Review

A Review of Epidemiological Approaches for Answering Causal Questions using the Korean National Health Insurance Claims Database

Juhee Ahn[1,2,*]

[1]Devision of AI Data Science, College of Intelligent Software Convergence, University of Suwon, Gyeonggi-do, Republic of Korea

[2]Institute of Health and Environment, Seoul National University, Seoul, Republic of Korea

Abstract

The Korean National Health Insurance Database (KNHID) has become a resource for comparative effectiveness research since 2014. However, studies based on the KNHID are inherently vulnerable to bias since claims data are not generated for research purposes. As a result, alignment of the research questions, operational definitions, and time zero is essential. This narrative and focused review summarizes common threats to validity in claims-based studies, and discusses epidemiological approaches and analytic strategies applicable to KNHID research: comparative cohort study, self-controlled case series study, and target trial emulation framework. Comparative cohort study provides a practical default for comparisons of treatment groups, and the landmark method can mitigate immortal time bias. Self-controlled case series study effectively controls for time-invariant confounders through within-person comparisons. More recently, target trial emulation framework has been proposed to approximate randomized controlled trials using observational data, and analytic methods should be selected according to the emulated protocol. In conclusion, this review provides practical guidance for selecting and implementing appropriate approaches in KNHID-based research, tailored to threats to validity and corresponding solutions. The credibility of evidence generated from the KNHID also depends on transparent reporting and robustness checks.

keywords: *Korean National Health Insurance Database, comparative cohort study, self-controlled case series study, target trial emulation framework*

Introduction

Since the release of the National Sample Cohort from National Health Insurance Service (NHIS) of Korea in 2014, the Korean Health Insurance Database (KNHID) has been widely used for comparative effectiveness research [1]. Notably, nearly 97% of South Korea's population is covered under a single-payer national health insurance system, enabling population-wide coverage of healthcare utilization [2]. This feature makes the KNHID a crucial resource for nationwide epidemiologic research. Nevertheless, these data are automatically generated for reimbursement purposes rather than research, and therefore require data extraction, operational definitions, and selection of study designs that align with the specific research question.

Randomized controlled trials (RCTs) are commonly regarded as the preferred methodology for assessing the causal relationships between exposure and its impact on the outcome, providing unbiased estimates [3]. They involve random allocation of individuals to different treatment groups to address confounders and to ensure comparability between them. Consequently, RCTs can provide high-quality scientific evidence through these characteristics. However, when RCTs are economically, ethically, or practically infeasible, observational studies using administrative claims data often represent the viable approaches to answer important clinical and public health questions. Regardless of the data source, however, observational studies are inherently susceptible to multiple forms of bias [4, 5].

*** Corresponding author:** Juhee Ahn (jhahn@suwon.ac.kr, 031-220-8831)
Division of AI Data Science, College of Intelligent Software Convergence, University of Suwon, 17 Wauan-gil, Bongdam-eup, Hwaseong-si, Gyeonggi-do, Republic of Korea.

To minimize these limitations, a range of epidemiologic methodologies has been proposed [6-9]. This review focuses on a selected set of methodological approaches that are implementable using the KNHID and directly linked to well-defined epidemiologic estimands. These approaches are discussed in a practice-oriented manner to help researchers more effectively leverage the KNHID to address public health concerns.

This narrative review was informed by targeted literature searches of PubMed and Google Scholar for methodological papers and book, with emphasis on KNHID and common threats to validity. Search terms included combinations of “Korea National Health Insurance data” OR “cohort study” OR “immortal time bias” OR “landmark method” OR “self-controlled case series” OR “target trial emulation” OR “clone-censor-weight” as well as reporting guidance. Reference lists of key articles were also screened to identify additional influential publications. Articles were selected based on relevance to the scope of this review, methodological contribution, and applicability to KNHID-based research.

Common threats and considerations in Korean National Health Insurance claims studies

KNHID provides information on patients’ demographics, healthcare utilization/claims, prescription records, characteristics of healthcare provider, biometric data, and diagnoses. Diagnostic information is recorded based on the Korean Classification of Diseases (KCD), a nationally modified version of the International Classification of Diseases. These data are used to operationally define treatment strategies, outcomes, covariates, and follow-up rules in epidemiologic studies.

A primary threat to validity is misclassification in claims-based studies, attributable to imperfect agreement between definitions derived from the administrative data and actual clinical practice, as well as restricted clinical granularity [10-12]. As KNHID is not designed primarily for research, features are necessarily dependent on researcher-derived definition, which may vary in sensitivity and specificity across research questions. For instance, dyslipidemia can be identified based on at least one inpatient or outpatient record with a KCD code of E78 in combination with the use of lipid-lowering medications, or alternatively defined using laboratory data such as total cholesterol or low-density lipoprotein cholesterol levels. Even with well-constructed definitions based on prior literature or researchers, true clinical diagnoses and medication adherence may differ from what is captured in claims data. Transparent reporting of coding algorithms and sensitivity analyses using alternative definitions are therefore essential.

A second threat relates to temporal alignment, particularly when the definition of time zero and initiation of treatment strategy are misaligned. In pharmacoepidemiologic studies, treatment initiation often occurs after diagnosis or cohort entry, and dates of claims may reflect dispensing or prescribing processes, rather than the exact timing of medication intake. For example, among patients diagnosed with dyslipidemia, some are prescribed lipid-lowering medications on the same day as diagnosis, whereas others initiate treatment at a later time. Failure to properly align treatment assignment, eligibility assessment, and the start of follow-up can introduce immortal time bias (Figure 1A), leading to biased effect estimates in either direction [13, 14]. Related issues may also arise in case of treatment switching or discontinuation if not handled appropriately [15]. Furthermore, operational definitions should specify how treatment initiation, discontinuation, switching, and allowable gaps are handled and should align with the research questions.

Third, claims-based research in Korea is commonly conducted using data provided by two national public institutions, the NHIS and the Health Insurance Review and Assessment Service (HIRA). The availability of linked mortality information differs by source. NHIS data can be provided with death information, whereas HIRA data typically do not include mortality data. When death data unavailable, researchers may need to link to mortality records from Statistics Korea, or alternatively, define loss of follow-up in eligibility of health insurance. Therefore, researchers should clearly report how follow-up and censoring rules were operationalized given structure of the data.

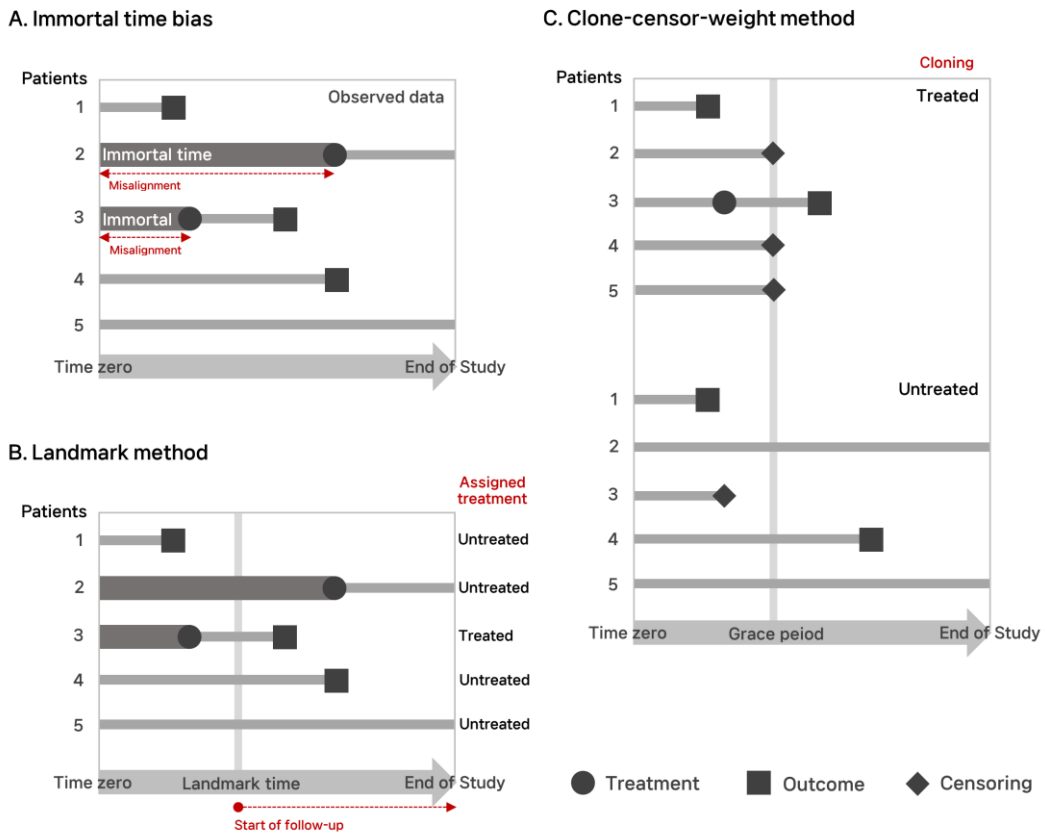

Figure 1. Immortal time bias and temporal alignment strategies

Finally, unmeasured and residual confounding remains a limitation of observational studies using administrative data, especially for factors not captured in claims databases, such as disease severity, education, health behaviors, socio-economic context, or provider-level characteristics [16]. Where possible, use of proxy variables—for example, patterns of healthcare utilization—or sensitivity analyses are important to assess the credibility of causal interpretations.

Given these KNHID-specific features, transparent reporting of key elements is essential across approaches. RECORD (REporting of studies Conducted using Observational Routinely-collected Data) [17] or STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) [18] provide general guidance for conventional epidemiological approaches. The recently proposed TARGET (TrAnsparent ReportinG of observational studies Emulating a Target trial) guideline provides structured recommendations for reporting target trial emulation studies [19].

Epidemiological approaches in claim-based database

Table 1 summarizes complementary epidemiological approaches discussed in this review and highlights their analytic strategies. Design-specific pitfalls and practical solutions are provided in the corresponding “Pitfalls & Solutions” for each section.

Comparative cohort study

In observational epidemiologic research, cohort studies are an established design. They are used to determine the association between a specific treat-

Table 1. Matrix for epidemiologic methodological approaches in claims-based research

	Comparative cohort study	Self-controlled case series study	Target trial emulation
Conceptual level	Study design	Study design	Causal framework
Primary purpose	Comparing outcomes between two or more groups defined by treatment strategy in observational studies		
Analytic strategy	- Landmark method	-	- Clone-censor-weight method
Key assumptions	- Exchangeability - Positivity - Consistency	- Rare/non-recurrent outcomes or independent/recurrent outcomes - Outcome-independent treatment probability - Outcome-independent observation periods	- Exchangeability - Positivity - Consistency - Correct model specification for treatment and censoring mechanisms
Strengths	- Traditional and established method - Conceptually straightforward ** Landmark method - Avoids immortal time bias - Improved interpretability and simplicity	- Effectively control for time-fixed covariates, even unmeasured - No control group needed	- Conceptually powerful framework - Avoids immortal time bias (clone-censor-weight) - Emulation of hypothetical randomized controlled trial ** Clone-censor-weight method - Avoids immortal time bias
Limitations	- Unmeasured confounding - Selection bias - Time-related bias (Immortal time bias) - Reverse causation ** Landmark method - Survivor bias - Reduced sample size - Dependent on landmark time - Time-varying confounding	- Not suited if outcome strongly affects subsequent treatment or observation - Time-varying confounding - Event-dependent censoring	- Requires rich longitudinal data - Unmeasured confounding - Model misclassification - Sensitive to protocol choices ** Clone-censor-weight method - Dependent on grace period - Computationally intensive
Reporting guideline	RECORD* guideline, STROBE† statement		TARGET‡ guideline

*RECORD: Reporting of studies Conducted using Observational Routinely-collected Data
†STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
‡TARGET: TrAnsparent ReportinG of observational studies Emulating a Target trial

ment strategy and subsequent outcomes by following-up outcome-free individuals at baseline. In a comparative cohort study, outcomes are contrasted between two or more treatment groups—for example, treatment group vs. comparator group.

Cohort studies can be distinguished as prospective or retrospective [20]. Prospective cohort studies detect treated and untreated individuals at baseline and follow them forward in time to observe the incidence of outcome. On the other hand, retrospective cohort studies collect pre-existing data, such as administrative claims or electronic health records, to ascertain both treatment strategies and outcomes after they have occurred. Although the former is more time-consuming and costly than the latter in general, it provides a higher level of scientific evidence [7, 21].

Comparative cohort studies are dependent on several key assumptions for causal interpretation: exchangeability, positivity, and consistency [9]. First, exchangeability requires that treated and untreated groups are comparable conditional on measured confounders. Second, positivity assumption requires that the probability of receiving each treatment strategy, conditional on confounders, lies between 0 and 1. Lastly, consistency implies that the observed outcome under the observed treatment strategy corresponds to the potential outcome under that treatment strategy. In practice, exchangeability is often violated due to unmeasured confounding factors or selection bias in pharmacoepidemiologic studies [9, 22].

To address measured confounding effects, covariate adjustment or propensity score-based methods—adjustment, matching, inverse probability of treatment weighting, or stratification—are often used in comparative cohort studies [23]. Time-to-event outcomes are commonly evaluated using Cox proportional hazards models and yield adjusted hazard ratios. Increasingly, absolute measures such as restricted mean survival time are recommended to enhance clinical interpretability [24, 25].

A strength of comparative cohort studies is to evaluate real-world effectiveness in representative populations with long follow-up, especially when RCTs are not possible. However, there are some limitations. Even with advanced methods to control confounding effects, residual confounding from unmeasured covariates cannot be completely eliminated. In addition, selection bias, time-related biases including immortal time bias, and reverse causation may still threaten validity of research. These limitations emphasize the necessity of rigorous approaches and reporting guidelines, and assessment of robustness.

One practical analytic strategy to mitigate time-related bias in cohort-based analyses is the landmark method (Figure 1B), which has been proposed to address immortal time bias in observational studies [26]. Immortal time bias arises when treatment assignment, eligibility assessment, and the start of follow-up are not properly aligned (Figure 1A). This misalignment creates a period during which the outcome cannot occur, which leads to biased effect estimates in favor of the treated group [14].

A key feature of landmark method is that it compares outcome across treatment strategies conditional on survival until a prespecified landmark time. The landmark time is defined a priori and treatment status is determined using information accrued before that time [8]. Follow-up starts at the landmark time, not at cohort entry time or date of diagnosis. By synchronizing treatment assignment with the start of follow-up, landmark method can reduce immortal time bias and provide valid comparisons between treatment groups [8]. Standard time-to-event methods are applied, and hazard ratios or absolute measures are commonly estimated. These estimands should be interpreted as the effect of treatment among individuals who have survived up to the landmark time [8].

Landmark method has several practical advantages. Compared with time-varying exposure models as an alternative method to addressing immortal time bias, landmark method is often more intuitive to interpret in clinical settings [8, 27]. Moreover, conditional time-to-event probabilities can be estimated without immoral time bias [28]. Nevertheless, the landmark method is limited for causal interpretation when time-varying confounding is present, which is common in survival analyses. In such settings, landmark-based estimates may be biased and should be interpreted cautiously [29-31]. Alternative methods that address time-varying confounding—marginal structural model—may be more appropriate when

sufficiently rich longitudinal covariate data are available [15].

Additional considerations are required when implementing landmark-based methods. The landmark time should be prespecified based on clinical or biological rationales, and not be chosen post-hoc based on observed results [8, 26]. Different landmark times redefine the eligible study population and may yield different effect estimates. Especially, the use of longer landmark time can reduce statistical power due to loss of sample size and may limit generalizability to the target population [26]. These differences may also be influenced by treatment rates, event rates, and survival characteristics of the population, which may potentially result in treatment misclassification and selection bias [26, 32]. Thus, sensitivity analyses using multiple landmark times are recommended to evaluate the robustness of findings and to evaluate whether conclusions are driven by a specific choice of landmark [26].

Pitfalls & Solutions in comparative cohort study
Pitfalls (in additions to the common threats described above)

- **Confounding by indication:** Unmeasured clinical severity and health behaviors may bias effect estimation.
- **Time-zero misalignment:** Misalignment of eligibility, treatment strategy classification, and start of follow-up can induce immortal time bias.

Practical strategies

- Prefer a new-user, active-comparator design when feasible; implement robust confounding control (propensity score matching/weighting/stratification with covariance balance diagnostics), and consider negative control outcomes/treatments when appropriate.
- Explicitly define time zero and align it with treatment strategy classification; consider landmark method, and assess robustness to alternative definitions of time-zero or treatment episodes.

Self-controlled case series study

Case-only designs have been proposed as an alternative epidemiologic approach in order to overcome unmeasured confounders in traditional cohort studies [22]. Among these, the self-controlled case series (SCCS) study includes only individuals who have experienced the outcome [33]. By design, each individual serves as their own control across different risk periods. It effectively minimizes confounding effects by time-invariant factors, whether measured or unmeasured [34]. The SCCS study was originally developed to evaluate vaccine safety, a context in which outcomes are typically acute and the timing of treatment is well defined [34]. Accordingly, the standard SCCS study is suited for detecting short-term effects of treatments.

The SCCS study has been applied in post-approval pharmacoepidemiologic studies [6]. It compares the incidence of events during treated periods with that during untreated periods within the same individual using a time-varying structure. This feature allows the design to account for dynamic patterns of treatment strategies [22]. Relative incidences are statistically estimated as the primary effect using conditional Poisson regression. This measure refers to the relative incidence of outcomes during treated versus untreated periods.

Application of the SCCS study relies on several key assumptions [6]. First, recurrent outcomes arise independently, or the outcome should be rare if outcomes are non-recurrent. Second, most importantly, the occurrence of the outcome should not influence subsequent treatment strategy and period of observation. Third, although not specific to the SCCS study, treatment strategies should not influence the ascertainment of outcomes. The SCCS study has been extended to accommodate more complex settings in which the second assumptions may be violated including outcome-dependent treatments or outcome-dependent observation periods. Moreover, extensions for spline-based age or treatment effects, multi-type outcomes, quantitative treatments, and environmental exposures have been suggested [6]. These extensions of methods broaden the applicability of SCCS beyond vaccine safety research.

Despite these strengths, the SCCS study is still vulnerable to time-varying confounders such as changes in disease severity over time. In addition, reverse causation may occur if the outcome alters

future patterns of treatment strategy [35]. Finally, since SCCS includes only cases, it cannot estimate absolute risks or incidence rates at the population level [34].

Cohort and SCCS studies are often referred to as complementary in nature, rather than competing designs. Comparing findings from cohort studies and SCCS studies can strengthen robust interpretation [6], as concordant results across designs with different underlying assumptions increase confidence in the observed associations.

Pitfalls & Solutions in SCCS study

Pitfalls (in additions to the common threats described above)

- **Risk period misspecification:** Dates of claims may not represent exact clinical onset.
- **Key assumption violations:** Especially when outcomes influence subsequent treatment strategies and period observation.
- **Time effects:** Age, seasonality, calendar time trends, and time-variant confounders can confound within-person comparisons if not appropriately modeled.

Practical strategies

- Pre-specify risk periods and define treatment episodes transparently (e.g., allowable gaps) and assess robustness using alternative definitions of risk periods and episodes.
- Evaluate plausibility of SCCS assumptions in KNHID and perform sensitivity analyses (e.g., pre-risk periods).

Target trial emulation framework

To approximate the causal inference framework of RCTs using observational data, target trial emulation has been proposed, designed as if they were attempting to emulate a hypothetical randomized trial—the "target trial" [36]. Although randomization and blinding make RCTs the highest level of evidence, such designs are often infeasible or unethical for many clinical questions. The concept of the target trial originates from the framework of pragmatic trials, which aim to evaluate the effectiveness of interventions in real-world clinical settings rather than under highly controlled settings of explanatory RCTs [37]. Target trial emulation operationalizes this concept by specifying key protocol components—eligibility criteria, treatment strategies, random assignment, start and end of follow-up, outcomes, causal contrasts, and statistical methods [36].

Since target trial emulation is a framework, not a single analytic method, the choice of analytic implementation should be guided by the protocol, the causal contrasts, and the structure of the available observational data. One of the key components in of a hypothetical randomized trial is random assignment, which target trial emulation attempts to approximate using observational data. It is specified as the hypothetical allocation mechanism that would operate in the target RCT, even though treatment strategies in observational data is not randomized [36]. Another step is to define the causal contrasts, most commonly aligned with either an intention-to-treat effect or a per-protocol effect. The intention-to-treat corresponds to the causal effect of being assigned to a treatment strategy at time zero, regardless of subsequent treatment deviations. In contrast, the per-protocol effect corresponds to the causal effect that would be observed if individuals adhered to their assigned treatment strategy throughout follow-up.

As a practical analytic strategy of target trial emulation in Figure 1C, clone-censor-weight method has been proposed by Hernán and colleagues [36, 38]. An advantage of this method is its ability to address immortal time bias (Figure 1A), estimating per-protocol effects in observational settings [39]. The clone-censor-weight method consists of three core steps:

1. Cloning: At the time of eligibility, each individual is copied and simultaneously assigned to each treatment strategy (e.g., treatment vs. comparator). Consequently, an individual contributes analytical units corresponding to different hypothetical treatment strategies.
2. Censor: Each clone is censored at the time when its observed treatment deviates from the assigned treatment strategy during the follow-up. For example, if a clone assigned to the treatment strategy fails to initiate treatment within the grace period, that clone is censored at the end of the grace period. Here, the grace period is prespecified window after time zero which treatment

initiation is allowed to occur, which reflects real-world delays.

3. Weighting: This artificial censoring is generally informative since treatment deviations are often related to prognostic factors and can therefore introduce selection bias [32]. To address for this issue, inverse probability of censoring weights is used to reweight the uncensored clones so that they represent the full population that would have remained under each treatment strategy. These weights are incorporated into weighted regression models [32].

Several important assumptions are also required in target trial emulation, similar to those in cohort studies, including exchangeability, positivity and consistency. In addition, correct specification of the models used for treatment and censoring mechanisms is essential [40]. Violations of these assumptions may still lead to biased effect estimates.

Target trial emulation offers a conceptually powerful framework for causal inference in observational research, particularly when RCTs are infeasible. However, several practical limitations should be considered at two levels. First, at the framework level, valid emulation requires a clearly specified protocol and rich longitudinal data to operationalize key components of target trial emulation; unmeasured confounding and misclassification may still threaten validity. Second, at the analytic strategy level, implementing clone-censor-weight method can be computationally intensive and requires adequate longitudinal information to model censoring processes and treatment deviation; results may also be sensitive to the choice of the grace period, which should reflect clinical practice and plausible delays in treatment initiation [9].

Existing epidemiological study can benefit from articulating the key components of target trial emulation, as these elements substantially overlap with those of conventional observational designs. Even without a full emulation, aligning study design and analysis with the target trial framework can provide a transparent structure and strengthen the validity of inferences drawn from claims-based data.

Pitfalls & Solutions in target trial emulation framework

Pitfalls (in additions to the common threats described above)

- **Protocol infeasibility:** Eligibility, time zero, treatment strategies, outcomes, and deviations may be difficult to operationalize using claims data alone.
- **Estimand-implementation mismatch:** Misalignment between the estimand and the chosen analytic implementation can lead to biased or misinterpreted effect estimates.
- **Weight/model instability:** When weighting is used, limited time-varying covariate history and violation of the positivity assumption can yield extreme or unstable weights and amplify model misspecification.

Practical strategies

- Specify the target trial protocol and ensure each component is measurable in KNHID.
- Select an analytic implementation aligned with the target estimand and justify key protocol choices.
- When using weighting (e.g., clone-censor-weight method for per-protocol), use appropriate longitudinal covariates, report weight diagnostics, and consider stabilization/truncation.

Conclusion

The methodological approaches discussed in this review can cover a range of epidemiological questions that are feasible using claims-based data; however, valid KNHID-based inference requires alignment of research question, operational definitions, and time zero. A practical guideline for selecting an approach is to consider (i) causal contrast of interest, (ii) treatment dynamics, and (iii) data feasibility. Comparative cohort studies provide a practical default for group comparisons when confounding control and time-zero alignment are feasible, and the landmark method can be more straightforwardly applied within cohort analyses when treatment initiation is delayed. SCCS is useful when within-person comparisons are appropriate and risk periods can be defined reliably.

Target trial emulation offers an overarching framework to make the protocol and estimand explicit, with the clone-censor-weight method serving as one practical implementation for per-protocol effects when deviations must be handled. Overall, transparent reporting and robustness checks are essential for producing credible evidence from the KNHID. Where appropriate, articulating key components of the target trial emulation framework can strengthen causal interpretation and improve transparency in KNHID-based research.

Funding

This work was supported by grants from the National Research Foundation of Korea (NRF), funded by the Ministry of Education through the Basic Science Research Program (RS-2024-00450664) and by the Ministry of Science and ICT, Republic of Korea (RS-2026-25499279).

Conflicts of Interest

The author serves as an external committee member of the Institute of Health and Environment at Seoul National University. The author had no role in the editorial decision-making or peer-review process for this manuscript and has no potential conflicts of interest to disclose.

References

1. Kim DW. Statistical methods for baseline adjustment and cohort analysis in Korean National Health Insurance claims data: a review of PSM, IPTW, and survival analysis with future directions. Journal of Korean Medical Science. 2025;40(8):e110.
2. Seong SC et al. Data resource profile: the national health information database of the National Health Insurance Service in South Korea. International journal of epidemiology. 2017;46(3):799–800.
3. Burns PB, Rohrich RJ, Chung KC. The levels of evidence and their role in evidence-based medicine. Plastic and reconstructive surgery. 2011;128(1):305–310.
4. Hammer GP, du Prel JB, Blettner M. Avoiding bias in observational studies: part 8 in a series of articles on evaluation of scientific publications. Deutsches Ärzteblatt International. 2009;106(41):664–668.
5. Sterrantino, AF. Observational studies: practical tips for avoiding common statistical pitfalls. The Lancet Regional Health-Southeast Asia. 2024;25: 100415.
6. Farrington P, Whitaker H, Weldeselassie YG. Self-controlled case series studies: a modelling guide with R. Boca Raton: CRC Press, Taylor & Francis Group. 2018.
7. Rothman KJ, Greenland S, Lash TL. Modern epidemiology, Vol 3. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins. 2008.
8. Morgan CJ. Landmark analysis: a primer. Journal of Nuclear Cardiology. 2019;26(2):391–393.
9. Hernán MA, Robins JM. Causal Inference: What If. Boca Raton: Chapman & Hall/CRC. 2010.
10. Carl vW. A comparison of methods to correct for misclassification bias from administrative database diagnostic codes. International journal of epidemiology. 2018;47(2):605–616.
11. Jonsson Funk M, Suzanne N L. Misclassification in administrative claims data: quantifying the impact on treatment effect estimates. Current epidemiology reports. 2014;1(4):175–185.
12. Nicholls SG, Langan SM, Benchimol EI. Routinely collected data: the importance of high-quality diagnostic coding to research. CMAJ. 2017;189(33):E1054–E1055.
13. Suissa S. Immortal time bias in pharmacoepidemiology. American journal of epidemiology. 2008;167(4):492–499.
14. Lévesque LE et al. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. BMJ. 2010;340:b5087.
15. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology. 2000;11(5):550–560.
16. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. Journal of clinical epidemiology. 2005;58(4):323–337.
17. Benchimol EI et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS medicine. 2015;12(10):e1001885.

18. Von Elm E et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Journal of clinical epidemiology. 2008;61(4):344–349.
19. Cashin AG et al. Transparent reporting of observational studies emulating a target trial—the TARGET statement. JAMA. 2025;334(12):1084–1093.
20. Wang X, Kattan MW. Cohort studies: design, analysis, and reporting. Chest. 2020;158(1S):S72–S78.
21. Rochon PA et al. Reader's guide to critical appraisal of cohort studies: 1. Role and design. BMJ. 2005;330(7496):895–897.
22. Nordmann S et al. Case-only designs in pharmacoepidemiology: a systematic review. PLOS ONE. 2012;7(11):e49444.
23. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983;70(1):41–55.
24. Kim DH, Uno H, Wei LJ. Restricted Mean Survival Time as a Measure to Interpret Clinical Trial Results. JAMA Cardiology. 2017;2(11):1179–1180.
25. Royston P, Parmar MKB. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. BMC Medical Research Methodology. 2013;13:152.
26. Mi X et al. Use of the landmark method to address immortal person-time bias in comparative effectiveness research: a simulation study. Statistics in Medicine. 2016;35(26):4824–4836.
27. Fisher LD, Lin DY. Time-dependent covariates in the Cox proportional-hazards regression model. Annual review of public health. 1999;20:145–157.
28. Dafni U. Landmark analysis at the 25-year landmark point. Circulation: Cardiovascular Quality and Outcomes. 2011;4(3):363–371.
29. Weymann D, Krebs E, Regier DA. Addressing immortal time bias in precision medicine: practical guidance and methods development. Health Services Research. 2025;60(1):e14376.
30. Pazzagli L et al. Methods for time-varying exposure related problems in pharmacoepidemiology: an overview. Pharmacoepidemiology and drug safety. 2018;27(2):148‐160.
31. Hernán MA et al. Observation plans in longitudinal studies with time-varying treatments. Statistical methods in medical research. 2009; 18(1):27–52.
32. Hernán MA et al. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. Journal of Clinical Epidemiology. 2016;79:70–75.
33. Farrington P et al. A new method for active surveillance of adverse events from diphtheria/tetanus/pertussis and measles/mumps/rubella vaccines. Lancet. 1995;345(8949):567–569.
34. Farrington CP, Relative incidence estimation from case series for vaccine safety evaluation. Biometrics. 1995;51(1):228–235.
35. Takeuchi Y, Shinozaki T, Matsuyama Y. A comparison of estimators from self-controlled case series, case-crossover design, and sequence symmetry analysis for pharmacoepidemiological studies. BMC Medical Research Methodology. 2018;18(1):4.
36. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. American Journal of Epidemiology. 2016;183(8):758–764.
37. Ford I, Norrie J. Pragmatic trials. New England journal of medicine. 2016;375(5):454–463.
38. Hernán MA, Wang W, Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. JAMA. 2022;328(24):2446–2447.
39. Danaei G et al. Electronic medical records can be used to emulate target trials of sustained treatment strategies. Journal of clinical epidemiology. 2018;96:12–22.
40. Reep CAT, Wils EJ, Heunks L. Opportunities, challenges and future perspectives for target trial emulation in critical care clinical research. Critical Care. 2025;29:484.