Journal of Clinical Epidemiology 185 (2025) 111876

Journal of Clinical Epidemiology

ORIGINAL RESEARCH

Adherence to RECORD reporting guidelines among observational studies using routinely collected health data published in general medical journals: a metaepidemiologic study

Heather J. Zhao[a,b,1], Inna Ushcatz[a,b,1], Chantel Walwyn[a], Megan S. Lowe[c], Kevin S. Kim[d], Eric I. Benchimol[e,f,g,h], Sinéad M. Langan[i], Aaron M. Drucker[a,j,*]

[a]*Women's College Research Institute and Department of Medicine, Women's College Hospital, Toronto, Ontario, Canada M5G 1N8*
[b]*Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada M5S 1A8*
[c]*Faculty of Medicine, Queens University, Kingston, Ontario, Canada K7L 3N6*
[d]*Department of Health Research Methodology, Evidence and Impact, McMaster University, Hamilton, Ontario, Canada L8S4K1*
[e]*Division of Gastroenterology, Hepatology and Nutrition, The Hospital for Sick Children, Toronto, Canada M5G 1X8*
[f]*Department of Paediatrics and Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Canada M5S 1A8*
[g]*ICES, Toronto, Canada M4N 3M5*
[h]*Child Health Evaluative Sciences, SickKids Research Institute, The Hospital for Sick Children, Toronto, Ontario, Canada M5G 1X8*
[i]*London School of Hygiene and Tropical Medicine, London, United Kingdom WC1E 7HT*
[j]*Division of Dermatology, Department of Medicine, University of Toronto, Toronto, Ontario, Canada M5T 2S8*

Accepted 19 June 2025; Published online 23 June 2025

Abstract

Objectives: The REporting of studies Conducted using Observational Routinely Collected Data (RECORD) tool was developed to address gaps around reporting routinely collected health data. The objective of this study was to assess adherence to RECORD in general medical journals and to evaluate its correlation with study quality.

Methods: We searched PubMed using a filter to identify studies using routinely collected health data published in 8 high impact medical journals between 2016 and 2023. Four journals endorsed RECORD, while 4 did not. For each journal, 24 articles were randomly selected, with 3 studies per year. Study characteristics, RECORD and quality assessments were completed in duplicate and described using proportions and means with SDs. Linear regression was used to estimate the association between journal and study characteristics with adherence to RECORD items.

Results: Studies reported a mean of 70.7% (SD 1.8%) of RECORD items. There was no substantial difference in adherence in RECORD-endorsing journals compared to non−RECORD-endorsing journals (1.8% lower adherence; 95% CI: -5.8, 2.2). Adherence of >80% was reported for RECORD items 1.1, 1.2, 6.1, 7.1, 19.1 and 22.1.

Conclusion: Studies in general medical journals had moderate adherence to RECORD, with no association between journals' endorsement of RECORD and reporting completeness. Other measures to improve adherence to RECORD should be explored, including refinements to the checklist itself. Authors and journals should be aware of and adhere to items required for RECORD reporting to improve the reproducibility of research using routinely collected health data.

Keywords: Observational; Routinely collected; RECORD guidelines; Adherence; Study quality; Reporting guidelines

1. Introduction

Routinely collected health data are increasingly important resources for medical research [1,2]. Data from sources like billing records and electronic health records give researchers access to information on millions of people, making it easier to answer important health-related questions efficiently. The REporting of studies Conducted using

[1] Equal contributions, shared first authorship.

* Corresponding author: Division of Dermatology, Department of Medicine, Women's College Research Institute, Women's College Hospital, Department of Medicine, University of Toronto, 76 Granville Street, Toronto, Ontario M5S 1B2, Canada.
E-mail address: aaron.drucker@wchospital.ca (A.M. Drucker).

https://doi.org/10.1016/j.jclinepi.2025.111876

What is new?

Key findings

- Mean adherence to REporting of studies Conducted using Observational Routinely Collected Data (RECORD) items was 70.7%, with no important difference between RECORD-endorsing journals (69.8%) and nonendorsing journals (71.6%).
- High adherence (>80%) was observed for items like data type description (1.1), study timeframe/geographic region (1.2), and inclusion/exclusion criteria (6.1).
- Critical gaps were noted in reporting database linkage details (1.3: 51%), validation of codes/algorithms (6.2: 49.5%), graphical linkage displays (6.3: 10.8%), and data access statements (12.1: 76% in nonendorsing vs 20.8% in endorsing journals).
- No association was found between RECORD adherence and study quality (Newcastle-Ottawa Scale scores) or journal characteristics (eg, impact factor, funding source).

What this adds to what is known?

- Journal endorsement of RECORD did not improve adherence, suggesting endorsement alone is insufficient to ensure compliance.
- The study identifies specific RECORD items (eg, data linkage transparency, validation methods) as consistent weaknesses, pinpointing areas for targeted improvement.
- Unlike findings for other reporting guidelines (e.g., Strengthening the Reporting of Observational Studies In Epidemiology), RECORD adherence was not correlated with study quality, emphasizing that reporting and methodological rigor are distinct issues.

What is the implication and what should change now?

- For research communities, it is important to investigate barriers to adherence (eg, space constraints, lack of incentives).
- For RECORD developers, it is important to refine the checklist to clarify ''suggested'' vs ''mandatory'' items, such as linkage and diagrams.
- For authors, the pitfalls for poor reporting include transparency if validating variables and if linkage was used, to clarify how databases were linked.

Observational Routinely Collected Data (RECORD) guidelines, published in 2016 as an extension of the Strengthening the Reporting of Observational Studies In Epidemiology (STROBE) guidelines, focused on items specific to routinely collected data, such as linkage of databases, data availability, and validation of codes and algorithms to identify subjects, exposures, outcomes, and other variables [3]. RECORD was developed in collaboration with over 100 international stakeholders across health care fields to promote transparency and improve the reliability of research that guides health-care practices [4].

Studies have shown that adherence to other reporting guidelines like Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and Consolidated Standards of Reporting Trials (CONSORT) is highly variable [5−7]. Poorly reported studies tend to be of lower quality and at higher risk of bias than studies that better adhere to reporting guidance [8−10]. Although a few studies have assessed RECORD adherence, they focused on narrow contexts such as pharmacoepidemiology or neurosurgery [11−13]. RECORD adherence in the general medical literature remains unexplored, including high impact general medical journals where studies have the most potential to widely influence clinical practice. Additionally, some general medical journals have endorsed RECORD whereas others have not, providing an opportunity to study the influence of journal endorsement on RECORD reporting.

The objective of this study was to assess adherence to RECORD reporting guidance among observational studies using routinely collected health data published in general medical journals, and whether study characteristics, including study quality and the influence of journal endorsement of RECORD, are associated with better reporting.

2. Materials and methods

We drafted and posted a protocol before initiating this study (https://osf.io/t3sgp/).

2.1. Study identification

To focus on observational studies likely to have the most influence on clinical practice, we searched for studies published in the top 4 highest-impact medical journals that have not endorsed RECORD reporting (The Lancet, The New England Journal of Medicine (NEJM), JAMA, Nature Medicine), and the top four highest-impact medical journals that endorse RECORD (British Medical Journal (BMJ), Public Library of Science (PLoS) Medicine, Canadian Medical Association Journal (CMAJ), and International Journal of Epidemiology). Impact factors were derived from the Science Citation Index from January

2016 to December 2023 and were accessed on May 6, 2024.

BMJ mandates submission of a reporting guideline but only lists examples such as RECORD for observational studies, while stating that other guidelines may be used if relevant [14]. PLoS Medicine mandates use of STROBE extensions guidelines as appropriate for study design outlined on the EQUATOR website [15,16]. CMAJ's submission guideline encourages authors to follow RECORD in addition to STROBE for observational studies that uses routinely collected data but does not mandate a submission of RECORD with the article [17]. International journal of Epidemiology's general instructions page requires all studies using routinely collected data to follow the RECORD statement but does not state if submission of the RECORD checklist with the article is mandatory [18].

We included observational studies conducted using routinely collected health data. We included cohort, case-control, and cross-sectional studies reporting at least one health-related outcome. Studies must be available as full-text articles (ie, not abstracts). Exclusion criteria included studies with medications as an exposure, because there are separate reporting guidelines for pharmacoepidemiology studies using routinely collected health data [19], and prediction or validation studies.

First, we searched PubMed using a filter to identify studies using routinely collected health data for studies published in 8 highest impact journals between January 1, 2016, and December 31, 2023. We adapted the search strategy from previous systematic reviews that identified studies that used routinely collected data [20−22]. Our full search is accessible in the supplement. We then screened titles and abstracts and reviewed full-text articles independently and in duplicate. Of included studies, we randomly selected three studies per year totalling 24 studies from each journal. If we could not find three studies per year for a given year in a given journal, we selected studies from the previous year. If an insufficient number of studies were available from the previous year, we selected studies from another journal in same RECORD endorsing journal group. If we could not find enough studies from the same RECORD endorsing journal group from the same year, we selected studies from the previous year of the same RECORD endorsing group.

2.2. Data extraction and quality assessment

Two reviewers completed data extraction, RECORD and quality assessment independently, in duplicate. For the first three studies, three authors extracted data using the RECORD statement, assessed study quality using the Newcastle-Ottawa Scale (NOS), then held a consensus meeting to ensure harmonization. This process was repeated three times until all authors' initial evaluations reached >80% agreement. Two new authors underwent training and assessment on their first ten studies to ensure >80% agreement with one of the original three authors. For subsequent studies, scores from three authors were compared and discrepancies were resolved through discussion. If discrepancies could not be resolved, a third author was consulted.

For each included study, we recorded study characteristics including the geographic region of the data source, number of authors, sample size, and funding source [9,23]. In addition, we recorded the year of study publication, journal name, and if the study reported using the RECORD guideline or any other reporting guideline [9].

Each included study was assessed by two independent authors as to whether it provided the required information to meet each item of the RECORD checklist per the RECORD guideline. Each study's quality was assessed using the NOS for case-control, cohort, or cross-sectional studies as applicable [10].

The RECORD tool contains 13 individual items across five sections: title and abstract, methods, results, discussion, and supplementary information. Studies were granted points for 3 RECORD items in the title and abstract based on identification of data type, geographic location, and time period of the study. For all other RECORD items, points were granted if the necessary information was present anywhere throughout the paper or supplementary material. Each RECORD item was assessed using a binary scoring system, with one point awarded if the reporting criterion was met and zero if it was not. When evaluating reporting of RECORD items, studies were divided into 2 categories, single-database or linked database. Studies that used a single database were evaluated out of a denominator of 11 items because RECORD items 6.3 (linkage of database graphical representation) and 12.3 (related to database linkage) were not applicable. For item 6.2 (validation studies for selection of study population), a point was provided to all studies that mentioned or referenced a validation study for selection of their study population. When evaluating each RECORD item, we referred to the RECORD statement that provided examples and rationale for sufficient RECORD item reporting [24]. We presented RECORD adherence as a percentage because studies had different denominators depending on their design.

2.3. Statistical analysis

Study characteristics, RECORD reporting and NOS scores are described using proportions and means with SDs.

We used unadjusted and adjusted linear regression to estimate the association between journal and study characteristics with adherence to RECORD items. We included journals' RECORD-endorsing status (yes vs no), year of publication (continuous), geographic region of dataset (Europe, North America, Asia, Oceania, South America, Africa, or Intercontinental), primary funding source (government grant, not for profit, no funding, industry, or not specified), and whether the study reported use of

RECORD or other reporting guidelines (yes vs no) as predictor variables. We hypothesized that publication in a RECORD-endorsing journal and use of reporting guidelines would be associated with better RECORD reporting. We report effect estimates as absolute difference with their 95% CI. Because RECORD item 6.3 suggests, rather than mandates, a figure demonstrating database linkage, we conducted sensitivity analyses removing that item from the RECORD adherence proportion.

In a secondary analysis, we conducted unadjusted and adjusted linear regression to assess the association between the NOS and RECORD adherence. The adjusted model included the same covariates as the main analysis. We also calculated the Pearson correlation coefficient between NOS and RECORD adherence. We used the NOS to evaluate the methodological quality of observational studies, and hypothesized that better RECORD adherence would be associated with higher study quality. While the validity of NOS as a measure of study quality has been questioned [25,26], we used it for its relative simplicity and because it generates a summary score suitable for regression analysis.

All analyses were conducted using Stata v.16 (StataCorp LLC).

3. Results

3.1. Characteristics of included studies

The PubMed search identified 2903 publications, with two duplicates removed before screening (Fig 1). A total of 2901 studies were screened, and 473 articles were deemed eligible to be randomized for data extraction and analysis after abstract and full text review.

After the 473 articles were stratified by journal and year, 192 were randomized for inclusion in the study. There were 96 articles from RECORD-endorsing journals and 96 from

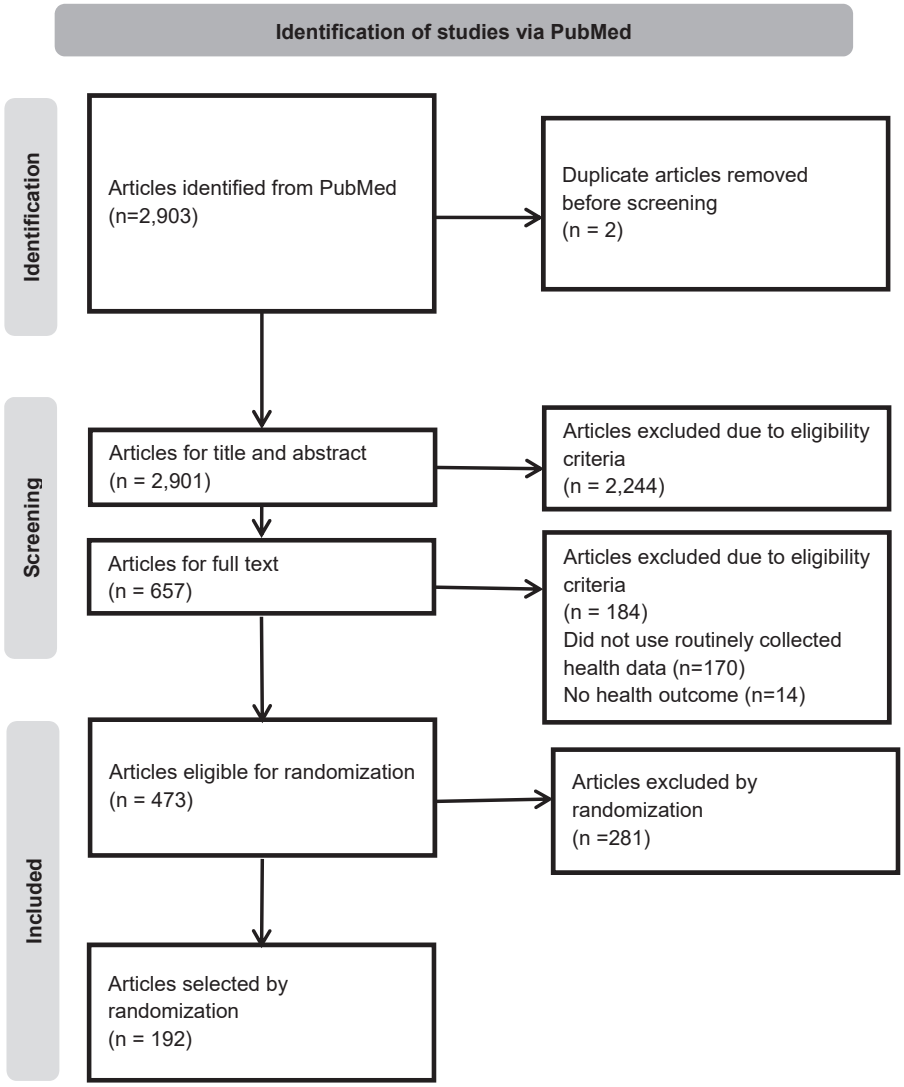

Figure 1. PRISMA diagram of included articles published from 2016 to 2023.

non−RECORD-endorsing journals (Table 1). We included 24 articles per journal, selecting three per year from 2016 to 2023 for *JAMA*, *The Lancet*, *International Journal of Epidemiology*, and *BMJ*. However, for *NEJM*, *Nature Medicine*, *CMAJ*, and *PLoS Medicine*, not every year had sufficient articles for inclusion. In these instances, we first aimed to select articles from a different year. If this was not feasible due to article availability, an article from a different journal was utilized (Fig 2).

Twenty-two of the 192 included studies (11.5%) reported used RECORD either independently or as an extension to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist. Thirty-six of 96 studies (37.5%) in the RECORD-endorsing journal group reported using any reporting guidelines, compared to 19 of 96 studies (19.8%) in the non−RECORD-endorsing journal group.

In total, 23 studies (12.0%) included COVID-19 as the main exposure or outcome. Primary funding included a combination of government grants and not-for-profit funding. European countries were the most common source of data (106 of 192, 55.0%).

3.1.1. Reporting of RECORD items

Studies reported a mean of 70.7% (SD = 1.8%) RECORD items, and reporting was similar between RECORD-endorsing journals (69.8%, SD = 1.6%) and non−RECORD-endorsing journals (71.6%, SD = 1.6%).

3.1.2. Assessing study quality

The 166 cohort studies had a NOS mean score of 7.9 (SD = 0.7%) out of 9 NOS items. There were 13 case control studies identified with a NOS mean score of 6.9 (SD = 0.7%) out of 9 NOS items which is consistent with previously outlined thresholds for high quality studies [27]. There were 13 cross-sectional studies with a NOS mean score of 5.8 (SD = 0.8%) out of 10 NOS items. Based on previous research, these were of satisfactory quality [28].

3.1.3. Adherence to specific record items

Adherence in >80% of studies was found for RECORD items 1.1, 1.2, 6.1, 7.1, 19.1, and 22.1 (Table 2). Almost all (90.1%) studies reported data type, study timeframe, and geographic region (RECORD 1.1 and 1.2). Similarly, most studies reported their inclusion and exclusion criteria (97.4%, RECORD 6.1), outlined the codes and algorithms used to define their variables (83.4%, RECORD 7.1), described implications, limitations and biases of data within the discussion (99.0%, RECORD 19.1), and provided references to supplementary materials (96.9%, RECORD 22.1).

There were 4 notable RECORD items, 1.3, 6.2, 6.3, and 12.1 reported by approximately 50% of studies. More than half (51.0%) reported on linkage of databases in the abstract or title. Similarly, almost half (49.5%) reported or referenced on validation studies of the codes or algorithm used to select the population. Few studies (10.8%) provided a graphical display or diagram to demonstrate the number of individuals included in a data linkage process. Lastly, while reporting was similar between RECORD-endorsing and nonendorsing journals, they differed on reporting of data access (RECORD 12.1); 76.0% of non−RECORD-endorsing journals reported item 12.1 compared to 20.8% of RECORD endorsing journals.

3.1.4. Association of journal and study characteristics with RECORD reporting

There is a weak difference in the proportion of reported items between RECORD-endorsing and non-RECORD endorsing journals (1.8% lower; 95% CI: −5.8, 2.2, Table 3). Similarly, year of publication, geographic region of the study population, primary funding source, and reported use of RECORD or other guidelines were not associated with RECORD reporting.

The association of journal and study characteristics with RECORD adherence was unchanged when excluding RECORD item 6.3 (Supplementary Tables 1 and 2).

3.1.5. Association between RECORD adherence and study quality

There was no apparent association between RECORD adherence and NOS in univariable and multivariable linear regression (Table 4). Similarly, the Pearson correlation coefficient between RECORD adherence and NOS scores was weak overall (r = 0.19, 95% CI: 0.02, 0.35; Fig 3) and for each of cohort studies (r = 0.14, 95% CI −0.01, 0.29), case control studies (r = 0.38, 95% CI: −0.16, 0.79) and cross-sectional studies (r = 0.12, 95% CI: −0.38, 0.59). The association of NOS scores with RECORD adherence was unchanged when excluding RECORD item 6.3 (Supplementary Tables 1 and 2).

4. Discussion

Among 196 studies using routinely collected health data published in high-impact general medical journals, we found that adherence to RECORD checklist was moderate, with gaps in reporting. RECORD adherence was not associated with journals' endorsement of RECORD or other factors such as geographic region, funding source or year of publication. Moreover, RECORD reporting was not associated with study quality as measured by the NOS. While studies published in RECORD-endorsing journals were more likely to mention a reporting guideline, endorsement alone was insufficient to ensure optimal adherence.

Routinely collected health data enable the longitudinal study of large, representative populations [2,29]. However, because health-data collection varies globally, transparent reporting is critical for reproducibility and comparability. Despite RECORD's availability since 2015, we found gaps

Table 1. Characteristics of included studies by RECORD-endorsing journal groups

Characteristics	RECORD-endorsing articles (n = 96)	Non–RECORD-endorsing articles (n = 96)
Journal	BMJ (n = 24) CMAJ (n = 24) Int Epi (n = 25) PLoS Med (n = 23)	JAMA (n = 40) Lancet (n = 34) Nature Medicine (n = 5) NEJM (n = 17)
Articles reporting use of guideline (%)		
Any	36 (37.5%)	19 (19.8%)
RECORD only	7 (7.3%)	1 (1.0%)
STROBE only	25 (26.0%)	7 (7.3%)
RECORD and STROBE	4 (4.2%)	1 (1.0%)
Other	0 (0%)	10 (10.4%)
None	60 (62.5%)	77 (80.2%)
COVID-19 as main exposure or outcome (%)	6 (6.3%)	17 (17.7%)
Number of authors		
≤5	20 (20.8%)	22 (22.9%)
6–10 authors	58 (60.6%)	36 (37.5%)
11–24 authors	16 (16.7%)	29 (30.2%)
>25	2 (2.1%)	9 (9.4%)
Primary Funding (%)[a]		
Government Grant	61 (63.5%)	54 (56.3%)
Not for Profit	24 (25.0%)	26 (27.1%)
Industry	3 (3.1%)	2 (2.1%)
No funding	3 (3.1%)	6 (6.3%)
Not specified	5 (5.2%)	8 (8.3%)
Total sample size (SD)	1 355 655 (2 594 868)	2 648 747 (6 201 842)
Geographic region of dataset (%)		
Europe	57 (59.4%)	49 (51.0%)
North America	27 (28.1%)	35 (36.5%)
Asia	5 (5.2%)	2 (2.1%)
Oceania	2 (2.1%)	1 (1.0%)
Africa	0 (0.0%)	1 (1.0%)
South America	2 (2.1%)	1 (1.0%)
Intercontinental	3 (3.1%)	7 (7.3%)
Mean number of reported RECORD items (n, SD)		
Study does not involve dataset linkage (score out of 11)	8.3 (n = 18, 1.5)	9.2 (n = 30, 1.4)
Study involves dataset linkage (score out of 13)	8.9 (n = 78, 1.6)	8.6 (n = 66, 1.7)
Mean number of reported Newcastle-Ottawa (n, SD)		
Cohort (score out of 9)	7.9 (n = 81,0.6)	7.9 (n = 85, 0.8)
Case-control (score out of 9)	6.8 (n = 8, 1.0)	7.0 (n = 5, 0.0)
Cross-sectional (score out of 10)	5.7 (n = 7, 0.8)	5.8 (n = 6, 0.8)

BMJ, British Medical Journal; CMAJ, Canadian Medical Association Journal; JAMA, Journal of the American Medical Association; NEJM, New England Journal of Medicine; RECORD, REporting of studies Conducted using Observational Routinely Collected Data; STROBE, Strengthening the Reporting of Observational Studies In Epidemiology.

[a] Primary funding indicates the funding is from the first listed funding source.

in adherence across high-impact journals, with only 70.7% of items fully reported. While this adherence rate compares favorably to other widely adopted guidelines—such as CONSORT [6] and PRISMA [5] where adherence often falls below 50% for key items—it still highlights room for improvement. Our findings are consistent with a

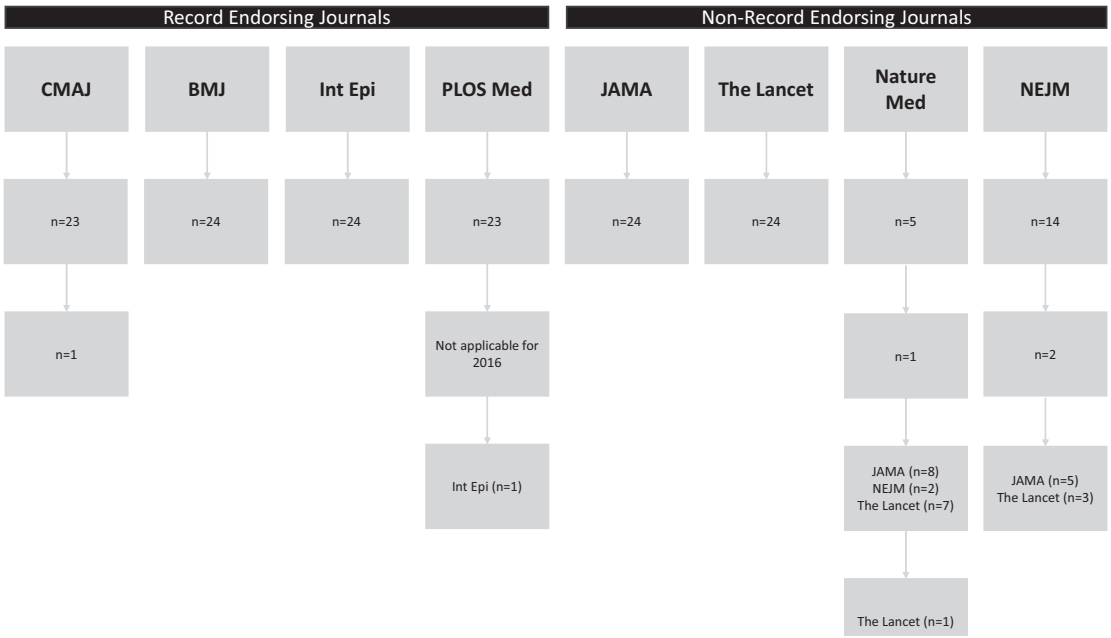

Figure 2. Article selection process. BMJ, British Medical Journal; CMAJ, Canadian Medical Association Journal; NEJM, New England Journal of Medicine.

previous systematic review which found 88% of studies claiming to use reporting guidelines such as CONSORT, PRISMA, and STROBE still exhibited suboptimal reporting [30]. This is despite reporting guidelines being associated with better study quality across different research areas [6,7,13,30]. Our study adds to the literature suggesting that implementation of reporting guidelines remains inconsistent, even in highly influential general medical journals.

The three most frequently missed RECORD items were validation of codes or algorithms used to define the study population (6.2), graphical representation of the linking process (6.3, a suggested but not mandatory item) and reporting the extent of authors' access to data (12.1). Low adherence to item 6.2 may be related to challenges validating coding algorithms in administrative data, which are well-documented in prior literature, particularly for noncardiac outcomes [31]. These challenges are compounded by changes in coding systems over time, such as the transition to International Classification of Diseases, Tenth Revision in the United States in 2015—the same year RECORD was released—which may have limited opportunities to conduct validation studies. [32]. Still, it is important for authors to report on the validation status of variables used, even it is to report that no validation data are available. In our study, RECORD items were assessed as binary, either fully met or not met to reduce subjectivity. Studies that partially validated their data (eg, confounding variables but not exposure) were scored as nonadherent, contributing to lower scores for item 6.2. Item 6.3 involves a visual representation of database linkage, which is important for assessing the validity of studies using linked datasets. This item is suggested rather than mandatory, and it remains unclear whether a textual description of linkage is sufficient, or in which situations a visual representation would be most helpful or should be required. Item 12.1, authors' access to data, was often reported when journals mandated standardized disclosures. For example, JAMA's strict requirements led to high adherence; however, the generic nature of these statements often made it difficult to determine what level of access authors had. This variability underscores the limitation of comparing adherence across journals and suggests that standardized templates do not necessarily guarantee improved transparency.

Our finding that RECORD adherence and study quality measured by NOS score are not strongly related is counterintuitive and different from the findings of metaepidemiologic studies assessing reporting of other study types [3,5,33]. However, RECORD and all reporting guidelines aim to improve reporting transparency and not study methodology, so the lack of a relationship is understandable. Further, we only included studies from high impact journals whose study quality is generally high.

Our findings have implications for several key groups. Research communities should investigate barriers to adherence, such as lack of awareness of such guidelines and space constraints. RECORD developers should clarify mandatory vs suggested items, especially frequently omitted items like 6.3 linkage diagrams, which may increase compliance in reporting. Authors should clearly state whether variables were validated, and whether linkage was performed. While passive endorsement of RECORD (eg, referencing it in author guidelines) was not associated with higher adherence, reporting may improve if journals actively align manuscript formatting and submission requirements with RECORD.

Table 2. RECORD adherence percentages per RECORD item by groups

Record items	1.1[a] N = 192	1.2[b] N = 192	1.3[c] N = 192	6.1[d] N = 192	6.2[e] N = 192	6.3[f] N = 144	7.1[g] N = 192	12.1[h] N = 192	12.2[i] N = 192	12.3[j] N = 144	13.1[k] N = 192	19.1[l] N = 192	21.1[m] N = 192	Total N = 192
Journals														
CMAJ	91.7	91.7	66.7	100.0	83.3	0.0	91.7	25.0	37.5	58.8	50.0	100.0	95.8	69.0
BMJ	95.8	95.8	50.0	100.0	58.3	20.0	83.3	41.7	79.2	70.0	79.2	100.0	95.8	74.4
Int Epi	88.0	88.0	48.0	96.0	48.0	21.7	80.0	12.0	60.0	69.6	76.0	100.0	92.0	66.8
PLOS Med	95.7	95.7	43.5	100.0	34.8	16.7	87.0	4.4	69.6	72.2	69.6	95.7	100.0	69.2
JAMA	82.5	97.5	57.5	97.5	47.5	0.0	85.0	100.0	77.5	73.9	75.0	10.0	100.0	78.8
Lancet	88.2	94.1	50.0	91.2	44.1	7.7	79.4	91.2	55.9	46.2	71.8	97.1	94.1	70.4
Nature Med	100.0	20.0	40.0	100.0	40.0	0.0	40.0	20.0	40.0	33.3	60.0	100.0	100.0	56.5
NEJM	94.1	70.6	35.3	100.0	29.4	14.3	88.2	5.9	41.2	50.0	64.7	100.0	100.0	61.7
RECORD-endorsing journals	92.7	92.7	52.1	99.0	56.3	15.4	85.4	20.8	61.5	68.0	68.8	99.0	95.8	69.8
Non-endorsing-RECORD journals	87.5	87.5	50.0	95.8	42.7	6.1	81.3	76.0	61.5	56.1	67.7	99.0	97.9	71.6

BMJ, British Medical Journal; CMAJ, Canadian Medical Association Journal; NEJM, New England Journal of Medicine; RECORD, REporting of studies Conducted using Observational Routinely Collected Data.

[a] RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included.

[b] RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract.

[c] RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.

[d] RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided.

[e] RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided.

[f] RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.

[g] RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.

[h] RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population.

[i] RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.

[j] RECORD 12.3: State whether the study included person -level, institutional -level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.

[k] RECORD 13.1: Describe in detail the selection of the persons included in the study (ie, study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.

[l] RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research questions. Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.

[m] RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.

Table 3. Association of journal and study characteristics with adherence to RECORD items in 192 articles

Characteristics	Unadjusted model: RECORD item adherence percentage[a] (95% CI)	Adjusted model: RECORD item adherence percentage[b] (95% CI)
RECORD-endorsing vs nonendorsing journals	−1.8 (−5.8, 2.2)	−1.7 (−5.8, 2.5)
Year of publication (continuous, starting from 2016)	0.4 (−0.4, 1.3)	0.6 (−0.3, 1.6)
Geographic region of dataset (%)		
Europe (n = 106)	1.0 (ref)	1.0 (ref)
North America (n = 62)	1.1 (−3.4, 5.5)	0.5 (−4.0, 5.1)
Asia (n = 7)	2.6 (−8.2, 13.5)	3.1 (−8.0, 14.1)
Oceania (n = 3)	−6.3 (−22.6, 10.0)	−5.6 (−21.8, 10.6)
South America (n = 3)	−1.2 (−17.3, 14.9)	−1.3 (−17.6, 15.1)
Africa (n = 1)	−32.0 (−59.6, −4.3)	−34.5 (−62.5, −6.5)
Intercontinental (n = 10)	3.0 (−6.1, 12.1)	3.1 (−6.2, 12.3)
Primary Funding source		
Government Grant/Not for Profit/No funding (n = 174)	1.0 (ref)	1.0 (ref)
Industry (n = 5)	1.1 (−11.5, 13.7)	0.3 (−12.4, 13.0)
Not specified (n = 13)	0.6 (−7.4, 8.6)	−0.1 (−8.4, 8.2)
Article reported use of RECORD or other guidelines (Yes vs no)	−1.2 (−5.6, 3.2)	−2.0 (−6.8, 2.8)

RECORD, REporting of studies Conducted using Observational Routinely Collected Data.
[a] Listed variables adjusted separately for outcome of adherence to RECORD items.
[b] Listed variables adjusted together for outcome of adherence to RECORD items.

For instance, journals could require submission of a completed RECORD checklist as supplementary material rather than treating it as optional [34]. The RECORD checklist itself could also be modified to include example references, as found in the explanatory paper [24], or by developing an online tool with clickable examples for each

Table 4. Association between Newcastle-Ottawa scale and RECORD adherence in 192 articles

Characteristics	RECORD item adherence percentage increase per point increase of Newcastle-Ottawa scale for cohort studies (95% CI) n = 166	RECORD item adherence percentage increase per point increase of Newcastle-Ottawa scale for case-control studies (95% CI) n = 13	RECORD item adherence percentage increase per point increase of Newcastle-Ottawa scale for cross-sectional studies (95% CI) n = 13
Unadjusted	2.6 (−0.3, 5.5)	5.6 (−3.5, 14.8)	2.8 (−12.6, 18.3)
Adjusted for article reported use of RECORD or other guidelines, Journal endorsement of RECORD, year of publication	2.6 (−0.3, 5.5)	1.2 (−12.6, 15.0)	2.9 (−14.0, 19.9)
Adjusted for article reported use of RECORD or other guidelines, Journal endorsement of RECORD, year of publication, funding source	2.6 (−0.3, 5.5)	N/A[a]	N/A[a]
Adjusted for article reported use of RECORD or other guidelines, Journal endorsement of RECORD, year of publication, funding source, geographic region of dataset	2.6 (−0.4, 5.7)	N/A[a]	N/A[a]

RECORD, REporting of studies Conducted using Observational Routinely Collected Data.
[a] Sample size too low to report valid adjusted difference.

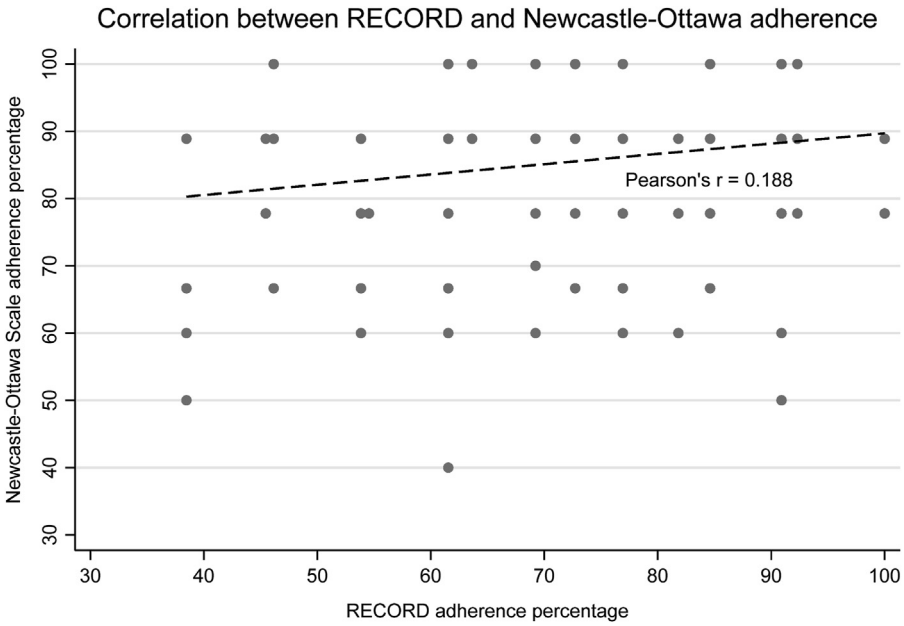

Figure 3. Correlation between RECORD adherence percentage and Newcastle-Ottawa Scale adherence percentage. RECORD, REporting of studies Conducted using Observational Routinely Collected Data.

item. These targeted actions many help improve the reliability and reproducibility of studies using routinely collected data.

Our study is limited by focusing on high-impact general medical journals, which may not be generalizable to specialty or lower-impact journals. Second, while we used independent duplicate extraction in RECORD and NOS assessments, some subjectivity remains in evaluating checklist items. Third, we used the NOS to assess study quality, despite concerns about its validity [25,26]. Although ROBINS-I is the recommended tool for evaluating risk of bias in nonrandomized studies, it does not provide a summary score suitable for regression modeling [35]. Fourth, there was variation in what ''endorsement'' meant across journals, from optional checklist to mandatory submission. Additionally, we used the RECORD elaboration publication to guide our scoring of RECORD items [24], but we did not create a fully operationalized scoring framework before conducting our assessments. As such, some of the assessments are subjective, reducing transparency and reproducibility. Sixth, we could not determine when each journal began endorsing RECORD, which may have affected our results. Finally, our categorization of journals into endorsing and nonendorsing journals could introduce confounding because nonendorsing journals in our sample had, on average, higher impact factors. For instance, *JAMA, The Lancet, Nature Medicine,* and *NEJM* had journal impact factors of 63.5, 98.4, 58.7, 96.3 respectively whereas RECORD endorsing journals like *BMJ, CMAJ, PLoS Medicine* and *the International Journal of Epidemiology* had impact factors of 93.7, 12.9, 10.5, 6.4, respectively [36]. Non−RECORD-endorsing journal's higher average impact factor could reflect differences in editorial processes and verification of reporting items compared to RECORD endorsing journals.

5. Conclusion

We found moderate reporting of items in RECORD guidelines among studies using routinely collected health data. While we did not find an association between reporting and study quality, missing items reduce the reproducibility and trustworthiness of the medical literature. Future measures are needed to improve the implementation of RECORD, including refining the RECORD checklist, supporting authors through clearer expectations and examples, and aligning journal policies with reporting standards.

CRediT authorship contribution statement

Heather J. Zhao: Writing − review & editing, Writing − original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Inna Ushcatz:** Writing − review & editing, Writing − original draft, Validation, Methodology, Investigation, Data curation, Conceptualization. **Chantel Walwyn:** Writing − review & editing, Validation, Project administration, Methodology, Data curation, Conceptualization. **Megan S. Lowe:** Writing − review & editing, Validation, Data curation. **Kevin S. Kim:** Writing − review & editing, Validation, Data curation. **Eric I. Benchimol:** Writing − review & editing, Validation, Supervision, Resources, Methodology, Conceptualization. **Sinéad M.**

Langan: Writing − review & editing, Validation, Supervision, Resources, Methodology, Conceptualization. **Aaron M. Drucker:** Writing − review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Conceptualization.

Declaration of competing interest

E.B. holds the Northbridge Financial Corporation Chair in Inflammatory Bowel Disease, a joint Hospital-University Chair between the University of Toronto, The Hospital for Sick Children, and the SickKids Foundation, is the co-Chair of the RECORD Steering Committee, the Editor-in-Chief of a RECORD-endorsing journal (Journal of the Canadian Association of Gastroenterology), and has acted as a consultant for McKesson Canada, the Dairy Farmers of Ontario, and the Canadian Drug Agency for matters unrelated to the topic of this study. S.M.L is co-Chair of the RECORD Steering Committee. There are no competing interests for any other author.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jclinepi.2025.111876.

Data availability

Data will be made available on request.

References

[1] Nicholls SG, Langan SM, Benchimol EI. Routinely collected data: the importance of high-quality diagnostic coding to research. CMAJ 2017;189(33):E1054−5.

[2] Hemkens LG, Contopoulos-Ioannidis DG, Ioannidis JPA. Routinely collected data and comparative effectiveness evidence: promises and limitations. C Can Med Assoc J 2016;188(8):E158.

[3] Harron K, Benchimol E, Langan S. Using the RECORD guidelines to improve transparent reporting of studies based on routinely collected data. Int J Popul Data Sci 2018;3(1):2.

[4] McGrail KM, Gutteridge K, Meagher NL. In: Gkoulalas-Divanis A, Loukides G, editors. Medical Data Privacy Handbook1 Chapter. Switzerland: Springer International Publishing; 2015:737−64.

[5] Pussegoda K, Turner L, Garritty C, Mayhew A, Skidmore B, Stevens A, et al. Systematic review adherence to methodological or reporting quality. Syst Rev 2017;6(1):1−14.

[6] Audet S, Doyle C, Lemieux C, Tardif MA, Gauvreau A, Simonyan D, et al. Adherence to CONSORT guidelines and reporting of the determinants of external validity in clinical oncology randomized controlled trials: a review of trials published in four major journals between 2013 and 2015. Curr Oncol 2023;30(2):2061−72.

[7] Kshirsagar K, Bhate K, Shete A, Kakodkar P. Adherence to the consolidated standards of reporting trials (CONSORT) guidelines for reporting randomized controlled trials related to mandibular third molars. J Oral Maxillofac Surg 2021;79(6):1207−13.

[8] Yin Y, Shi F, Zhang Y, Zhang X, Ye J, Zhang J. Evaluation of reporting quality of randomized controlled trials in patients with COVID-19 using the CONSORT statement. PLoS One 2021;16(9):e0257093.

[9] Samaan Z, Mbuagbaw L, Kosa D, Borg Debono V, Dillenburg R, Zhang S, et al. A systematic scoping review of adherence to reporting guidelines in health care literature. J Multidiscip Healthc 2013;6:169.

[10] Ottawa hospital research institute [Internet]. Available at: https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed April 10, 2024.

[11] Yolcu Y, Wahood W, Alvi MA, Kerezoudis P, Habermann EB, Bydon M. Reporting methodology of neurosurgical studies utilizing the American College of surgeons-national surgical quality improvement program database: a systematic review and critical appraisal. Neurosurg 2020;86(1):46−59.

[12] Zhao R, Zhang W, Zhang ZD, He C, Xu R, Tang XD, et al. Evaluation of reporting quality of cohort studies using real-world data based on RECORD: systematic review. BMC Med Res Methodol 2023;23(1):152.

[13] Wiehn E, Ricci C, Alvarez-Perea A, Perkin MR, Jones CJ, Akdis C, et al. Adherence to the strengthening the reporting of observational studies in Epidemiology (STROBE) checklist in articles published in eaaci journals: a bibliographic study. Allergy 2021;76(12):3581−8.

[14] Reporting guidelines - BMJ author hub [Internet]. Available at: https://authors.bmj.com/before-you-submit/reporting-guidelines/. Accessed January 2, 2025.

[15] Reporting guidelines | EQUATOR network [Internet]. Available at: https://www.equator-network.org/?post_type=eq_guidelines&eq_guidelines_study_design=0&eq_guidelines_clinical_specialty=0&eq_guidelines_report_section=0&s=+STROBE+extension&btn_submit=Search+Reporting+Guidelines. Accessed January 2, 2025.

[16] Submission guidelines | PLOS Med [Internet]. Available at: https://journals.plos.org/plosmedicine/s/submission-guidelines#loc-guidelines-for-specific-study-types. Accessed January 2, 2025.

[17] Submission guidelines | CMAJ [Internet]. Available at: https://www.cmaj.ca/submission-guidelines. Accessed January 2, 2025.

[18] General Instructions | International Journal of Epidemiology | Oxford Academic [Internet]. Available at: https://academic.oup.com/ije/pages/general_instructions. Accessed January 2, 2025.

[19] Langan SM, Schmidt SA, Wing K, Ehrenstein V, Nicholls SG, Filion KB, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). BMJ 2018;363:3532.

[20] MEDLINE/PubMed search strategy & electronic health record information resources [internet]. Available at: https://www.nlm.nih.gov/services/queries/ehr_details.html. Accessed January 2, 2025.

[21] Systematic reviews filter strategy [Internet]. Available at: https://www.nlm.nih.gov/bsd/pubmed_subsets/sysreviews_strategy.html. Accessed January 2, 2025.

[22] Observational designs - systematic reviews and meta analysis - research guides at harvard library [Internet]. Available at: https://guides.library.harvard.edu/c.php?g=309982&p=2079546. Accessed January 2, 2025.

[23] Innocenti T, Feller D, Giagio S, Salvioli S, Minnucci S, Brindisino F, et al. Adherence to the PRISMA statement and its association with risk of bias in systematic reviews published in rehabilitation journals: a meta-research study. Braz J Phys Ther 2022;26(5):100450.

[24] Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Peteresen I, et al. The REporting of studies conducted using observational routinely-collected health data (RECORD) statement. Plos Med 2015;12(10):e1001885.

[25] Stang A, Jonas S, Poole C. Case study in major quotation errors: a critical commentary on the Newcastle−Ottawa scale. Eur J Epidemiol 2018;33(11):1025−31.

[26] Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. Eur J Epidemiol 2010;25(9):603−5.

[27] Forte AJ, Guliyeva G, McLeod H, Dabrh AMA, Salinas M, Avila FR, et al. The impact of optimism on cancer-related and postsurgical cancer pain: a systematic review. J Pain Symptom Manage 2022;63(2):e203−11.
[28] de Oliveira Almeida K, Nogueira Alves IG, de Queiroz RS, de Castro MR, Gomes VA, Santos Fontoura FC, et al. A systematic review on physical function, activities of daily living and health-related quality of life in COVID-19 survivors. Chronic Illn 2022;19(2):279.
[29] Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. Annu Rev Public Health 2011;32:91−110.
[30] Jin Y, Sanger N, Shams I, Luo C, Shahid H, Li G, et al. Does the medical literature remain inadequately described despite having reporting guidelines for 21 years? - a systematic review of reviews: an update. J Multidiscip Healthc 2018;11:495−510.
[31] Carnahan RM. Mini-Sentinel's systematic reviews of validated methods for identifying health outcomes using administrative data: summary of findings and suggestions for future research. Pharmacoepidemiol Drug Saf 2012;21(SUPPL. 1):90−9.
[32] Hirsch JA, Nicola G, McGinty G, Liu RW, Barr RM, Chittle MD, et al. ICD-10: history and context. AJNR Am J Neuroradiol 2016;37(4):596.
[33] Adams AD, Benner RS, Riggs TW, Chescheir NC. Use of the STROBE checklist to evaluate the reporting quality of observational research in obstetrics. Obstet Gynecol 2018;132(2):507−12.
[34] Cobo E, Cortés J, Ribera JM, Cardellach F, Selva-O'Callaghan A, Kostov B, et al. Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: masked randomised trial. BMJ 2011;343(7833):d6783.
[35] Chapter 25. Assessing risk of bias in a non-randomized study | Cochrane Training [Internet]. Available at: https://training.cochrane.org/handbook/current/chapter-25. Accessed June 10, 2025.
[36] Journal citation reports - home [Internet]. Available at: https://jcr-clarivate-com.myaccess.library.utoronto.ca/jcr/home?app=jcr&Init=Yes&authCode=null&SrcApp=IC2LS. Accessed January 27, 2025.