Tazare John (Orcid ID: 0000-0002-7194-2615)

a) Manuscript Title

Prevalent new user designs: a literature review of current implementation practice

b) Running Title

Prevalent new user design review

c) Authors

John Tazare[1], Daniel C Gibbons[2], Marleen Bokern[1], Elizabeth J Williamson[1 3], Iain A. Gillespie[2], Marianne Cunnington[4], John Logie[2], Ian J Douglas[1 3]

d) Institutional Affiliations

[1]Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom.

[2]GSK, London, United Kingdom.

[3]Health Data Research United Kingdom (HDR-UK), London, United Kingdom.

[4]Analysis Group Inc., London, United Kingdom

e) Corresponding Author Information

John Tazare
john.tazare1@lshtm.ac.uk
London School of Hygiene & Tropical Medicine,
Keppel St,
Bloomsbury,
London,
WC1E 7HT.

f) Key Words

Prevalent new user designs, database research, literature review

g) Key Points

1. Prevalent new user designs (PNU) are increasingly used to study the real-world effects of medications and provide the opportunity to include the clinically important subgroup of patients who switch from a (typically older) treatment to a newer study drug.
2. The uptake in PNU designs has led to methodological development and extensions of the initial proposal, including more complicated exposure set definitions and the incorporation of the high-dimensional propensity score algorithm.
3. We found that the PNU design has been successfully applied across a number of disease and therapeutic areas, as well as across both electronic health record and administrative claims databases.
4. However, a lack of clarity in the approaches used in many of the studies highlights the need for improved reporting and sharing of analytical code (or software packages) to allow for more widespread application of these designs.

h) Word Count excluding abstract, tables, figures and references

3,005

i) A statement about prior postings and presentations, name of any sponsor of the research contained in the paper, along with grant numbers

This work was funded by an unrestricted grant from GlaxoSmithKline (GSK).

Plain Language Summary

Prevalent new user (PNU) designs extend established study designs, such as the active comparator new user design, by including patients who switch from an older comparator drug to a newer study drug. The PNU design aims to provide an assessment of effectiveness and safety that is not limited to first-time users of the study or comparator drug. We performed a literature review summarising published studies implementing this study design. From the 15 studies (incorporating a class of switchers) included we highlight several trends, for example, surrounding extensions to the initial exposure set definition proposals and incorporation of complex analytical methods within the PNU framework. Overall, we highlight that PNU designs have been applied to answer research questions in a range of therapeutic and disease areas. Wider uptake of the design may be achieved through increased sharing of analytical code and clearer guidance surrounding the reporting of key implementation details.

Abstract

Purpose

Prevalent new user (PNU) designs extend the active comparator new user design by allowing for the inclusion of initiators of the study drug who were previously on a comparator treatment. We performed a literature review summarising current practice.

Methods

PubMed was searched for studies applying the PNU design since its proposal in 2017. The review focused on three components. Firstly, we extracted information on the overall study design, including the database used. We summarised information on implementation of the PNU design, including key decisions relating to exposure set definition and estimation of time-conditional propensity scores. Finally, we reviewed the analysis strategy of the matched cohort.

Results

Nineteen studies met the criteria for inclusion. Most studies (73%) implemented the PNU design in electronic health record or registry databases, with the remaining using insurance claims databases. Of 15 studies including a class of prevalent users, 40% deviated from the original exposure set definition proposals in favour of a more complex definition. Four studies did not include prevalent new users but used other aspects of the PNU framework. Several studies lacked details on exposure set definition (n=2), time-conditional propensity score model (n=2) or integration of complex analytical techniques, such as the high-dimensional propensity score algorithm (n=3).

Conclusion

PNU designs have been applied in a range of therapeutic and disease areas. However, to encourage more widespread use of this design and help shape best practice, there is a need for improved accessibility, specifically through the provision of analytical code alongside guidance to support implementation and transparent reporting.

1. Background

The active comparator new user (ACNU) study design has become the accepted standard to examine the real-world safety and effectiveness of medications.[1] The ACNU compares initiators of two therapies, that are both indicated and prescribed for the same indication, in patients with no prior use of the drugs of interest. Restricting to new users is an appealing characteristic of the ACNU design since prevalent users are more likely to have been influenced by prior treatment.[2] However, when comparing a newer study drug to an older and established comparator, the ACNU has several shortcomings. Firstly, patients receiving the newer drug can be quite different to those initiating the comparator, leading to challenges around confounding by indication or channeling bias.[2,3] Secondly, the exclusion of the subgroup of patients receiving the study drug who have switched from the comparator drug in the ACNU can lead to selection bias. For example, if this represents a large proportion of the patient population, it is unclear how generalisable these results are to a target population of all patients initiating a drug (external validity).

The prevalent new user (PNU) design was proposed by Suissa *et al* in 2017 to address this limitation and builds upon the ACNU by also including initiators of the new drug who were previously on the older comparator (i.e. individuals adding or switching); thereby aiming to provide a more comprehensive assessment of relative drug effects.[3] The PNU requires a number of steps (summarised in Figure 1):

- Step 1, a base cohort is formed containing all users of the comparator and study drugs.[3] The study drug users identified will include both those who were first treated with the study drug and those who switched from the comparator.
- Step 2, for each user of the study drug, an exposure set, comprising of comparator drug users with a similar prior cumulative exposure to the comparator drug, is formed. There are several proposals for how prior cumulative exposure to the comparator should be defined, including time-based (time since initial exposure), prescription-based (number of drug prescriptions since initial exposure) and hybrid exposure sets (a combination of time-based and prescription-based). Full details described in Figure 2.[3,4]

- Step 3, for each study drug user, time-conditional propensity scores (TCPS) are used to identify the most similar comparator drug user from their exposure set. TCPSs are typically estimated using one logistic regression model fitted (conditional on exposure set) on the stacked exposure sets. The positivity assumption is checked within exposure set and exposure sets violating this (i.e. where the TCPS for the study drug user is not contained within the range of TCPSs for the comparator users within that exposure set) are dropped from future steps.[3]
- Step 4, within each exposure set, the study drug user is matched to a comparator drug user with the most similar TCPS. As proposed by Suissa *et al*, matching is typically performed chronologically and without replacement (i.e. once matched into the comparator group, a patient is not considered for subsequent study drug users).[3] Alternatively, matching with replacement (where comparators can be reused in the matching process) might be considered in settings where there is a relatively small pool of suitable potential comparator users.
- Steps 1-4 result in a matched cohort containing incident and prevalent new users of the study drug matched with a comparator drug user.[3] Follow-up is defined dependent on the question of interest (i.e. intent-to-treat or as-treated) and treatment effects are estimated using standard methods.[3,5] Investigation of effect modification by prevalent/incident user status is strongly encouraged to explore whether it is most appropriate to report an overall treatment effect or separate treatment effects by prevalent/incident user status.

In this study, we review published, peer-reviewed studies applying the PNU design to understand how it has been implemented in practice.[3] We aim to summarise current practice surrounding design choices (for example, the choice of exposure set), describe any developments to, and deviations from, the original proposal and highlight possible areas for future methodological work.

2. Methods

2.1 Search strategy and study selection

A PubMed search was conducted on 9th May 2022 to identify articles applying the PNU design. The search strategy identified articles with 'prevalent new user' or 'time conditional' in the title, abstract or as a Medical Subject Heading (MeSH). Given the publication of Suissa *et al.*'s seminal PNU paper in 2017, only articles published afterwards were considered (the text search string is provided in Supplementary Information 1).[3]

For inclusion, studies had to be peer-reviewed and use the PNU design to address a clinical question of interest. Commentaries, reviews or methodological studies were excluded.

2.2 Data extraction

For all eligible studies, an initial review of the full-text was performed to assess suitability before data extraction. The information extracted focused on three components: the overall study design, the PNU design implementation, and the analysis strategy.

The study design component assessed information such as the treatment decision being studied (e.g. switching to, or adding, the study drug versus remaining on the comparator drug), therapies under investigation, the disease/therapeutic area and the database used. The PNU design component focused on specific investigator decisions surrounding: exposure set definition, estimation of the TCPS, matching, and software used. Finally, the analysis strategy component considered the follow-up approach (e.g. as-treated versus intent-to-treat), outcome model and whether effect modification by incident- or prevalent-user status were investigated. The full data extraction table is provided in Supplementary Information 2.

Data extraction was conducted primarily by JT. A random sample (N=3) of the selected papers was independently reviewed by MB to check the consistency of the data extraction and discuss possible ambiguities. Online supplementary materials were only accessed if the main text made direct reference to information concerning the three data extraction components.

3. Results

The PubMed search identified 51 studies, of which 19 met the inclusion criteria (Figure 3). A list of the 19 studies included is provided in Supplementary Information 3.

3.1 Characteristics of studies included

Table 1 summarises the characteristics of the studies included in this review.

PNU designs were applied to answer clinical questions in a broad range of disease areas, including type 2 diabetes mellitus (T2DM) (42%), rheumatoid arthritis (21%) and chronic obstructive pulmonary disease (16%).

The PNU design has been implemented in both electronic health records systems (for example, in the UK and Hong Kong)[7–9] and administrative databases (for example, in Taiwan, the US and Canada)[4,10–13]. Most studies (73%) included use of an electronic health record (EHR) or detailed longitudinal registry database, rather than administrative insurance claims database.

Most studies (79%) included at least one co-author from the McGill University research group who developed the PNU design.[3]

We found that most studies (79%) investigated switching to, or adding, an alternative therapy. Of these studies, all compared the study drug to an active treatment comparison group.

During the initial review of studies, we also noted a group of studies which focused only on initiation (i.e. restricting to the incident-new user subgroup of the PNU design)[33-35] or discontinuation treatment decisions (N=3 and N=1, respectively).[36]

3.2 Prevalent new user design implementation

The next sections solely focus on the group of studies (N=15) that include a prevalent subgroup; however, we include these other studies at this stage to highlight the potential for PNU designs to study these additional types of treatment questions.

The reporting of implementation details is summarised in Table 2.

There were examples of each of the three proposed exposure set types (summarised in Figure 2) being applied (47% of studies; hybrid-based (N=4), time-based (N=2), prescription-based (N=1)). Furthermore, 40% of studies deviated from the use of only prescription count, time since cohort entry and calendar time to define exposure sets. These studies typically favoured more complex definitions, likely resulting in exposure sets which identify a more restricted number of users of the comparator drug. For example, one study investigated the T2DM treatment classes, sodium-glucose cotransporter 2 inhibitors (SGLT2i) and dipeptidyl peptidase-4 inhibitors (DPP4i), and the risk of below-knee amputation. Information on overall level of antidiabetic treatment, prior use of glucagon-like peptide-1 (GLP-1) receptor agonists (another 2nd line anti-diabetic medication) and calendar period was used to define exposure sets.[17] None of the studies included reported sensitivity analyses varying the exposure set definition. Finally, in two instances (13%), it was difficult to ascertain whether exposure sets had been used and which definition had been implemented.[28,29]

Suissa *et al* highlight the use of sampling to reduce the computational burden resulting from many large exposure sets.[3] Of the studies included, none explicitly mentioned the use of exposure set sampling. However, it is possible that the use of more restrictive exposure set definitions reduces the computational burden in many cases enough to avoid the need for sampling. A study by Yang *et al* did take a 10% random sample of comparator use sets to reduce computation difficulties.[18] This was performed in the context of the three-step matching procedure (an alternative proposal for deriving a PNU cohort) which matches on index date within a time interval, medication possession ratio, and propensity score.[19]

Step 3 of the PNU requires fitting the TCPS model. We found that most studies (66%) used either conditional logistic regression or Cox proportional hazards regression. Three studies reported using logistic regression, however, no justification for the deviation from the proposal by Suissa *et al* was provided.[8,18,29] Four studies (27%) fitted the TCPS model separately in prevalent and incident users[10,17,20,24] This approach has the advantage of avoiding estimated regression parameter coefficients representing the pooled effect across incident and prevalent

users. Furthermore, in the case where the number of incident users far outweighs the number of prevalent users, this avoids these coefficients being dominated by the incident users. However, in the setting where the proportion of prevalent users is relatively small it might be challenging to fit the TCPS model, especially when including a large set of covariates in the TCPS model. Finally, in 73% of studies there was evidence of TCPS positivity assumption checking, with exposure sets violating this assumption dropped from the matching and analysis steps.

Three studies also used the high-dimensional propensity score (HDPS) algorithm to supplement the investigator set of covariates in the TCPS model.[9, 20, 22]

The final step of the PNU process involves creation of a matched cohort based on matching study drug users to comparator drug users based on the TCPSs. Whilst the majority of studies (60%) implemented one-to-one (1:1) matching without replacement (which is consistent with the Suissa *et al* proposal), studies also matched one-to many (1:m; m ranging from 2 to 4).[7,8,20,22,23] Of the studies matching 1:m, it was not always clear how this was achieved, for example, if investigators attempted to obtain matches iteratively or all at once. However, one approach repeated the matching procedure to try to obtain another match for each study drug user.[8] Finally, when greater than 10% of exposure sets were dropped because of violation of the positivity assumption matching with replacement was used as an alternative analytical decision to improve the proportion of successfully matched study drug users, for example, see Fisher *et al* and Fillion *et al*.[11,24]

All the included studies presented the characteristics of the matched cohort and there was evidence of formal covariate balance checking (for example, based on calculation of standardised differences) in 80% of studies. Conversely, only 40% of studies presented pre-matching characteristics. This was typically provided as either the selection of a random comparator (sampled from each exposure set),[22] or the overall number of potential comparators.[18]

3.3 Analytical strategy

Information surrounding the analysis of the matched cohort and software/analytical code availability is presented in Table 3.

All included studies analysed the matched cohort in a survival analysis framework. The majority of studies (87%) conducted an 'as-treated' analysis, censoring individuals at the time of treatment switch or discontinuation. The remaining two studies performed an 'intent-to-treat' analysis where individuals were not censored at time of treatment switch or discontinuation. Forty percent of studies performed a sensitivity analysis changing their follow-up approach. Despite this, none of the studies performing an 'as-treated' analysis looked to more formally investigate the practical implications of possible selection bias due to dependent censoring, for example through the use of inverse probability of censoring weighting.[5]

As well as reporting the overall effect estimate including both prevalent and incident users, 53% of studies also investigated effect modification by incident/prevalent user status (reporting separate treatment effects by subgroup).

Most of the studies included (87%) commented on the statistical software package used. Six studies used SAS (40%), two (13%) used R, five (33%) used a combination of SAS and R, and two (13%) did not state the statistical software used. Currently, there is not a publicly available package or macro implementing PNU designs. Furthermore, no analytical code was available from any of the studies included in this review.

4. Discussion

In this review, we summarised studies implementing the PNU design and highlighted trends in current implementation practice. We identified all peer-reviewed studies applying the PNU design and extracted information surrounding the overall study design, PNU implementation, and analysis strategy of the resulting matched cohort.

Most of the studies included were conducted, at least in part, using an EHR or detailed longitudinal registry database. One reason for this may be the increased availability of detailed longitudinal prescription data in EHR databases which is necessary for establishing historical treatment patterns when applying the PNU design.[14–16] This review also indicates that the PNU design has been used to study questions across a wide range of therapeutic and disease areas. For T2DM management in adults, whether to stop, switch or add treatments is a key research recommendation in the latest UK National Institute for Health and Care Excellence (NICE) guideline and has not been comprehensively studied in clinical trials.[6] The number of studies in this area highlights the potential for PNU designs to contribute meaningfully to an ongoing research agenda and improvements in clinical practice where the underlying question of interest does not naturally align with more traditional designs, such as the ACNU. However, when applying the PNU it is important to clearly state the effect being studied and interpret this appropriately, especially when directly comparing study results with an ACNU design.[5] We also found extensions of the initial PNU proposal, for example, use of additional criteria to define exposure sets and integration of the HDPS algorithm.

As summarised in Figure 1, implementation of the PNU design requires key decisions on the exposure set definition, TCPS model and matching procedure to obtain the analysis cohort. We briefly summarise our findings of each of these aspects below.

Forty percent of studies used more complex definitions for exposure sets than those based on a combination of prescription count, time since cohort entry or calendar (i.e. as proposed by Suissa *et al* and Lin *et al*).[3,4] Whilst there are potential computational efficiencies to be gained from more restrictive exposure set definitions (since the resulting exposure sets will be smaller), it is unclear whether this is always the optimal approach in terms of the resulting matched cohort. Future work could explore the impact of exposure set definition choice on the characteristics of the matched cohort and possible consequences for estimated treatment effects.

The TCPS model was typically fitted using conditional logistic regression, as proposed by Suissa *et al*.[3] We highlighted several studies which estimated the TCPS separately in incident and

prevalent users.[17,20] However, further work would be needed to establish in which circumstances this approach may be preferred to fitting a single TCPS model.

Another extension was the incorporation of the HDPS algorithm.[21] Briefly, the HDPS is a data-driven approach to defining covariates based on the frequency of recording of codes in a period prior to cohort entry (full details in Schneeweiss *et al*).[21] When applied within the PNU design, in all instances 500 covariates were selected for inclusion in the TCPS model. One motivation for applying HDPS to PNU designs is to capture hard-to-measure differences between incident and prevalent users which might contribute to residual confounding. Confounding bias is likely more of a challenge in PNU studies since an implicit assumption is that the investigator-chosen variables also characterise the switching-outcome relationship.[5] However, even with HDPS, capturing the complex thought process behind a decision to switch versus continue is likely to be difficult with routine data. Furthermore, use of the HDPS within PNU designs is not straightforward (due to the use of time-conditional covariate information) and current applications have not described in sufficient detail any necessary modifications to the HDPS. This is particularly important since the HDPS has its own important investigator decisions and can significantly contribute to the computational burden of the PNU.[21,25,26]

We observed that one-to-one matching (without replacement) was predominantly performed (60% of studies), as proposed by Suissa *et al*.[3] We also highlighted the use of matching with replacement when greater than 10% of exposure sets were dropped because of violation of the positivity assumption. Recent simulation work by Webster-Clark *et al* has demonstrated that matching with replacement can yield unbiased results in the PNU design.[26] This work also highlights analytical strategies which using weighting schemes rather than matching and have the potential to reduce computational burden compared the initial PNU proposal. [26]

The provision of available sample data and analytical code (or software packages) has been shown to aid successful diffusion of innovative designs and methods in pharmacoepidemiological research.[27] However, our work has highlighted a lack of available analytical code and software packages for implementing PNU designs, with none of the studies here using code that is publicly available. Availability of these materials alongside clear

reporting of implementation steps (as outlined in Figure 1) would remove possible ambiguity translating the proposal into practice and potentially facilitate the more widespread use of these designs. However, as discussed elsewhere, this is important for all pharmacoepidemiology studies.[30-32]

This review has several limitations. Firstly, journals often have a page or word limit and this can restrict the information investigators can convey in the main text which then has implications for the information extracted in this review. We did consider supplemental material where it contained relevant information but we acknowledge our ability to fully capture implementation details may not be complete. Supplemental materials and external sources (for example, online posting of study protocols) could be more widely used or referenced to provide further details of the implementation of methods. This is particularly important when applying novel and complex proposals, such as the PNU design, as the main text is unlikely to allow space to articulate adequately the nuances of the method. Secondly, this review is limited by the relatively recent introduction of this method which may initially increase variability in information reported on study design application. Finally, whilst our review focused on the applications of PNU where the research question surrounded treatment switching, the PNU design has started to be used for questions surrounding treatment initiation (including settings without an active comparator) and treatment discontinuation.[33-36]

Amidst a fast-evolving study design methodology, this review has captured a snapshot of pioneering practice and highlights key areas for future guidance development, which given the relative complexity around implementing PNU designs, might lead to wider uptake.

5. Conclusion

This review highlights the application of PNU designs across a diverse range of study questions and data sources, and underlines the ability of this design to complement ACNU designs to usefully contribute to the assessment of real-world effectiveness and safety.[3,5] However, to encourage more widespread use of this design, there is a need for improved accessibility (e.g. through software packages or provision of example analytical code) and guidance for implementation and transparent reporting to help shape best practice.

Authors' contributions

JT identified the relevant papers, performed data extraction and analysis, and wrote the initial manuscript draft. IJD reviewed a draft version of the data extraction table and MB conducted a pilot check of data extraction. All authors provided comments on the finalised manuscript.

Funding

This work is funded by an unrestricted grant from GlaxoSmithKline (GSK).

Competing interests

JT is funded as a Research Fellow at LSHTM through an unrestricted grant from GSK. DCG, IAG, and JL are employees of, and own shares in, GSK; MC was an employee of GSK at time of study implementation. MB is funded by a PhD studentship provided by GSK. ID is funded by an unrestricted grant from, and holds stock in, GSK.

Availability of data and materials

The data extraction dataset analysed in this study is available in the following GitHub repository: https://github.com/johntaz/PNU-Review

Supplementary Information

Supplementary Information 1: Search string used to identify articles

Supplementary Information 2: Data extraction table

Supplementary Information 3: List of included studies

References

1. Lund JL, Richardson DB, Stürmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. *Curr Epidemiol Rep*. 2015;2(4):221-228. doi:10.1007/s40471-015-0053-5
2. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol*. 2003;158(9):915-920. doi:10.1093/aje/kwg231
3. Suissa S, Moodie EEM, Dell'Aniello S. Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores. *Pharmacoepidemiol Drug Saf*. 2017;26(4):459-468. doi:10.1002/pds.4107
4. Lin HMD, Lai CL, Dong YH, Tu YK, Chan KA, Suissa S. Re-evaluating Safety and Effectiveness of Dabigatran Versus Warfarin in a Nationwide Data Environment: A Prevalent New-User Design Study. *Drugs - Real World Outcomes*. 2019;6(3):93-104. doi:10.1007/s40801-019-0156-2
5. Webster-Clark M, Ross RK, Lund JL. Initiator Types and the Causal Question of the Prevalent New-User Design: A Simulation Study. *Am J Epidemiol*. 2021;190(7):1341-1348. doi:10.1093/aje/kwaa283
6. Recommendations for research | Type 2 diabetes in adults: management | Guidance | NICE. Accessed May 23, 2022. https://www.nice.org.uk/guidance/ng28/chapter/Recommendations-for-research
7. Au PCM, Tan KCB, Cheung BMY, Wong ICK, Wong Y, Cheung CL. Association Between SGLT2 Inhibitors vs DPP-4 Inhibitors and Risk of Pneumonia Among Patients With Type 2 Diabetes. *J Clin Endocrinol Metab*. 2022;107(4):e1719-e1726. doi:10.1210/clinem/dgab818
8. Ozen G, Dell'Aniello S, Pedro S, Michaud K, Suissa S. Reduction of Cardiovascular Disease and Mortality versus Risk of New Onset Diabetes with Statin Use in Patients with Rheumatoid Arthritis. *Arthritis Care Res*. Published online February 4, 2022. doi:10.1002/acr.24866
9. Douros A, Dell'Aniello S, Yu OHY, Filion KB, Azoulay L, Suissa S. Sulfonylureas as second line drugs in type 2 diabetes and the risk of cardiovascular and hypoglycaemic events: population based cohort study. *BMJ*. 2018;362:k2693. doi:10.1136/bmj.k2693
10. Douros A, Lix LM, Fralick M, et al. Sodium-Glucose Cotransporter-2 Inhibitors and the Risk for Diabetic Ketoacidosis : A Multicenter Cohort Study. *Ann Intern Med*. 2020;173(6):417-425. doi:10.7326/M20-0289
11. Filion KB, Lix LM, Yu OH, et al. Sodium glucose cotransporter 2 inhibitors and risk of major adverse cardiovascular events: multi-database retrospective cohort study. *BMJ*. 2020;370:m3342. doi:10.1136/bmj.m3342
12. Li G, Keenan A, Daskiran M, et al. Relapse and Treatment Adherence in Patients with Schizophrenia Switching from Paliperidone Palmitate Once-Monthly to Three-Monthly Formulation: A Retrospective Health Claims Database Analysis. *Patient Prefer Adherence*. 2021;15:2239-2248. doi:10.2147/PPA.S322880
13. Hudson M, Dell'Aniello S, Shen S, Simon TA, Ernst P, Suissa S. Comparative safety of biologic versus conventional synthetic DMARDs in rheumatoid arthritis with COPD: a real-world population study. *Rheumatol Oxf Engl*. 2020;59(4):820-827. doi:10.1093/rheumatology/kez359
14. Herrett E, Gallagher AM, Bhaskaran K, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol*. 2015;44(3):827-836. doi:10.1093/ije/dyv098

15. Wolf A, Dedman D, Campbell J, et al. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. *Int J Epidemiol*. 2019;48(6):1740-1740g. doi:10.1093/ije/dyz034
16. Hetland ML. DANBIO: a nationwide registry of biological therapies in Denmark. *Clin Exp Rheumatol*. 2005;23(5 Suppl 39):S205-207.
17. Yu OHY, Dell'Aniello S, Shah BR, et al. Sodium-Glucose Cotransporter 2 Inhibitors and the Risk of Below-Knee Amputation: A Multicenter Observational Study. *Diabetes Care*. 2020;43(10):2444-2452. doi:10.2337/dc20-0267
18. Yang CT, Yang CY, Ou HT, Kuo S. Comparative cardiovascular safety of GLP-1 receptor agonists versus other glucose-lowering agents in real-world patients with type 2 diabetes: a nationwide population-based cohort study. *Cardiovasc Diabetol*. 2020;19(1):83. doi:10.1186/s12933-020-01053-0
19. Yang CY, Kuo S, Lai ECC, Ou HT. Three-step matching algorithm to enhance between-group comparability and minimize confounding in comparative effectiveness studies. *Sci Rep*. 2022;12(1):214. doi:10.1038/s41598-021-04014-z
20. Suissa S, Hudson M, Dell'Aniello S, Shen S, Simon TA, Ernst P. Comparative safety of abatacept in rheumatoid arthritis with COPD: A real-world population-based observational study. *Semin Arthritis Rheum*. 2019;49(3):366-372. doi:10.1016/j.semarthrit.2019.03.007
21. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiol Camb Mass*. 2009;20(4):512-522. doi:10.1097/EDE.0b013e3181a663cc
22. Douros A, Dell'Aniello S, Yu OHY, Suissa S. Comparative cardiovascular and hypoglycaemic safety of glimepiride in type 2 diabetes: A population-based cohort study. *Diabetes Obes Metab*. 2020;22(2):254-262. doi:10.1111/dom.13893
23. Khosrow-Khavar F, Bouganim N, Filion KB, Suissa S, Azoulay L. Cardiotoxicity of Use of Sequential Aromatase Inhibitors in Women With Breast Cancer. *Am J Epidemiol*. 2020;189(10):1086-1095. doi:10.1093/aje/kwaa065
24. Fisher A, Fralick M, Filion KB, et al. Sodium-glucose co-transporter-2 inhibitors and the risk of urosepsis: A multi-site, prevalent new-user cohort study. *Diabetes Obes Metab*. 2020;22(9):1648-1658. doi:10.1111/dom.14082
25. Tazare J, Wyss R, Franklin JM, et al. Transparency of high-dimensional propensity score analyses: Guidance for diagnostics and reporting. *Pharmacoepidemiol Drug Saf*. 2022;31(4):411-423. doi:10.1002/pds.5412
26. Webster-Clark M, Panagiotis M, Garry EM, et al. Alternative Analytic and Matching Approaches for the Prevalent New-User Design: A Simulation Study. *Pharmacoepidemiol Drug Saf*. 2022;31(7):796-803. doi:10.1002/pds.5446
27. Cadarette SM, Ban JK, Consiglio GP, et al. Diffusion of Innovations model helps interpret the comparative uptake of two methodological innovations: co-authorship network analysis and recommendations for the integration of novel methods in practice. *J Clin Epidemiol*. 2017;84:150-160. doi:10.1016/j.jclinepi.2016.12.006
28. Li G, Keenan A, Daskiran M, et al. Relapse and Treatment Adherence in Patients with Schizophrenia Switching from Paliperidone Palmitate Once-Monthly to Three-Monthly Formulation: A Retrospective Health Claims Database Analysis. *Patient Prefer Adherence*. 2021;15:2239-2248. doi:10.2147/PPA.S322880
29. Abtahi S, Cordtz R, Dreyer L, Driessen JHM, Boonen A, Burden AM. Biological disease-modifying antirheumatic drugs and osteoporotic fracture risk in patients with rheumatoid

arthritis: a Danish cohort study. *Am J Med*. Published online February 5, 2022:S0002-9343(22)00079-1. doi:10.1016/j.amjmed.2022.01.017

30. 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

31. Wang SV, Verpillat P, Rassen JA, Patrick A, Garry EM, Bartels DB. Transparency and Reproducibility of Observational Cohort Studies Using Large Healthcare Databases. *Clin Pharmacol Ther*. 2016;99(3):325-332. doi:10.1002/cpt.329

32. Wang S, Pottegård A, Crown W, et al. HARmonized Protocol Template to Enhance Reproducibility (HARPER) of Hypothesis Evaluating Real-World Evidence Studies on Treatment Effects: A Good Practices Report of a Joint ISPE/ISPOR Task Force. Published online May 6, 2022. Accessed September 27, 2022. https://osf.io/6qxpf/

33. Suissa S, Dell'Aniello S, Ernst P. Comparative Effects of LAMA-LABA-ICS vs LAMA-LABA for COPD: Cohort Study in Real-World Clinical Practice. *Chest*. 2020;157(4):846-855. doi:10.1016/j.chest.2019.11.007

34. Suissa S, Dell'Aniello S, Ernst P. Comparative Effectiveness and Safety of LABA-LAMA vs LABA-ICS Treatment of COPD in Real-World Clinical Practice. *Chest*. 2019;155(6):1158-1165. doi:10.1016/j.chest.2019.03.005

35. Tran T, Assayag D, Ernst P, Suissa S. Effectiveness of Proton Pump Inhibitors in Idiopathic Pulmonary Fibrosis: A Population-Based Cohort Study. *Chest*. 2021;159(2):673-682. doi:10.1016/j.chest.2020.08.2080

36. Suissa S, Dell'Aniello S, Ernst P. Discontinuation of Inhaled Corticosteroids from Triple Therapy in COPD: Effects on Major Outcomes in Real World Clinical Practice. *COPD*. 2022;19(1):133-141. doi:10.1080/15412555.2022.2045265

Tables

Table 1: Characteristics of studies included (n=19)

Description	Number	(%)
Publication Year		
2018	1	(5)
2019	3	(16)
2020	9	(47)
2021	2	(11)
2022+	4	(21)
Treatment Decision		
Switching and/or Adding	15	(79)
Initiate	3	(16)
Discontinue	1	(5)
Comparator		
Active Treatment	16	(84)
Non-users	2	(11)
Continuers	1	(5)
Disease Area		
Breast Cancer	1	(5)
Chronic Obstructive Pulmonary Disease	3	(16)
Idiopathic Pulmonary Fibrosis	1	(5)
Non-valvular Atrial Fibrillation	1	(5)
Rheumatoid Arthritis	4	(21)
Schizophrenia	1	(5)
Type 2 Diabetes	8	(42)
Database+		
Linked UK Electronic Health Records	12	(63)
US Administrative Claims Databases	3	(16)
Canadian Administrative Healthcare Databases	4	(21)

Taiwan Health Insurance Databases	2	(11)
Hong Kong Electronic Health Records	1	(5)
Danish Registry Databases	1	(5)
Multi-site Studies		
Single-site	15	(79)
Multi-site	4	(21)

*Percentages do not sum to 100 as there is overlap with some studies using more than one database.
[+]Indicates a partial year.

Table 2: Implementation of prevalent new user design in studies where the treatment decision under investigation is switching and/or adding a medication (n=15)

	Switching and/or Adding Studies	
Description	**Number**	**(%)**
Primary Exposure Set Type		
Prescription	1	(7)
Time	2	(13)
Hybrid	4	(27)
Other	6	(40)
Not Clear	2	(13)
Time-Conditional Propensity Score (TCPS) Model		
Conditional Logistic Regression	8	(53)
Cox Proportional Hazards Regression	2	(13)
Logistic Regression	3	(20)
Not Clear	2	(13)
TCPS Model Fitted Separately in Prevalent/Incident Users	4	(27)
Evidence of TCPS Positivity Assumption Checking	11	(73)
High-dimensional Propensity Score Covariates Included In TCPS	3	(20)
Matching Ratio		
1:1	9	(60)
1:2	3	(20)
1:3	1	(7)
1:4	2	(13)
Covariate Balance		
Pre-matching characteristics were provided	6	(40)
Post-matching characteristics were provided	15	(100)

Evidence of Formal Covariate Balance Checking	12	(80)

Table 3: Analytical approach in studies where the treatment decision under investigation is switching and/or adding a medication (n=15)

	Switching and/or Adding Studies	
Description	**Number**	**(%)**
Follow-up Approach		
As-treated	13	(87)
Intent-to-treat	2	(13)
Sensitivity Analysis	6	(40)
Outcome model		
Cox Proportional Hazards Regression	14	(93)
Poisson regression	1	(7)
Effect Modification by Incident/Prevalent User Status Investigated	8	(53)
Software & Code Availability		
SAS	6	(40)
R	2	(13)
Combination of SAS & R	5	(33)
Not Stated	2	(13)
Analytical Code Available	0	(0)

Figures

Figure 1: Summary of steps required to implement the prevalent new user design

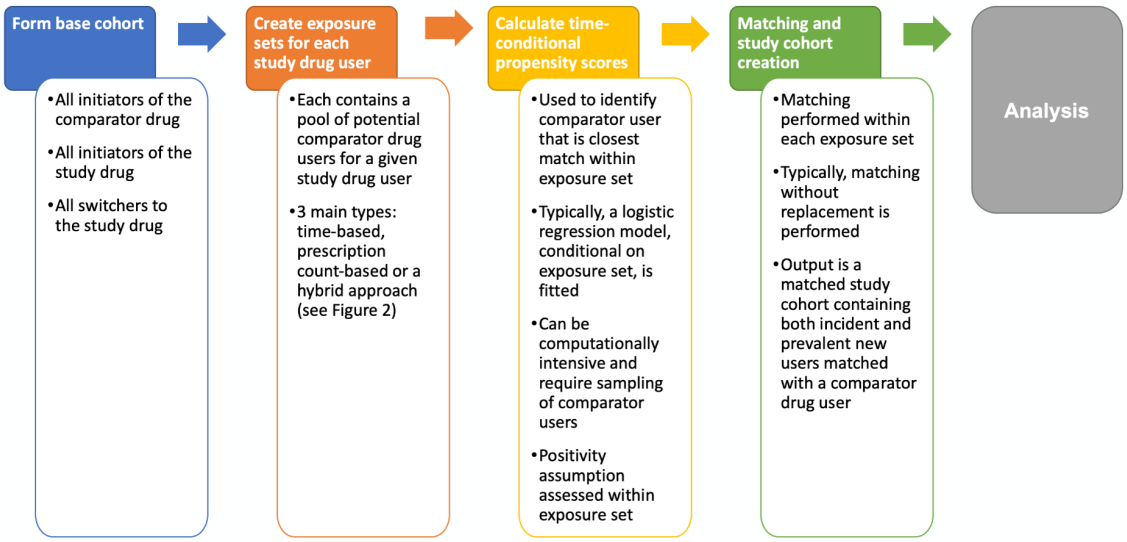
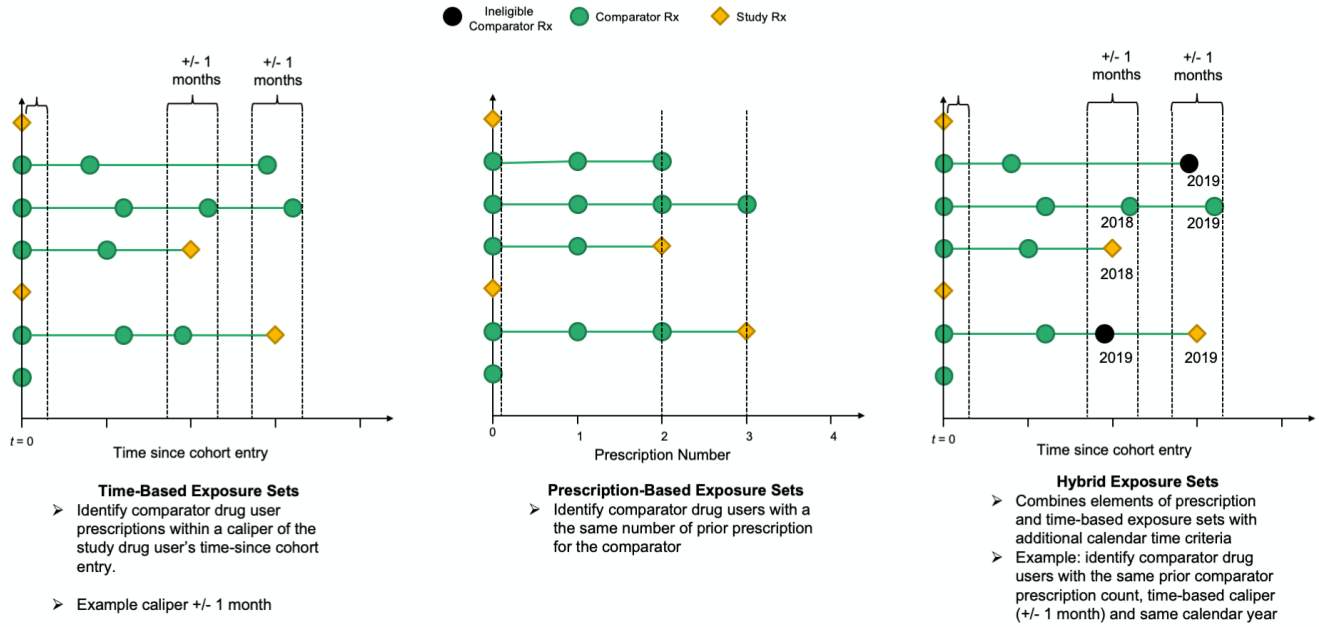


Figure 2: Exposure set definitions for time-based, prescription-based and hybrid exposure sets. Note that the same patient can appear in many exposure sets and therefore have many potential index dates.



Time-Based Exposure Sets

- Identify comparator drug user prescriptions within a caliper of the study drug user's time-since cohort entry.
- Example caliper +/- 1 month

Prescription-Based Exposure Sets

- Identify comparator drug users with a the same number of prior prescription for the comparator

Hybrid Exposure Sets

- Combines elements of prescription and time-based exposure sets with additional calendar time criteria
- Example: identify comparator drug users with the same prior comparator prescription count, time-based caliper (+/- 1 month) and same calendar year

Figure 3: Flowchart of study identification and inclusion