast moedigen we lezers van farmacoepidemiologische onderzoeken, zoals regelgevers, artsen en apothekers aan om te beoordelen hoe de blootstelling wordt gedefinieerd en wat de beperkingen van die methode zijn, om zo de onderzoeksresultaten juist te kunnen interpreteren. Dit stelt hen vervolgens in staat de juiste farmacotherapeutische beslissingen te nemen en patiënten goed te informeren over de effectiviteit en veiligheid van geneesmiddelen in relatie tot specifieke blootstellingspatronen.

Als laatste identificeren we gebieden waar uitdagingen zullen blijven bestaan rondom het bepalen van blootstelling. Dit betreft bijvoorbeeld het meten van geneesmiddelblootstelling in onderzoeken waarin meerdere databases worden gebruiken. Een andere uitdaging is hoe om te gaan met verschillen tussen verschillende databases wat betreft het type en detailniveau van de blootstellingsinformatie. Daarnaast zal ook de bepaling van blootstelling aan nieuwe generatie geneesmiddelen een uitdaging vormen. Deze middelen worden vaak niet in de eerste lijn verstrekt. Informatie over de blootstelling hieraan wordt daarom vaak verkregen door middel van lokale of regionale databases. Zowel het verkrijgen van toegang als ook de wijze waarop deze resultaten gecombineerd kunnen worden met resultaten van andere (inter) nationale of regionale databanken, zal methodologische uitdagingen opleveren voor de toekomst. Tot slot kunnen pandemieën, zoals de COVID-19-pandemie, impact hebben op de manier waarop geneesmiddelen worden voorgeschreven en verstrekt. Er traden veranderingen op in de gebruikspatronen en de therapietrouw van veel geneesmiddelen. De impact hiervan op geneesmiddelengebruik en effecten daarvan zal moeten worden beoordeeld en op basis daarvan zal moeten worden bepaald hoe gegevens die tijdens dergelijke omstandigheden gegenereerd zijn op de juiste wijze gebruikt kunnen worden in farmacoepidemiologisch onderzoek.

Kortom, bepaling van blootstelling in farmacoepidemiologische onderzoeken en het rapporteren hiervan kunnen worden verbeterd, om zo bij te dragen aan meer betrouwbare informatie ter onderbouwing van effectief en veilig gebruik van geneesmiddelen.



6.

APPENDICES

6.1. DANKWOORD

De afgelopen zes jaar heb ik aan dit proefschrift gewerkt. Dat heb ik uiteraard niet alleen gedaan. Ik ben daar direct of indirect door veel mensen in begeleid en ondersteund. Of het nu ging om inhoudelijke begeleiding, een luisterend oor of een kopje koffie. Ik wil iedereen die een rol heeft gespeeld in het tot stand komen van mijn proefschrift hartelijk bedanken. Een paar mensen wil ik specifiek bedanken:

Mijn grootste dank gaat natuurlijk uit naar mijn promotieteam bestaande uit prof. dr. Olaf Klungel, prof. dr. Ton de Boer en dr. Helga Gardarsdottir.

Olaf, jij kwam met het idee voor de richting van mijn proefschrift: iets met longitudinale bepaling van geneesmiddelblootstelling. Je gaf mij binnen dit onderwerp de ruimte een eigen pad te kiezen en zelf aan de slag te gaan, wat ik erg fijn vond. Tegelijkertijd bewaakte je de kwaliteit en de focus van het geheel. Onder jouw supervisie kon ik er op vertrouwen dat de studies methodologisch gezien goed in elkaar zaten.

Ton, dit proefschrift was niet de eerste keer waarin je mij begeleidde. Ook tijdens mijn bachelorwerkstuk ben ik door jou begeleid. En in beide trajecten heb ik veel van je geleerd. Je stelde kritische vragen bij gemaakte keuzes, speelde soms advocaat van de duivel en liet mij zo met een frisse blik naar mijn eigen keuzes kijken. Hierbij gaf je me altijd de ruimte en het vertrouwen om die eigen keuzes te maken.

Helga, vanaf 2017 was jij mijn dagelijkse begeleider. Onder het genot van goede kopjes koffie hebben wij veel over mijn studies gesproken, gebrainstormd en gediscussieerd. Je nam de tijd om mijn stukken goed te lezen om me zo met gedegen feedback verder te helpen. Je daagde me uit op een andere manier te kijken, keer op keer wees je me om de lezer voor ogen te houden en niet te methodologisch te schrijven. Zo heb je mij geholpen om dit voor leken taaie onderwerp veel toegankelijker te maken.

Rolf, je hebt een belangrijke bijdrage geleverd aan dit proefschrift, waarvan in het begin als dagelijkse begeleider. Bijzonder vond ik het dat je betrokken bleef als coauteur bij alle studies ook toen je hoogleraar werd in Leiden en niet meer formeel mijn begeleider was. Je wist de studies snel focus te geven wanneer die er nog niet was. Had ik een vraag dan kwam er altijd antwoord, helder en eenvoudig. Ik heb veel van je methodologische kennis en je feedback kunnen leren.

Kim, Patrick, Simone en Karin ook jullie wil ik graag bedanken voor jullie rol in dit proefschrift. Kim, bedankt voor de gezellige tijd die we hadden tijdens het schrijven van de review in hoofdstuk 2.1. Patrick, bedankt voor jouw waardevolle hulp bij het analyseren van de data in hoofdstuk 3.3 en al je hulp bij vele andere hoofdstukken. Simone, dank je wel voor je input op de studie in hoofdstuk 3.1. Karin, bedankt voor alle hulp en feedback die ik van je mocht ontvangen bij hoofdstuk 3.1, ook al waren de deadlines zo aan het einde van mijn proefschrift soms erg kort.

Fijne collega's, bedankt voor de gezelligheid en het prettige samenwerken waar ik nog steeds erg van geniet. Lieve mededocenten, dank jullie wel voor de ruimte die ik aan het einde van dit

traject kreeg om het proefschrift af te ronden, en jullie betrokkenheid als ik dagen lang achter elkaar zat te schrijven.

Lieve Lenneke, als vriendin, roomie, collega en paranimf wil ik jou bedanken voor de afgelopen jaren waarin we samen optrokken. Dank je wel voor alle momenten dat we samen konden sparren, dat je met me meedacht en dat je nu mijn paranimf bent! Joanne, al sinds de basisschool zijn wij vriendinnen, en delen we de kleine en grote dingen in het leven. Bedankt voor je altijd oprechte aandacht en interesse, ook tijdens het schrijven van mijn proefschrift. Ik vind het fijn dat je nu mijn paranimf bent! Emilie, begeleid door vele koppen thee en bagels hebben wij de afgelopen jaren onze kijk op de wereld kunnen delen en zo ook over mijn proefschrift kunnen doorpraten. Het is altijd fijn jouw genuanceerde mening hierover te horen, dank je wel daarvoor. Jennine, wat fijn dat je nu weer in Nederland woont en we elkaar weer vaker in het echt kunnen zien. Jij schreef je proefschrift in Stockholm, ik ondertussen in Utrecht. Het was mooi om dat, ondanks de afstand, te kunnen delen. Ook alle vrienden die ik nu niet bij naam heb genoemd wil ik graag bedanken voor alle gezellige momenten die we de afgelopen jaren hebben mogen delen.

Lieve familie en schoonfamilie, ook jullie wil ik heel erg bedanken voor alle betrokkenheid de afgelopen jaren. Het was fijn alle vorderingen met jullie te kunnen delen. Pap, mam, jullie hebben het traject vanaf het begin met aandacht gevolgd, dank jullie wel daarvoor. Mam, Eline en Simon, heel erg bedankt voor alle extra oppas in de laatste maanden van het schrijven, zo kreeg ik tijd en focus om mijn general discussion te kunnen schrijven.

En tot slot wil ik mijn eigen gezin bedanken voor alles de afgelopen jaren. Lieve Pepijn en Floris, jullie hebben de afgelopen jaren mijn leven verrijkt en daar ben ik heel blij mee! Dank jullie wel voor alle lieve knuffels, het spelen met de Lego en al jullie verhalen. Lieve Rieko, heel erg bedankt voor al je mentale support die je me de afgelopen jaren hebt gegeven. Jij was er altijd, als er iets tegenzat of als er juist iets te vieren was. Mijn dank daarvoor is onbeschrijfelijk groot! En natuurlijk ook bedankt voor je grafische vormgeving, ook voor dit boekje.

Lieve lezers, jullie wil ik bedanken voor het lezen van dit dankwoord. Ik hoop dat jullie met net zoveel plezier de andere delen van dit proefschrift zullen lezen!

6.2. LIST OF CO-AUTHORS

Anthonius de Boer
Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Utrecht University;
Medicines Evaluation Board of the Netherlands.

Helga Gardarsdottir
Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Utrecht University;
Department of Clinical Pharmacy, Division Laboratory and Pharmacy, University Medical Center Utrecht;
Faculty of Pharmaceutical Sciences, School of Health Sciences, University of Iceland.

Rolf H.H. Groenwold
Department of Clinical Epidemiology, Leiden University Medical Center

Karin Hek
Department of Integrated Primary Care, Nivel (Netherlands Institute for Health Services Research), Utrecht, The Netherlands

Olaf H. Klungel
Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Utrecht University;
Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht

Kim Luijken
Department of Clinical Epidemiology, Leiden University Medical Center

Simone Rijken
Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Utrecht University

Patrick C. Souverein
Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Utrecht University

6.3. LIST OF PUBLICATIONS

Hempenius M, Groenwold RHH, de Boer A, Klungel OH, Gardarsdottir H. Amiodarone use and the risk of acute pancreatitis: Influence of different exposure definitions. *Pharmacoepidemiol Drug Saf* 2019;28(12):1563-1571.

Hempenius M, Luijken K, de Boer A, Klungel O, Groenwold R, Gardarsdottir H. Quality of reporting of drug exposure in pharmacoepidemiological studies. *Pharmacoepidemiol Drug Saf* 2020;29(9):1141-1150.

Hempenius M, Groenwold RHH, Souverein PC, de Boer A, Klungel OH, Gardarsdottir H. Impact of anticoagulant exposure misclassification on the bleeding risk of direct oral anticoagulants. *Br J Clin Pharmacol* 2021;87(9):3508-3517.

Hempenius M, Groenwold RHH, de Boer A, Klungel OH, Gardarsdottir H. Drug exposure misclassification in pharmacoepidemiology: Sources and relative impact. *Pharmacoepidemiol Drug Saf* 2021 Sep 7. Online ahead of print.

6.4. ABOUT THE AUTHOR

Mirjam Hempenius was born in 1988 in Kampen, the Netherlands. In 2011, she obtained her Bachelor's, and in 2014 her Master's degree in Pharmacy at the Utrecht University, both cum laude. During her studies, she was awarded the FACF Went prize for an excellent bachelor thesis and the "*KNMP Studentenprijs*."

After her graduation, she worked as a non-dispensing clinical pharmacist in a general practice as part of the POINT study. The aim of this new function was to reduce hospitalizations linked to medication and medication-related problems.

In 2015, she was selected for a tenure track position at the Division of Pharmacoepidemiology and Clinical Pharmacology at Utrecht University. Her tenure track was divided equally between teaching and performing research.

In her role as a teacher, she has been involved in several activities, including tutoring and supervising students, teaching, as well as coordinating and being the examinator of a master's course. She is member of the board of examiners of the Master of Pharmacy and a member of the learning line "pharmacotherapeutic patient care," which is responsible for safeguarding education regarding pharmacotherapeutic patient care. Mirjam is also responsible for the development of (new) educational content within the Master of Pharmacy program. She obtained her Basis Qualification for Academic Teaching in 2017.

Mirjam carried out her PhD project under the supervision of Olaf Klungel, Ton de Boer, and Helga Gardarsdottir. Her research focused on exposure misclassification in pharmacoepidemiologic database studies. During her PhD, she obtained a postgraduate master's degree in epidemiology from the Julius Centre, University Medical Centre Utrecht in 2021. She presented her work at national and international conferences.

Mirjam currently lives in Zeist, together with her husband Rieko and their children Pepijn and Floris. After completion of this PhD thesis, she will continue working as a teacher.