

Review

# Pharmacoepidemiologic Research Based on Common Data Models: Systematic Review and Bibliometric Analysis

Yongqi Zheng<sup>1</sup>; Meng Zhang<sup>1</sup>; Conghui Wang<sup>1</sup>; Ling Gao<sup>1</sup>; Junqing Xie<sup>2</sup>; Peng Shen<sup>3</sup>; Yexiang Sun<sup>3</sup>; Mengling Feng<sup>4</sup>; Seng Chan You<sup>5</sup>; Feng Sun<sup>1,6,7,8</sup>

<sup>1</sup>Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China

<sup>2</sup>NDORMS, The Botnar Research Centre, University of Oxford, Oxford, United Kingdom

<sup>3</sup>Yinzhou District Center for Disease Control and Prevention, Ningbo, China

<sup>4</sup>Saw Swee Hock School of Public Health and the Institute of Data Science, National University of Singapore, Singapore, Singapore

<sup>5</sup>Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Republic of Korea

<sup>6</sup>Institute of Traditional Chinese Medicine, Xinjiang Medical University, Urumqi, China

<sup>7</sup>Department of Ophthalmology, Peking University Third Hospital, Beijing, China

<sup>8</sup>School of Medicine, Shihezi University, Shihezi, China

## Corresponding Author:

Feng Sun

Department of Epidemiology and Biostatistics

School of Public Health, Peking University

38 Xueyuan Road, Haidian District

Beijing 100191

China

Phone: 86 13811155891

Email: [sunfeng@bjmu.edu.cn](mailto:sunfeng@bjmu.edu.cn)

## Abstract

**Background:** The adoption of common data models (CDMs) has transformed pharmacoepidemiologic research by enabling standardized data formatting and shared analytical tools across institutions. These models facilitate large-scale, multicenter studies and support timely real-world evidence generation. However, no comprehensive global evaluation of CDM applications in pharmacoepidemiology has been conducted.

**Objective:** This study aimed to conduct a systematic review and bibliometric analysis to map the landscape of CDM usage in pharmacoepidemiology, including publication trends, institutional authors and collaborations, and citation impacts.

**Methods:** In total, 5 English databases (PubMed, Web of Science, Embase, Scopus, and Virtual Health Library) and 4 Chinese databases (CNKI, Wan-Fang Data, VIP, and SinoMed) were searched for studies applying CDMs in pharmacoepidemiology from database inception to January 2024. Two reviewers independently screened studies and extracted information about basic publication details, methodological details, and exposure and outcome information. The studies were categorized into 2 groups according to their Total Citations per Year (TCpY), and a comparative analysis was conducted to examine the differences in characteristics between the 2 groups.

**Results:** A total of 308 studies published between 1997 and 2024 were included, involving 1580 authors across 32 countries and 140 journals. The United States led in both publication volume and citation counts, followed by South Korea. Among the 10 most cited studies, 7 used the Vaccine Safety Datalink, 2 used Sentinel, and one used Observational Medical Outcomes Partnership. Studies were stratified by TCpY to reduce citation bias from publication timing. Comparative analysis showed that high-TCpY studies were significantly more associated with multicenter collaboration ( $P=.008$ ), United States-based institutions ( $P=.04$ ), and vaccine-related research ( $P=.009$ ). These studies commonly featured larger sample sizes, cross-regional data, and enhanced generalizability. International collaborations primarily occurred among North America, Europe, and East Asia, with limited involvement from limited-income countries.

**Conclusions:** This study presents the first bibliometric overview of CDM-based pharmacoepidemiologic research. The consistent output from United States institutions and increasing engagement from South Korea underscore their central roles in this field. High-TCpY studies tend to be multicenter, collaborative, and vaccine-focused, reflecting structural factors linked to research visibility and influence. Stratified citation analysis supports the value of real-world data integration and international

cooperation in producing impactful studies. The dominance of limited-income countries in collaboration networks highlights a need for broader inclusion of underrepresented regions. These findings can help researchers identify key contributors, guide partner selection, and target appropriate journals. As CDM-based methods continue to expand, fostering diverse and collaborative research efforts will be crucial for advancing pharmacoepidemiologic knowledge globally.

*JMIR Med Inform* 2025;13:e72225; doi: [10.2196/72225](https://doi.org/10.2196/72225)

**Keywords:** common data model; pharmacoepidemiology; bibliometric analysis; systematic review

Introduction

Pharmacoepidemiology applies epidemiological principles and methods to study the effects of drugs, vaccines, and devices, focusing on adverse event monitoring, health economic evaluation, and quality of life assessment [1]. Ongoing pharmacoepidemiologic research aims to ensure drug safety, standardize postmarket surveillance methods, and provide scientific evidence for intervention decisions [1]. To achieve these objectives, especially in postmarketing contexts, robust real-world evidence is essential. However, many current studies are limited by small sample sizes and insufficient population representativeness, making it difficult to detect rare or long-term adverse events [2]. To address these limitations, conducting multicenter pharmacoepidemiologic evaluations has proven to be more reflective of actual clinical settings, enhancing statistical power, improving detection efficiency, and enabling early identification of vaccine safety concerns [3,4]. However, multicenter studies also face challenges related to data heterogeneity, personally identifiable information leakage, lack of standardized protocols for data integration, and terminological harmonization difficulties [5,6].

Common data model (CDM) is a standardized data model designed to facilitate the exchange, integration, sharing, or storage of data from multiple sources [7]. Its development has provided a practical and scalable solution for enabling multicenter studies by harmonizing heterogeneous data sources across institutions and regions. Currently, various CDMs are widely used, including the Observational Medical Outcomes Partnership (OMOP) developed by the Observational Health Data Sciences and Informatics [8], PCORnet [9], and Fast Healthcare Interoperability Resources [10]. Other CDMs, such as the Vaccine Safety Datalink (VSD) [11] and Sentinel [12], are specifically designed for active vaccine safety monitoring. Each CDM emphasizes distinct aspects, with differing methodologies and application areas tailored to specific research needs [13].

Currently, CDMs are widely applied to generate new evidence in clinical practice and drug selection. Some

studies leverage real-world data to produce scalable evidence, aiding in the understanding of population diversity as well as the similarities and differences in clinical characteristics and treatment pathways across regions [14,15]. In addition, certain studies based on the OMOP have introduced a large-scale, comprehensive approach to evaluate the effectiveness and safety of various drugs across the world, offering strong evidence to support clinical drug selection [16,17]. Despite the growing application of CDMs in pharmacoepidemiology, no study has systematically reviewed the global landscape of CDM use in this field across all major models and contexts. Previous reviews have often focused on a single CDM or specific application scenarios [6].

In recent years, bibliometrics has advanced rapidly and is widely used to explore the characteristics of academic publications in specific research fields, including influential countries, journals, institutions, and authors, as well as trends in frequently cited references and keywords [18,19]. Since Garfield’s seminal work in 1955 on identifying the most cited scientific papers in the Institute for Scientific Information Web of Knowledge (now known as Web of Science) database [20,21], numerous scholars have conducted bibliometric analysis comparing highly cited and less cited papers [22–24]. Building on previous studies that highlighted key features of highly cited papers, this study uses bibliometric analysis to systematically characterize publication patterns, collaboration networks, and thematic trends in CDM-based pharmacoepidemiologic research.

This study aimed to fill that gap by systematically reviewing and visually analyzing global literature on CDM-based pharmacoepidemiology. By grouping studies by citation impact, we sought to identify the characteristics of highly cited studies and offer references for future research design, collaboration strategy, and infrastructure development. Our findings aim to inform researchers and policy makers on emerging trends, methodological priorities, and opportunities to enhance the global reach and inclusiveness of CDM-enabled pharmacoepidemiologic studies. The statement of significance is provided in Table 1.

**Table 1.** Statement of significance.

Content	Statement
Problem or issue	There is a critical challenge in integrating heterogeneous, multicenter, real-world data for pharmacoepidemiologic research, which hampers the detection of rare and long-term adverse events and limits the standardization of postmarket surveillance.

Content	Statement
What is already known	Previous studies have demonstrated that CDMs <sup>a</sup> facilitate data standardization and enhance multicenter collaborations in pharmacoepidemiology. However, the global adoption, comparative impact, and evolution of these models remain insufficiently explored.
What this paper adds	This paper provides a comprehensive bibliometric analysis of the application of CDMs in pharmacoepidemiology, revealing publication trends, key contributors, and emerging research themes across different countries. It also identifies significant differences between high and low citation groups, underscoring the pivotal role of multi-center research and CDM-based methodologies.
Who would benefit from the new knowledge in this paper?	Researchers, policymakers, and healthcare practitioners involved in pharmacoepidemiologic studies and health informatics will benefit, as the findings offer insights into effective collaboration, selection of appropriate research models, and strategic dissemination for enhancing drug and vaccine safety monitoring.

<sup>a</sup>CDM: common data models.

Methods

Overview

This study was conducted according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta Analyses) guidelines (the PRISMA checklist is provided in [Checklist 1](#)) to identify studies that applied CDM in pharmacoepidemiologic research [25].

Search Strategy

Five English databases: PubMed, Web of Science, Embase, Scopus, and Virtual Health Library, and 4 Chinese databases: China National Knowledge Infrastructure, Wan-Fang Data, VIP Database, and SinoMed were searched from the inception to January 22, 2024. In preliminary searches, the names of currently used CDMs worldwide were identified and used as search terms in both their full and abbreviated forms, in both English and Chinese, including terms like “common data model,” “Observational Health Data Sciences and Informatics,” “Observational Medical Outcomes Partnership,” and “Clinical Data Interchange Standards Consortium,” among others. Logical operators such as “OR” truncation symbols, and subject term matching were applied based on the search rules and syntax of each platform to form search strategies. All search terms were listed in [Multimedia Appendix 1](#).

Eligibility Criteria

The inclusion criteria were as follows: (1) studies using CDMs (any CDM identified in the preliminary search, with inclusion based on the use of any single model) to address issues in the drug, vaccine, or medical device fields; (2) the study scope included safety, efficacy, usage, and economic evaluations for drugs, vaccines, and medical devices; and (3) drugs, vaccines, or devices must be the primary exposure, research focus, or outcomes of the study.

The exclusion criteria were as follows: (1) studies that did not apply CDMs, or where the CDM was incomplete; (2) studies that did not include drugs, vaccines, or devices as primary exposure or influence factors, or as main research content or outcomes; (3) editorial materials, including letters, editorials, comments, responses, editorial opinions, advertise-

ments, and unpublished studies; (4) duplicate publications; and (5) studies in languages other than Chinese or English.

Data Extraction

Two reviewers independently screened studies and extracted information about basic publication details, methodological details, and exposure and outcome information. Basic information from the literature was extracted using Excel, including study title, authors, publication year, country, sample size, number of centers, type of CDMs, use of subgroup analysis, whether sensitivity or subgroup analysis was conducted, study exposure, study outcomes, and whether adherence to reporting guidelines. English studies were exported in BibTeX format and merged into a single XLSX file. Some English studies that could not be imported directly into the Bibliometrix package and all of the Chinese studies were exported in XLSX format. These files were modified into compatible tables based on the Bibliometrix data frame structure for analysis.

Statistical Analysis

Statistical analysis was conducted using R (version 4.4.0; R Core Team). The bibliometrix package was used to generate standard bibliometric indicators, including annual publication trends, total and average citation counts, and H-index values, as well as to construct coauthorship and keyword co-occurrence networks. To ensure consistency and data quality, all bibliographic records were cleaned and standardized before analysis. This process included deduplication, harmonization of author names and affiliations, normalization of journal titles, and manual review of keywords [26]. Total citation, Total Citations per Year (TCpY), and the H-index were used as standard metrics for academic influence of studies and authors. The total citation of all studies published in a specific journal was used to represent the journal’s overall citation count [27]. TCpY was defined as the average number of citations a study has received per year since its publication, calculated as follows:

$$TCpY = \frac{Total\ Citations}{Years\ since\ Publication}$$

The H-Index is defined as *h* if an author has *h* publications, each of which has been cited at least *h* times [28].

Citation analysis was conducted on 285 studies which could get citations from the Web of Science Core Collection (WoSCC). Analysis of these 285 studies was conducted using R version 4.4.0. The *t* test was used to compare means under the assumption of normal distribution, while the Wilcoxon test served data that did not meet normality assumptions. Chi-square or Fisher exact tests were used for categorical variables. A 2-tailed *P* value of <.05 was considered to indicate statistical significance. To visualize interdisciplinary citation relationships, we conducted a dual-map overlay analysis using CiteSpace (version 6.4.1; Chaomei Chen, Drexel University). This method enables the mapping of citation trajectories between disciplines, highlighting the knowledge flow from citing to cited journals. The map was constructed using WoSCC data, with parameters set to the default configuration for dual-map overlays.

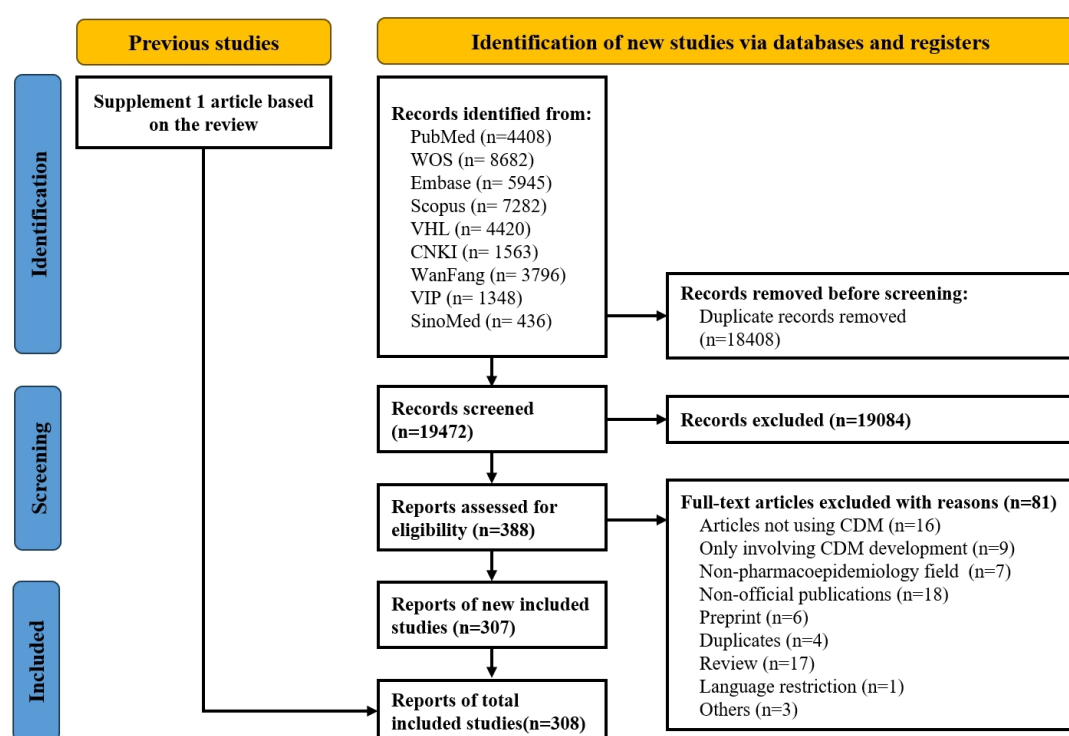
in Chinese (Figure 1). The publications originated from 32 countries, involving 1580 authors and published across 140 journals. In total, 590 keywords were identified, covering 12 types of CDMs: OMOP (76 studies), VSD (163 studies), PCORnet (14 studies), PEDSnet (4 studies), Sentinel (19 studies), Mini-Sentinel (14 studies), ConcePTION (3 studies), K-CDM (3 studies), Asian Pharmacoepidemiology Network (1 study), Cancer Research Network's Virtual Data Warehouse (1 study), Intensive Care Unit Medications (1 study), Health Maintenance Organization Research Network (1 study), along with 1 study that modified an existing CDM and 11 studies that created new CDMs. The research topics were categorized as follows: vaccines (173/308, 56.17%), drugs (132/308, 42.85%), and medical devices (3/308, 0.97%).

## Results

### Study Selection

A total of 37,880 studies were identified, of which 308 met the inclusion criteria, comprising 307 in English and 1

**Figure 1.** PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of the included studies.

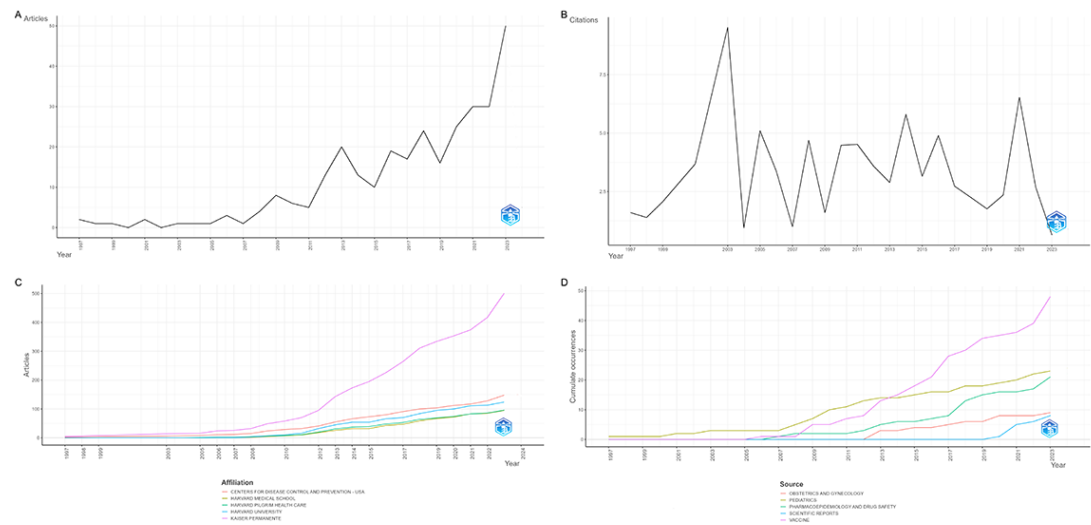


### Temporal Distribution of Publications

The volume of publications in this field has shown a continuous upward trend. The earliest identified study dates back to 1997. From 1997 to 2007, the volume of studies remained low with minimal growth, maintaining fewer than

5 studies per year. Between 2008 and 2016, the number of studies began to increase, albeit with some fluctuation. Following 2019, study volume exhibited steady growth, reaching a peak of 49 studies in 2023 (Figure 2A).

**Figure 2.** General trends of included studies in CDM-based pharmacoepidemiologic research through 2024. **(A)** Temporal distribution of publications in pharmacoepidemiologic research based on CDM through 2024. **(B)** Average annual citation frequency in CDM-based pharmacoepidemiologic research through 2024. **(C)** Temporal distribution of publications for the top 5 most productive institutions in CDM-based pharmacoepidemiologic research through 2024. **(D)** Annual publication trends across top 5 journals in CDM-based pharmacoepidemiologic research through 2024. CDM: Common Data Model.



Citation Information of the Studies

Among the 308 studies, the total citation count was 8666, with an average citation frequency of 28 per study and a median of 13.00 (IQR 3.00-30.00) citations. The citation distribution by publication year was shown in Figure 2B, with peaks in 2003 and 2021.

The top 10 most cited studies originated entirely from the United States, comprising 7 studies on VSD, 2 on Mini-Sentinel, and one using OMOP. Regarding the focus of exposure, 8 studies investigated vaccines, while 2 examined drugs. Safety outcomes were predominant, accounting for 90% (9/10) of the research endpoints. Multimedia Appendix 2 provides detailed information on the top 10 most cited studies.

The citation analysis was conducted on 285 studies according to citations from the WoSCC. We compared key characteristics between the 285 WoSCC-indexed studies and the 23 non-WoSCC studies. As shown in Multimedia Appendix 3, there were no statistically significant differences

between the 2 groups. The included studies were divided into 2 groups based on total citation counts, and their basic characteristics were compared (Table 2). The median citation count for the low citation group was 4.00 (IQR 1.00-8.00), while for the high-citation group it was 35.00 (IQR 20.00-64.00). Compared to the low-citation group, studies in the high-citation group generally had larger sample sizes ( $P=.049$ ), involved more research centers ( $P<.001$ ), and predominantly originated from the United States ( $P<.001$ ). Regarding exposure, studies in the high-citation group were more likely to focus on vaccines than those in the low-citation group (72.34%, 102/141 vs 40.97%, 59/144;  $P<.001$ ). A comparison between studies with vaccine as the exposure and those with drug indicated that high-citation studies were more frequently focused on vaccines ( $P<.001$ ). Studies with high citation counts also tended to use the VSD. Differences were observed in research direction, with a higher proportion of safety studies in the high-citation group compared to the low-citation group (75.18%, 106/141 vs 64.58%, 93/144).

**Table 2.** Comparison of characteristics between high and low total citation study groups in CDM<sup>a</sup>-based pharmacoepidemiologic research through 2024.

Variable	Low total citations (N=144)	High total citations (N=141)	P value
Total citations, median (IQR)	4 (1-8)	35 (20-64)	<.001
Sample size, median (IQR)	153,438 (10,674-1,600,000)	381,807 (92,357-1,505,381)	.049
Number of centers, median (IQR)	6.00 (3.00-8.00)	7.00 (6.00-8.00)	<.001
Data from the United States, n (%)	88 (61.11)	123 (87.23)	<.001
Research exposure, n (%)			<.001
Drug	83 (57.64)	38 (26.95)	
Vaccine	59 (40.97)	102 (72.34)	
Medical device	2 (1.39)	1 (0.71)	
Types of CDMs n (%)			<.001
OMOP <sup>b</sup>	49 (34.03)	17 (12.06)	



Variable	Low total citations (N=144)	High total citations (N=141)	P value
VSD <sup>c</sup>	54 (37.50)	99 (70.21)	
Sentinel or mini-Sentinel	15 (10.42)	14 (9.93)	
Others	26 (18.06)	11 (7.8)	
Research directions, n (%)			.02
Safety	93 (64.58)	106 (75.18)	
Effectiveness	12 (8.33)	2 (1.42)	
Usage	28 (19.44)	19 (13.48)	
Others	11 (7.64)	14 (9.93%)	
Including sensitive analysis, n (%)	46 (32.86)	40 (28.57)	.52
Including subgroup analysis, n (%)	75 (52.45)	54 (38.30)	.02
Including reporting guideline, n (%)	12 (8.33)	4 (2.84)	.07

<sup>a</sup>CDM: Common Data Model.

<sup>b</sup>OMOP: Observational Medical Outcomes Partnership.

<sup>c</sup>VSD: Vaccine Safety Datalink.

We further divided the included studies into 2 groups based on TCpY to compare their characteristics (Table 3). The median annual citation for the low TCpY group was 1.00 (IQR 0.50-1.50), while for the high TCpY group it was 4.00 (IQR 2.75-6.28). Compared to the low TCpY group, studies in the high TCpY group generally involved a greater number of research centers ( $P=.008$ ) and were predominantly conducted in the United States ( $P=.045$ ). Regarding exposure, studies in the high TCpY group were more likely to focus on

vaccines than those in the low TCpY group (63.70%, 93/146 vs 48.92%, 68/139;  $P=.003$ ). A comparison between studies with vaccine as the exposure and those with drug indicated that high TCpY studies were more frequently focused on vaccines ( $P=.009$ ). The proportion of the high TCpY group using reporting guidelines was higher, although the difference was not statistically significant (6.16%, 9/146 vs 5.04%, 7/139;  $P=.88$ ).

**Table 3.** Comparison of characteristics between high and low total citation per year study groups in CDM<sup>a</sup>-based pharmacoepidemiologic research through 2024.

Variable	Low total citation per year (N=139)	High total citation per year (N=146)	P value
Total citations per year, median (IQR)	1.00 (0.50-1.50)	4.00 (2.75-6.28)	<.001
Sample size, median (IQR)	168,046 (13,484-1,368,976)	376,677 (91,692-1,721,186)	.057
Number of centers, median (IQR)	6.00 (3.00-8.00)	7.00 (5.00-8.00)	.008
Data from the United States, n (%)	95 (68.35)	116 (79.45)	.045
Research exposure, n (%)			.003
Drug	71 (51.08)	50 (34.25)	
Vaccine	68 (48.92%)	93 (63.70)	
Medical device	0 (0.00)	3 (2.05)	
Types of CDM n (%)			.24
OMOP <sup>b</sup>	36 (25.90)	30 (20.55)	
VSD <sup>c</sup>	66 (47.48)	87 (59.59)	
Sentinel or Mini-Sentinel	16 (11.51)	13 (8.90)	
Others	21 (15.11)	16 (10.96)	
Research directions, n (%)			.01
Safety	89 (64.03)	110 (75.34)	
Effectiveness	9 (6.47)	5 (3.42)	
Usage	32 (23.02)	15 (10.27)	
Others	9 (6.47)	16 (10.96)	
Including sensitive analysis, n (%)	36 (26.67)	50 (34.48)	.20
Including subgroup analysis n (%)	72 (52.17)	57 (39.04)	.04
Including reporting guideline n (%)	7 (5.04)	9 (6.16)	.88

<sup>a</sup>CDM: Common Data Model.

<sup>b</sup>OMOP: Observational Medical Outcomes Partnership.

<sup>c</sup>VSD: Vaccine Safety Datalink.

Analysis of Authors and Institutions

Among the top 10 productive authors, Dr Nicola P Klein, director of the Vaccine Study Center at Kaiser Permanente Northern California, held the highest rank with 69 publications. Her work has an H-index of 29, and all publications were based on VSD. The second-ranked author, Dr Matthew F Daley, a senior clinical researcher at Kaiser Permanente Colorado, has 53 publications. The third position is held by Dr Allison L Naleway from Kaiser Permanente Northwest Center for Health Research, with a total of 48 publications. All of the top 10 authors are from the United States, with five affiliated with different Kaiser Permanente institutions. The publications by these 5 authors are all based on VSD (Multimedia Appendix 4).

The top 10 most productive institutions are listed in Multimedia Appendix 5, and the temporal trends of the top 5 are visualized in Figure 2C. Among these, eight institutions are based in the United States, and two are located in South Korea. Kaiser Permanente is the institution with the highest publication volume and the most rapid growth, with all

publications based on VSD. South Korea’s Hallym University and Seoul National University began publishing around 2020, primarily utilizing OMOP and K-CDM in their studies.

Analysis of Journals

The top 10 productive journals were listed in Table 4. Vaccine ranks first, with 49 publications, accounting for 16.0% (49/308) of the total publications in this field. Among these, studies were based on the following CDM: VSD (46/49, 93.90%), ConcePTION (2/49, 4.10%), and Sentinel (1/49, 2.00%). Pediatrics ranks second, with 24 publications, representing 7.79% (24/308) of the total. The CDMs used in these studies include VSD (21/24, 87.50%), PEDSnet (1/24, 4.17%), PEDSnet + OMOP (1/24, 4.17%), and PCORnet (1/24, 4.17%). *Pharmacoepidemiology and Drug Safety* published 22 studies, accounting for 7.20% (22/308) of the total publications. The CDM used in these studies included VSD (9/22, 40.91%), Sentinel (5/22, 22.73%), Mini-Sentinel (2/22, 9.09%), OMOP (2/22, 9.09%), Asian Pharmacoepidemiology Network (1/22, 4.55%), and newly created CDMs (3/22, 13.64%).

Table 4. Top 10 journals with most studies in pharmacoepidemiologic research based on CDMs<sup>a</sup> through 2024.

Rank	Journal title	Country	Counts, n	IF <sup>b</sup> (2023)	JCR <sup>c</sup> (2023)	H-index	Total citations, n
1	Vaccine	United Kingdom	49	4.5	Q3	19	1207
2	Pediatrics	United States	24	6.2	Q2	18	1167
3	Pharmacoepidemiology and Drug Safety	United Kingdom	22	2.4	Q3	10	310
4	Obstetrics and Gynecology	United States	9	5.7	Q2	8	322
5	Scientific Reports	United Kingdom	8	3.8	Q3	4	38
6	JAMA-Journal of the American Medical Association	United States	7	63.1	Q1	7	925
7	Jama Network Open	United States	6	10.5	Q1	5	147
8	American Journal of Epidemiology	United States	6	5.0	Q2	4	187
9	Drug Safety	New Zealand	5	4.0	Q2	3	62
10	Academic Pediatrics	United States	4	3.0	Q3	4	34

<sup>a</sup>CDM: Common Data Model.

<sup>b</sup>IF: impact factor.

<sup>c</sup>JCR: journal citation reports.

As illustrated in Figure 2D, the number of pharmacoepidemiologic research studies based on CDMs in *Vaccine* shows notable variability, with marked increases in 2016 and 2022. In contrast, publication volume in *Pediatrics* and *JAMA Network Open* has remained relatively stable within this field.

Analysis of Countries and Regions

The top 10 data source countries were listed in Table 5. An analysis of the 308 pharmacoepidemiology publications

based on CDMs reveals that the leading countries in research output are the United States with 286 (92.86%) studies, South Korea with 53 (17.21%) studies, the United Kingdom with 11 (3.57%) studies.

Table 5. Top 10 data source countries in pharmacoepidemiologic research based on CDMs<sup>a</sup> through 2024.

Rank	Country	Counts, n	Average citation counts per study	Average TCpY <sup>b</sup> counts per study
1	United States	286	34.53	4.13
2	South Korea	53	9.26	2.15

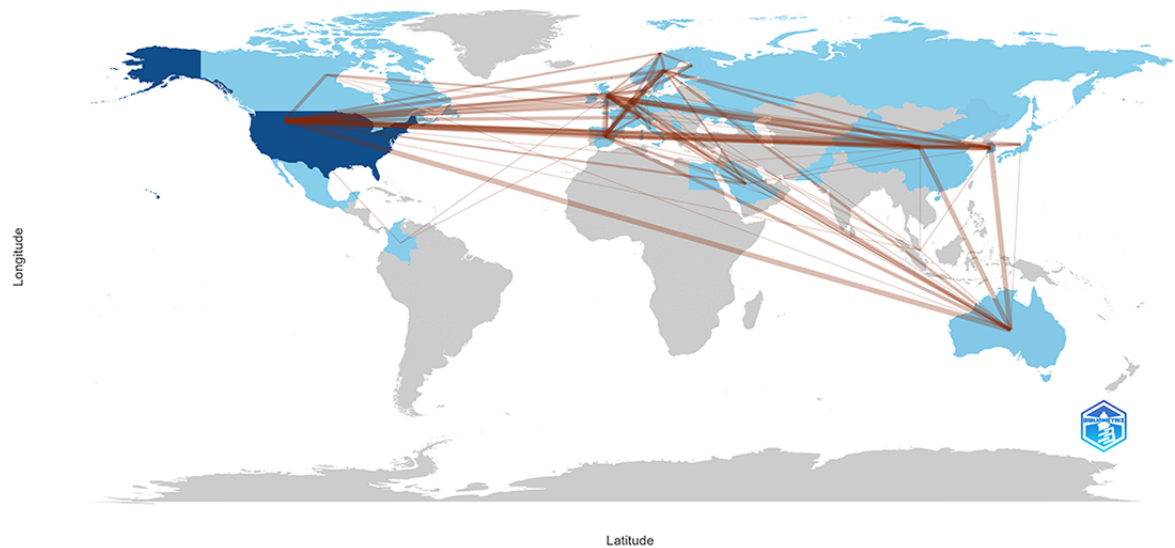
Rank	Country	Counts, n	Average citation counts per study	Average TCpY <sup>b</sup> counts per study
3	United Kingdom	11	16.36	3.85
4	France	9	19.56	4.52
5	China	9	15.67	2.73
6	Spain	8	25.75	6.78
7	Germany	7	23.43	6.31
8	Japan	7	23.14	4.52
9	Netherlands	7	15.86	4.20
10	Denmark	6	10.83	1.86

<sup>a</sup>CDM: Common Data Model.

<sup>b</sup>TCpY: Total Citations per Year. Data from multiple countries were counted separately for each contributing country.

As illustrated in Figure 3, international collaboration is strong among the United States, Europe, Australia, and East Asia. The United States acts as a central hub for multinational collaborations, with substantial partnerships involving China, the United Kingdom, Spain, South Korea, and Australia. South Korea is also actively engaged in various international projects, maintaining particularly close connections with the United States and the United Kingdom.

**Figure 3.** Global collaboration patterns among countries conducting CDM-based pharmacoepidemiologic research. CDM: Common Data Model; The thickness of the lines represents the strength of collaboration.



Analysis of Keywords

The included publications encompass a total of 590 keywords. Figure 4 illustrates the temporal trends of popular research topics in CDMs based on pharmacoepidemiology. Since 2012, topics such as vaccine safety, immunization, and

influenza have remained consistently active. Starting in 2020, research on COVID-19, stroke, and diabetes has increased significantly, with extensive use of electronic health record (EHR) data.



**Figure 4.** Trends in research topics over time in pharmacoepidemiologic research based on CDM through 2024. Dot: Represents the appearance of a specific topic or keyword. Dot size: Indicates the influence or research volume of the topic. Horizontal line: the duration or active period of the topic. CDM: Common Data Model.

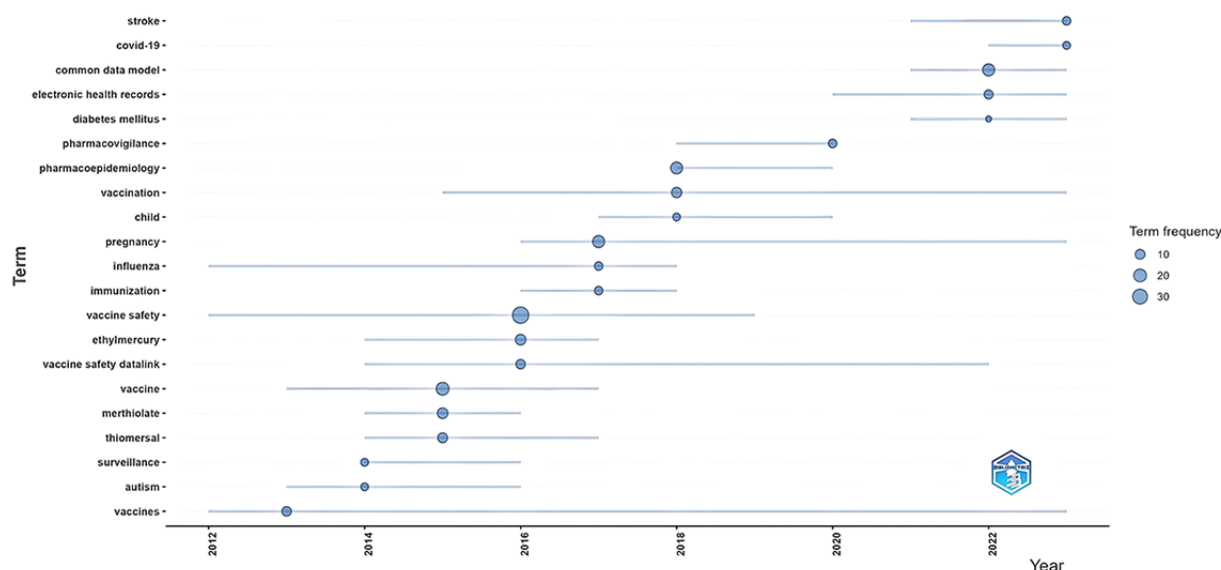
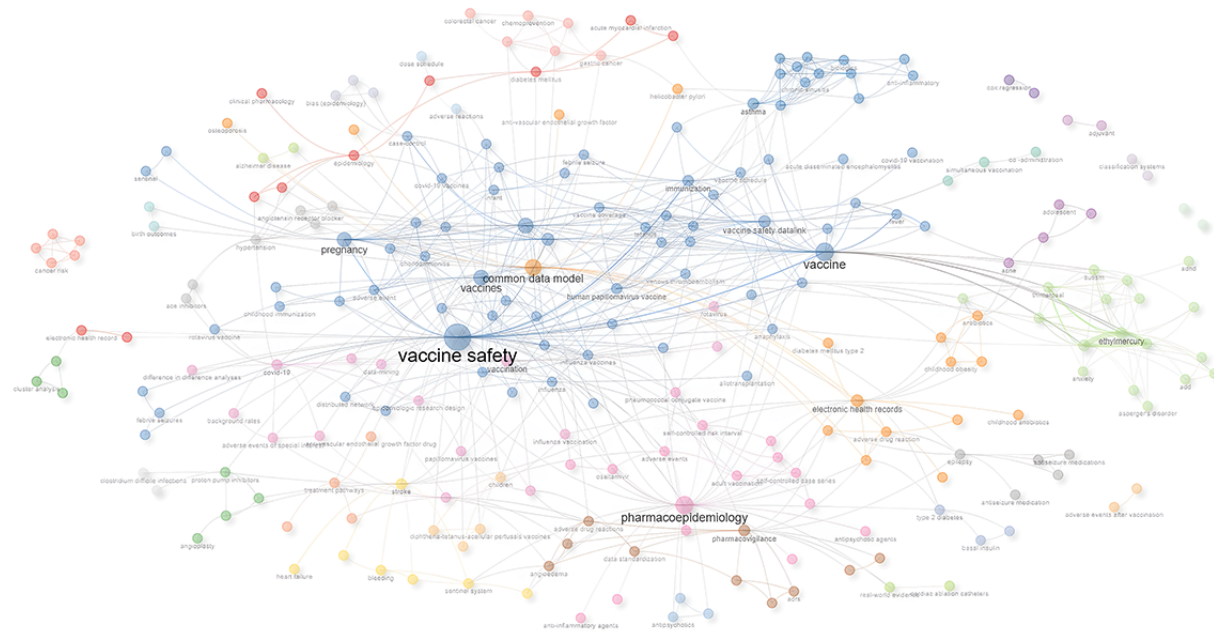


Figure 5 illustrates the trends in research topics. The nodes of the same color within a cluster represent closely related co-occurrences, with node size and link width varying based on the degree and strength of co-occurrence. “Vaccine safety” and “pharmacoepidemiology” occupy central positions in the network, showing strong connections with numerous surrounding topics. “Common data model” serves

as another key topic, closely linked to “electronic health records” and “observational research.” The nodes in different colors indicate topic clusters, with the blue cluster primarily focusing on vaccine-related research, the orange cluster on drug monitoring and adverse reactions, and the brown cluster on drug regulation and risk assessment.

**Figure 5.** Network map of keywords in pharmacoepidemiologic research based on CDM. CDM: Common Data Model.



The dual-map overlay analysis illustrated the interdisciplinary citation structure of CDM-based pharmacoepidemiologic research. Most publications were clustered in the “Medicine, Medical, Clinical” domains and predominantly cited journals from the fields of “Health, Nursing, Medicine” and “Molecular Biology, Genetics,” suggesting strong disciplinary integration ([Multimedia Appendix 6](#)).

## Discussion

### *Principal Findings*

This study used bibliometric methods to analyze research trends in the application of CDMs in the field of pharmacoepidemiology, while also exploring research hot spots among

highly cited studies in this field. The findings revealed that highly cited studies were predominantly originated from the United States, reflecting the central role of the United States researchers and institutions in this field, while South Korea's influence has been rising rapidly in recent years. Included studies were often multicenter, interorganizational collaborations, leveraging extensive sample sizes and cross-regional data coverage to enhance the generalizability and scientific value of research outcomes, thereby increasing citation frequency [29]. Vaccine-related studies were a primary focus, highlighting the significance of the CDM in public health, particularly in supporting real-time assessment of drugs and vaccines and adverse event monitoring during global health crises by integrating data across regions and organizations [30]. Our findings provide essential guidance for researchers embarking on studies in this field, assisting in the identification of potential collaborators, influential studies, and exemplary work for reference, as well as in the selection of suitable journals for publication. It enhances the efficiency of establishing research networks and supports the effective dissemination and academic impact of new studies.

### ***Characteristics of Highly Cited Studies***

The most cited publications globally have provided critical insights into adverse events associated with various vaccines, drug safety, and drug usage [14,30-33]. One particularly influential study, titled "Surveillance for Adverse Events After COVID-19 mRNA Vaccination," monitored severe adverse events related to COVID-19 vaccines and has emerged as the most cited and impactful publication in this field [30]. This highlights the rapid response and proactive surveillance capabilities of CDM during public health emergencies, enabling timely evaluation of new vaccines.

We conducted a comparative analysis between globally highly cited studies and less cited ones. To mitigate the advantage of earlier publications due to absolute citation counts, we also grouped studies based on TCpY. The findings indicate that studies in the high-citation group tend to involve more institutions. Cross-institutional collaboration and data sharing provide significant practical guidance for research and policy-making. Multicenter studies, with larger participant numbers from diverse regions, enhance statistical power and ensure that findings are more generalizable for broader populations. Such studies allow researchers to capture a wider range of variables and cases, thus providing a robust data foundation for investigating rare diseases, uncommon adverse events, or smaller subpopulations [29,34,35].

Vaccine studies are more frequently highly cited compared to drug research, due to factors related to their public health impact and resource allocation. First, the high citation rates of vaccine research are closely tied to its critical role in public health. Global public health crises, such as the COVID-19 pandemic, have further elevated the priority of vaccine research, increasing its visibility and citation rates, as such studies are essential for policy-making and health interventions [36,37]. In addition, vaccine development often involves substantial public funding and data-sharing efforts, enabling wider access and greater citation of vaccine-related findings.

In contrast, drug research is typically driven by pharmaceutical companies, with more restricted data availability and limited openness, which can constrain its citation impact [38].

### ***Key Contributors and Institutional Insights***

The analysis of authors and institutions highlights the contributions of key researchers to the application of CDMs in pharmacoepidemiology. Core authors such as Nicola P Klein and Matthew F Daley demonstrate their extensive expertise and consistent productivity in this field. Among the top 10 most prolific authors, 5 are affiliated with different regions of Kaiser Permanente, a leading health care provider and nonprofit health plan in the United States since its establishment in 1945, currently serving members across 8 states and the District of Columbia [39]. In South Korea, Hallym University and Seoul National University have conducted multiple studies in this domain around 2020, using OMOP and K-CDM. Between 2016 and 2018, the Ministry of Food and Drug Safety and the Korea Institute of Drug Safety and Risk Management converted EHR of over 9 million patients into CDM, reflecting Korea's progress in integrating and sharing cross-cohort data resources [40]. Subsequently, Korea developed the K-CDM specifically for pharmacovigilance systems, leveraging multi-center EHRs to monitor adverse drug reactions [41,42].

### ***Global Landscape and Collaborative Trends***

The majority of studies in the high-citation group originated from the United States, which is the birthplace of several CDMs. As early as 2001, the United States established the VSD to track adverse events across multiple regions and populations [43]. Subsequently, the OMOP was released in 2007, the Sentinel was launched by the FDA in 2008, and the PCORnet initiative began in 2014, positioning the United States at the forefront of CDM development and usage worldwide [43]. Moreover, the United States excels in drug and vaccine surveillance through the VSD and Sentinel systems. VSD provides real-time updates and comprehensive vaccine data, while Sentinel builds on VSD with tools enabling direct, standardized analyses, streamlining research processes [11].

In addition, the United States is the primary contributor in terms of publication volume and frequency, followed by South Korea. These countries have made significant contributions to advancing knowledge in this field. Notably, co-occurrence analysis highlights active participation from other countries in collaborative networks, with strong connections observed among North America, Europe, and East Asia. However, we observed that most international collaborations occur primarily among limited-income countries, with limited participation from other nations, particularly in limited-income countries. Given that CDMs are designed to harmonize data from multiple databases and facilitate standardized analyses [44], future research should focus on fostering broader and more impactful international collaborations from countries at different levels

of development. Furthermore, the analysis of institutional productivity provides practical insights that may directly benefit researchers seeking potential collaborators. Our results highlight that Kaiser Permanente and CDC in the United States, both major users of VSD, represent key contributors to CDM-based vaccine research in the United States. Meanwhile, institutions such as Hallym University and Seoul National University in South Korea have emerged as important centers of CDM implementation, primarily using OMOP and K-CDM. Recognizing these leading institutions and their methodological preferences may facilitate strategic planning of future collaborative studies, enabling researchers to align their research agendas, optimize resource allocation, and ultimately enhance the global impact of CDM-based pharmacoepidemiologic research.

### **Implications for Future Research and Practice**

To foster more equitable global participation, particularly from low- and middle-income countries, future efforts should prioritize improving access to CDM infrastructure. This includes the provision of open-source tools, standardized implementation workflows, and multilingual technical documentation. In parallel, the establishment of dedicated funding mechanisms and long-term capacity-building programs—such as technical training and data transformation support—will be essential to enable meaningful international collaboration. These strategies aim to bridge existing disparities and promote a more inclusive and diverse global CDM research ecosystem.

Furthermore, the findings of this study offer practical implications for clinical decision-making and real-world health policy. By identifying highly cited, multicenter studies—many of which focus on vaccine safety and effectiveness—this analysis highlights how CDM-based research can rapidly generate actionable evidence during public health emergencies. The global collaborative networks and methodological practices identified in this study may serve as a reference for designing future pharmacoepidemiologic studies that directly inform clinical guidelines, pharmacovigilance strategies, and regulatory decision-making. Strengthening the application of CDMs in such contexts can ultimately improve patient safety, treatment optimization, and health system responsiveness.

### **Strengths and Limitations**

This study has some strengths. To our knowledge, this bibliometric study is the first of its kind in the field of application of CDMs in pharmacoepidemiology, to summarize the publication characteristics. Using the systematic review, we screened studies that met the inclusion criteria, ensuring the reliability of the selected studies. Furthermore, we imposed no restrictions on publication date or on the types of CDMs, which allowed us to include studies

globally. This approach enables comparisons across different periods, countries, and types of CDMs. It must be acknowledged that our research has several modest limitations. First, the most significant limitation of our study lies in the inherent biases of citation analysis. These include the tendency for earlier publications to accumulate higher citation counts, journal and author self-citations, incomplete citation data, and omissions, all of which may affect citation rates [23,45,46]. To mitigate such biases, we attempted to use TCpY as an alternative metric. Second, we selected WoSCC based on previous studies [18,36], as it allows standardized citation-based analysis. However, WoSCC does not index all eligible studies. Among the 308 included studies, 285 were found in WoSCC and therefore included in citation-related analyses. The remaining 23 studies lacked citation data and were excluded from this component. To evaluate whether their exclusion introduced selection bias, we compared key characteristics between the 2 groups. As shown in [Multimedia Appendix 3](#), no statistically significant differences were found. Therefore, the potential bias introduced by excluding non-WoSCC studies is likely minimal. Third, we included only English and Chinese studies, which means that studies published in other languages were excluded. This may lead to the omission of some research. However, we searched the Virtual Health Library to ensure that literature published in Latin America was also covered to the greatest extent possible. Fourth, the use of the bibliometrix R package, while powerful for large-scale bibliometric analysis, may introduce methodological limitations. For example, its analytical outputs rely on the quality and structure of the input data and specific co-occurrence or clustering algorithms [26]. To minimize these effects, we carefully cleaned and standardized the dataset and interpreted the results in combination with manual validation and descriptive statistics. In addition, the release dates of different CDMs vary, with earlier CDMs having accumulated a higher publication volume and citations compared to those released more recently.

### **Conclusions**

This bibliometric analysis provides a comprehensive overview of the application of CDMs in pharmacoepidemiology. The findings indicate a significant increase in publications over time, with the United States leading in both publication volume and citation counts. Notably, high-citation studies often involve multicenter studies, particularly focusing on vaccine, underscoring the importance of collaborative research efforts. These insights emphasize the critical role of CDMs in facilitating large-scale, collaborative pharmacoepidemiologic research, as well as provide researchers with insights for selecting collaborators and choosing journals for publication. To enhance global equity and clinical relevance, further efforts should support infrastructure accessibility, international cooperation, and the integration of CDM-based evidence into public health and regulatory decision-making.

### **Acknowledgments**

Supported by the National Natural Science Foundation of China (grants 72474008 and 72074011), the third batch of Key Projects of Scientific Act for Drug Regulation of China (grants RS2024X006 and RS2024Z008), Special Project for Director,

China Center for Evidence Based Traditional Chinese Medicine (grant 2020YJSZX-2), Research Project of China Society for Drug Regulation (2025-Y-Y-012), and Project of Center for Drug Reevaluation of China National Medical Products Administration (CDR2024R01001).

### Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

### Authors' Contributions

YZ: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Formal analysis, Data curation, Conceptualization. Meng Zhang: Writing – review & editing, Data curation, Conceptualization. Conghui Wang: Data curation. LG: Data curation. JX: Writing – review & editing. PS: Writing – review & editing. YS: Writing – review & editing. MF: Writing – review & editing. SCY: Writing – review & editing. FS: Writing – review & editing, Validation, Funding acquisition, Conceptualization.

### Conflicts of Interest

None declared.

### Multimedia Appendix 1

Search strategies.

[\[DOCX File \(Microsoft Word File\), 24 KB-Multimedia Appendix 1\]](#)

### Multimedia Appendix 2

Top 10 most cited articles in pharmacoepidemiologic research based on Common Data Models through 2024.

[\[DOCX File \(Microsoft Word File\), 20 KB-Multimedia Appendix 2\]](#)

### Multimedia Appendix 3

Comparison of study characteristics between articles included and those not included in the Web of Science Core Collection.

[\[DOCX File \(Microsoft Word File\), 20 KB-Multimedia Appendix 3\]](#)

### Multimedia Appendix 4

Top 10 authors with most articles in pharmacoepidemiologic research based on Common Data Models through 2024.

[\[DOCX File \(Microsoft Word File\), 19 KB-Multimedia Appendix 4\]](#)

### Multimedia Appendix 5

Top 10 institutions with most articles in pharmacoepidemiologic research based on Common Data Models through 2024.

[\[DOCX File \(Microsoft Word File\), 18 KB-Multimedia Appendix 5\]](#)

### Multimedia Appendix 6

Dual-map overlay of journal citation trajectories in Common Data Model-based pharmacoepidemiologic research.

[\[DOCX File \(Microsoft Word File\), 2345 KB-Multimedia Appendix 6\]](#)

### Checklist 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 checklist.

[\[DOCX File \(Microsoft Word File\), 276 KB-Checklist 1\]](#)

### References

1. Liu X, Hou J. Pharmacoepidemiology. In: Wang C, Liu F, editors. Textbook of Clinical Epidemiology: For Medical Students. Springer Nature; 2023:265-280. [doi: [10.1007/978-981-99-3622-9\\_15](https://doi.org/10.1007/978-981-99-3622-9_15)]
2. Moro PL, Haber P, McNeil MM. Challenges in evaluating post-licensure vaccine safety: observations from the Centers for Disease Control and Prevention. Expert Rev Vaccines. Oct 2019;18(10):1091-1101. [doi: [10.1080/14760584.2019.1676154](https://doi.org/10.1080/14760584.2019.1676154)] [Medline: [31580725](https://pubmed.ncbi.nlm.nih.gov/31580725/)]
3. Platt A, Wilson J, Hall R, et al. Comparative effectiveness of alternative treatment approaches to secondary hyperparathyroidism in patients receiving maintenance hemodialysis: An observational trial emulation. Am J Kidney Dis. Jan 2024;83(1):58-70. [doi: [10.1053/j.ajkd.2023.05.016](https://doi.org/10.1053/j.ajkd.2023.05.016)] [Medline: [37690631](https://pubmed.ncbi.nlm.nih.gov/37690631/)]
4. Romio S, Weibel D, Dieleman JP, et al. Guillain-Barré syndrome and adjuvanted pandemic influenza A (H1N1) 2009 vaccines: a multinational self-controlled case series in Europe. PLoS ONE. 2014;9(1):e82222. [doi: [10.1371/journal.pone.0082222](https://doi.org/10.1371/journal.pone.0082222)] [Medline: [24404128](https://pubmed.ncbi.nlm.nih.gov/24404128/)]
5. Chang Q, Yan Z, Zhou M, et al. Mining multi-center heterogeneous medical data with distributed synthetic learning. Nat Commun. 2023;14(1):5510. [doi: [10.1038/s41467-023-40687-y](https://doi.org/10.1038/s41467-023-40687-y)]



6. Reinecke I, Zoch M, Reich C, Sedlmayr M, Bathelt F. The usage of OHDSI OMOP - a scoping review. *Stud Health Technol Inform.* Sep 21, 2021;283:95-103. [doi: [10.3233/SHTI210546](https://doi.org/10.3233/SHTI210546)] [Medline: [34545824](https://pubmed.ncbi.nlm.nih.gov/34545824/)]
7. Garza M, Del Fiore G, Tenenbaum J, Walden A, Zozus MN. Evaluating common data models for use with a longitudinal community registry. *J Biomed Inform.* Dec 2016;64:333-341. [doi: [10.1016/j.jbi.2016.10.016](https://doi.org/10.1016/j.jbi.2016.10.016)] [Medline: [27989817](https://pubmed.ncbi.nlm.nih.gov/27989817/)]
8. Overhage JM, Ryan PB, Reich CG, Hartzema AG, Stang PE. Validation of a common data model for active safety surveillance research. *J Am Med Inform Assoc.* Jan 2012;19(1):54-60. [doi: [10.1136/amiajnl-2011-000376](https://doi.org/10.1136/amiajnl-2011-000376)]
9. Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. *J Am Med Inform Assoc.* 2014;21(4):578-582. [doi: [10.1136/amiajnl-2014-002747](https://doi.org/10.1136/amiajnl-2014-002747)] [Medline: [24821743](https://pubmed.ncbi.nlm.nih.gov/24821743/)]
10. Pfaff ER, Champion J, Bradford RL, et al. Fast Healthcare Interoperability Resources (FHIR) as a Meta model to integrate common data models: Development of a tool and quantitative validation study. *JMIR Med Inform.* Oct 16, 2019;7(4):e15199. [doi: [10.2196/15199](https://doi.org/10.2196/15199)] [Medline: [31621639](https://pubmed.ncbi.nlm.nih.gov/31621639/)]
11. Fahey KR. The Pioneering Role of the Vaccine Safety Datalink Project (VSD) to advance collaborative research and distributed data networks. *EGEMS (Wash DC).* 2015;3(1):1195. [doi: [10.13063/2327-9214.1195](https://doi.org/10.13063/2327-9214.1195)] [Medline: [26793736](https://pubmed.ncbi.nlm.nih.gov/26793736/)]
12. Ball R, Robb M, Anderson SA, Dal Pan G. The FDA's sentinel initiative--A comprehensive approach to medical product surveillance. *Clin Pharmacol Ther.* Mar 2016;99(3):265-268. [doi: [10.1002/cpt.320](https://doi.org/10.1002/cpt.320)] [Medline: [26667601](https://pubmed.ncbi.nlm.nih.gov/26667601/)]
13. Ahmadi N, Zoch M, Kelbert P, Noll R, Schaaf J, Wolfien M, et al. Methods used in the development of common data models for health data: Scoping review. *JMIR Med Inform.* Aug 3, 2023;11:e45116. [doi: [10.2196/45116](https://doi.org/10.2196/45116)]
14. Hripcsak G, Ryan PB, Duke JD, et al. Characterizing treatment pathways at scale using the OHDSI network. *Proc Natl Acad Sci U S A.* Jul 5, 2016;113(27):7329-7336. [doi: [10.1073/pnas.1510502113](https://doi.org/10.1073/pnas.1510502113)] [Medline: [27274072](https://pubmed.ncbi.nlm.nih.gov/27274072/)]
15. Willame C, Dodd C, Durán C, et al. Background rates of 41 adverse events of special interest for COVID-19 vaccines in 10 European healthcare databases - an ACCESS cohort study. *Vaccine (Auckl).* Jan 2023;41(1):251-262. [doi: [10.1016/j.vaccine.2022.11.031](https://doi.org/10.1016/j.vaccine.2022.11.031)] [Medline: [36446653](https://pubmed.ncbi.nlm.nih.gov/36446653/)]
16. Ryan PB, Buse JB, Schuemie MJ, et al. Comparative effectiveness of canagliflozin, SGLT2 inhibitors and non-SGLT2 inhibitors on the risk of hospitalization for heart failure and amputation in patients with type 2 diabetes mellitus: A real-world meta-analysis of 4 observational databases (OBSERVE-4D). *Diabetes Obesity Metabolism.* Nov 2018;20(11):2585-2597. [doi: [10.1111/dom.13424](https://doi.org/10.1111/dom.13424)] [Medline: [29938883](https://pubmed.ncbi.nlm.nih.gov/29938883/)]
17. Suchard MA, Schuemie MJ, Krumholz HM, et al. Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis. *Lancet.* Nov 16, 2019;394(10211):1816-1826. [doi: [10.1016/S0140-6736\(19\)32317-7](https://doi.org/10.1016/S0140-6736(19)32317-7)] [Medline: [31668726](https://pubmed.ncbi.nlm.nih.gov/31668726/)]
18. Zhang L, Zheng H, Jiang ST, et al. Worldwide research trends on tumor burden and immunotherapy: a bibliometric analysis. *Int J Surg.* Mar 1, 2024;110(3):1699-1710. [doi: [10.1097/JS9.0000000000001022](https://doi.org/10.1097/JS9.0000000000001022)]
19. Ai S, Li Y, Zheng H, et al. Collision of herbal medicine and nanotechnology: a bibliometric analysis of herbal nanoparticles from 2004 to 2023. *J Nanobiotechnology.* Apr 1, 2024;22(1):140. [doi: [10.1186/s12951-024-02426-3](https://doi.org/10.1186/s12951-024-02426-3)] [Medline: [38556857](https://pubmed.ncbi.nlm.nih.gov/38556857/)]
20. Garfield E. Journal impact factor: a brief review. *CMAJ.* Oct 19, 1999;161(8):979-980. [Medline: [10551195](https://pubmed.ncbi.nlm.nih.gov/10551195/)]
21. Garfield E. How can impact factors be improved? *BMJ.* Aug 17, 1996;313(7054):411-413. [doi: [10.1136/bmj.313.7054.411](https://doi.org/10.1136/bmj.313.7054.411)] [Medline: [8761234](https://pubmed.ncbi.nlm.nih.gov/8761234/)]
22. Pagni M, Khan NR, Cohen HL, Choudhri AF. Highly cited works in radiology: the top 100 cited articles in radiologic journals. *Acad Radiol.* Aug 2014;21(8):1056-1066. [doi: [10.1016/j.acra.2014.03.011](https://doi.org/10.1016/j.acra.2014.03.011)] [Medline: [24833569](https://pubmed.ncbi.nlm.nih.gov/24833569/)]
23. Brandt JS, Hadaya O, Schuster M, Rosen T, Sauer MV, Ananth CV. A bibliometric analysis of top-cited journal articles in obstetrics and gynecology. *JAMA Netw Open.* Dec 2, 2019;2(12):e1918007. [doi: [10.1001/jamanetworkopen.2019.18007](https://doi.org/10.1001/jamanetworkopen.2019.18007)] [Medline: [31860106](https://pubmed.ncbi.nlm.nih.gov/31860106/)]
24. Li T, Kang C, Ren G, et al. Top 100 cited articles related to EUS: A bibliometric analysis. *Endosc Ultrasound.* 2024;13(4):259-268. [doi: [10.1097/eus.0000000000000081](https://doi.org/10.1097/eus.0000000000000081)] [Medline: [39318756](https://pubmed.ncbi.nlm.nih.gov/39318756/)]
25. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med.* Jul 21, 2009;6(7):e1000097. [doi: [10.1371/journal.pmed.1000097](https://doi.org/10.1371/journal.pmed.1000097)]
26. Aria M, Cuccurullo C. bibliometrix: An R-tool for comprehensive science mapping analysis. *J Informetr.* Nov 2017;11(4):959-975. [doi: [10.1016/j.joi.2017.08.007](https://doi.org/10.1016/j.joi.2017.08.007)]
27. Garfield E. Citation analysis as a tool in journal evaluation. *Science.* Nov 3, 1972;178(4060):471-479. [doi: [10.1126/science.178.4060.471](https://doi.org/10.1126/science.178.4060.471)] [Medline: [5079701](https://pubmed.ncbi.nlm.nih.gov/5079701/)]
28. Hirsch JE. An index to quantify an individual's scientific research output. *Proc Natl Acad Sci U S A.* Nov 15, 2005;102(46):16569-16572. [doi: [10.1073/pnas.0507655102](https://doi.org/10.1073/pnas.0507655102)] [Medline: [16275915](https://pubmed.ncbi.nlm.nih.gov/16275915/)]



29. Lane K, Palm ME, Marion E, et al. Approaches for enhancing the informativeness and quality of clinical trials: Innovations and principles for implementing multicenter trials from the Trial Innovation Network. *J Clin Trans Sci.* 2023;7(1):e131. [doi: [10.1017/cts.2023.560](https://doi.org/10.1017/cts.2023.560)]
30. Klein NP, Lewis N, Goddard K, et al. Surveillance for Adverse Events After COVID-19 mRNA Vaccination. *JAMA.* Oct 12, 2021;326(14):1390-1399. [doi: [10.1001/jama.2021.15072](https://doi.org/10.1001/jama.2021.15072)] [Medline: [34477808](https://pubmed.ncbi.nlm.nih.gov/34477808/)]
31. McNeil MM, Weintraub ES, Duffy J, et al. Risk of anaphylaxis after vaccination in children and adults. *J Allergy Clin Immunol.* Mar 2016;137(3):868-878. [doi: [10.1016/j.jaci.2015.07.048](https://doi.org/10.1016/j.jaci.2015.07.048)] [Medline: [26452420](https://pubmed.ncbi.nlm.nih.gov/26452420/)]
32. Yih WK, Lieu TA, Kulldorff M, et al. Intussusception risk after rotavirus vaccination in U.S. infants. *N Engl J Med.* Feb 6, 2014;370(6):503-512. [doi: [10.1056/NEJMoa1303164](https://doi.org/10.1056/NEJMoa1303164)] [Medline: [24422676](https://pubmed.ncbi.nlm.nih.gov/24422676/)]
33. Toh S, Reichman ME, Houstoun M, et al. Comparative risk for angioedema associated with the use of drugs that target the renin-angiotensin-aldosterone system. *Arch Intern Med.* Nov 12, 2012;172(20):1582-1589. [doi: [10.1001/2013.jamainternmed.34](https://doi.org/10.1001/2013.jamainternmed.34)] [Medline: [23147456](https://pubmed.ncbi.nlm.nih.gov/23147456/)]
34. Wu Q, Tong J, Zhang B, et al. Real-World Effectiveness of BNT162b2 against infection and severe diseases in children and adolescents. *Ann Intern Med.* Feb 2024;177(2):165-176. [doi: [10.7326/M23-1754](https://doi.org/10.7326/M23-1754)] [Medline: [38190711](https://pubmed.ncbi.nlm.nih.gov/38190711/)]
35. Ahmadi N, Zoch M, Guengoeze O, et al. How to customize common data models for rare diseases: an OMOP-based implementation and lessons learned. *Orphanet J Rare Dis.* Aug 14, 2024;19(1):298. [doi: [10.1186/s13023-024-03312-9](https://doi.org/10.1186/s13023-024-03312-9)] [Medline: [39143600](https://pubmed.ncbi.nlm.nih.gov/39143600/)]
36. Zhang Y, Quan L, Xiao B, Du L. The 100 top-cited studies on vaccine: a bibliometric analysis. *Hum Vaccin Immunother.* 2019;15(12):3024-3031. [doi: [10.1080/21645515.2019.1614398](https://doi.org/10.1080/21645515.2019.1614398)] [Medline: [31112440](https://pubmed.ncbi.nlm.nih.gov/31112440/)]
37. Stehlik P, Dowsett C, Camacho X, Falster MO, Lim R, Nasreen S, et al. Evolution of the data and methods in real-world COVID-19 vaccine effectiveness studies on mortality: a scoping review protocol. *BMJ Open.* Mar 19, 2024;14(3):e079071. [doi: [10.1136/bmjopen-2023-079071](https://doi.org/10.1136/bmjopen-2023-079071)] [Medline: [38508618](https://pubmed.ncbi.nlm.nih.gov/38508618/)]
38. How often do vaccine trials hit payday? MIT News. URL: <https://news.mit.edu/2020/how-often-vaccine-trials-succeed-0527> [Accessed 2025-07-18]
39. Kaiser Permanente exists to provide high-quality, affordable health care services and to improve the health of our members and the communities we serve. Kaiser Permanente. URL: <https://about.kaiserpermanente.org/who-we-are/our-mission> [Accessed 2025-07-18]
40. Son N, Kim B, Chung S, Han S. Korean pharmacovigilance system based on EHR-CDM. *Stud Health Technol Inform.* Aug 21, 2019;264:1592-1593. [doi: [10.3233/SHTI190550](https://doi.org/10.3233/SHTI190550)] [Medline: [31438247](https://pubmed.ncbi.nlm.nih.gov/31438247/)]
41. Choe S, Lee S, Park CH, et al. Development and application of an active pharmacovigilance framework based on electronic healthcare records from multiple centers in Korea. *Drug Saf.* Jul 2023;46(7):647-660. [doi: [10.1007/s40264-023-01296-2](https://doi.org/10.1007/s40264-023-01296-2)] [Medline: [37243963](https://pubmed.ncbi.nlm.nih.gov/37243963/)]
42. Choi YI, Park SJ, Chung JW, et al. Development of machine learning model to predict the 5-year risk of starting biologic agents in patients with inflammatory bowel disease (IBD): K-CDM Network study. *JCM.* Nov 2020;9(11):3427. [doi: [10.3390/jcm9113427](https://doi.org/10.3390/jcm9113427)]
43. Weeks J, Pardee R. Learning to share health care data: A brief timeline of influential common data models and distributed health data networks in U.S. health care research. *EGEMS (Wash DC).* Mar 25, 2019;7(1):4. [doi: [10.5334/egems.279](https://doi.org/10.5334/egems.279)] [Medline: [30937326](https://pubmed.ncbi.nlm.nih.gov/30937326/)]
44. Lai EC, Stang P, Yang YH, Kubota K, Wong ICK, Setoguchi S. International multi-database pharmacoepidemiology: Potentials and pitfalls. *Curr Epidemiol Rep.* Dec 2015;2(4):229-238. [doi: [10.1007/s40471-015-0059-z](https://doi.org/10.1007/s40471-015-0059-z)]
45. Ogdan TL, Bartley DL. The ups and downs of journal impact factors. *Ann Occup Hyg.* Mar 2008;52(2):73-82. [doi: [10.1093/annhyg/men002](https://doi.org/10.1093/annhyg/men002)] [Medline: [18316351](https://pubmed.ncbi.nlm.nih.gov/18316351/)]
46. Mishra S, Fegley BD, Diesner J, Torvik VI. Self-citation is the hallmark of productive authors, of any gender. *PLoS ONE.* 2018;13(9):e0195773. [doi: [10.1371/journal.pone.0195773](https://doi.org/10.1371/journal.pone.0195773)] [Medline: [30256792](https://pubmed.ncbi.nlm.nih.gov/30256792/)]

## Abbreviations

**AsPEN:** Asian Pharmacoepidemiology Network  
**CDM:** Common Data Model  
**EHR:** electronic health record  
**OMOP:** Observational Medical Outcomes Partnership  
**PRISMA:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses  
**TCpY:** Total Citations per Year  
**VSD:** Vaccine Safety Datalink  
**WoSCC:** Web of Science Core Collection

*Edited by Andrew Coristine; peer-reviewed by Ali Jafarizadeh, Xiongfeng Huang; submitted 05.02.2025; final revised version received 06.05.2025; accepted 20.05.2025; published 28.07.2025*

*Please cite as:*

Zheng Y, Zhang M, Wang C, Gao L, Xie J, Shen P, Sun Y, Feng M, You SC, Sun F

*Pharmacoepidemiologic Research Based on Common Data Models: Systematic Review and Bibliometric Analysis*

*JMIR Med Inform 2025;13:e72225*

URL: <https://medinform.jmir.org/2025/1/e72225>

doi: [10.2196/72225](https://doi.org/10.2196/72225)

©Yongqi Zheng, Meng Zhang, Conghui Wang, Ling Gao, Junqing Xie, Peng Shen, Yexiang Sun, Menglin Feng, Seng chan You, Feng Sun. Originally published in JMIR Medical Informatics (<https://medinform.jmir.org>), 28.07.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <https://medinform.jmir.org/>, as well as this copyright and license information must be included.