

Research Advance of Causal Inference in Clinical Medicine: A Bibliometrics Analysis via Citespace

Guoqiang Qin¹, Jianxiang Wei^{1,2}, Yuehong Sun³, Wenwen Du¹

¹School of Management, Nanjing University of Posts and Telecommunications, Nanjing, People's Republic of China; ²Library, Nanjing University of Posts and Telecommunications, Nanjing, People's Republic of China; ³School of Mathematical Sciences, Nanjing Normal University, Nanjing, People's Republic of China

Correspondence: Jianxiang Wei, School of Management, Nanjing University of Posts and Telecommunications, Nanjing, 210003, People's Republic of China, Email jxwei@njupt.edu.cn; Yuehong Sun, School of Mathematical Sciences, Nanjing Normal University, Nanjing, 210023, People's Republic of China, Email 05234@njnu.edu.cn

Objective: Causal inference in clinical medicine provides scientific evidence for precision medicine and individualized treatment by revealing the true associations between interventions and health outcomes. This study aims to conduct a comprehensive bibliometric analysis to identify current research trends, primary themes, and future directions for the application of causal inference in clinical medicine.

Methods: We conducted a literature search in the Web of Science database using causal inference and medical terminology as subject keywords, covering the period from January 1986 to December 2024. After screening, we obtained 4,316 documents for analysis. Utilizing CiteSpace to generate network diagrams, we analyzed data related to the number of publications, citation analysis, collaboration relationships, keyword co-occurrence, and highlighted terms to illustrate the knowledge map and collaboration network in this field.

Results: Publications on medical causal inference shows a fluctuating growth trend over time. The United States was the top contributors to this field. Harvard University is the leading research institution. George David Smith is the most prolific author, Robbins JM is the most cited scholar. The major research hotspots concentrated in fields such as epidemiology, coronary heart disease and health. Notably, marginal structural models, counterfactual forecasting, and Mendelian randomization have consistently been key methodologies in research. The burstness of keywords reveals that big data, DNA methylation, and robust estimation are emerging research directions.

Conclusion: In clinical research, counterfactual forecasting provides prospective guidance for optimizing clinical strategies; Mendelian randomization helps uncover potential therapeutic targets; and marginal structural models enhance the accuracy of causal effect estimation in clinical studies. The future integration of various data sources to improve causal inference methods is anticipated to enhance the sensitivity and specificity of trials, ultimately elucidating the complex mechanisms of diseases and drug effects. The literature retrieve strategy and the metrics of the tools adopted may have a certain impact on the results of this study.

Keywords: causal inference, counterfactual, marginal structural model, Mendelian randomization, bibliometrics

Introduction

In clinical medicine, understanding causal relationships between drugs and health outcomes is essential for evaluating medication efficacy and safety, as well as informing clinical decision-making and public health policy. Traditional correlation analysis can identify relationships between variables; however, it often fails to differentiate true causal relationships and is susceptible to confounding factors.¹ Causal inference has emerged as a vital research topic, particularly in medicine, where it effectively adjusts for confounders and conducts counterfactual analyses to enhance clinical trial design and decision support.² As these methods evolve, causal modeling has potential to further reveal the occurrence mechanism of diseases and provide a more scientific basis for disease prevention and treatment.

For a long time, “Problems involving causal inference have dogged at the heels of statistics since its earliest days”.³ Before robust theories of causal relationships were established, statisticians frequently oscillated between correlation and

causation. This simplifies data analysis but obscures important causal relationships due to the focus on correlation. Unlike correlation, which is symmetric, causation is asymmetric and offers a clearer depiction of event relationships.⁴ Causal inference quantifies causal effects using observational data, enabling conclusions about causal relationships based on the conditions of effect occurrence.⁵ Initially, the integration of causal inference with various disciplines has led to numerous inferential methods, such as matching-based approaches,⁶ tree-based techniques,⁷ and dynamic perturbation analysis.⁸ As causal relationships have evolved, the potential outcomes model⁹ and the structured causal model¹⁰ have gained prominence due to their robust theories, broader applications, and higher recognition in the field of causal inference.

Understanding causal relationships between events allows researchers to predict risks and formulate risk management plans. While Randomized Controlled Trials (RCT) is considered the gold standard for evaluating treatment effects, their exclusion of certain patients can limit the applicability of results to real-world populations.¹¹ Causal inference leverages observational data to help physicians assess treatment effects in broader contexts and apply these findings in clinical practice.¹² Currently, research in causal inference and counterfactual prediction within machine learning is a prominent direction in healthcare.² By simulating clinical trials with observational data, the effectiveness of treatment allocation schemes can be evaluated for optimality, aiding precise drug management.¹³ The “Thalidomide Disaster” notably spurred reforms in drug approval and management systems, illustrating the importance of understanding causality for drug repurposing and developing new chemical entities.¹⁴ Given the complex relationships among drugs, diseases, and individual characteristics, graph-based causal inference methods offer advantages in modeling these intricate connections.¹⁵ Furthermore, causal inference methods can identify and control selection and confounding biases, improving research accuracy.¹⁶ As technology advances, causal inference now extends beyond text data to include images and audio.¹⁷ While the potential for causal inference continues to grow, challenges remain, including evaluating the validity of inference results and addressing Simpson’s Paradox.¹⁸

A search for “Causal inference” in the Web of Science yields over 10,000 results across various disciplines, including medicine, biology, and economics. This extensive body of literature can overwhelm researchers, making it challenging to identify research hotspots. Subjective judgments are often constrained by personal experiences and existing literature. Consequently, bibliometric reviews—grounded in quantitative changes in academic outputs—provide a more objective and comprehensive understanding of the field, encompassing historical overviews, research hotspots, and development trends. CiteSpace, a Java-based application, employs econometrics, co-occurrence analysis, and clustering analysis to visualize trends in scientific literature within a specific discipline.^{19,20}

In this study, bibliometric analysis was performed using the CiteSpace tool to screen articles from the Web of Science database. Further cluster analysis of keywords and co-cited references was conducted based on the network analysis of authors, institutions, and countries. The primary objectives are: (a) to explore the distribution and collaboration among authors, institutions, and countries in causal inference research within medicine; (b) to identify major research themes and the knowledge structure of causal inference in medicine; (c) to pinpoint research hotspots and systematically summarize and forecast the future of causal inference in the medical field.

Methods

Data Sources

Data for the bibliometric analysis were obtained from Clarivate Analytics’ Web of Science Core Collection, including SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, and IC.

Search Strategies

In this study, we retrieved literature from the Web of Science Core Collection from its inception to December 2024, for the following terms: causal inference, drug, epidemiology, pharmacology, prevalence, therapy, diagnosis, health, disease, clinical and medicines. We used “causal inference” as the main keyword and combined it with other terms for subject searches. The specific search expressions were: (TS=(“causal inference”) AND TS=(drug)), (TS=(“causal inference”) AND TS=(epidemiology)), (TS=(“causal inference”) AND TS=(therapy)), (TS=(“causal inference”) AND TS=

(diagnosis)), (TS=("causal inference") AND TS=(health)), (TS=("causal inference") AND TS=(disease)), (TS=("causal inference") AND TS=(clinical)), (TS=("causal inference") AND TS=(pharmacology)), (TS=("causal inference") AND TS=(prevalence)), and (TS=("causal inference") AND TS=(medicines)). The results of the respective searches were then all imported into a unified data tagging result set, totaling 5,856 documents, for the next screening step.

Inclusion and Exclusion Process

This article conducts a literature selection and systematic analysis of research related to digital feedback, following the PRISMA 2020 systematic review process (Figure 1). Initially, only peer-reviewed, causally driven original articles published in the medical field were chosen, excluding review articles. Only articles of a specific type were included, and the language was restricted to English. The titles and abstracts were further examined to exclude unrelated disciplines, research content, erratum documents, unpublished works, and duplicate articles. Finally, the reference lists were manually checked to identify sources that were missed during the search. This process ensured that the selected literature closely relates to causal inference in the medical field, encompassing the years from 1986 to 2024, resulting in a total of 4,316 documents that will be used for CiteSpace analysis.

Bibliometrics and Visual Analytics

The study exported the retrieved articles in plain text format with complete records and references, named "download_XXX.txt", and then imported them into CiteSpace 6.4.R4 for further bibliometric and visualization analysis. When drawing the visual knowledge graph, we followed the main procedural steps of CiteSpace, including time slicing, threshold segmentation, modeling, pruning, merging, and mapping.²¹ The core concepts of CiteSpace include burst detection, betweenness centrality, and heterogeneous networks, which help to visualize the current state, hotspots, and frontiers of research in a timely manner. Nodes in the different maps represent authors, institutions, countries, or keywords. Node size indicates the frequency of occurrence or citation, and node color indicates the frequency of occurrence or year of citation. CiteSpace provides two indicators based on network structure and clustering clarity: the module value (Q value) and the average silhouette value (S value).^{22,23} These can serve as benchmarks for assessing the quality of the generated maps. Generally, the Q value falls within the interval [0, 1). A Q value greater than 0.3 indicates a significant community structure; when the S value is at 0.7, the clustering is considered highly efficient and convincing,

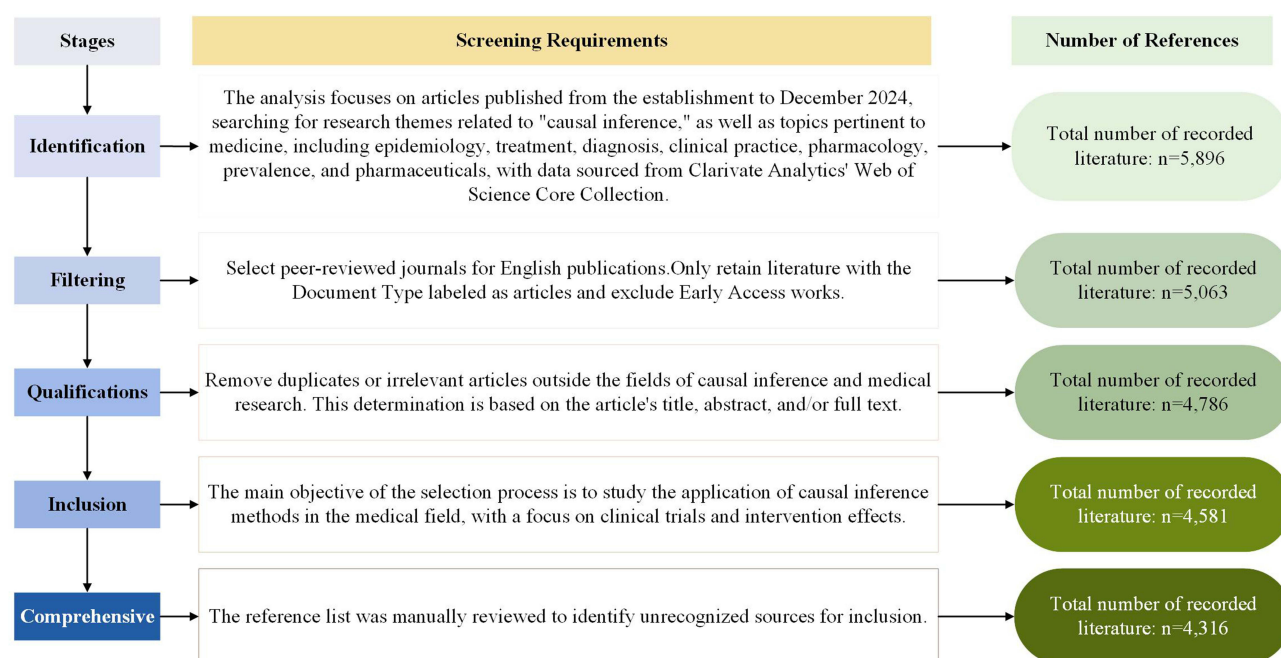


Figure 1 PRISMA literature search flowchart.

while a value above 0.5 is generally deemed reasonable. If the S value approaches infinity, the number of clusters usually is one, suggesting that the selected network may be too small and only represents a single research theme. Creating knowledge graphs requires selecting different thresholds and generating multiple drawings, ultimately choosing the most ideal graph based on the Q and S values as the final result.

Results

Using Citespace, we will analyze the countries, institutions, authors, journals, and publications involved in causal inference research in the field of medicine.

Bibliometric Analysis of Publication year

Literature related to the application of causal inference in the field of medicine has shown a tendency to grow with time fluctuations (Figure 2). The earliest document in the Web of Science Core Collection is James Robins’ “A new approach to causal inference in mortality studies with a sustained exposure period - application to control of the healthy worker survivor effect”. (Published in 1986) This work has inspired subsequent research by scholars in the field of causal inference, further exploring the relationship between mortality rates and treatment regimens, thereby providing an important theoretical foundation and practical framework for future causal inference studies. The citation of this research has exceeded 1,500 times, establishing it as a foundational study in the application of causal inference to epidemiology. Due to its pioneering value, we will consider the year of its publication as the starting year for our investigation. However, during its early stages, causal inference was only regarded as a method of data analysis in statistics and did not form a systematic disciplinary structure. As a result, few scholars specialized in this field, and many important theories were primarily based on statistical research. For example, Jerzy Neyman, who proposed the “potential outcomes” model, made outstanding contributions to hypothesis testing, and his research played a fundamental role in the theory of causal inference. Therefore, for a long time, very few applied studies were associated with causal inference, and even fewer were specifically known in the field of medicine.

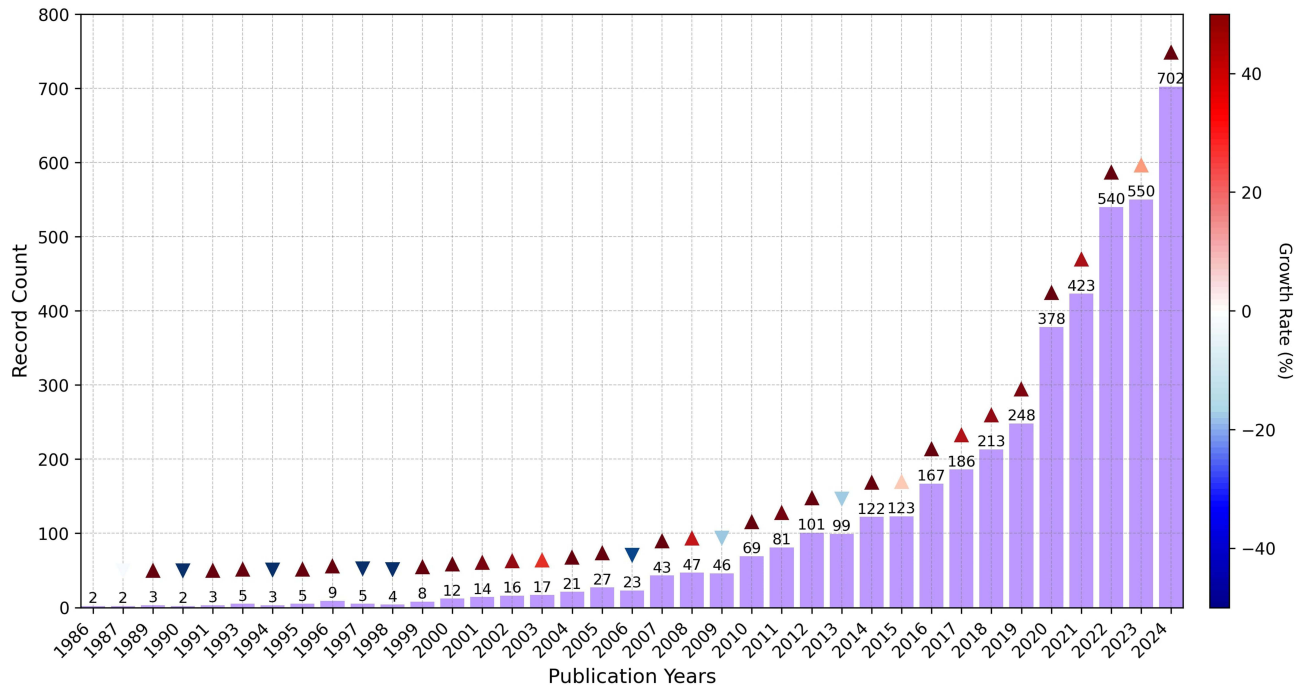


Figure 2 Annual trend map of publications.

This study employs bibliometric methods to analyze the growth patterns of causal inference in the medical field from 1986 to 2024. From 1986 to 2000, related research grew slowly; it saw rapid expansion from 2000 to 2018, followed by an explosive growth post-2019.

Around the year 2000, the emergence of computer science provided enhanced technological support for processing large datasets and conducting causal inference. In this context, Robins et al published “Marginal Structural Models and Causal Inference in Epidemiology” which advanced causal modeling by improving the treatment of confounding factors, thereby facilitating the prospective application of causal inference in epidemiology.²⁴ Concurrently, Hernández-Díaz and others²⁵ contributed significantly to the application of causal inference in medicine through studies like the case-control method. Subsequently, related research rapidly proliferated.

In 2019, the outbreak of the COVID-19 pandemic and the increasing public focus on public health led to evolving clinical features and management of COVID-19 patients, necessitating changes in prognostic assessment methods.²⁶ The influences on COVID-19 are often complex and interrelated, making it essential to clarify these intricate relationships and unveil potential causal mechanisms to adapt public health responses to the dynamically changing pandemic situation.²⁷ Research combining causal inference with epidemiology demonstrated its effectiveness in addressing such large-scale medical events, leading topics like vaccine safety, public management, and mental health to gain prominence in subsequent years.^{28,29}

By 2021, generative artificial intelligence began to showcase exceptional capabilities in data processing and causal inference. Concepts such as smart healthcare, digital health, and personal physicians continued to gain traction, driving an explosive increase in research on causal inference in medicine. Through the development and evaluation processes grounded in causal inference, AI has the potential to achieve better clinical outcomes and enhance overall healthcare quality.³⁰ Furthermore, combining causal inference methods with advanced representation learning techniques allows for more effective handling of large-scale and high-dimensional heterogeneous data, thereby improving model interpretability.³¹

Bibliometric Analysis of Institutions

CiteSpace software determines collaboration primarily based on the co-occurrence frequency matrix. The size of the nodes represents the number of papers published by each institution, and the thickness of the connecting lines between the nodes indicates the degree of collaboration between institutions. The top three institutions contributing the most articles are Harvard University (203), Harvard T.H. Chan School of Public Health (193) and University of Pennsylvania (179). Among these, Harvard University has a centrality greater than 0.1, indicating significant contributions to the field. Overall, there is a very close connection between institutions, especially Harvard University and the University of Bristol, which not only dominate in terms of the number of published papers but have also established close cooperation with other institutions. (Figure 3) However, it is also evident that some institutions have less cooperation or are even isolated, indicating that the sense of collaboration among these institutions needs to be further strengthened.

Figure 4 illustrates the clustering of collaboration areas among institutions, with a Q value of 0.5244 and an S value of 0.8513, indicating that the clustering results are convincing. Harvard University is closely collaborating with the University of Pennsylvania on HIV research, while in the fields of child health and targeted trials, Harvard has also shown active cooperation with institutions like Columbia University. Additionally, the University of Copenhagen and the University of Bristol have formed a strong partnership in real-world evidence and Mendelian randomization studies. The French National Institute of Health and Medical Research has also established significant collaborations with the Massachusetts Institute of Technology in the area of gut microbiome research.

Bibliometric Analysis of Countries

Figure 5 presents the country-specific results of research on causal inference in the applied domain. The size of each circle represents the number of papers published in that country. The top three countries contributing the most articles are the United States (2,452), the People's Republic of China (739) and England (676). The England has a centrality greater than 0.1, indicating its significant contribution to the field. Additionally, although Sweden (182 articles) and South Africa (49 articles) contributed fewer articles, their centrality values also exceeded 0.1, with Sweden at 0.15 and South Africa at 0.24. North

CiteSpace v. 5.4.R1 (64-bit) Advanced
 March 17, 2025, 7:55:52 PM CST
 WoS: E:\OneDrive - njust.edu.cn\美国\数据\2025-3-17\data
 Timespan: 1999-2025 (Slice Length=1)
 Selection Criteria: g-index (k=25), LRF=2.5, L/N=10, LBY=5, e=1.0
 Network: N=172, E=1923 (Density=0.9985)
 Largest CCs: 548 (81%)
 Nodes Labeled: 1.0%
 Pruning: Pathfinder
 Excluded:

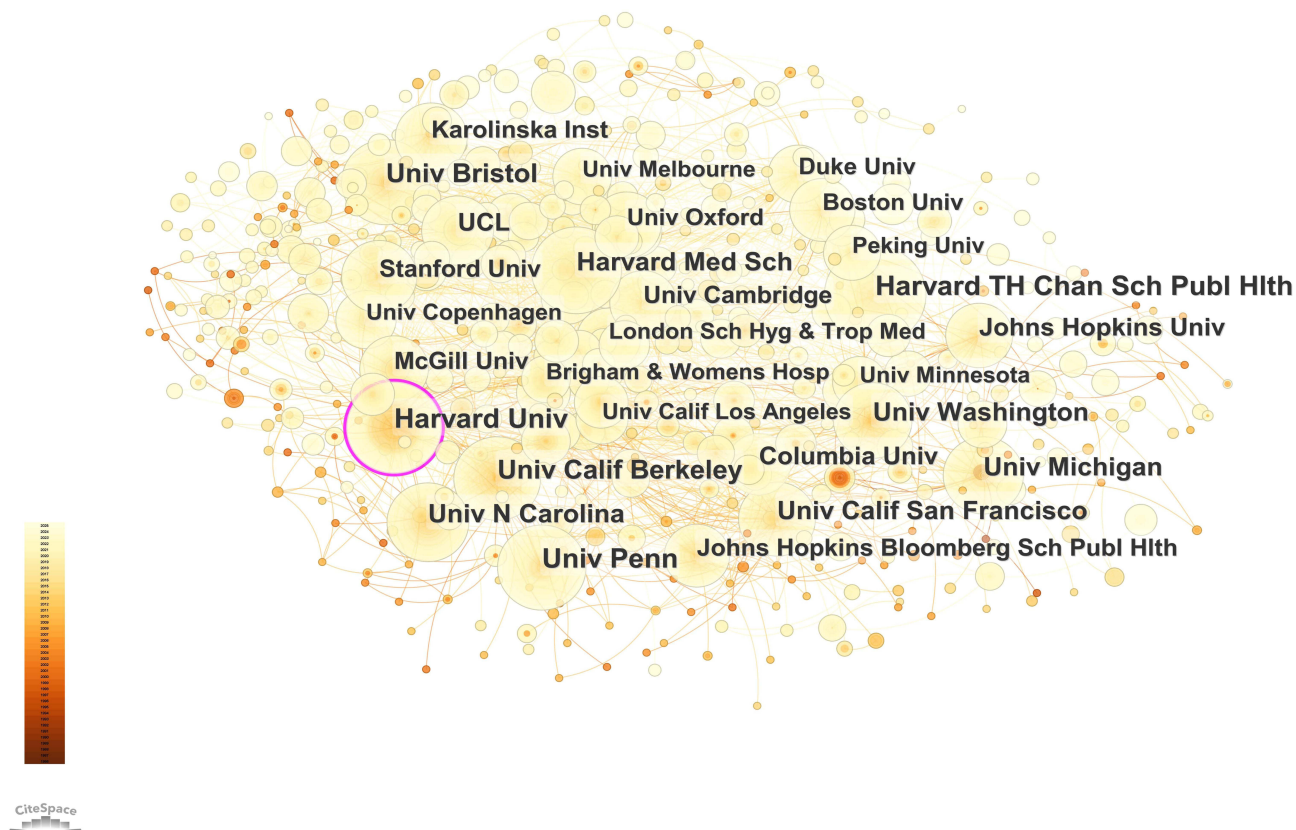


Figure 3 Network of cooperating institutions.

America and Europe dominate, while Asia is rapidly growing, and Africa is also showing its unique contributions to research. [Figure 5](#) also clearly shows that the connectivity between the nodes is robust and the cooperation between individual countries is active.

Bibliometric Analysis of Authors

The co-authorship network is illustrated in [Figure 6](#), where the size of the circles represents the number of studies published by each author. The shorter the distance between two circles, the more collaboration exists between the respective authors. The colors of the circles denote authors within the same cluster. Gray nodes represent earlier published studies, while red nodes indicate more recently published studies. In [Figure 6](#), it is evident that many authors tend to work with a relatively stable set of collaborators, forming several major clusters of authors. Each cluster typically contains two or more core authors. Additionally, prolific authors generally maintain stable relationships with other authors. As seen in [Figure 6](#), the most represented author is Smith, George Davey with 54 publications, followed by Cole, Stephen R (48), and Kaj Hernan, Miguel A (34). The research of the three individuals focuses on public environmental occupational health, with varying degrees of involvement in genetics, heredity, statistics, probability, and medical informatics. This analysis clearly visualizes the collaboration between authors and aids in tracking research questions and identifying cutting-edge research.

CiteSpace v. 5.4.R1 (64-bit) Advanced
 March 29, 2025, 4:05:15 PM CST
 Voci: 5, timeSlice: 10, stepSize: 10, weight: 1, pruning: Pathfinder
 Timespan: 1986-2024 (Since Largest)
 Selection Criteria: g-index=0.95, L=0.2, Q=0.1, M=0.1, LBV=0, w=1.0
 Network: N=657, E=1901 (Density=0.0086)
 Nodes Labeled: 126
 Pruning: Pathfinder
 Modularity Q=0.9244
 Weighted Mean Silhouette S=0.8513
 Harmonic Mean(Q, S)=0.8849
 Executed:

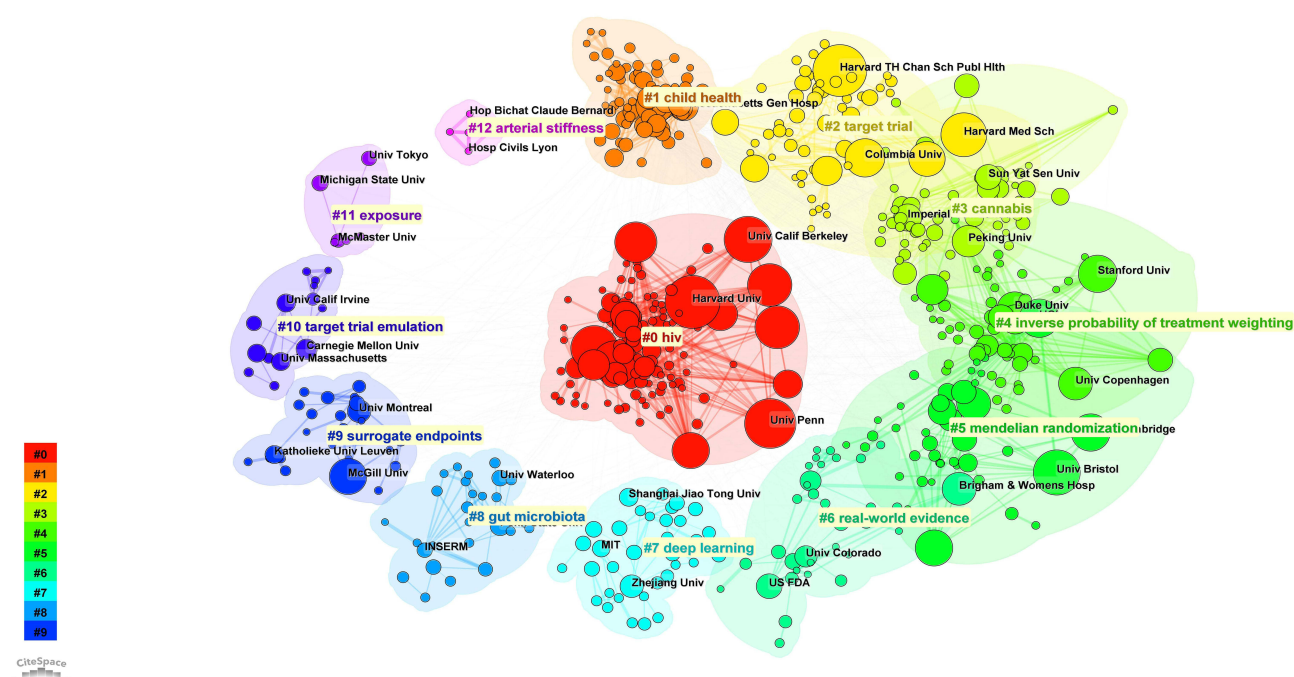


Figure 4 Clustering of collaborative areas among institutions.

Bibliometric Analysis of Author and Journal Co-Citation

Figures 7 and 8 show the co-citation network of authors and journals. The most cited authors were Robins JM (1,148 citations) and Rubin DB (1,100 citations), followed by Hernán MA (977 citations) and Rosenbaum PR (778 citations). Robins JM's significant contributions include marginal structural models,²⁴ selection bias,²⁵ and simulated target trials.³² These methods have laid the groundwork for the advancement of practical clinical research, providing tools for understanding and handling complex data, promoting theoretical progress in causal inference, and improving evaluations of treatment effects and policy impacts.

The top-cited journal was the American Journal of Epidemiology with 1,984 citations, followed by Statistics in Medicine (1,983 citations) and Epidemiology (1,980 citations). Additionally, the American Journal of Epidemiology received both high frequency and high centrality, indicating its key role in the field. Research published in the American Journal of Epidemiology focuses on the application of causal inference in epidemiology, including techniques such as inverse probability weighting and multivariable Mendelian randomization, which are crucial for evaluating treatment effects and understanding causal relationships. The exploration of genetic epidemiology in these studies also demonstrates how genetic information can be utilized to comprehend the causal relationships of diseases, contributing to the advancement of personalized medicine.

Bibliometric Analysis of Co-Occurring Keywords and Clustering

Keywords are high-level summaries. High-frequency, high-centrality keywords tend to reflect hot research topics in the field. We analyzed publications in 1-year time slices and identified the top 35 most-cited or most frequently occurring keywords in each slice. The co-occurring keyword network consisted of 986 nodes and 4,470 links.

CiteSpace v. 5.4.R1 (64-bit) Advanced
 March 17, 2025, 3:27:36 PM CST
 WoS: E:\OneDrive - njupt.edu.cn\桌面\2025-3-17\data
 Timespan: 1986-2025 (Slice Length=1)
 Selection Criteria: g-index (k=25), LRF=2.5, L/N=10, LBY=5, e=1.0
 Network: N=108, E=557 (Density=0.9964)
 Largest CCs: 98 (90%)
 Nodes Labeled: 1.0%
 Pruning: Pathfinder
 Excluded:

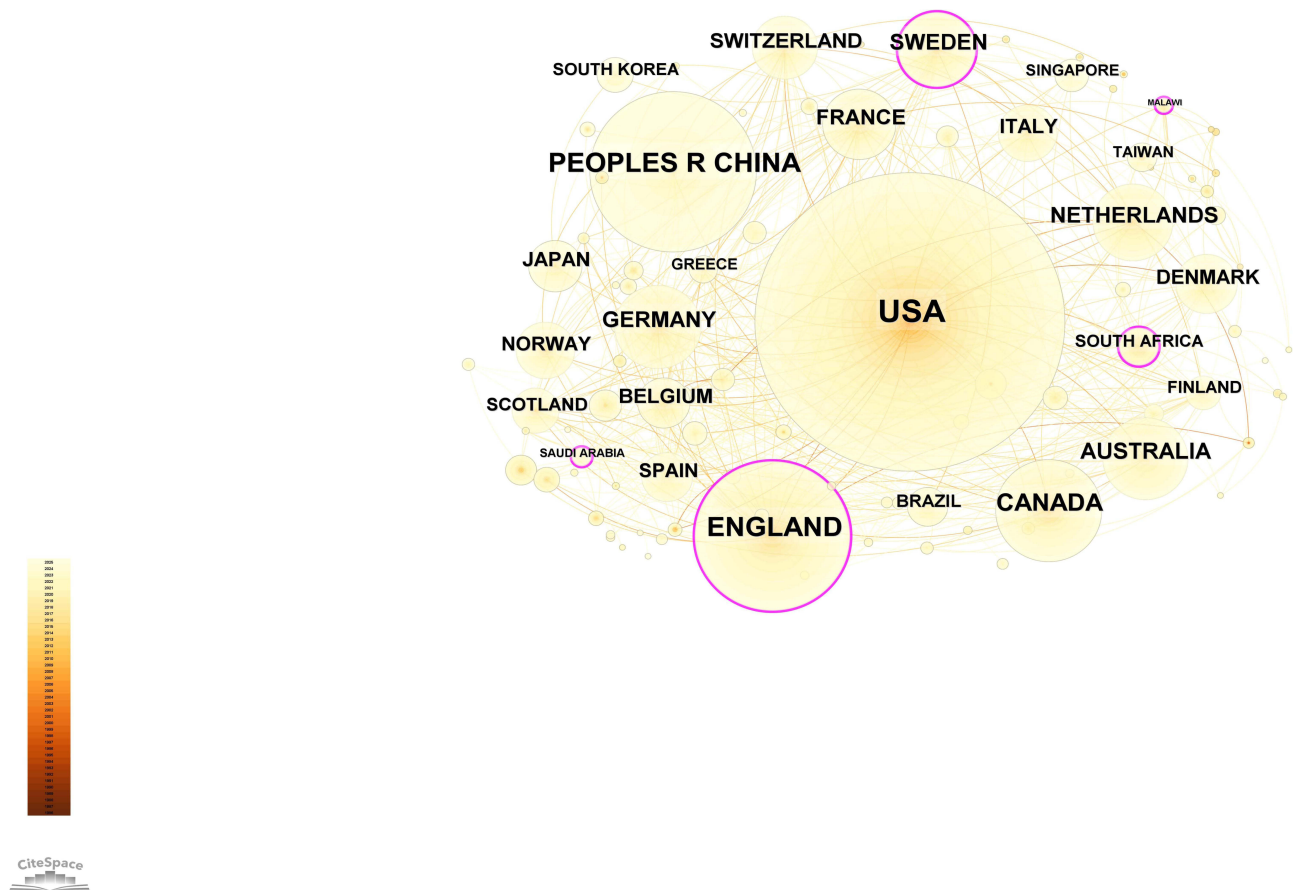


Figure 5 Networks in cooperating countries.

As shown in [Table 1](#), the three keywords with the highest centrality are: epidemiology (0.52), coronary heart disease (0.14), health (0.14) and model (0.13). The keywords with the highest frequency are: causal inference (2831 times), risk (559 times), and health (336 times). Additionally, [Figure 9](#) illustrates the keyword network. Epidemiology and coronary heart disease emerge as the most central keywords, highlighting the importance of cardiovascular diseases in public health management. Furthermore, health-related keywords such as mortality, air pollution, and Mendelian randomization underscore the crucial role of causal inference in disease research, particularly in controlling for potential biases. The most frequently mentioned keywords—including causal inference, risk, and propensity scores—indicate a growing emphasis in modern research on evidence-based decision-making, which is increasingly vital for understanding complex health issues and implementing effective intervention strategies.

Ten clusters were obtained by CiteSpace. The clustering modularity value (Modularity Q) was 0.4581, which is greater than 0.3, indicating a significant clustering structure. Additionally, the silhouette value (Silhouette) of each cluster was above 0.7, demonstrating that the results were credible and significant ([Figure 10](#) and [Table 2](#)).

CiteSpace, v. 5.4.R1 (64-bit) Advanced
 March 17, 2025, 8:42:46 PM CST
 WoS: E:\OneDrive - njupt.edu.cn\数据\图论\2025-3-17\data
 Timespan: 1989-2025 (Slice Length=1)
 Selection Criteria: g-index (k=25), LRF=2.5, LN=10, LBY=5, w=1.0
 Network: N=210, E=1232 (Density=0.0029)
 Largest CCs: 231 (24%)
 Nodes Labeled: 1.0%
 Pruning: Pathfinder
 Excluded:

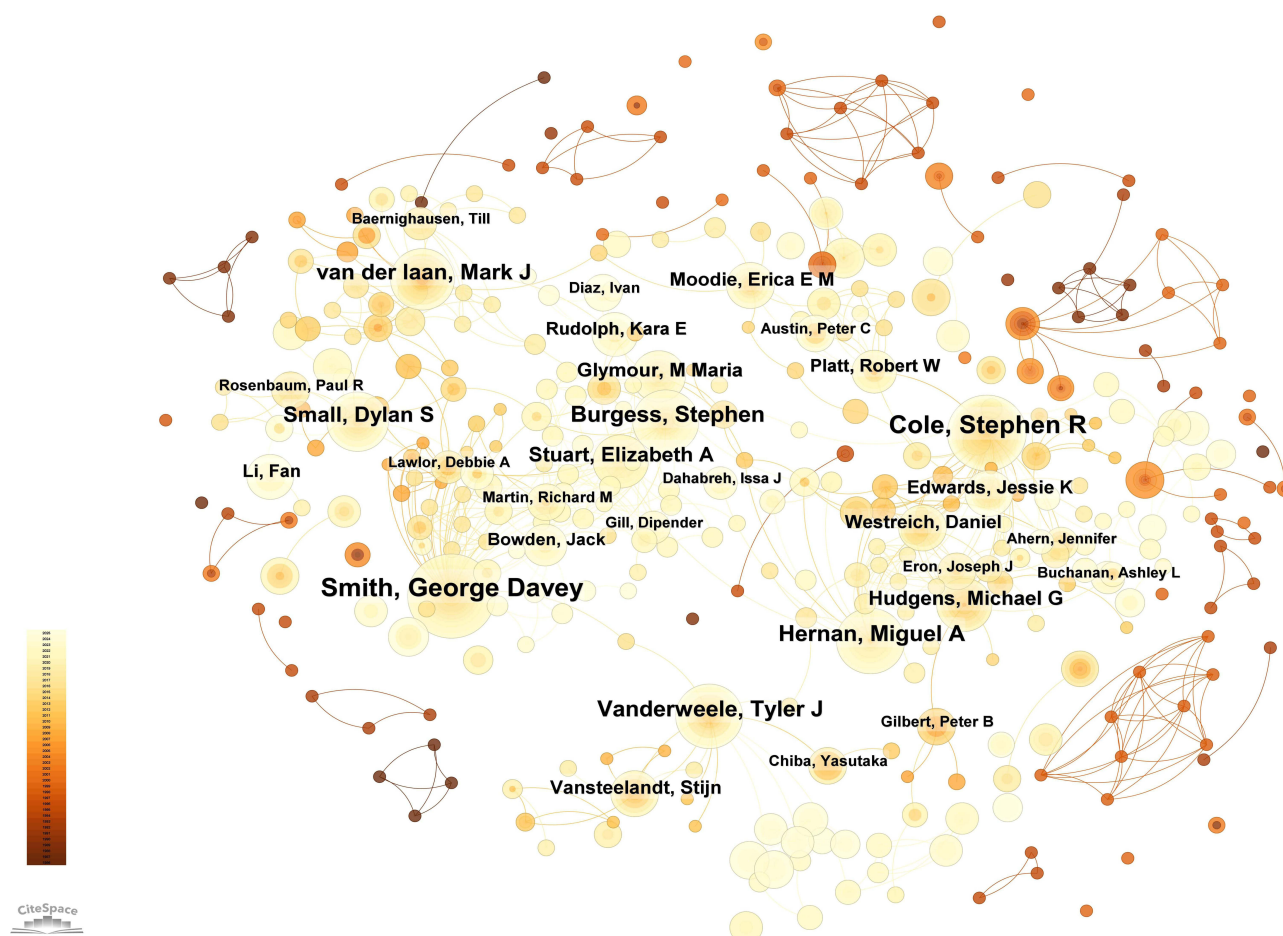
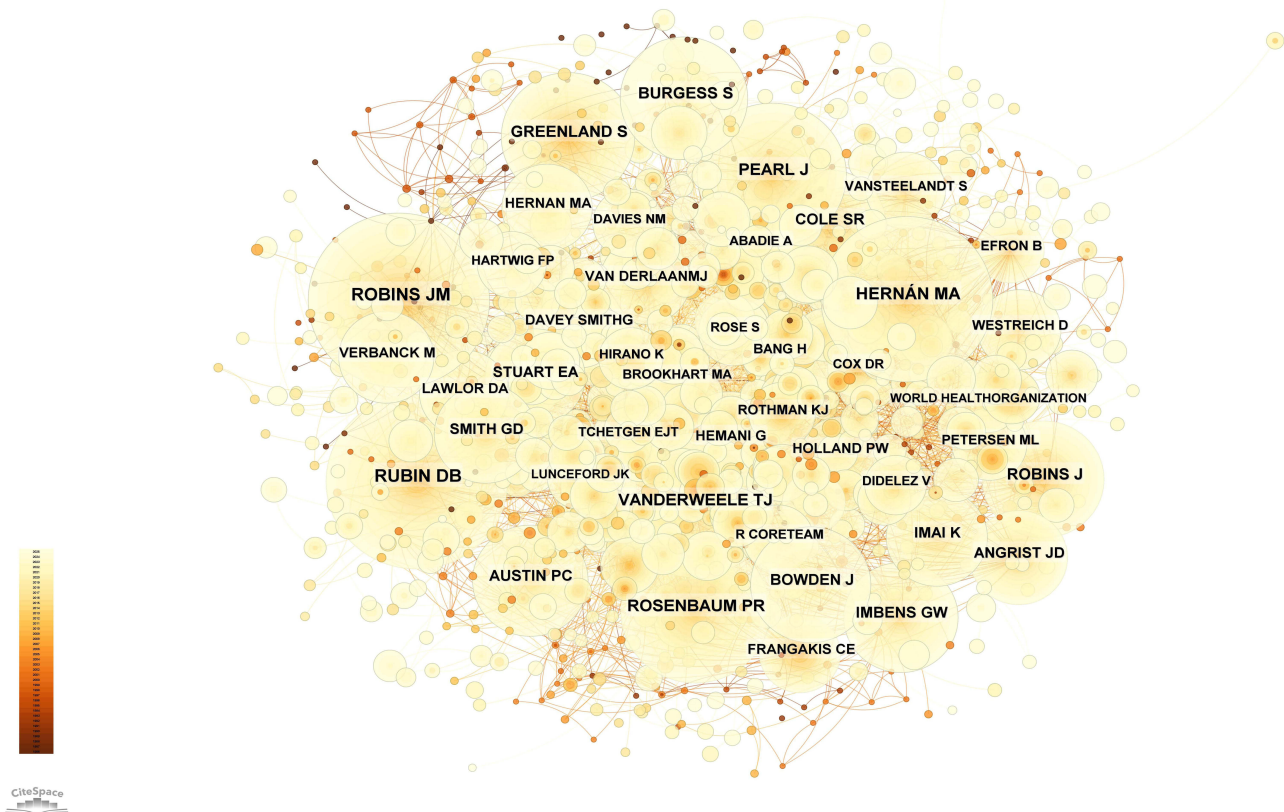


Figure 6 Network of collaborating authors.

Keywords with Citation Explosion

Figure 11 shows the top 40 citation outbreak keywords. The deep green line indicates the time interval, and the red line indicates the time of the keyword outbreak.²⁰ The keyword “marginal structural model” had the strongest citation explosion in 2007, which lasted until 2016, indicating the importance of the marginal structural model in causal inference. Although the marginal structural model was proposed early by Robins JM and has since been utilized in epidemiology,²⁴ its broader adoption and application have taken time to percolate. To date, the marginal structural model has evolved into a trusted approach for causal inference. The latest keywords to see a spike in citations in 2022 include “instruments”, “anxiety” and “prediction”.

Moreover, the high intensity of “noncompliance” and “principal stratification” underscores the importance of these concepts in research, suggesting they may be key elements in understanding complex causal relationships and controlling biases. “Epidemiologic methods” have persisted from 1995 to 2017, reflecting the enduring relevance and significance of epidemiological approaches in related fields. Both “epidemiologic methods” and “randomized trials” exhibit strong citation surges, possibly indicating an increasing prevalence of randomized trial methodologies in epidemiological studies, which reflects a shift in scientific research practices and a greater emphasis on evidence-based medicine. Conversely, certain keywords such as “big data”, “DNA methylation”, “robust estimation”, and “prediction” have relatively short presence durations, potentially indicating that these methods or topics are emerging frontiers in recent research, illustrating researchers’ rapid response to new challenges.



Bibliometric Analysis of Co-Cited References

2612 <https://doi.org/10.2147/JMDH.S516826>

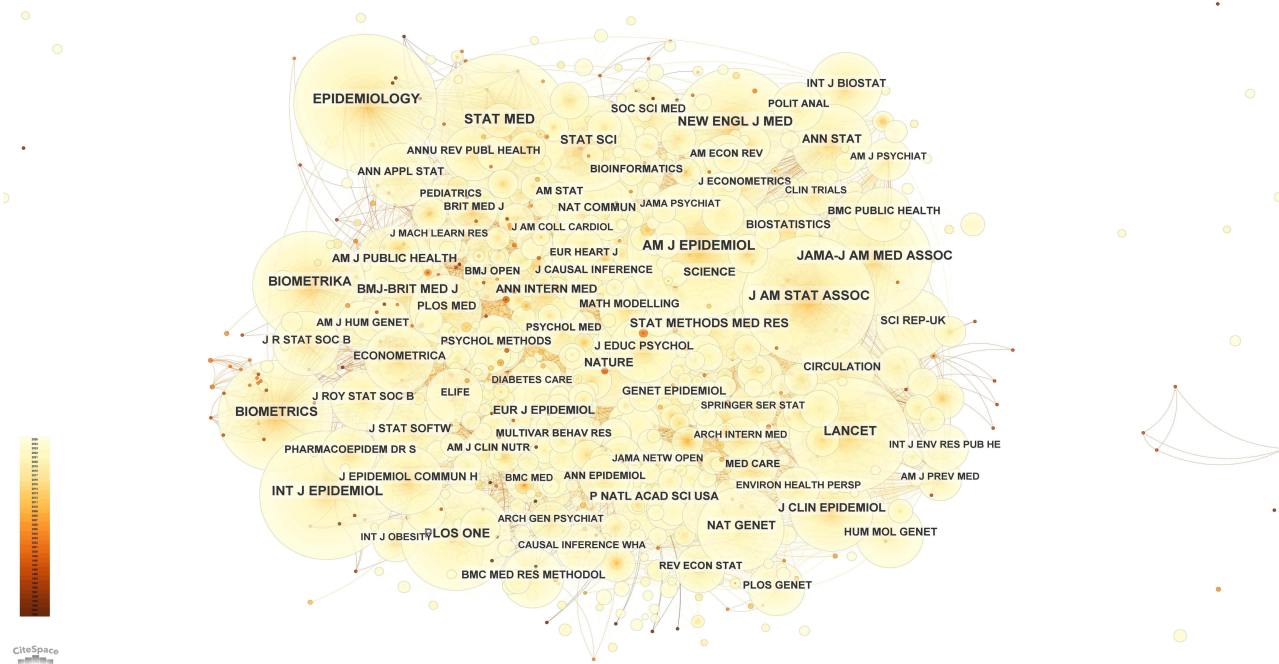


Figure 8 Journal co-citation network.

Discussion

Summary of Findings

We reviewed the latest research status of medical causal inference using CiteSpace, revealing relevant research hotspots and frontiers. From January 1986 to December 2024, we obtained 4,316 documents related to medical causal inference from the Web of Science Core Collection, showing a growth trend that fluctuates over time. The United States ranks first in related research, contributing 2,452 publications, which accounts for over 50% of the total, dominating the field of medical causal inference. In contrast, while countries like Canada and the People's Republic of China also contributed a significant number of documents, their impact remains low, indicating the need for further exploration of the value of their research projects. Harvard University is the leading institution with the highest number of publications (203), while George David Smith is the most prolific author (54 publications). Robbins JM is the most cited scholar with 1,148

Table 1 Top 10 Keywords in Terms of Frequency and Centrality

Rank	Count	Keywords	Centrality	Keywords
1	2,831	Causal inference	0.09	Epidemiology
2	559	Risk	0.08	Coronary heart disease
3	447	Mendelian randomization	0.08	Health
4	389	Propensity score	0.07	Inference
5	361	Mortality	0.07	Models
6	336	Health	0.06	Mortality
7	332	Association	0.06	Air pollution
8	331	Models	0.05	Mendelian randomization
9	294	Inference	0.05	Bias
10	292	Bias	0.04	Breast cancer

CiteSpace v. 5.4.R1 (64-bit) Advanced
 March 10, 2025, 1:08:35 PM CST
 View: 0.200/0.100 - Weighted Mean Q=0.9925, S=0.9746
 Timespan: 1995-2025 (Slice Length=1)
 Network: N=886, E=4470 (Density=0.0092)
 Largest CC=1 (Size=16.95%)
 Modularity Q=0.976
 Pruning: Pathfinder
 Excluded:

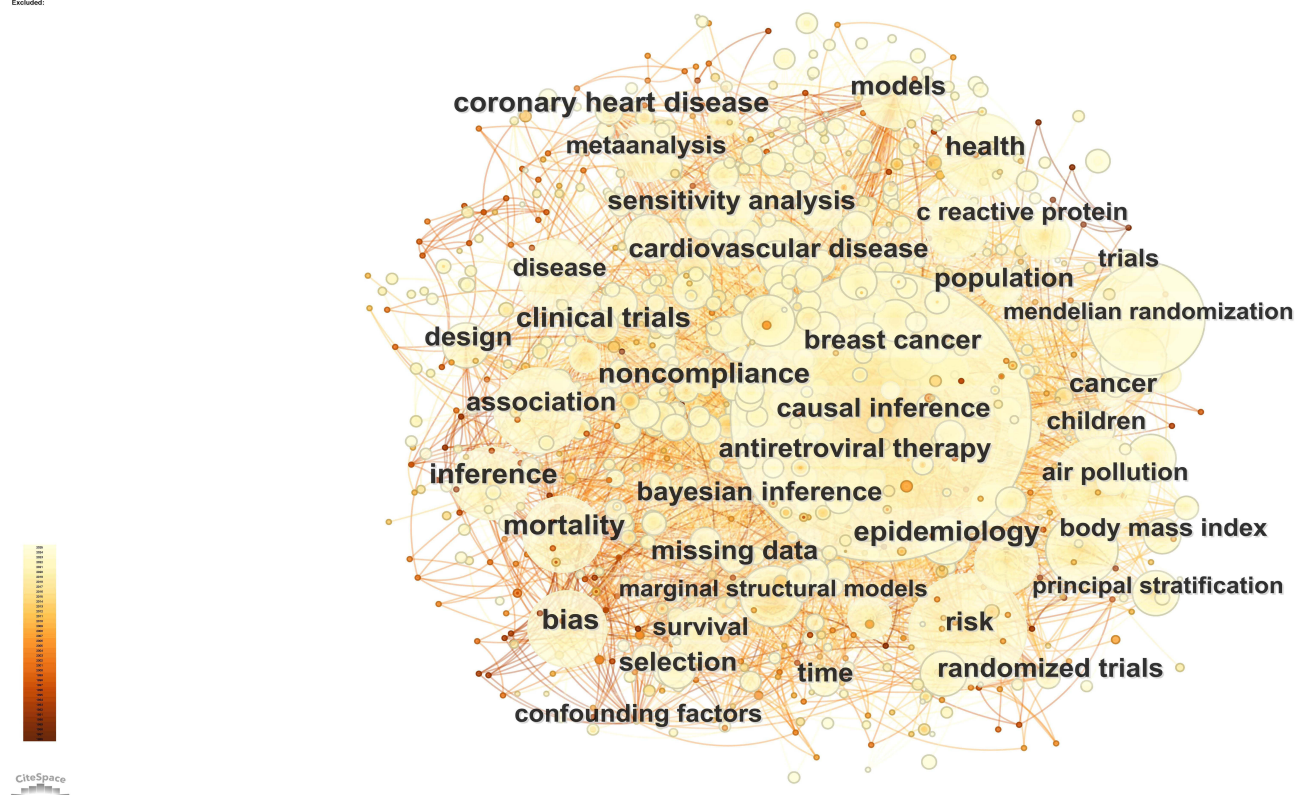


Figure 9 Keyword co-occurrence network.

citations for relevant papers, and the American Journal of Epidemiology is the most cited journal, having received 1,984 citations. The most frequently cited authors and references typically focus on epidemiology and Mendelian randomization. The top five keywords ranked by frequency and centrality are epidemiology, coronary heart disease, health, inference, and models. The keyword “marginal structural models” ranks first with the highest citation burst rate. Additionally, “DNA methylation”, “instruments”, “anxiety”, and “robust estimation” are emerging research hotspots.

Causal Inference on Hot Issues in Medicine

Keywords represent a high degree of generalization and conciseness of the topic, and commonly used keywords are often used to identify research hotspots during the analysis process. The results of the co-occurring keywords and cluster analysis indicate that the major current research trends include counterfactual prediction, Mendelian randomization, and marginal structural models.

Counterfactual Forecast

In causal inference, counterfactuals attempt to answer the question, “What would have happened if things had been different?” In medical research, the counterfactual question for a patient who receives treatment is what would have happened to the patient’s health in the absence of the treatment. Causal inference is the process of estimating the causal effect of a particular treatment or intervention by comparing the actual situation with a counterfactual scenario. According to Prosperi et al, counterfactual prediction is important in actionable healthcare.² Machine learning and causal inference models enable researchers to assess the potential outcomes of different healthcare strategies and provide personalized treatment plans for patients. This approach emphasizes the crucial role of counterfactual prediction in optimizing therapeutic decision-making and improving the quality of patient care.

CiteSpace, v. 5.4.R1 (64-bit) Advanced
 March 19, 2025, 4:00:46 PM CST
 VOS: E-Undirected - weighted (g-index=0.0005-1.0)
 Timespan: 1986-2025 (Slice Length=1)
 Selection Criteria: g-index (q=0.95, LRF=1.0, LBY=5, v=1.0)
 Network: N=946, E=4470 (Density=0.0092)
 Largest CC: 554 (59%)
 Nodes Labeled: 1.0%
 Pruning: Pathfinder
 Modularity Q=0.4381
 Weighted Mean Silhouette S=0.7256
 Harmonic Mean(Q, S)=0.5819
 Excluded:

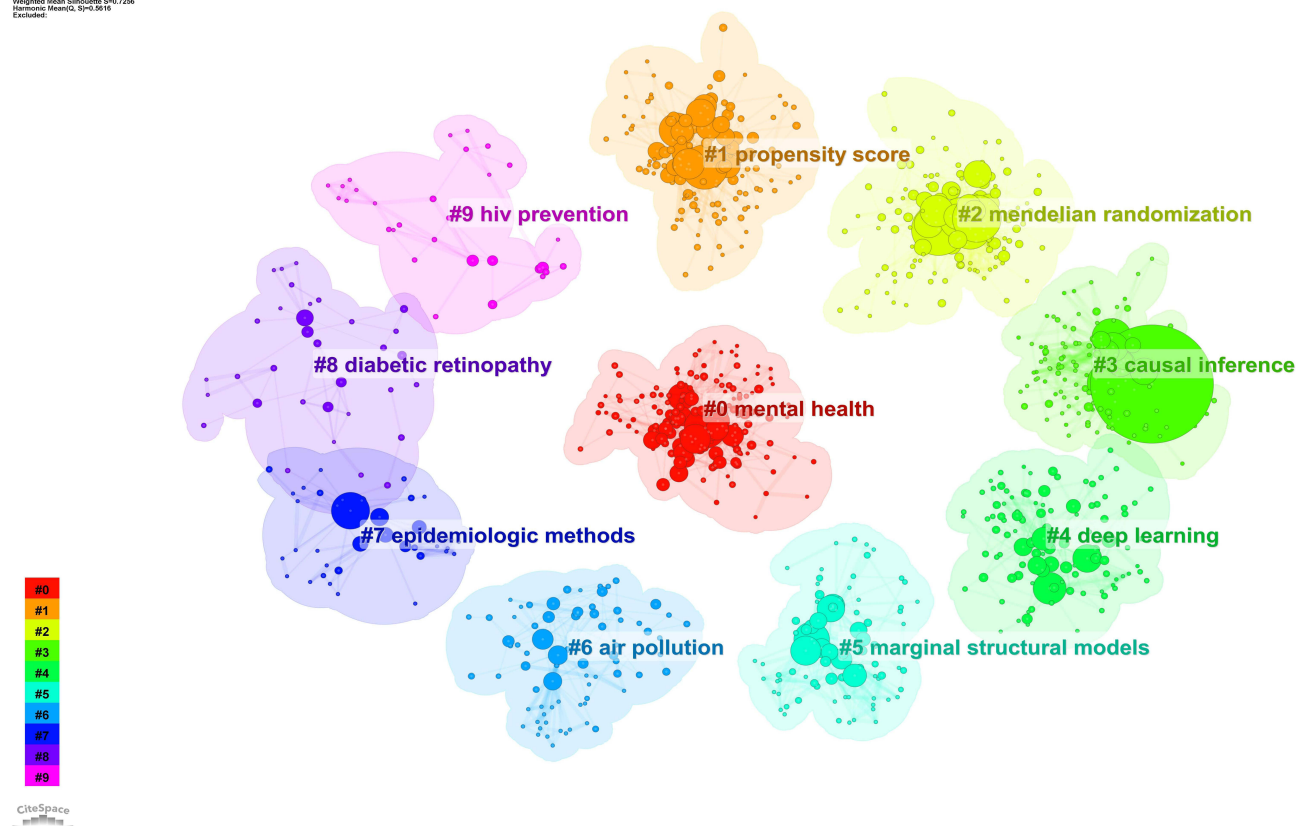


Figure 10 Keyword co-occurrence clustering network.

Electronic Health Records

Researchers often prefer to use a data source known as electronic health records (EHRs), which serves as a systematized, electronic version of a patient's health information. EHRs contain detailed information about a patient's medical history,

Table 2 Cluster Analysis of Keywords

Cluster ID	Silhouette	Size	Label (LLR)	Included Keywords (top 5)	Mean (Year)
0	0.721	152	Mental health	Mental health; depression; physical activity; anxiety; depressive symptoms	2012
1	0.703	137	Propensity score	Propensity score; Mendelian randomization; machine learning; observational studies; inverse probability weighting	2012
2	0.758	137	Mendelian randomization	MENDELIAN randomization; body mass index; propensity score; genome-wide association study; obesity	2012
3	0.79		Causal inference	Causal inference; principal stratification; missing data; noncompliance; sensitivity analysis	2004
4	0.721	107	Deep learning	Deep learning; gut microbiota; artificial intelligence; risk factors; regression discontinuity	2014
5	0.756	83	Marginal structural models	Marginal structural models; Mendelian randomization; hiv; longitudinal data; antiretroviral therapy	2007
6	0.764	74	Air pollution	Air pollution; particulate matter; chronic obstructive pulmonary disease; mediation analysis; lung cancer	2010

(Continued)

Table 2 (Continued).

Cluster ID	Silhouette	Size	Label (LLR)	Included Keywords (top 5)	Mean (Year)
7	0.894	38	Epidemiologic methods	Epidemiologic methods; instrumental variable; causal effect; disease; model selection	2008
8	0.861	34	Diabetic retinopathy	Diabetic retinopathy; immune cell; vitamin d; mr analysis; immunity	2014
9	0.929	30	HIV prevention	HIV prevention; definition; Mendelian randomization analysis; causal inference; social determinants	2011

treatment history, laboratory and imaging results, medication information, allergic reactions, hospitalization records, outpatient follow-up information, and personal identity and lifestyle habits.⁴⁰ Researchers have used EHRs to assess their potential impact on patient health outcomes by simulating different intervention or treatment scenarios. Recognized results include confirming that there is no causal relationship between gout and Alzheimer’s disease⁴¹ as well as revealing a relationship between anaerobic exercise and hyperglycemia.⁴² As it has evolved, the counterfactual analysis of EHRs is no longer limited to revealing relationships between specific diseases or health conditions, but also extends to assessing the effectiveness of medical interventions and optimizing treatment regimens.⁴³ While EHRs have received much attention from researchers, the accuracy, completeness, and consistency of the EHRs data itself are critical for effective counterfactual analysis. In recent years, artificial intelligence has made remarkable progress, and AI models that integrate multi-omics data and EHRs have effectively combined medical data from diverse sources, enhancing the potential for precision medicine.⁴⁴ However, further advancements in this field require continuous efforts to overcome challenges related to data integration, privacy protection, and algorithmic trust.

Propensity Score Matching

The propensity score is the probability that an individual will receive a specific intervention (eg, treatment) given the observed covariates.⁴⁵ Propensity score matching in counterfactual prediction mimics the effects of RCT in observational studies by matching individuals with similar propensity scores.⁶ Researchers can also create a balanced dataset between treatment and control groups, where confounders are statistically similar, thus reducing confounding bias. However, this approach requires a large sample size and can only adjust for differences in observed variables. As a result, bias due to unobserved covariates may still exist.⁴⁶ Furthermore, propensity score-based methods extend unnaturally when there are more than two treatment levels.⁴⁷ Shu Yang et al combined weak unbiasedness and generalized propensity scores for counterfactual prediction in the presence of multiple treatment levels,⁴⁸ extending propensity score matching (PSM) to the broader problem of counterfactual prediction. This approach is particularly important for studies involving intervention diversification. PSM is also an important tool for addressing sparse data and high dimensionality problems. Researchers have proposed many improvements for PSM, including post-dual selection methods,⁴⁹ regularization methods,⁵⁰ and multivariate matching methods.⁵¹ Although PSM still faces challenges in terms of data balancing, heterogeneity of treatment effects, and time dependency, combining it with machine learning and big data techniques shows great potential. Through this combination, PSM has found significant applications in enhanced matching techniques,⁵² dynamic propensity score matching,⁵³ and multidimensional and network matching.⁵⁴ It can also be combined with artificial intelligence algorithms to better handle complex data structures and improve the accuracy of propensity score estimation.⁵⁵ Additionally, it can adapt to dynamic and long-term data sources for analyzing the effects of treatments using EHRs,⁵⁶ and develop new methods for time-series data and for dealing with time-varying intervention effects.

Observational Studies

Observational studies hold potential value for investigating medical causal effects that are difficult to answer through randomized trials. With the promotion of causal paradigms, reliance on observational population data for causal inference

Top 40 Keywords with the Strongest Citation Bursts



Figure 11 Top 40 most cited keywords.

may be necessary when suitable target experiments are not available.⁵⁷ Early research indicated that the principle of unconfoundedness supports the basis for causal inference in observational studies; therefore, under this condition, the Average Causal Effect (ACE) can be identified and estimated using observational data.⁵⁸ Based on this principle, employing more robust methods such as conditional exchangeability, multivariable regression analysis, sensitivity analysis, and dynamic tracking and data collection can enable more effective estimation of comparative treatment benefits from observational data.⁵⁹ However, in the process of selecting estimators, observational studies must avoid confounding of outcome variables to ensure scientific rigor and reliability in the analysis. To address this issue, Janie Coulombe et al⁶⁰ proposed a new multiple robust estimator designed for estimating causal effects in observational studies, particularly addressing confounding factors and irregular observation times in electronic health records. Additionally, when describing observational studies, it is essential to use appropriate language and communication methods to ensure accurate reporting of study results. Issa J. Dahabreh and colleagues⁶¹ developed a new causal inference framework that effectively utilizes causal language in observational studies, significantly improving

CiteSpace, v. 5.4.R1 (64-bit) Advanced
 March 18, 2025, 5:31:01 PM CST
 Web: E:\Oncotree - jingyi\citespace\2025-3-17\data
 Diagnostics: 1989-2025 (Date: LongRun=1)
 Selection Criteria: p=0.01, q=0.05, L=1, M=1, N=10, LBY=5, e=1.0
 Network: g=0.9839 (Density=0.9825)
 Largest CCs: 1576 (72%)
 Nodes Labeled: 1.0%
 Pruning: Pathfinder
 Modularity Q=0.4581
 Weighted Mean Silhouette S=0.7256
 Harmonic Mean(Q, S)=0.5918
 Executed:

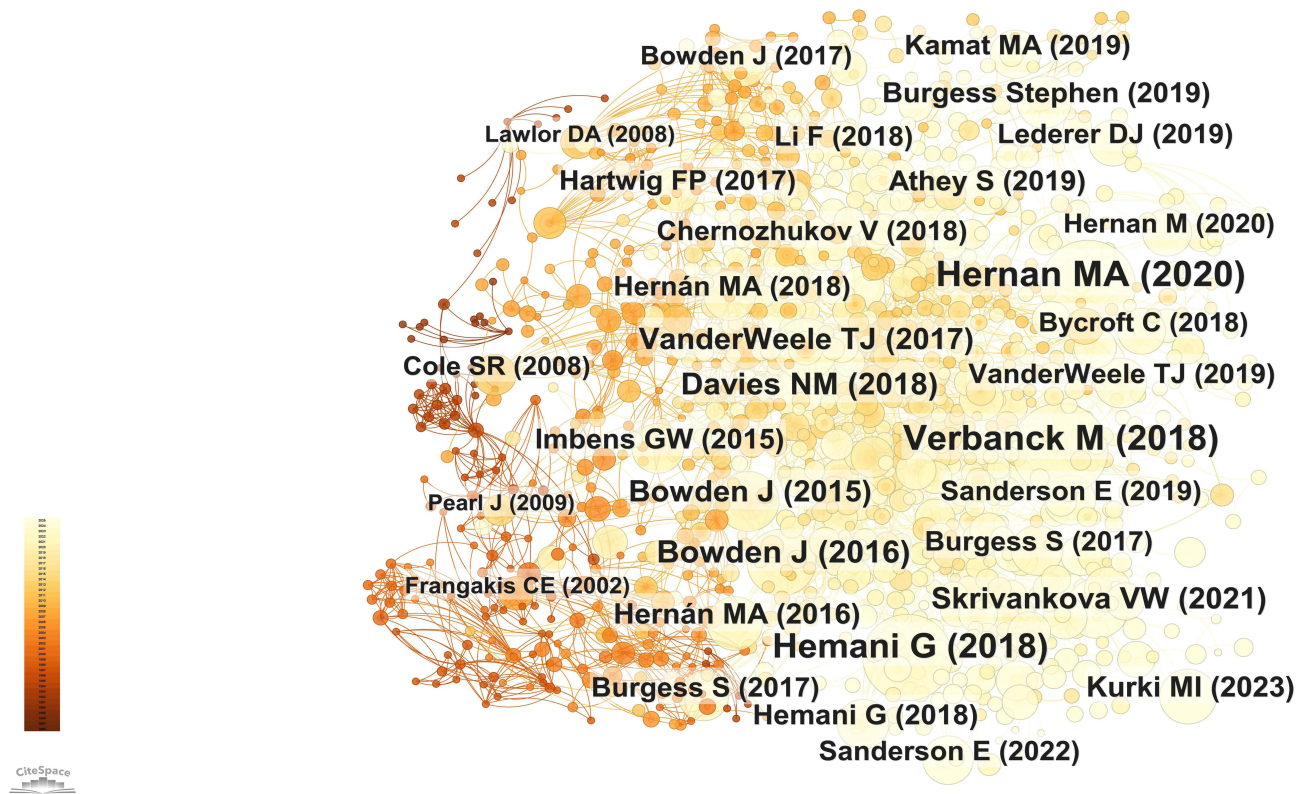


Figure 12 Network of co-cited references.

communication among researchers and enhancing research transparency. Rheanna M. Mainzer et al⁶² conducted a scoping review of observational studies involving causal issues, discovering gaps in the use and reporting of multiple imputation for causal questions, emphasizing the importance of rigorous reporting frameworks and methodologies for accurate causal inference in medical journals. Future research should continue to explore and refine these frameworks to enhance the quality and reliability of observational study outcomes.

Table 3 Top 7 Co-Cited References in Order of Citations

Rank	First Authors	Count	Year	Co-Cited References
1	Marie Verbanck	163	2018	Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases
2	Gibran Hemani	146	2018	The MR-Base platform supports systematic causal inference across the human phenome
3	Jack Bowden	100	2016	Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator
4	Jack Bowden	77	2015	Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression
5	Neil M Davies	75	2018	Reading Mendelian randomization studies: a guide, glossary, and checklist for clinicians
6	Tyler J. VanderWeele	70	2017	Sensitivity Analysis in Observational Research: Introducing the E-Value
7	Stephen Burgess	54	2017	Interpreting findings from Mendelian randomization using the MR-Egger method

Counterfactual prediction, which focuses on evaluating events that are likely to occur under different conditions, is an important aspect of causal inference. In the healthcare field, this method is particularly critical, as it helps healthcare professionals understand the potential impact of a particular treatment or intervention on a patient's health. Counterfactual prediction is undergoing a revolution as technology advances, particularly in the areas of big data analytics and artificial intelligence. In the future, counterfactual prediction will increasingly rely on sophisticated computational models and algorithms capable of processing massive data sets and extracting valuable insights from them. Machine learning methods, such as Random Forests, Support Vector Machines, and Deep Neural Networks,^{63,64} have begun to be used for propensity score estimation. This indicates that future counterfactual predictions may increasingly rely on these advanced technologies to address nonlinear relationships and high-dimensional features in data. Causal artificial intelligence enhances machine decision-making capabilities in complex situations by mimicking human reasoning processes, enabling machines to understand and explain events based on causal relationships.⁶⁵ By generating counterfactual data and altering labels, models can more accurately identify causal features.⁶⁶ Furthermore, with the development of personalized medicine, counterfactual predictions will not only focus on the general treatment effects but also on how to tailor treatment plans based on individuals' specific characteristics. This suggests that future research may place greater emphasis on understanding how inter-individual heterogeneity affects treatment outcomes.

Mendelian Randomization

Mendelian randomization (MR) is a technique that uses genetic variation to assess whether risk factors (eg, biomarkers) have a causal effect on disease outcomes in non-experimental (observational) settings.^{37,67} Combining causal inference with observational epidemiology often yields controversial findings due to challenges such as potential confounders, reverse causation, and other biases.⁶⁸ Meanwhile, RCT, which is the most reliable method for causal inference in epidemiological studies, is also usually difficult to use directly to study the etiology of diseases due to demanding design and implementation conditions, strict controls, challenges in execution, and medical ethical considerations.⁶⁹ Dorothea Nitsch et al discussed the challenges and implications of using Mendelian randomization studies to estimate the effect of gene products on disease outcomes. They drew analogies between Mendelian randomization and randomized controlled trials, highlighting differences in the interpretation of causality between the two, particularly regarding the biological effects of the intended treatment versus the biological effects of the treatment received. Furthermore, they emphasized the need for caution when interpreting MR results due to the wide range of assumptions required for causal inference.⁷⁰

MR-Egger

Early conventional MR used instrumental variable methods, assuming that all genetic variants met the conditions of the conditional instrumental variable. Traditional methods are plausible in some cases, such as when the risk factor is a protein biomarker and the genetic variant is localized to the coding region of that protein. However, these assumptions may be less plausible for cases involving polygenic risk factors (eg, body mass index or blood glucose). Building on the traditional approach, Bowden et al discussed the use of Mendelian randomization with multiple genetic variants as a meta-analysis to estimate causal effects. They focused on the problem of invalid instrumental variables due to pleiotropic effects and introduced the concept of Egger regression as a tool to detect pleiotropic bias and provide consistent estimates of causal effects. This method, named MR-Egger, allows for a sensitivity analysis of the robustness found in Mendelian randomization studies.³⁸ The following year, Bowden et al introduced a novel weighted median estimator for Mendelian randomization studies, which addressed the problem of obtaining reliable results when not all genetic variants used as instrumental variables are valid. They defined a weighted median to handle invalid instrumental variables and later demonstrated the improved performance of the model in simulation analysis compared to the MR-Egger method.³⁷ Later, Stephen Burgess et al used summarized genetic data to interpret the results of Mendelian randomization using the MR-Egger method. They provided an example demonstrating the difference in estimates between the traditional method and MR-Egger.⁷¹ Unlike traditional methods of analysis, MR-Egger assesses whether genetic variation has multiple effects on the outcome and provides consistent estimates of causal effects under the InSIDE assumption. However, this assumption is still considered weak compared to traditional methods. MR-Egger can

be used as a valuable sensitivity analysis for detecting violations of instrumental variable assumptions and for assessing situations where estimates differ between traditional methods and MR-Egger.

Genome-Wide Association Study

With the accumulation of Genome-Wide Association Study (GWAS) data and the popularization of multi-omics technology, MR studies have been increasingly used in causal inference, not only to validate well-established causal associations, such as LDL-C and obesity increasing the risk of coronary heart disease,^{72,73} but also to draw more reliable conclusions on long-standing controversial causal associations, such as the inability of small amounts of alcohol to prevent cardiovascular disease.⁷⁴ Although GWAS provides MR with data on genetic and trait associations, issues such as pleiotropy, linkage disequilibrium, and weak instruments cannot be avoided. Pleiotropy may affect the estimation of associations between risk factors and outcomes; linkage disequilibrium may lead to an overestimation of the precision of methods for pooling data; and weak instruments may result in biased causal estimates, especially when multiple weak instruments are involved.⁷⁵ Studies have shown that weak instrument bias occurs when the strength of the relationship between the instrumental variable and the phenotype is low.⁷⁶ Researchers can reduce this bias by using parsimonious models and adjusting for covariates. Additionally, as the F-statistic decreases, bias increases, but adjusting for the F-statistic can minimize weak instrument bias. Increasing the sample size can also help reduce this bias. Relying on the underlying assumptions of ZEMPA, Davey and colleagues proposed a new method called model-based estimation (MBE) for obtaining a single causal effect estimate from multiple genetic instruments. They also provided a weighting strategy to enhance this method, thereby more comprehensively improving the validity and interpretability of MR results.⁷⁷ Currently, for the existence of a single problem, researchers have corresponding means to solve it. However, with the continuous advancement of genetic engineering, the superposition of multiple issues has become the norm in modern research. Researchers now need to integrate multiple statistical methods and gradually combine various MR analysis techniques. Such as MR-Egger mentioned above, or the Mendelian Randomization Pleiotropy RESidual Sum and Outlier (MR-PRESSO), which has recently received much attention for correcting for pleiotropy bias⁷⁸ and the widely used multivariate MR (MVMR),⁷⁹ each of which has different advantages for addressing various types of bias. Meanwhile, researchers are also attempting to utilize advanced genetic analysis techniques, including but not limited to the application of fine-mapping techniques and the integration of functional genomics data, to identify and confirm true causal variants.^{80,81} These approaches will significantly aid in more accurately localizing genetic variants that affect specific traits and understanding their functions.

Two-Sample MR

Specific methods and frameworks for two-sample MR were developed around 2000, and the approach has become increasingly popular in epidemiology and other fields due to its potential to address confounders and reverse causality. Two-sample MR can utilize a large amount of existing public GWAS data, avoiding the sample size limitations that single-sample MR may face, and reducing the reliance on a single dataset, thereby minimizing potential confounding bias. Although two-sample MR reduces data requirements, it is limited by the lack of novelty due to the use of publicly available data. Hartwig et al discussed the concept and application of two-sample Mendelian randomization techniques, emphasizing the importance of effective and accurate data coordination to avoid issues that can arise in two-sample MR analyses. The authors pointed out that although two-sample MR is a powerful and widely applicable technique, improper handling can lead to misleading conclusions. Specifically, the process of data harmonization to ensure that the same genetic variants and effect alleles are used is crucial for the accuracy in the analysis.⁸² This study provides a detailed set of guidelines and recommendations for two-sample MR to avoid potential biases in the analysis and to ensure the reliability and accuracy of the findings. Using two-sample multivariate Mendelian randomization (MVMR), Sanderson et al demonstrated how reliable estimates of causal effects can be obtained in the presence of weak instruments and pleiotropy by introducing a two-sample conditional F-statistic to test the strength of genetic variation in predicting each exposure in an MVMR model.⁸³ Composite models constructed by incorporating diverse approaches are expected to demonstrate significant utility in assessing the validity of instrumental variables, adjusting for the effects of pleiotropy, and enhancing the adaptability and breadth of application of two-sample Mendelian randomization.

Utilizing genetic variation as an instrumental variable, the MR approach bypasses the problem of confounding in traditional epidemiologic studies and provides a unique way to explore and test causal hypotheses. The advantage of this approach is that it relies on the random distribution of genetic variants that biologically predate the emergence of the disease, thus minimizing the issues of confounding factors and reverse causation. MR is not limited to assessing the impact of a single risk factor on a specific disease; it is also beginning to be used to explore a wider range of biological processes and complex disease mechanisms, such as metabolic pathways, immune responses, and cell signaling.^{84–86} With advances in computational biology and bioinformatics technologies, MR methods are incorporating more biological data, such as transcriptomics, proteomics, and metabolomics data.^{87,88} This multi-omics integration provides new perspectives for revealing the multilevel causal mechanisms of diseases. MR has demonstrated a strong ability to address bias caused by confounding and reverse causation problems. It also provides reliable evidence for inferring causal relationships between exposure factors and outcomes. Related studies are expected to guide clinical trials and drug development, and provide a theoretical basis for clinical and public health decision-making.⁸⁹

Marginal Structural Model

The Marginal structural model (MSM) serves as a causal model for estimating the effects of transient exposures in the presence of time-dependent covariates based on observational data.²⁴ Its effective adjustment for time-dependent confounders affecting time-varying treatments or treatment regimens allows for unbiased estimation of causal effects.⁹⁰ Some published results often estimate time-varying exposure effects in conjunction with weighted estimates of the parameters of the marginal structural model. Unlike standard statistical methods, these models can appropriately adjust for time-varying confounders affected by prior exposures, providing considerable flexibility.^{91–93}

Time-Dependent Confounders

Time-dependent factors are prevalent in modern medical research, and these factors implicitly influence treatment effects, patient adherence, and disease progression.⁹⁴ Robins and colleagues found in their early research that in observational studies, when time-dependent confounders are present, the traditional method of confounding adjustment may lead to bias, resulting in incorrectly estimated treatment effects and confounding the inference of true causality.²⁴ Based on these issues, they proposed the Marginal Structural Model (MSM). In the aftermath, Daniel Westreich et al applied the parameter g-formula to appropriately adjust for time-varying confounders influenced by prior treatments, avoiding the introduction of stratification bias and enabling the estimation of marginal rather than conditional effects. This approach avoids problems associated with non-collapsible effect scales,⁹⁵ providing a new way of thinking about effectively dealing with time-dependent confounding in long-term observational studies. In addition, Maya L. Petersen et al took an alternative approach to explore the importance of positivity assumptions and the implications of their violation. By identifying violations of positivity assumptions, researchers can more accurately assess the applicability of the model and the reliability of the estimation, which are crucial for the proper treatment of time-dependent confounding.⁹⁶ This study serves as an important guide for researchers using MSM to address time-dependent confounding and estimate time-varying treatment effects, enabling them to more effectively design studies and interpret results. As time-varying factors continue to expand, VanderWeele et al describe a simple technique for estimating direct and indirect effect ratios by combining logistic and linear regression, which is applicable when outcomes are rare and mediators are continuous.⁹⁷ Consequently, the application of mediation analysis can be extended through appropriate methodological improvements when conducting research on time-dependent factors, leading to more accurate causality and effect decompositions in a wider range of research contexts. In addition, propensity score methods are often misapplied when estimating the effect of treatment on time-to-event outcomes. To address this, Austin et al describe how two different propensity score methods—matching and treatment-weighted inverse probability—can be used to estimate effect measures commonly reported in randomized controlled trials.⁹⁸ Propensity score methods can be effective in estimating the marginal effects of treatments, which is particularly important for the analysis of time-to-event data. With propensity score matching and weighting, it is possible to simulate the conditions of a randomized trial and generate estimates of marginal effects, thereby improving the assessment of long-term treatments or interventions.

Inverse Probability Weighting

In the early days, there was no unified reference for weighting and factor selection in MSM. Cole et al developed an informal and easy-to-implement weight truncation method considering Inverse Probability Weighting (IPW) for marginal structural modeling. This method addresses the trade-off between bias and accuracy, while IPW also provides a powerful methodological tool to reveal the impact of causal exposure, thus helping to determine whether it is masked or not.⁹⁹ Inverse probability weighting (IPW) is an important tool for adjusting for confounders and selection bias. Since the relationship between exposure and outcome may be affected by confounders, IPW balances these confounders by assigning weights to study participants, allowing for a more direct estimation of the causal effect of the treatment on the outcome. Additionally, IPW adjusts for the impact of confounders on the relationship between treatments and outcomes by considering confounders at each point in time, further enhancing MSM's ability to deal with confounding variables that vary over time.¹⁰⁰ Previous studies have shown that the marginal effects of IPW estimation are highly sensitive to model misspecification in constructing the weights.¹⁰¹ To address this issue, researchers have either self-balanced the covariates via covariate-balanced propensity score methods to automatically balance the covariates and thus improve the construction of the weights,¹⁰² constructed the MSM weights based on residual balancing methods,¹⁰³ or used the targeted maximum likelihood estimation (TMLE) method with a doubly robust estimation technique¹⁰⁴ to mitigate the effects of propensity score model misspecification to a certain extent. In addition, the use of stabilized weights in IPW estimation helps to mitigate the problem of high variability due to extreme weights. These weights estimate the conditional means of individual observations and treatment histories by balancing the expected value of past confounders at each time point. However, extreme weights can lead to highly variable estimates, which need to be mitigated by stabilizing weights.¹⁰⁵ Combined with the need for personalized medicine, IPW methods can deepen our understanding of disease mechanisms and optimize treatment plans. With the integration of big data and machine learning technologies, more detailed causal inference is expected to be realized in the future, providing real-time and precise support for clinical decision-making.¹⁰⁶ This will help discover new treatment pathways, achieve more efficient health management, and improve patient outcomes, driving the field of medicine towards a more personalized and scientific approach.

The Marginal Structural Model (MSM) is uniquely suited to dealing with time-dependent confounding in observational studies, especially when exposure variables and covariates vary over time. This makes MSM a powerful tool for evaluating the effects of long-term interventions or treatments. To this day, MSM is constantly being improved to accommodate more complex research scenarios, such as improving propensity score estimation by incorporating machine learning methods¹⁰⁷ and developing more flexible models to handle complex temporal dynamics.⁷ However, MSM has its limitations, especially in cases of “treatment saturation”, where all observed subjects receive the same treatment in a certain state. In such scenarios, MSM may struggle to estimate accurate causal effects, as it cannot learn from the data the potential outcomes under different treatment conditions.¹⁰⁸ To overcome these limitations, cutting-edge research is exploring how to combine other statistical techniques and theories to enhance the applicability and flexibility of MSM. For example, researchers have attempted to combine sequence analysis methods and multi-state modeling to better handle time-dependent confounding and multi-stage decision-making processes.¹⁰⁹ Additionally, research is exploring how causal inference networks and structural equation modeling can be utilized to provide a deeper understanding of causality and to address more complex time-dependent intervention effects.¹¹⁰ These approaches offer more insightful and nuanced analytical tools for medical research. These advances have not only expanded the scope of MSM applications but also driven innovations in medical research methodologies that help accurately assess the long-term benefits and potential risks of therapeutic interventions.

Limitations

This study has several limitations. First, the constraints of CiteSpace limited our data collection to the WoSCC, which may not fully represent all available literature. Second, our selection of specific keywords might have excluded relevant and significant terms, potentially affecting the comprehensive understanding of the literature and the identification of research trends. Third, our analysis focused solely on English-language articles and restricted the document type to dissertations. Finally, our use of CiteSpace for visualization analysis primarily relied on frequency and centrality

measures. Centrality is influenced by the number of nodes selected, which could impact the accuracy of our conclusions. Additionally, the presence of synonyms may have led to some overlap between categories during keyword clustering. Our co-citation network focuses only on the first authors, which does not fully reflect the influence of all authors. Additionally, various co-citation networks may not necessarily capture recent trends, as recent publications have not been sufficiently cited.

Conclusion

This study conducts a bibliometric analysis of causal inference research in the medical field from 1986 to 2024. The publication trend shows fluctuations but a sustained increase, particularly an explosive growth since 2019. The United States leads in epidemiological research based on causal inference, although other countries are rapidly advancing. Harvard University has contributed the most and actively promotes collaboration. The American Journal of Epidemiology has published the highest number of relevant studies, emphasizing causal inference applications in epidemiology. Key research topics identified by keyword analysis include epidemiology, coronary heart disease, and health, with frequent terms such as causal inference, risk, and health. Recent terms like “big data”, “DNA methylation”, “robust estimation”, and “prediction” reflect the field’s responsiveness to emerging trends. As big data and artificial intelligence advance, researchers increasingly highlight that inter-individual heterogeneity affects treatment outcomes. This necessitates a more refined approach to causal inference by integrating diverse data sources to enhance trial sensitivity and specificity, ultimately revealing the complex mechanisms underlying diseases and drug actions.

Author Contributions

G.Qin: Formal analysis, Investigation, Resources, Visualization, Writing – original draft, Writing – review & editing. J. Wei: Supervision, Methodology, Formal analysis, Investigation, Writing – review & editing. Y.Sun: Investigation, Methodology, Writing – review & editing. W.Du: Investigation, Resources, Writing – review & editing. All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This work was supported by the National Social Science Foundation of China (23ATQ009), the Project of Social Science Foundation of Jiangsu Province (23TQB007), the Major Project of Philosophy and Social Science Research in Jiangsu Universities (2020SJZDA102), and the Humanities and Social Sciences Research Fund Project of Nanjing University of Posts and Telecommunications (NYP223004).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Janse RJ, Hoekstra T, Jager KJ. et al. Conducting correlation analysis: important limitations and pitfalls. *Clin Kidney J.* **2021**;14(11):2332–2337. doi:10.1093/cjkj/sfab085
2. Prosperi M, Guo Y, Sperrin M, et al. Causal inference and counterfactual prediction in machine learning for actionable healthcare. *Nat Mach Intell.* **2020**;2(7):369–375. doi:10.1038/s42256-020-0197-y
3. Holland PW. Statistics and causal inference. *J Am Stat Assoc.* **1986**;81(396):945–960. doi:10.1080/01621459.1986.10478354
4. Sassower R. Causality and correlation. In: Turner BS, editors. *The Wiley-Blackwell Encyclopedia of Social Theory.* John Wiley & Sons, Ltd; **2017**:1–4. doi:10.1002/9781118430873.est0585
5. Igelström E, Craig P, Lewsey J, Lynch J, Pearce A, Katikireddi SV. Causal inference and effect estimation using observational data. *J Epidemiol Community Health.* **2022**;76(11):960–966. doi:10.1136/jech-2022-219267
6. Stuart EA. Matching methods for causal inference: a review and a look forward. *Statist Sci.* **2010**;25(1):1. doi:10.1214/09-STS313
7. Athey S, Imbens GW. Machine learning methods for estimating heterogeneous causal effects. *stat.* **2015**;1050(5):1–26.
8. Chakraborty B, Murphy SA. Dynamic Treatment Regimes. *Annu Rev Stat Appl.* **2014**;1(1):447–464. doi:10.1146/annurev-statistics-022513-115553

9. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol.* 1974;66(5):688. doi:10.1037/h0037350
10. Pearl J. Causal diagrams for empirical research. *Biometrika.* 1995;82(4):669–688. doi:10.1093/biomet/82.4.669
11. Gale RP, Zhang MJ, Lazarus HM. The role of randomized controlled trials, registries, observational databases in evaluating new interventions. *Best Pract Res Clin Haematol.* 2023;36(4):101523. doi:10.1016/j.beha.2023.101523
12. van Amsterdam WAC, Elias S, Ranganath R. Causal inference in oncology: why, what, how and when. *Clin Oncol.* 2025;38. doi:10.1016/j.clon.2024.07.002
13. Béal J, Latouche A Causal inference with multiple versions of treatment and application to personalized medicine. Available from Published online May 25, 2020: <http://arxiv.org/abs/2005.12427>. Accessed July 2, 2024.
14. Michael T, Zhang JD. Causal inference in drug discovery and development. *Drug Discovery Today.* 2023;28(10):103737. doi:10.1016/j.drudis.2023.103737
15. Castro DC, Walker I, Glocker B. Causality matters in medical imaging. *Nat Commun.* 2020;11(1):3673. doi:10.1038/s41467-020-17478-w
16. Tan L, Wang X, Yuan K, et al. Structural and temporal dynamics analysis on drug-eluting stents: history, research hotspots and emerging trends. *Bioact Mater.* 2023;23:170–186. doi:10.1016/j.bioactmat.2022.09.009
17. Zhao Y, Yu Y, Wang H, et al. Machine learning in causal inference: application in pharmacovigilance. *Drug Saf.* 2022;45(5):459–476. doi:10.1007/s40264-022-01155-6
18. Bonovas S, Piovani D. Simpson's paradox in clinical research: a cautionary tale. *J Clin Med.* 2023;12(4):1633. doi:10.3390/jcm12041633
19. Chen C. CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. *J Am Soc Inf Sci.* 2006;57(3):359–377. doi:10.1002/asi.20317
20. Chen C, Dubin R, Kim MC. Emerging trends and new developments in regenerative medicine: a scientometric update (2000 – 2014). *Expert Opin Biol Ther.* 2014;14(9):1295–1317. doi:10.1517/14712598.2014.920813
21. Chen C. Searching for intellectual turning points: progressive knowledge domain visualization. *Proc Natl Acad Sci USA.* 2004;101(suppl_1):5303–5310. doi:10.1073/pnas.0307513100
22. Chen C. Science mapping: a systematic review of the literature. *J Data Inf Sci.* 2017;2(2):1–40. doi:10.1515/jdis-2017-0006
23. Chen C, Ihekwe-SanJuan F, Hou J. The structure and dynamics of cocitation clusters: a multiple-perspective cocitation analysis. *J Am Soc Inf Sci Technol.* 2010;61(7):1386–1409. doi:10.1002/asi.21309
24. Robins JM. Marginal structural models versus structural nested models as tools for causal inference. In: *Statistical Models in Epidemiology, the Environment, and Clinical Trials*. Springer; 2000:95–133. https://link.springer.com/content/pdf/10.1007/978-1-4612-1284-3_2?pdf=chapter%20toc.
25. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology.* 2004;15(5):615–625. doi:10.1097/01.ede.0000135174.63482.43
26. Kurtz P, Bastos LSL, Dantas LF, et al. Evolving changes in mortality of 13,301 critically ill adult patients with COVID-19 over 8 months. *Intensive Care Med.* 2021;47(5):538–548. doi:10.1007/s00134-021-06388-0
27. Bonvini M, Kennedy EH, Ventura V, Wasserman L. Causal inference for the effect of mobility on COVID-19 deaths. *Ann Appl Stat.* 2022;16(4):2458–2480. doi:10.1214/22-AOAS1599
28. Goodman-Bacon A, Marcus J. Using difference-in-differences to identify causal effects of COVID-19 policies. *Surv Res Methods.* 2020;14(2):153–158. doi:10.18148/srm/2020.v14i2.7723
29. Luo L, Risk M, Shi X. Online causal inference with application to near real-time post-market vaccine safety surveillance. *Stat Med.* 2024;43(14):2734–2746. doi:10.1002/sim.10095
30. Joshi S, Urteaga I, van Amsterdam WAC, et al. AI as an intervention: improving clinical outcomes relies on a causal approach to AI development and validation. *J Am Med Inf Assoc.* 2025;32(3):589–594. doi:10.1093/jamia/ocae301
31. Cui P, Shen Z, Li S, et al. Causal inference meets machine learning. In: *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*; 2020 July 6–10. San Diego, CA, USA: KDD' 20, Association for Computing Machinery; 2020:3527–3528. doi:10.1145/3394486.3406460.
32. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Ame J Epidemiol.* 2016;183(8):758–764. doi:10.1093/aje/kwv254
33. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nature Genet.* 2018;50(5):693–698. doi:10.1038/s41588-018-0099-7
34. Welter D, MacArthur J, Morales J, et al. The NHGRI GWAS catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* 2014;42(D1):D1001–D1006. doi:10.1093/nar/gkt1229
35. Hemani G, Zheng J, Elsworth B, et al. The MR-base platform supports systematic causal inference across the human genome. *elife.* 2018; 7: e34408.
36. Smith GD, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. *Int J Epidemiol.* 2004;33(1):30–42. doi:10.1093/ije/dyh132
37. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol.* 2016;40(4):304–314. doi:10.1002/gepi.21965
38. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through egger regression. *Int J Epidemiol.* 2015;44(2):512–525. doi:10.1093/ije/dyv080
39. Davey Smith G, Ebrahim S. Mendelian randomization: can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol.* 2003;32(1):1–22. doi:10.1093/ije/dyg070
40. Nelson R, Staggars N. *Health Informatics: An Interprofessional Approach*. Elsevier Health Sciences; 2013. Available from: <https://books.google.com/books?hl=zh-CN&lr=&id=v7XQjy0377cC&oi=fnd&pg=PP1&dq=Health+informatics:+An+interprofessional+approach&ots=04NhXF5Yzy&sig=ksO2AlfXidi70ea4Za5zLfOKci0>. Accessed July 2, 2024.
41. Fu M; Group UPHDDRW, Group UPHAW. Phenome-wide association study of polygenic risk score for Alzheimer's disease in electronic health records. *Front Aging Neurosci.* 2022;14:800375. doi:10.3389/fnagi.2022.800375

42. Heintzman N, Kleinberg S. Using uncertain data from body-worn sensors to gain insight into type 1 diabetes. *J Biomed Informat.* 2016;63:259–268. doi:10.1016/j.jbi.2016.08.022
43. Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-world evidence — what is it and what can it tell us? *N Engl J Med.* 2016;375(23):2293–2297. doi:10.1056/NEJMsbl609216
44. Tong L, Shi W, Isgut M, et al. Integrating multi-omics data with EHR for precision medicine using advanced artificial intelligence. *IEEE Rev Biomed Engine.* 2024;17:80–97. doi:10.1109/RBME.2023.3324264
45. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika.* 1983;70(1):41–55. doi:10.1093/biomet/70.1.41
46. Schober P, Vetter TR. Propensity score matching in observational research. *Anesthesia Analg.* 2020;130(6):1616–1617. doi:10.1213/ANE.0000000000004770
47. Zhao QY, Luo JC, Su Y, Zhang YJ, Tu GW, Luo Z. Propensity score matching with R: conventional methods and new features. *Ann Translat Med.* 2021;9(9). <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8246231/>.
48. Yang S, Imbens GW, Cui Z, Faries DE, Kadziola Z. Propensity score matching and subclassification in observational studies with multi-level treatments. *Biometrics.* 2016;72(4):1055–1065. doi:10.1111/biom.12505
49. Wei J, Wang X, Schuurmans D, et al. Chain of thought prompting elicits reasoning in large language models. *ArXiv.* 2024. doi:10.1117/12.2511699
50. Hahn PR, Carvalho CM, Puelz D, He J. Regularization and confounding in linear regression for treatment effect estimation. Available from Published online 2018: <https://projecteuclid.org/journals/bayesian-analysis/volume-13/issue-1/Regularization-and-Confounding-in-Linear-Regression-for-Treatment-Effect-Estimation/10.1214/16-BA1044.short>. Accessed May 2, 2025. Accessed July 2, 2024.
51. Iacus SM, King G, Porro G. Multivariate matching methods that are monotonic imbalance bounding. *J Am Stat Assoc.* 2011;106(493):345–361. doi:10.1198/jasa.2011.tm09599
52. Karim ME, Pang M, Platt RW. Can we train machine learning methods to outperform the high-dimensional propensity score algorithm? *Epidemiology.* 2018;29(2):191–198. doi:10.1097/EDE.0000000000000787
53. Bondonio D, Greenbaum RT. Natural disasters and relief assistance: empirical evidence on the resilience of U.S. counties using dynamic propensity score matching. *J Reg Sci.* 2018;58(3):659–680. doi:10.1111/jors.12379
54. Iacus SM, King G, Porro G. Causal inference without balance checking: coarsened exact matching. *Political Anal.* 2012;20(1):1–24. doi:10.1093/pan/mpr013
55. Farrell MH, Liang T, Misra S. Deep neural networks for estimation and inference. *Econometrica.* 2021;89(1):181–213. doi:10.3982/ECTA16901
56. Franklin JM, Schneeweiss S. When and how can real world data analyses substitute for randomized controlled trials? *Clin Pharmacol Therap.* 2017;102(6):924–933. doi:10.1002/cpt.857
57. Hernán MA. Methods of public health research — strengthening causal inference from observational data. *N Engl J Med.* 2021;385(15):1345–1348. doi:10.1056/NEJMp2113319
58. Hammerton G, Munafò MR. Causal inference with observational data: the need for triangulation of evidence. *Psychol Med.* 2021;51(4):563–578. doi:10.1017/S0033291720005127
59. Zwahlen M, Salanti G. Causal inference from experiment and observation. *Evid Based Mental Health.* 2018;21(1):34–38. doi:10.1136/eb-2017-102859
60. Coulombe J, Yang S. Multiply robust estimation of marginal structural models in observational studies subject to covariate-driven observations. *Biometrics.* 2024;80(3):ujae065. doi:10.1093/biomtc/ujae065
61. Dahabreh IJ, Bibbins-Domingo K. Causal inference about the effects of interventions from observational studies in medical journals. *JAMA.* 2024;331(21):1845–1853. doi:10.1001/jama.2024.7741
62. Mainzer RM, Moreno-Betancur M, Nguyen CD, Simpson JA, Carlin JB, Lee KJ. Gaps in the usage and reporting of multiple imputation for incomplete data: findings from a scoping review of observational studies addressing causal questions. *BMC Med Res Method.* 2024;24(1):193. doi:10.1186/s12874-024-02302-6
63. Shalit U, Johansson FD, Sontag D. Estimating individual treatment effect: generalization bounds and algorithms. In: *International Conference on Machine Learning*. PMLR; 2017 Aug 6–11. Sydney, NSW, Australia; 2017:3076–3085. <https://proceedings.mlr.press/v70/shalit17a.html>.
64. Westreich D, Lessler J, Funk MJ. Propensity score estimation: neural networks, support vector machines, decision trees (CART), and meta-classifiers as alternatives to logistic regression. *J Clin Epidemiol.* 2010;63(8):826–833. doi:10.1016/j.jclinepi.2009.11.020
65. Sahoh B, Kliangkhlao M, Haruehansapong K, Yeranee K, Punsawad Y. A personal thermal comfort model based on causal artificial intelligence: a physiological sensor-enabled causal identifiability. *IEEE J Biomed Health Inform.* 2024;28(12):7565–7576. doi:10.1109/JBHI.2024.3432766
66. Yang Z, Liu Y, Ouyang C, Ren L, Wen W. Counterfactual can be strong in medical question and answering. *Inform Process Manage.* 2023;60(4):103408. doi:10.1016/j.ipm.2023.103408
67. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med.* 2008;27(8):1133–1163. doi:10.1002/sim.3034
68. Smith GD, Data dredging, bias, or confounding: they can all get you into the BMJ and the Friday papers. *BMJ.* 2002;325(7378):1437–1438. doi:10.1136/bmj.325.7378.1437
69. Wang LN, Zhang Z. [Mendelian randomization approach, used for causal inferences]. *Zhonghua Liu Xing Bing Xue Za Zhi.* 2017;38(4):547–552. doi:10.3760/cma.j.issn.0254-6450.2017.04.027
70. Nitsch D, Molokhia M, Smeeth L, DeStavola BL, Whittaker JC, Leon DA. Limits to causal inference based on Mendelian randomization: a comparison with randomized controlled trials. *Am J Epidemiol.* 2006;163(5):397–403. doi:10.1093/aje/kwj062
71. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-egger method. *Eur J Epidemiol.* 2017;32(5):377–389. doi:10.1007/s10654-017-0255-x
72. Benn M, Tybjærg-Hansen A, Stender S, Frikke-Schmidt R, Nordestgaard BG. Low-density lipoprotein cholesterol and the risk of cancer: a Mendelian randomization study. *JNCI.* 2011;103(6):508–519. doi:10.1093/jnci/djr008

73. Larsson SC, Bäck M, Rees JMB, Mason AM, Burgess S. Body mass index and body composition in relation to 14 cardiovascular conditions in UK biobank: a Mendelian randomization study. *Eur Heart J*. 2020;41(2):221–226. doi:10.1093/eurheartj/ehz388
74. Lv K, Yang G, Wu Y, et al. The causal effect of metabolic syndrome and its components on benign prostatic hyperplasia: a univariable and multivariable Mendelian randomization study. *Prostate*. 2023;83(14):1358–1364. doi:10.1002/pros.24598
75. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genetic Epidemiol*. 2013;37(7):658–665. doi:10.1002/gepi.21758
76. Burgess S, Thompson SG. CRP CHD genetics collaboration. avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol*. 2011;40(3):755–764. doi:10.1093/ije/dyr036
77. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol*. 2017;46(6):1985–1998. doi:10.1093/ije/dyx102
78. Ong J, MacGregor S. Implementing MR-PRESSO and GCTA-GSMR for pleiotropy assessment in Mendelian randomization studies from a practitioner's perspective. *Genetic Epidemiol*. 2019;43(6):609–616. doi:10.1002/gepi.22207
79. Burgess S, Thompson SG. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Ame J Epidemiol*. 2015;181(4):251–260. doi:10.1093/aje/kwu283
80. LaPierre N, Taraszka K, Huang H, He R, Hormozdiari F, Eskin E. Identifying causal variants by fine mapping across multiple studies. *PLoS Genetics*. 2021;17(9):e1009733. doi:10.1371/journal.pgen.1009733
81. Schaid DJ, Chen W, Larson NB. From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nat Rev Genet*. 2018;19(8):491–504. doi:10.1038/s41576-018-0016-z
82. Hartwig FP, Davies NM, Hemani G, Davey Smith G. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. *Int J Epidemiol*. 2016;45(6):1717–1726. doi:10.1093/ije/dyx028
83. Sanderson E, Spiller W, Bowden J. Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization. *Stat Med*. 2021;40(25):5434–5452. doi:10.1002/sim.9133
84. Folkersen L, Gustafsson S, Wang Q, et al. Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931 individuals. *Nat Metab*. 2020;2(10):1135–1148. doi:10.1038/s42255-020-00287-2
85. Richardson TG, Zheng J, Smith GD, et al. Mendelian randomization analysis identifies CpG sites as putative mediators for genetic influences on cardiovascular disease risk. *Am J Hum Genet*. 2017;101(4):590–602. doi:10.1016/j.ajhg.2017.09.003
86. Zheng J, Baird D, Borges M-C, et al. Recent developments in Mendelian randomization studies. *Curr Epidemiol Rep*. 2017;4(4):330–345. doi:10.1007/s40471-017-0128-6
87. Gamazon ER, Wheeler HE, Shah KP, et al. A gene-based association method for mapping traits using reference transcriptome data. *Nature Genet*. 2015;47(9):1091–1098. doi:10.1038/ng.3367
88. Suhre K, Arnold M, Bhagwat AM, et al. Connecting genetic risk to disease end points through the human blood plasma proteome. *Nat Commun*. 2017;8(1):14357. doi:10.1038/ncomms14357
89. Walker VM, Davey Smith G, Davies NM, Martin RM. Mendelian randomization: a novel approach for the prediction of adverse drug events and drug repurposing opportunities. *Int J Epidemiol*. 2017;46(6):2078–2089. doi:10.1093/ije/dyx207
90. Hernán MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*. 2000;11(5):561–570. doi:10.1097/00001648-200009000-00012
91. Cole SR, Hernán MA, Anastos K, Jamieson BD, Robins JM. Determining the effect of highly active antiretroviral therapy on changes in human immunodeficiency virus type 1 RNA viral load using a marginal structural left-censored mean model. *Ame J Epidemiol*. 2007;166(2):219–227. doi:10.1093/aje/kwm047
92. Yokoyama K, Matsuo N, Yoshida H. Middle molecular uremic substances retention itself might influence the bioincompatibility of PD solution. *Kidney Int*. 2008;74(6):822–831. doi:10.1038/ki.2008.283
93. Petersen ML, Wang Y, Van Der Laan MJ, Bangsberg DR. Assessing the effectiveness of antiretroviral adherence interventions: using marginal structural models to replicate the findings of randomized controlled trials. *J Acquir Immune Defic Syndr*. 2006; 43:S96–S103.
94. Rizopoulos D. Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics*. 2011;67(3):819–829. doi:10.1111/j.1541-0420.2010.01546.x
95. Westreich D, Cole SR, Young JG, et al. The parametric g-formula to estimate the effect of highly active antiretroviral therapy on incident AIDS or death. *Stat Med*. 2012;31(18):2000–2009. doi:10.1002/sim.5316
96. Petersen ML, Porter KE, Gruber S, Wang Y, Van Der Laan MJ. Diagnosing and responding to violations in the positivity assumption. *Stat Methods Med Res*. 2012;21(1):31–54. doi:10.1177/0962280210386207
97. VanderWeele TJ, Vansteelandt S. Odds ratios for mediation analysis for a dichotomous outcome. *Ame J Epidemiol*. 2010;172(12):1339–1348. doi:10.1093/aje/kwq332
98. Austin PC. The use of propensity score methods with survival or time-to-event outcomes: reporting measures of effect similar to those used in randomized experiments. *Stat Med*. 2014;33(7):1242–1258. doi:10.1002/sim.5984
99. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Ame J Epidemiol*. 2008;168(6):656–664. doi:10.1093/aje/kwn164
100. Nab L, Groenwold RH, van Smeden M, Keogh RH. Quantitative bias analysis for a misclassified confounder: a comparison between marginal structural models and conditional models for point treatments. *Epidemiology*. 2020;31(6):796–805. doi:10.1097/EDE.0000000000001239
101. Kurz CF. Augmented inverse probability weighting and the double robustness property. *Med Decis Mak*. 2022;42(2):156–167. doi:10.1177/0272989X211027181
102. Imai K, Ratkovic M. Covariate balancing propensity score. *J Royal Statist Soc Series B*. 2014;76(1):243–263. doi:10.1111/rssb.12027
103. Wodtke GT, Zhou X. Effect decomposition in the presence of treatment-induced confounding: a regression-with-residuals approach. *Epidemiology*. 2020;31(3):369–375. doi:10.1097/EDE.0000000000001168
104. Schuler MS, Rose S. Targeted maximum likelihood estimation for causal inference in observational studies. *Ame J Epidemiol*. 2017;185(1):65–73. doi:10.1093/aje/kww165
105. Zhou X, Wodtke GT. Residual balancing: a method of constructing weights for marginal structural models. *Political Anal*. 2020;28(4):487–506. doi:10.1017/pan.2020.2

106. Rose S. Targeted learning in data science: causal inference for complex longitudinal studies. Published online. 2018.
107. Polley E, van der Laan M. Super learner in prediction. *UC Berkeley Division of Biostatistics Working Paper Series*. Published online May 3, 2010. <https://biostats.bepress.com/ucbbiostat/paper266>. Accessed May 2, 2025.
108. Neugebauer R, van der Laan M. Why prefer double robust estimators in causal inference? *J Stat Plan Inference*. 2005;129(1–2):405–426. doi:10.1016/j.jspi.2004.06.060
109. Meira-Machado L, De Uña-álvarez J, Cadarso-Suárez C, Andersen PK. Multi-state models for the analysis of time-to-event data. *Stat Methods Med Res*. 2009;18(2):195–222. doi:10.1177/0962280208092301
110. Pearl J, Mackenzie D. *The Book of Why: The New Science of Cause and Effect*. Basic books; 2018. Available from: <https://books.google.com/books?hl=zh-CN&lr=&id=9H0dDQAAQBAJ&oi=fnd&pg=PT7&dq=The+book+of+why:+the+new+science+of+cause+and+effect&ots=ZitDhRqrgZ&sig=XkeBfL2cIVkZjfUEec9ZS1RoM8A>. Accessed July 2, 2024.

Journal of Multidisciplinary Healthcare

Publish your work in this journal

The Journal of Multidisciplinary Healthcare is an international, peer-reviewed open-access journal that aims to represent and publish research in healthcare areas delivered by practitioners of different disciplines. This includes studies and reviews conducted by multidisciplinary teams as well as research which evaluates the results or conduct of such teams or healthcare processes in general. The journal covers a very wide range of areas and welcomes submissions from practitioners at all levels, from all over the world. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/journal-of-multidisciplinary-healthcare-journal>

Dovepress
Taylor & Francis Group