



Methods for constructing treatment episodes and impact on exposure-outcome associations

Laura Pazzagli¹ · Lena Brandt¹ · Marie Linder¹ · David Myers² · Panagiotis Mavros³ · Morten Andersen^{1,4} · Shahram Bahmanyar^{1,5}

Received: 22 March 2019 / Accepted: 10 October 2019 / Published online: 22 November 2019
© The Author(s) 2019

Abstract

Purpose To assess the impact on exposure time and outcome misclassifications, and consequent impact on exposure-outcome associations from treatment episode construction. We investigated the dosage assumptions of 1 unit per day, and 1 DDD per day, versus actual prescribed dosage under different handling of gaps and overlaps of prescriptions.

Methods Data on mirtazapine and citalopram exposure (years 2006–2014) from the Swedish Prescribed Drug register were used. Via a within individuals design we compared method A, based on actual dosage, with methods B and C based on 1 unit of drug per day and 1 DDD per day assumptions, respectively, including consideration of gaps and overlaps. Four outcomes were used, hospitalizations and outpatient visits for all and for psychiatric causes.

Results Relative to method A, both alternative methods lead to misclassification of exposure time. With regard to outcome misclassifications, method B overestimates the effect of the exposure on the outcome in 77% and 100% of exposure definition comparisons for mirtazapine and citalopram respectively, while 23% of the comparisons for mirtazapine results in underestimation of exposure-outcome associations. Conversely, treatment episodes based on DDD (method C) result in underestimation of the exposure-outcome association in 100% and 87.5% of exposure definition comparisons for mirtazapine and citalopram respectively, while 12.5% of the comparisons for citalopram results in overestimation of the exposure-outcome associations.

Conclusions The study provides results that have consistent clinical relevance. We have showed that a non-accurate construction of exposure time may lead to errors on outcome detection during exposed time, and consequently affect conclusions on safety or efficacy profile of a treatment.

Keywords Treatment episode · Defined daily dose · Gaps · Overlaps · Treatment duration · Pharmacoepidemiology

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00228-019-02780-4>) contains supplementary material, which is available to authorized users.

✉ Laura Pazzagli
laura.pazzagli@ki.se

¹ Centre for Pharmacoepidemiology, Department of Medicine Solna, Karolinska Institutet, S-171 76 Stockholm, Sweden

² The Janssen Pharmaceutical Companies of Johnson & Johnson, Janssen Cilag AB, Stockholm, Sweden

³ Janssen Scientific Affairs, Titusville, NJ, USA

⁴ Department of Drug Design and Pharmacology, University of Copenhagen, Copenhagen, Denmark

⁵ The Center for Psychiatry Research, Karolinska Institutet, Stockholm, Sweden

Introduction

In the analysis of observational data, the choice of the method to construct treatment episodes is crucial to avoid misclassification of exposure time and to reduce the impact on the estimation of exposure-outcome associations, defined as any difference in the occurrence of the outcome between groups of individuals that differ by exposure. In studies using individual level data from prescription registers for evaluating the association of a pharmacological treatment with an outcome of interest, the treatment episode's construction is based on available data on dispensed quantity, dosage, or duration of the treatment [1]. However, in many prescription registers information about the actual prescribed dosage or duration of exposure is not available and the approximation of an individual's treatment

episode length is based on assumptions about the prescribed dosage. Moreover, the use of prescription registers to construct treatment episodes is based on the assumption that individuals will use the medication included in the filled prescriptions.

A common assumption to approximate the actual prescribed dosage is based on the defined daily dose (DDD), a statistical measure introduced by the WHO Collaborating Centre for Drugs Statistics Methodology. Formally, “the DDD is the assumed average maintenance dose per day for a drug used for its main indication in adults” [2]. Accordingly, the dosage is assumed to be one DDD per day [3–6]. The use of the average maintenance dosage to approximate an individual’s treatment episode duration may lead to differing conclusions [7–10]. Another commonly used dosage assumption is a fixed number of units of drug per day, as for example one unit per day, essentially equating days of supply per prescription fill to the number of prescribed units (tablets/pills/capsules) [11, 12].

When constructing treatment episodes, in addition to dosage assumptions, the presence of gaps and overlaps among prescriptions should be carefully considered [13, 14]. Overlaps refer to time spans under which a patient has supplies from two or more prescriptions of the same drug. Gaps refer to the temporal distance between two prescriptions that is in excess of the days of supply of the first prescription. In the presence of multiple prescriptions, the treatment episode duration is affected by the way prescription overlaps are handled and by assumptions about allowable gaps.

In the current study, we seek to assess the impact of the assumptions of one unit per day, and one DDD per day, versus actual prescribed dosage, on treatment episode durations, while at the same time we systematically vary the length of the allowable gaps between prescriptions, and the approach for handling overlaps. Resulting differences in treatment episode durations constructed under different assumptions may contribute to misclassification of exposure time and affect epidemiologic measures of association between exposure and outcomes. The current study also assesses the odds of observing the outcome of interest during the exposed time under different assumptions for treatment episode construction. The hypothesis behind the current study is that the use of different dosage assumptions results in different treatment episode durations, and thereby potential over- or underestimation of measures of rates of outcomes due to misclassification of the exposure time. To illustrate, we use mirtazapine and citalopram, two antidepressant medications commonly used to treat major depressive disorders.

Methods

Data sources

This was a non-interventional observational study utilizing population data from the Swedish Prescribed Drug Register (PDR), the National Patient Register (NPR), and the Total Population Register (TPR). Data sources are described in Appendix A of the supplementary material.

Study population

All individuals in the PDR with at least one prescription for mirtazapine or citalopram in tablet form from January 2006 to July 2014 were included. Included individuals were required to be new users, defined as no treatment with mirtazapine or citalopram for at least 6 months prior to the first observed prescription, and with available information on actual prescribed dosage in unstructured free-text format (information on medication dosages are included in Appendix A of the supplementary material). Additionally, included individuals had no record of migration registered in the TPR, and did not have dispensing of pre-packed daily supplies of the drugs of interest.

Text-mining algorithm for extraction of the actual prescribed dosage

In the PDR, the amount of dispensed drug is available as number of units of drug, unit strength and number of dispensed DDDs. A semi-manual, iterative method was used to ascertain the actual prescribed dosage for the particular drug from available unstructured free-text format data. Accordingly, a proportion of the dosage texts was read to create a look-up table associating dose text and prescribed dosage (numerical, number of units) of medication per day. Text that did not contain any information on dosage or was otherwise uninformative was omitted. The remaining prescriptions were merged with the look-up table of the manually extracted dosages. The result was manually proofread [15]. Information on dosages for mirtazapine and citalopram are reported in Appendix A.

Treatment episode construction

When constructing treatment episodes using register data, the main source is often an administrative record of prescription fills, which allows to aggregate the days of supply over all eligible prescription fills, to approximate the episode durations. Prescription eligibility is determined based on pre-specified assumptions regarding allowable maximum gaps between prescriptions. Aggregation may or may not account for overlaps between prescriptions.

In this study, the duration of each prescription filled was approximated and compared under three different dosage determination methods: actual prescribed dosage (method A), one unit per day assumption (method B), and one DDD per day assumption (method C). Method A provides the reference method in determining the actual duration of each prescription dispensed. The duration for each prescription under each of these approaches is given by:

$$\begin{array}{l} \text{Method A } D = \frac{p \times u}{r} \\ \text{Method B } D = p \times u \\ \text{Method C } D = p \times d \end{array}$$

where D is the duration of a prescription fill, p is the number of the prescribed packages, u is the number of units of drug per package, r is the actual prescribed number of units per day, and d is the number of DDDs per package as recorded in the PDR.

Under each of the three dosage determination methods, different sub-methods to account for gaps and overlaps between prescriptions were evaluated. The variability in treatment episodes duration was assessed for allowable gaps of different lengths, specifically 0, 10, 30, 60, 90, and 180 days. Allowing for a gap of a certain number of days means that two consecutive prescriptions will be included in the same treatment episode if the gap between them is less than or equal to the allowed gap. Two different approaches were applied to evaluate the impact of overlaps on the duration of treatment episodes. One approach accounts for overlapping days between prescriptions by adding them at the end of the duration of the episode, while the other ignores any overlapping days. Prescriptions filled on the same date were handled as overlaps. All the investigated methods (A, B, and C) assume as starting time of a treatment episode, the date when individuals first collect the drug from the pharmacy.

For each included individual, the duration of the first treatment episode was approximated by aggregating days of supply over all eligible prescriptions while systematically varying the length of the allowed gap and the handling of overlaps. The study design gave rise to a total of 36 (3 dosage methods \times 6 allowable gaps \times 2 ways for handling overlaps) alternative constructions of the first treatment episode for each of the two included medications. The application of the methods is illustrated in Appendix F of the supplementary material via R scripts used to produce the results.

Descriptive analysis

Box and whisker plots were used to display descriptive statistics (median, 1st, and 3rd quartiles) for the duration of the first treatment episode for each of the 36 combinations.

Study design and statistical modeling

To assess the impact of misclassification of exposure time from treatment episode construction, we performed a within individuals study which allows to compare individuals with themselves under different exposure definitions. To assess whether treatment episodes constructed under different assumptions impact the odds of observing the outcome of interest during exposed time, unadjusted conditional logistic regression models were fitted using generalized estimating equations conditioning at the individual level. Any misclassification of exposure (relative to the reference method) will lead to estimates of odds ratios (ORs) different than one. Conditioning at the individual level, the OR of observing the outcome as exposed was assessed for different pairs of treatment episodes constructed under different assumptions. In Appendix F of the supplementary material is illustrated as the models have been fitted via R scripts. Cluster-robust standard errors were calculated using the R package *drgee* [16–18].

Exposure definition

The exposure was defined as the first treatment episode constructed with method A (exposure equal 0) or the alternative method B or C (exposure equal 1). The contrast was among treatment episodes constructed based on either method B or C, with a pre-specified gap and method for handling overlaps, versus the corresponding treatment episode constructed with method A (reference method) with the same pre-specified gap and method for handling overlaps. This gave rise to 192 assessments (24 pairs \times 4 outcomes \times 2 treatments).

Outcomes

The four outcomes of interest were hospitalization for all causes, hospitalizations for psychiatric causes, outpatient visits for all causes, and outpatient visits for psychiatric causes identified in the NPR via the International Classification of Diseases 10th revision (ICD-10) codes described in Appendix B of the supplementary material.

Interpretation of the results

An OR less (or more) than one indicates that the alternative method B or C would result in an under-estimation (or over-estimation) of the outcome rate during exposed time relative to the reference method A. It can be interpreted as a measure of the impact of exposure time misclassification on measures of the outcomes of interest and consequently on misclassification impact on measures of association between exposure and outcomes.

Results

Descriptive statistics

The citalopram cohort included 400,388 individuals who filled 2,537,045 prescriptions and the mirtazapine cohort included 292,136 individuals with 1,482,411 prescriptions filled during the observation period. The distribution over strengths for the included filled prescriptions was 632,077 (42.64%) for 15 mg, 723,281 (48.79%) for 30 mg, and 127,053 (8.57%) for 45 mg for mirtazapine, and 481,427 (18.98%) for 10 mg, 1,786,705 (70.42%) for 20 mg, 130,306 (5.14%) for 30 mg, and 138,607 (5.46%) for 40 mg for citalopram.

Duration of prescriptions

The median number of days of supply for mirtazapine prescriptions based on the actual prescribed dosage (method A) was approximately 96 days. This was similar to the days of supply (100 days) based on the one unit per day assumption (method B), but was nearly two times longer than the median duration of a prescription (50 days) based on DDDs (method C), implying a significant misclassification of exposure duration. For citalopram, the median prescription duration under methods B and C was nearly identical (100 days and 98 days respectively) to the days of supply based on the actual prescribed dosage (98 days). However, there were sizable differences in the interquartile range of the days of supply based on the DDD, leading to misclassification of exposure by underestimating exposed time relative to the true distribution. The distributional characteristics of the durations of prescriptions using different dosage assumptions for mirtazapine and citalopram users are showed in Fig. 1.

Duration of treatment episodes

Figure 2 shows the distribution of the first treatment episode duration approximated with methods A, B and C, allowing for gaps of different lengths, and accounting or not for overlaps, for both mirtazapine and citalopram users. The top row of Fig. 2 presents distributional characteristics of treatment episode durations obtained using the actual prescribed dosage (method A) at different allowed gaps and handling of overlaps. In general, at each allowed gap, accounting for overlaps causes the distribution of the treatment episode duration to be more skewed to the right, while lower quartiles are rather unaffected. With respect to gaps, increasing the allowable length of the gap consistently shifts the distribution of the treatment episode duration to the right and increases the interquartile range. These patterns were seen for both medications although they were more pronounced for citalopram.

The second and third rows of Fig. 2 display the distributional characteristics of the first treatment episode duration under methods B and C. At each gap length and overlap handling, they can be contrasted with the distributional characteristics of the first treatment episode duration obtained using the actual dosage (method A). For mirtazapine users, employing the one unit per day assumption (method B) when constructing treatment episodes produces a nearly identical treatment episode duration distribution as the one produced under the actual dosage (method A), showing only slight differences in terms of median length, and first and third quartile (Q1 and Q3), of the first episode. However, using the one DDD per day assumption (method C) leads to an underestimation of the treatment episode duration distribution relative to the one obtained based on the actual dosage (method A). These differences become smaller as the length of allowable gap between mirtazapine prescriptions becomes larger. The same patterns

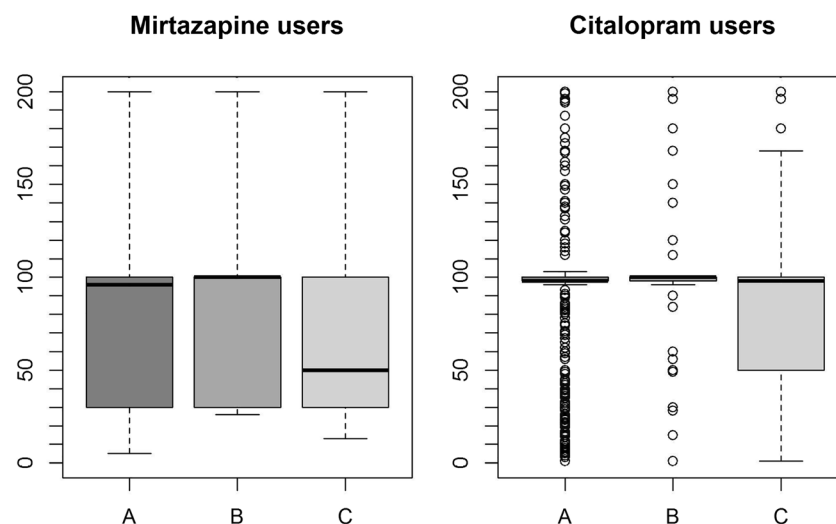


Fig. 1 Distributional characteristics of prescription durations in number of days for mirtazapine and citalopram under different dosage assumptions. Method A—actual prescribed dosage; method B—one unit per day assumption; method C—one DDD per day assumption

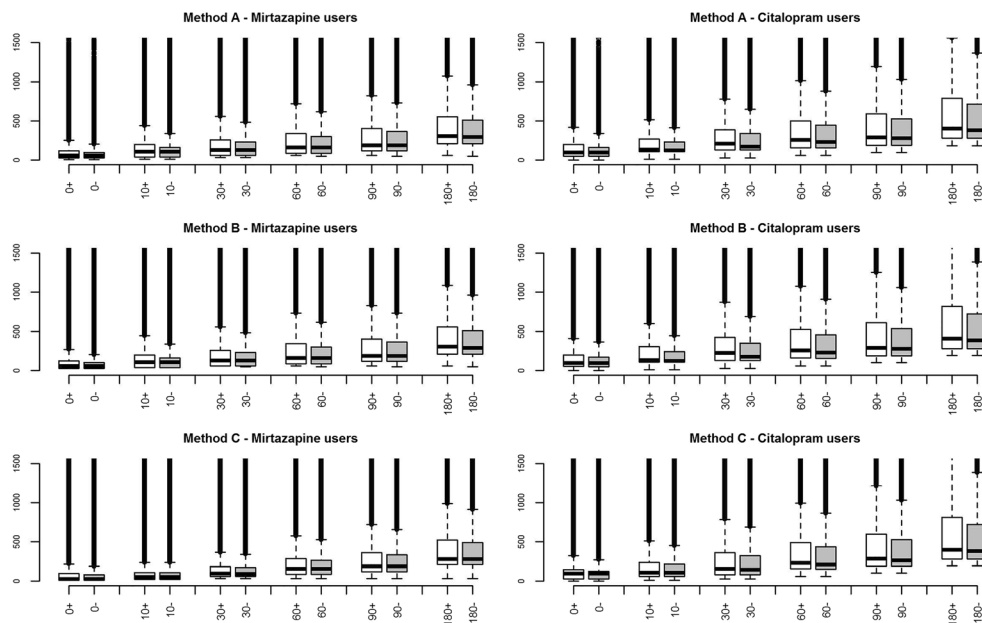


Fig. 2 Distributional characteristics of the duration of the first episode in number of days for mirtazapine and citalopram users allowing for gaps of different lengths (0, 10, 30, 60, 90, 180 days), and accounting or not for

overlaps. In the labels of the *x*-axes, the number indicates the number of days of the allowed gap, and +/- indicates respectively accounting or not for overlaps

were observed for citalopram users, with small or no differences in the treatment episode duration characteristics when constructed using the actual dosage and the one unit per day assumption, and an underestimation of exposure when using the one DDD per day assumption relative to actual dosage. For each medication, detailed values of descriptive statistics of the duration of the first treatment episode for the 36 alternatives are presented in tables 1 and 2 of Appendix C of the supplementary material.

Impact of the treatment episode construction on the odds of observing the outcome during exposed time

Summary results from the 192 models assessing the odds of observing an outcome of interest during exposed time constructed under either method B or C versus method A and a given set of pre-specified gaps and methods for handling of overlaps are reported in Table 1. The results (192 ORs) of using alternative construction methods for treatment episodes on the four investigated outcomes are illustrated in tables 4-19 in Appendix D of the supplementary material, and the count of the ORs greater or less than 1 is reported in Table 1. Each table in Appendix D has 12 comparisons, which consider 6 different lengths for the allowed gap (0, 10, 30, 60, 90, 180), and 2 ways of handling overlaps (accounting or not). Consequently in each cell of Table 1 are reported the directions of the 12 ORs (which could be $OR > 1$ or $OR < 1$) and the respective count (number of $OR > 1$ or $OR < 1$), of each table in Appendix D.

Table 1 indicates that potential exposure time misclassification associated with method B results in higher odds of observing the outcome during exposed time in 77% (37/48) and 100% (48/48) of the models for mirtazapine and citalopram respectively. Of the remaining alternative comparisons, 23% (11/48) for mirtazapine results in lower odds of observing the outcome during exposed time. On the other hand, approximating treatment episode durations based on the DDD (methods C) reduces the odds of observing the outcome during exposed time in 100% and 87.5% (42/48) of the models for mirtazapine and citalopram respectively, while 12.5% (6/48) of the models for citalopram result in higher odds of observing the outcome during exposed time. Table 20 in Appendix E shows the differences in number of events during each alternative approximation of a treatment episode duration.

Discussion

The present study sought to assess the impact of assumptions regarding dosage, allowable gaps, and handling of overlaps, on treatment episode durations, the impact of exposure time misclassification on measures of the outcomes of interest during exposed time, and consequently on the estimation of the exposure-outcome associations. The study demonstrated that relative to the actual prescribed dosage retrievable from the dose text, alternative assumptions significantly impact treatment episode durations leading to misclassification of the exposure time. The results indicate that this would affect estimates of measures of the outcomes of interest. This study

Table 1 Impact of treatment episode construction on the odds of capturing the outcome during exposed time. Frequency of the direction of the ORs relative to 1

| | Mirtazapine users | | Citalopram users | |
|--|-------------------|--------------|-------------------|-------------|
| | OR | Count | OR | Count |
| Hospitalization for all causes | | | | |
| Method B versus method A | ORs > 1 (OR < 1) | 11/12 (1/12) | ORs > 1 | 12/12 |
| Method C versus method A | ORs < 1 | 12/12 | ORs < 1 | 12/12 |
| Hospitalizations for psychiatric causes | | | | |
| Method B versus method A | ORs > 1 | 12/12 | ORs > 1 | 12/12 |
| Method C versus method A | ORs < 1 | 12/12 | ORs < 1 | 12/12 |
| Outpatient visits for all causes | | | | |
| Method B versus method A | ORs > 1 (OR < 1) | 2/12 (10/12) | ORs > 1 | 12/12 |
| Method C versus method A | ORs < 1 | 12/12 | ORs < 1 | 12/12 |
| Outpatient visits for psychiatric causes | | | | |
| Method B versus method A | ORs > 1 | 12/12 | ORs > 1 | 12/12 |
| Method C versus method A | ORs < 1 | 12/12 | ORs < 1 (ORs > 1) | 6/12 (6/12) |

ORs < 1: the alternative approach (method B or C) underestimates the exposure-outcome association relative to the reference (method A), during the first treatment episode

ORs > 1 → the alternative approach (method B or C) overestimates the exposure-outcome association relative to the reference (method A), during the first treatment episode

Method A: actual dosage; method B: one unit of drug per day; method C: one DDD per day

demonstrates the exceptional value of utilizing the actual dosage information, as well as the superior value of data registries that contain such information when assessing exposure-outcome associations.

Foundational to the construction of treatment episodes is the calculation of a prescription's days of supply. The medications considered in this study, mirtazapine and citalopram are both once daily medications commonly used at dosages equal to the available unit strength. We consistently found that the days of supply at the prescription level under the one unit per day assumption were similar to the ones derived using the actual dosage for both medications. For once daily medications, the days of supply using the DDD will depend on the degree to which the DDD aligns with the prescribed dose. In the case of citalopram, we found that the median days of supply were close to those obtained using the actual dosage. Albeit the overall distributions of days of supply were skewed to the left due to the fact that 20% of patients were prescribed a dose lower than the DDD, thus artificially halving the days of supply for citalopram for the respective users. For mirtazapine, the actual dose was equal to the DDD in half of the prescriptions while a lower dose was seen in approximately 43% of the prescriptions. Using the DDD in this case results in significant underestimation of the days of supply and misclassification of exposure time.

The above observations have been documented elsewhere in the literature using other medications. A study [4] using the Finnish Prescription Register, obtained the actual prescribed dosage based on unstructured free text format and

subsequently assessed the validity of the one unit per day versus one DDD per day dosage assumption when calculating treatment duration of statin prescriptions. The study found that more than 95% of statin prescriptions were dosed at one unit per day, but only 10% at one DDD per day. The authors suggested that the one DDD per day cannot be used to generate valid days of supply, use or exposure measures. A study using US Medicaid claims data assessed the concordance of exposure to treatment using reported days of supply and DDDs in eight medication groups (statins, metformin, atypical antipsychotics, warfarin, proton pump inhibitors, angiotensin converting enzyme-inhibitors, non-steroidal anti-inflammatory drugs, and coxib) [19]. The study reported that the DDD consistently underestimated the exposure across all drug groups with the exception for non-steroidal anti-inflammatory drugs where DDD overestimated exposure duration relative to the reported days of supply. These findings are consistent with the fact that, differences between the actual days of supply from the ones calculated based on the one unit per day or one DDD per day assumptions depend on the frequency of dosing, and the degree to which the DDD resembles the actual prescribed dosage. For once daily medications, an assessment of the distribution of the prescribed medication strengths and the DDD could provide significant insights on the potential of DDD to generate reliable days of supply.

In constructing treatment episodes using the actual prescribed dose, the handling of overlaps has a significant impact on treatment episode durations. This study showed that not accounting for overlaps consistently underestimated the

duration of treatment episodes. When the actual prescribed dosage, and therefore days of supply, is known, adjusting for overlaps would eliminate one potential source of measurement error in the treatment episode construction. However, it is not clear what the approach should be in instances where the one unit per day or one DDD per day assumptions are used to calculate the days of supply of a prescription and consequently the level of overlap. Especially when the method leads to an overestimation of days of supply, accounting for overlaps will worsen the impact of the method on the overall treatment episode duration and misclassification of exposure time.

Allowable gaps between prescriptions may be thought of as assuming continuous treatment although a patient may not be 100% adherent to the treatment. Poor compliance “in the real world” actually leads to a right-skewed distribution of treatment durations, even though at lower daily doses. This is the case with antidepressants which are often associated with poor compliance [1, 13], and therefore during the analyses the assumption of 100% compliance is modified by the allowance of a gap between end of days of supply and the next prescription fill. The choice of allowable gap between prescriptions has a significant impact on treatment episode durations and the approach in this study used a fixed number of days as gaps. The risk with greater gaps is of constructing treatment episodes of greater duration by including time that a patient is not under treatment and thus affecting measures of outcomes during exposed periods.

These findings are consistent with what is reported in a study of the impact of gaps and overlaps on the median antidepressant treatment episode duration [1, 13]. In general handling of overlaps and longer gaps are associated with increased median treatment episode duration and greater interquartile range especially when gaps are a percentage of days of supply and when overlaps are accounted for. Similarly, a study of statin medication use using the Swedish PDR [20] showed that disregarding overlapping days resulted in estimates of reduced use while increasing the allowed gap from 30 to 60 days resulted in increasing use from 57 to 69%, and consequently longer treatment episode durations.

Perhaps the most significant contribution of this study was the assessment of the effect of treatment episode duration, constructed under alternative dosage assumptions relative to the actual dosage, on the odds of observing the outcomes of interest during exposed periods. The modeling approach allows us to infer whether there is a concordance in the odds of observing an outcome between the actual exposure (model A) and an alternative treatment episode construction (method B or C). The results provide stark evidence that in the overwhelming proportion of cases, both model B and C would lead to over- or under-estimation of estimates in measures of frequency of an outcome. This occurs even in instances where the alternative approaches yield treatment episode durations that are similar in median to the actual exposure. It is evident

that differences in other distributional characteristics of treatment episode duration adversely affect estimation of such measures. For example, it was observed that the interquartile range increased when the allowable gaps were greater and when accounting for overlaps. Greater interquartile range is associated with increased spread in treatment episode durations, which in turn translates to over- or under-estimated measures of outcome ratios and/or rates. These observations have been confirmed in an earlier study assessing prevalence of statin use at a pre-specified point in time. Under different scenarios of treatment episode construction that vary by the handling of gaps and overlaps, Mantel-Teeuwisse et al. [21] showed that the estimates of point prevalence vary by method used to construct treatment episodes. While this study benefits of the strength of a within individuals design, which by definition account for all background factors, using other types of study designs and statistical approaches, research questions on the effect of background factors on exposure misclassification may be of interest for future investigations. In general, relying on assumptions regarding dosage when constructing treatment episodes entails the potential of misclassification of exposure time. Misclassification of exposure time due to assumptions used when constructing treatment episodes will likely have an adverse impact on measures of medication exposure risk. A study among Dutch NSAID users [22] compared pre-specified fixed exposure times to the one implied by quantity supplied and dosage information assuming full compliance. The study found that misclassification of exposure resulting from longer than actual exposure times, resulted in lower incidence rates of peptic ulcer therapy.

Depending on medication characteristics such as the number of available strengths, dosing frequencies, distribution of prescribed strengths, but also medication type, adverse effect profiles and other characteristics, treatment episode durations may be affected not only by the assumptions used but also by the treatments involved. In the present study this can be seen in the differences in the distributional profiles of the treatment episode duration under method C for mirtazapine and citalopram relative to the one obtained under the actual dosage assumption. In such instances the misclassification of exposure time would give rise to misclassification of outcomes by exposure type. Such differential misclassification of outcomes would lead to over- or under-estimation of measures of the exposure-outcome association. This observation was also highlighted previously by van Staa et al. [22] where it was recognized that if variation in exposure risk across groups is not accounted for, then risk comparisons would not be valid. A study using [23] the Clinical Practice Research Datalink, a national UK primary care registry, assessed how various decisions in constructing treatment episodes affect estimates of cardiovascular risk among users of hypoglycemic and glucocorticoid medications. The decisions included the use of a fixed exposure versus recorded prescription duration. The

study found that the choice of the actual versus the fixed duration resulted in significantly different hazard ratios. While the authors attribute the variation in the effect estimate to the high proportion of missing values of the reported prescription duration, the finding could also be related to misclassification of exposure time associated with the different choices of treatment episodes.

In summary, the results of the current work show the importance of careful consideration of all information available when approximating the duration of the exposure of interest. This is particularly important when prescription registers are used to make inference on exposure-outcome associations. In such cases the lack of specific information regarding exposure (actual drug intake, the intended indication and often the dosage prescribed) necessitates the use of assumptions, which have a bearing on the estimated measures of the outcomes. In this study, using either the one unit per day or the one DDD per day assumption led to exposure time misclassification relative to the actual dosage. While the study contributed significant insights on the impact of assumptions on the construction of treatment episodes, it is not possible to provide an “optimal” method or general recommendation for the treatment episode construction because this is strongly connected with the particular therapeutic area, where different medications require different assumptions on dosage.

Conclusion

We have demonstrated that relative to the actual prescribed dosage, alternative assumptions as well as different approaches to handling gaps and overlaps, significantly impact treatment episode durations, which in turn affects measures of the outcomes of interest during exposed time. This exposure time misclassification may give rise to misclassification of outcomes by exposure type and therefore lead to over- or under-estimation of measures of the exposure-outcome associations. Therefore, it is recommended, to utilize actual dosage information when constructing treatment episodes. When true dosage is not available, we recommend sensitivity analyses of treatment episode construction, in order to assess the robustness of the derived measures of associations between outcomes and exposure under different assumptions. The study provides results which have a consistent clinical relevance. In particular when conducting post authorization studies on the safety or effectiveness of a treatment using observational data from prescription registers, avoiding exposure misclassification is crucial. We have showed that a non-accurate construction of exposure time may lead to errors on outcome detection during exposed time, and consequently affect conclusions on safety or efficacy profile of a treatment.

Acknowledgments Open access funding provided by Karolinska Institute. Laura Pazzagli gratefully acknowledges support from the Swedish Society for Medical Statistics (FMS) and Karolinska Institutet. A special gratitude is extended to Anders Sundström for his valuable and constructive suggestions during the development of this research work.

Authors' contributions LP had full access to all data in the study, and takes responsibility for the integrity of the data and accuracy of the data analysis. LP, ML, PM, MA and SB conceptualized the study and the design. LP performed the analysis and LP, LB, ML, DM, PM, and SB performed the interpretation of data. LB retrieved doses from free text for method A. LP drafted the manuscript. All co-authors performed critical revision of the manuscript for important intellectual content.

Funding information This project was funded via the public-private real world evidence collaboration between Karolinska Institutet and Janssen Pharmaceuticals (contract: 5-63/2015). Morten Andersen has further received funding from the Novo Nordisk Foundation (NNF15SA0018404).

Compliance with ethical standards

Conflict of interest The Centre for Pharmacoepidemiology receives grants from several entities (pharmaceutical companies, regulatory authorities, and contract research organizations) for performance of drug safety and drug utilization studies. Morten Andersen reports grants from Novartis, Pfizer, AstraZeneca and H. Lundbeck & Mertz outside the submitted work, and personal fees from Atrium, the Danish Pharmaceutical Industry Association, for leading and teaching pharmacoepidemiology courses. Panagiotis Mavros is employee of Janssen Scientific Affairs, Titusville, NJ, USA. David Myers is employee of The Janssen Pharmaceutical Companies of Johnson & Johnson, Janssen Cilag AB, Stockholm, Sweden. The authors declare no other conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Gardarsdottir H, Egberts TCG, Heerdink ER (2010) The association between patient-reported drug taking and gaps and overlaps in antidepressant drug dispensing. *Ann Pharmacother* 44:1755–1761. <https://doi.org/10.1345/aph.1P162>
2. Lunde P (1979) Studies in drug utilisation. In: Bergamn U (ed) *Methods and applications*. WHO regional Office for Europe, Copenhagen, pp 17–28
3. Merlo J, Wessling A, Melander A (1996) Comparison of dose standard units for drug utilisation studies. *Eur J Clin Pharmacol* 50:27–30. <https://doi.org/10.1007/s002280050064>
4. Romppainen T, Rikala M, Aarnio E, Korhonen MJ, Saastamoinen LK, Huupponen R (2014) Measurement of statin exposure in the absence of information on prescribed doses. *Eur J Clin Pharmacol* 70:1275–1276. <https://doi.org/10.1007/s00228-014-1737-3>
5. Sihvo S, Wahlbeck K, McCallum A, Sevón T, Arffman M, Haukka J, Keskimäki I, Hemminki E (2010) Increase in the duration of antidepressant treatment from 1994 to 2003: a nationwide

- population-based study from Finland. *Pharmacoepidemiol Drug Saf* 19:1186–1193. <https://doi.org/10.1002/pds.2017>
6. Bjerrum L, Rosholm JU, Hallas J et al (1997) Methods for estimating the occurrence of polypharmacy by means of a prescription database. *Eur J Clin Pharmacol* 53:7–11. <https://doi.org/10.1007/s002280050329>
 7. Gardarsdottir H, van Geffen ECG, Stolker JJ, Egberts TC, Heerdink ER (2009) Does the length of the first antidepressant treatment episode influence risk and time to a second episode? *J Clin Psychopharmacol* 29:69–72. <https://doi.org/10.1097/JCP.0b013e31819302b1>
 8. Gardarsdottir H, Egberts TC, Stolker JJ, Heerdink ER (2009) Duration of antidepressant drug treatment and its influence on risk of relapse/recurrence: immortal and neglected time bias. *Am J Epidemiol* 170:280–285. <https://doi.org/10.1093/aje/kwp142>
 9. Rosholm JU, Andersen M, Gram LF (2001) Are there differences in the use of selective serotonin reuptake inhibitors and tricyclic antidepressants? A prescription database study. *Eur J Clin Pharmacol* 56:923–929. <https://doi.org/10.1007/s002280000234>
 10. Gichangi A, Andersen M, Kragstrup J, Vach W (2006) Analysing duration of episodes of pharmacological care: an example of antidepressant use in Danish general practice. *Pharmacoepidemiol Drug Saf* 15:167–177. <https://doi.org/10.1002/pds.1160>
 11. Ruokoniemi P, Korhonen MJ, Helin-Salmivaara A, Lavikainen P, Jula A, Junnila SY, Kettunen R, Huupponen R (2011) Statin adherence and the risk of major coronary events in patients with diabetes: a nested case-control study. *Brit J Clin Pharmacol* 71:766–776. <https://doi.org/10.1111/j.1365-2125.2010.03895.x>
 12. Helin-Salmivaara A, Korhonen MJ, Lehenkari P et al. (2012) Statins and hip fracture prevention - a population based cohort study in women. *PLoS One* 7, ARTN e48095 doi: <https://doi.org/10.1371/journal.pone.0048095>.
 13. Gardarsdottir H, Sovereign PC, Egberts TCG, Heerdink ER (2010) Construction of drug treatment episodes from drug-dispensing histories is influenced by the gap length. *J Clin Epidemiol* 63:422–427. <https://doi.org/10.1016/j.jclinepi.2009.07.001>
 14. Pazzagli L, Linder M, Zhang M, Vago E, Stang P, Myers D, Andersen M, Bahmanyar S (2018) Methods for time-varying exposure related problems in pharmacoepidemiology: an overview. *Pharmacoepidemiol Drug Saf* 27:148–160. <https://doi.org/10.1002/pds.4372>
 15. Qvarnström M, Kahan T, Kieler H, Brandt L, Hasselström J, Boström KB, Manhem K, Hjerpe P, Wettermark B (2016) Persistence to antihypertensive drug classes: a cohort study using the Swedish Primary Care Cardiovascular Database (SPCCD). *Medicine (Baltimore)* 95:e4908. <https://doi.org/10.1097/MD.0000000000004908>
 16. Zetterqvist J, Vansteelandt S, Pawitan Y, Sjölander A (2016) Doubly robust methods for handling confounding by cluster. *Biostatistics* 17:264–276. <https://doi.org/10.1093/biostatistics/kxv041>
 17. Zetterqvist J, Sjölander A (2015) Doubly robust estimation with the R package drgee. *Epidemiol Methods* 4:69–86
 18. R Development Core Team (2008). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>
 19. Sinnott SJ, Polinski JM, Byrne S, Gagne JJ (2016) Measuring drug exposure: concordance between defined daily dose and days' supply depended on drug class. *J Clin Epidemiol* 69:107–113. <https://doi.org/10.1016/j.jclinepi.2015.05.026>
 20. Lesén E, Sandström TZ, Carlsten A, Jönsson AK, Mårdby AC, Sundell KA (2011) A comparison of two methods for estimating refill adherence to statins in Sweden: the RARE project. *Pharmacoepidemiol Drug Saf* 20:1073–1079. <https://doi.org/10.1002/pds.2204>
 21. Mantel-Teeuwisse AK, Klungel OH, Verschuren WMM, Porsius A, de Boer A (2001) Comparison of different methods to estimate prevalence of drug use by using pharmacy records. *J Clin Epidemiol* 54:1181–1186. [https://doi.org/10.1016/S0895-4356\(01\)00396-1](https://doi.org/10.1016/S0895-4356(01)00396-1)
 22. Vanstaa TP, Abenhaim L, Leufkens H (1994) A study of the effects of exposure misclassification due to the time-window design in pharmacoepidemiologic studies. *J Clin Epidemiol* 47:183–189. [https://doi.org/10.1016/0895-4356\(94\)90023-X](https://doi.org/10.1016/0895-4356(94)90023-X)
 23. Pye SR, Sheppard T, Joseph RM, Lunt M, Girard N, Haas JS, Bates DW, Buckridge DL, van Staa T, Tamblyn R, Dixon WG (2018) Assumptions made when preparing drug exposure data for analysis have an impact on results: An unreported step in pharmacoepidemiology studies. *Pharmacoepidemiol Drug Saf* 27:781–788. <https://doi.org/10.1002/pds.4440>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.