BMC Medicine

RESEARCH

Open Access

Confounder adjustment in observational studies investigating multiple risk factors: a methodological study

Yinyan Gao[1†], Linghui Xiang[1†], Hang Yi[2], Jinlu Song[1], Dingkui Sun[1], Boya Xu[1], Guochao Zhang[2*†] and Irene Xinyin Wu[1,3*†]

Abstract

Background Confounder adjustment is critical for accurate causal inference in observational studies. However, the appropriateness of methods for confounder adjustment in studies investigating multiple risk factors, where the factors are not simply mutually confounded, is often overlooked. This study aims to summarise the methods for confounder adjustment and the related issues in studies investigating multiple risk factors.

Methods A methodological study was performed. We searched PubMed from January 2018 to March 2023 to identify cohort and case–control studies investigating multiple risk factors for three chronic diseases (cardiovascular disease, diabetes and dementia). Study selection and data extraction were conducted independently by two reviewers. The study objectives were grouped into two categories: widely exploring potential risk factors and examining specific risk factors. The methods for confounder adjustment were classified based on a summarisation of the included studies, identifying six categories: (1) each risk factor was adjusted for potential confounders separately (the recommended method); (2) all risk factors were mutually adjusted (i.e. including all factors in a multivariable model); (3) all risk factors were adjusted for the same confounders separately; (4) all risk factors were adjusted for the same confounders with some factors being mutually adjusted; (5) all risk factors were adjusted for the same confounders with mutual adjustment among them being unclear; and (6) unable to judge. All data were descriptively analysed.

Results A total of 162 studies were included, with 88 (54.3%) exploring potential risk factors and 74 (45.7%) examining specific risk factors. The current status of confounder adjustment was unsatisfactory: only ten studies (6.2%) used the recommended method, all of which aimed at examining several specific risk factors; in contrast, mutual adjustment was adopted in over 70% of the studies. The remaining studies either adjusted for the same confounders across all risk factors, or unable to judge.

†Yinyan Gao and Linghui Xiang contributed equally to this work and are the first authors.

†Guochao Zhang and Irene Xinyin Wu contributed equally to this work and are the corresponding authors.

*Correspondence:
Guochao Zhang
18801038718@163.com
Irene Xinyin Wu
irenexywu@csu.edu.cn
Full list of author information is available at the end of the article

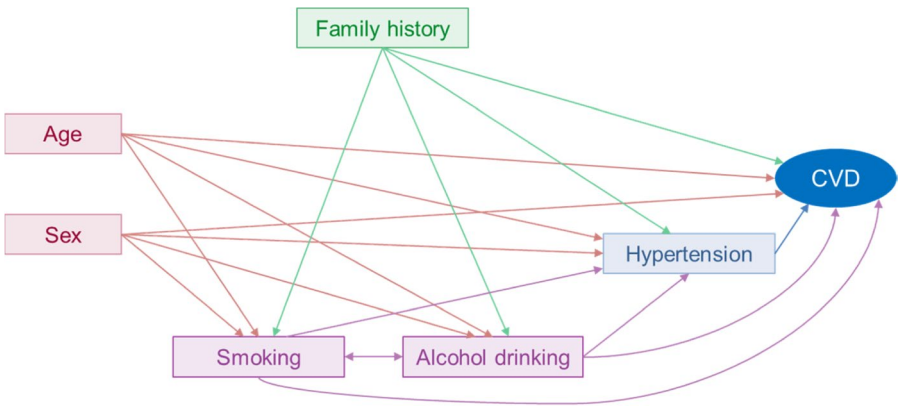
Conclusions There is substantial variation in the methods for confounder adjustment among studies investigating multiple risk factors. Mutual adjustment was the most commonly adopted method, which might lead to overadjustment bias and misleading effect estimates. Future research should avoid indiscriminately including all risk factors in a multivariable model to prevent inappropriate adjustment.

Keywords Confounder adjustment, Observational studies, Multiple risk factors, Methodological study

Background

Confounding bias significantly threatens the internal validity of causal inference research, especially in observational studies [1]. In an exposure-outcome relationship, confounders refer to a set of extraneous variables that are common causes of both the exposure and the outcome [2]. Once a focal exposure-outcome relationship is defined, it becomes easier to discern which variables act as confounders. However, in studies investigating multiple risk factors (or protective factors), multiple exposure-outcome relationships exist. Consequently, a risk factor may serve as a confounder, mediator or effect modifier in the relationships between other risk factors and the outcome. For example, in a study examining multiple risk factors for incident cardiovascular diseases (CVD) (Fig. 1), each factor plays a different role in other causal relationships of interest, rather than mutually confounding one another. Therefore, the set of confounders is specific to each risk factor-outcome relationship.

In observational studies, multivariable regression models are commonly used to estimate the exposure-outcome effects, or "independent associations", by including the interested factor and potential confounders in the model. Studies investigating multiple risk factors involve multiple risk factor-outcome relationships with different confounders; therefore, according to the principles of confounder identification and adjustment [3], it is appropriate to adjust for confounders specific to each relationship separately, thereby requiring multiple multivariable regression models. However, previous studies investigating multiple risk factors often overlooked the specific role of each risk factor in the associations between other risk factors and the outcome. For instance, substantial studies included all studied risk factors into a multivariable



The risk factors included in the confounder set of each risk factor-outcome relationship:
Sex→CVD: none
Age→CVD: none
Family history→CVD: none
Smoking→CVD: sex, age, family history, alcohol drinking
Alcohol drinking→CVD: sex, age, family history, smoking
Hypertension→CVD: sex, age, family history, smoking, alcohol drinking
Notes: A confounder is a common cause of both the exposure and the outcome, and is not a mediator that is caused by the exposure and in turn causes the outcome. Additionally, we did not present other potential confounders (such as genes, additional behaviours, and mental health) in the graph.

Fig. 1 Example of a causal graph for investigating multiple risk factors

model [4–7]. This approach means all risk factors were mutually adjusted, which might lead to coefficients for some factors measuring the "total effect" while others measure the "direct effect", potentially resulting in misleading effect estimates (i.e. the "Table 2 fallacy") [8]. Another common practice is to adjust the same confounders separately for all studied risk factors, which might also be inappropriate [9–12]. In contrast, adjusting for potential confounders for each risk factor separately was rarely seen in published studies [13, 14].

Inappropriate confounder adjustment may underestimate, overestimate or even reverse the effect size. We illustrated this with a comparison of the effect estimates between two methods of confounder adjustment: adjusting for potential confounders separately for each risk factor and mutual adjustment for all risk factors. We conducted this comparison using the data from our previously published study that explored potential factors associated with medication intake in essential tremor patients [15]. As shown in Additional file 1: Table S1, the two adjustment methods present differences in the effect estimates for certain variables (e.g. sex, education, intention tremor). In addition, Green and Popham also illustrated that mutual adjustment for multiple socioeconomic indicators (education, occupation and income) could lead to "mutual adjustment fallacy", making the mutually adjusted coefficients for each indicator incomparable [16]. While we acknowledge that these findings might be biased due to the cross-sectional design and limited number of studies, they highlighted the critical importance of appropriately adjusting for confounders to ensure accurate effect estimates in studies investigating multiple risk factors.

However, the appropriateness of confounder adjustment in studies investigating multiple risk factors was under-recognised. Previous studies have adopted various methods for confounder adjustment, and the potential issues these methods may introduce in causal inference still require clarification. In light of this, we conducted this methodological study of observational studies investigating multiple risk factors to (1) summarise and classify the methods used for confounder adjustment and their corresponding issues and (2) summarise the approaches used for confounder selection.

Methods

Table 1 provides a glossary of common terms used in this study.

Eligibility criteria

To enhance feasibility and reduce workload, we restricted the study outcomes to three major chronic diseases: CVD, diabetes and dementia.

We included all studies that met the following criteria: (1) the study design was cohort study or case–control study; (2) the study objective was to investigate multiple risk/protective factors (at least three factors), which were grouped into widely exploring potential risk factors (hypothesis-generating) and examining the associations of several specific risk factors with an outcome (hypothesis-driven) [24]; (3) the participants were adults (age ≥ 18 years); (4) multivariable regression models were adopted to estimate the effects, such as Cox regression, logistic regression, competing risk model and generalised linear regression; (5) studies published in journals listed in Science Citation Index Expanded (SCIE), as determined by the 2022 Journal Citation Reports (JCR).

Table 1 Glossary of common terms used in this study

Methods for confounder selection: The methods used to select confounders for the subsequent multivariable regression analysis, with the classification rules drawing on a previous review [17].
Methods for confounder adjustment: The methods employed to handle confounders in multivariable regression analysis, with the classification rules being developed based on the summarisation of included studies.
Total effect: The entire effect of an exposure through all causal pathways to the outcome [18].
Direct effect: The effect for certain specific pathways by blocking some causal pathways to understand the mechanism [18].
Directed acyclic graph (DAG): DAG is a non-parametric diagrammatic representation that provides a simple and transparent way to illustrate the causal paths between the exposure, outcome and other covariates, effectively aiding in the selection of confounders [19, 20].
Modified disjunctive cause criterion: This is a more practical method for confounder selection, which includes: controlling for variables that cause the risk factor, the outcome or both; excluding known instrumental variables; and including covariates that act as proxies for unmeasured variables that are common causes of both the risk factor and outcome [21].
Insufficient adjustment: The adjustment does not adequately account for all relevant confounders (unmeasured confounders were not considered in this study); therefore, the confounding bias is not adequately addressed [22]. Insufficient adjustment will cause residual confounding bias and can yield underestimates, overestimates and even sign-reversed estimates [23].
Overadjustment bias: A bias occurs when adjusting for a mediator or its downstream proxies, typically leading to a null-biased estimate of the causal effect [12].
Unnecessary adjustment: Adjusting for variables that do not impact the causal effect of interest (in expectation) but may reduce its statistical precision. It may occur when adjusting for variables that are completely outside the interested causal network (C1), only cause the exposure (C2), are the descendent variables of the exposure but not in the causal pathway (C3) or only cause outcome (C4) [12]. Of which, according to the modified disjunctive cause criterion [21], C2 and C4 can be selected for confounder adjustment.

Studies with their primary objective beyond causal inference were excluded, including prediction model development or validation, diagnostic test or methodological studies. We also excluded studies of high-dimensional exposure, such as those involving genomics, proteomics, metabolomics and gut microbiology, or those entirely focused on genes, proteins, imaging data and certain complex mixtures, because traditional methods for confounder adjustment are inapplicable to such research. Studies that described the trajectories of risk factors or outcomes were excluded. Additionally, we did not consider acute short-term outcomes (e.g. hospitalisation outcomes and acute postoperative adverse events), pre-conditions (e.g. pre-diabetes, hypertension, cognitive decline) and related death. Non-original research (e.g. letters, comments, conference abstracts and reviews) was excluded as well.

Search strategy and study selection

As this is not a systematic review, we only searched PubMed from 1 January 2018 to 31 March 2023 to identify eligible studies. The search strategy was developed using commonly used terms including "risk factors", "protective factors", "cardiovascular diseases", "diabetes", "dementia", "cohort study" and "case–control study". The detailed search strategies and results are provided in Additional file 1: Table S2. All records were imported to PICO Portal tool (available at www.picoportal.org) for study selection. Two authors independently screened the titles, abstracts and full texts based on the eligibility criteria. Disagreements were resolved by discussion or consulting a third author.

Data extraction

Data were extracted using a standard data extraction form (Additional file 1: Table S3). Two authors independently extracted data on study objectives and methods for confounder selection and adjustment. When discrepancies occurred, discussions were held with a third author to reach a consensus. The remaining data were extracted by one author and then checked by another.

Classification rules on methods for confounder adjustment

As there are no standard classification rules on methods for confounder adjustment in studies investigating multiple risk factors, we classified them based on a summarisation of the included studies, followed by group discussions and consultation with a senior epidemiologist. Ultimately, six categories were generated to classify the methods in the included studies for confounder adjustment. Detailed classification rules with typical examples are presented in Table 2. Since studies investigating multiple risk factors involve multiple factor-outcome relationships, confounders should be adjusted separately for each relationship based on the principles of confounder adjustment [3]; therefore, *category A*—"each risk factor was adjusted for potential confounders separately"—was designated as the "recommended method". A detailed explanation of each category is provided in Additional file 1: sMethods.

Data analysis

Data were summarised and presented descriptively. As all data were categorical, number (percentage) was used to describe the distribution. Given that the methods for confounder selection and adjustment may vary according to the study objectives—widely exploring potential risk factors (hypothesis-generating) and examining several specific risk factors (hypothesis-driven)—the characteristics, and methods for confounder selection and adjustment were described according to the study objective. Sankey diagram was adopted to describe the relationship between the methods for confounder selection and the methods for confounder adjustment. The characteristics of the recommended methods for confounder adjustment to others were compared. In addition, we further described the methods for confounder adjustment among the included studies published in the top 5% journals and high-impact medical and epidemiology journals (including *BMJ*, *Lancet*, *European Journal of Epidemiology* and *International Journal of Epidemiology*).

Results

Description of studies

A total of 162 studies were included. The flowchart of the literature screening and selection process is presented in Additional file 1: Fig. S1. A list of the 162 included studies with detailed extracted information is provided in Additional file 2.

The distributions of included studies by publication years, countries and journal categories are shown in Fig. 2. The most prevalent countries were China ($n=36$, 22.2%) and the USA ($n=31$, 19.1%). The journals covered 31 JCR categories, predominantly in *Medicine, General & Internal* ($n=29$, 17.9%), *Clinical Neurology* ($n=25$, 15.4%) and *Cardiac & Cardiovascular Systems* ($n=20$, 12.3%).

The basic characteristics of the included studies are displayed in Table 3. Of the 162 studies, 88 (54.3%) aimed to widely explore potential risk factors, while 74 (45.7%) focused on examining several specific risk factors. Both types of studies predominantly utilised cohort study design (≥75.0%), had first authors from

Table 2 Example statements of the methods used for confounder adjustment

A. Each risk factor was adjusted for potential confounders separately [recommended]

Example 1 (60[a]*)*: "According to the DAG, we assumed that none of the covariates has a causal relationship with age at menarche and CVD. We assumed that age, BMI, income, smoking status …were common causes of age at menopause and CVD (confounders) and were included in the model. Regarding the association between reproductive span and CVD, we additionally included age at menopause as a confounder. Other variables (e.g., hypertension) were mediators for the association between reproductive span and CVD, and not included in the model."
Example 2 (65): "For age at cancer diagnosis, the model adjusted for BMI, CCI, race/ethnicity, baseline tobacco use, education, and histology. For baseline BMI, CCI, the models adjusted for race/ethnicity, baseline tobacco use, education, rural residence, age at diagnosis, and family history of CVD. For education, the model adjusted for race/ethnicity, and rural residence. For baseline tobacco use, the model adjusted for race/ethnicity, education, and rural residence."

B. All risk factors were separately adjusted for the same confounders

B-a. All risk factors were separately adjusted for the same basic confounders[b] (sociodemographic variables)
Example 1 (14): "Logistic regression was used to calculate age-adjusted OR."
Example 2 (116): "We estimated the main effects of education, cardiovascular health, and APOE genotype on dementia risk in the full sample in separate models adjusting for baseline age, race, and sex."
Example 3 (126): "Demographics-adjusted association (adjusted for race, sex, income, and education)."
B-b. All risk factors were separately adjusted for the same potential confounders[c](basic confounders plus other variables)
Example 1 (145): "Each of the above factors was included in a separate multivariable regression model, adjusting for sex, hypertension, diabetes mellitus…"
Example 2 (155): "Four models were fitted for each of the birth characteristics: (1) without any covariates; (2) corrected for sex, year of birth, parity, and age of mother; (3) corrected for birth SES in addition to the covariates included in model 2; and (4) corrected for education level in addition to all covariates included in model 3."

C. All risk factors were mutually adjusted (i.e. including all factors in a multivariable model)

C-a. For studies that widely explored potential risk factors, all risk factors were mutually adjusted
Example 1 (13): "Multivariable Cox proportional hazards regression analysis was used to calculate hazard ratios with mutual adjustment for the included risk factors."
Example 2 (104): "Factors with $p < 0.25$ in the bivariate analysis were selected and used in the multiple logistic regression analysis to adjust the confounders."
Example 3 (151): "The adjusted model included all risk factors."
C-b. For studies that examined several specific risk factors, all risk factors were mutually adjusted on the basis of adjusting for basic confounders
Example 1 (17): "We included all the risk factors in addition to age."
Example 2 (20): "Risk factors were mutually adjusted, and we further adjusted for age, sex, and race."
Example 3 (118): "Any variable with $p < 0.25$ was then selected for inclusion in multivariate modelling with family history of dementia (parent or sibling) as a covariate."
C-c. For studies that examined several specific risk factors, all risk factors were mutually adjusted on the basis of adjusting for potential confounders
Example 1 (117): "Models were adjusted for age, sex, race and ethnicity, education, presence of health care insurance, and all risk factors included in this analysis."
Example 2 (127): "Each lifestyle factor adjusted for age, sex, education, body mass index, and history of hypertension, hypercholesterolemia and diabetes, and mutually adjusted for the other lifestyle factors and sleep duration."
Example 3 (148): "Factors with a p-value of <0.05 on the univariate analysis were included in a multivariate logistic regression analysis."

D. For studies that examined several specific risk factors, all risk factors were adjusted for the same confounders, with mutual adjustment among them being unclear

Example 1 (111): "Comparison between two categorical variables was compared with binary logistic regression, adjusted for age, sex, and nationality."
Example 2 (157): "The model was adjusted for sex, age, systolic and diastolic blood pressures, total and HDL cholesterol, diabetes, smoking, lipid-lowering, antihypertensive and rheumatic-specific drug use."

E. All risk factors were adjusted for the same variables, with some risk factors being mutually adjusted

Example (89): "The model adjusted for sex, age, education, BMI, alcohol consumption, ALT, physical activity, daily calories." (Sex, age, education status and BMI were the studied risk factors, and other studied risk factors were not listed in the set of adjusted variables)

F. Unable to judge

There is no sufficient information to judge the methods for confounder adjustment.

Notes: In the above classification of confounder adjustment methods, the meanings of specific terms are as follows: risk factors, the studied risk factors; confounders, the variables adjusted for in the multivariable regression model, thus the confounders might also be the studied risk factors, or other variables beyond these risk factors

BMI body mass index, *CCI* Charlson Comorbidity Index, *CVD* cardiovascular disease, *OR* odds ratio, *HDL* high-density lipoprotein, *ALT* alanine aminotransferase

[a] The number in parentheses is the study ID, and corresponding detailed information can be found in Additional file 2

[b] We defined basic confounders as those related to basic sociodemographic characteristics (e.g. sex, age, ethnicity, socioeconomic status, income, education, marital status)

[c] We defined potential confounders as those considering more variables in addition to basic sociodemographic characteristics

Asia (around 40.0%), held doctoral degrees (≥ 33.0%) and were affiliated with universities (> 80.0%). Compared with the studies that widely explored potential risk factors, those that examined several specific risk factors were more likely to have large sample sizes, be published in Q1 journals and include authors from the

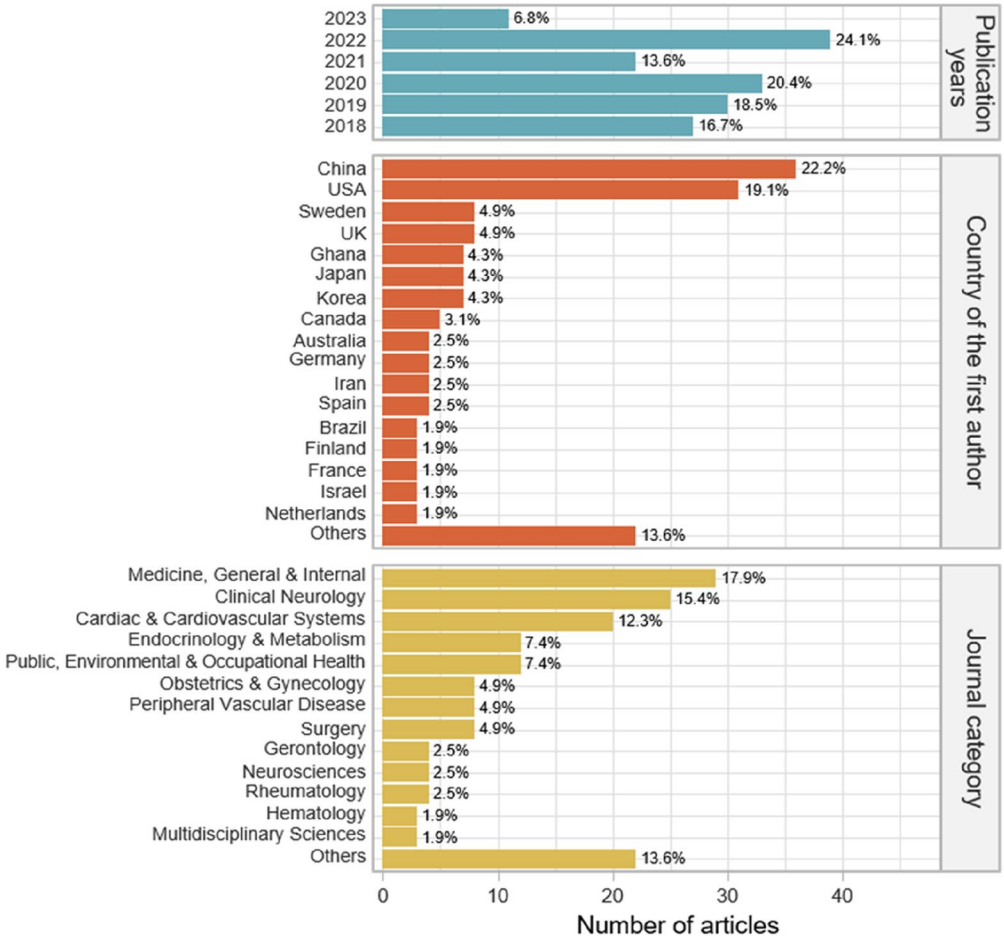


Fig. 2 Distributions of the 162 included studies by publication year, country of the first author and journal citation category

fields of epidemiology or biostatistics. In addition, most of the included studies (100% and 87.8%, respectively) exhibited unidirectional relationships between any two of the studied risk factors.

Methods for confounder selection

The summarised methods for confounder selection in the included studies are presented in Table 4. Among the 88 included studies exploring potential risk factors, 24 (27.3%) selected confounders based solely on prior knowledge, 28 (31.8%) adopted data-driven approaches (e.g. univariate analysis and stepwise method) only, and 20 (22.7%) utilised priori knowledge together with data-driven methods. Among the 74 studies examining several specific risk factors, 60 (81.1%) relied on prior knowledge only, two (2.7%) used data-driven methods only, and 10 (13.5%) used a combination of both approaches. In general, the studies examining several specific risk factors showed a significantly higher proportion of using prior knowledge compared to those exploring potential risk factors. Notably, in all included studies, only two utilised a causal graph to select confounders.

Methods for confounder adjustment

Table 5 demonstrates the methods for confounder adjustment in the included studies. Only 10 (6.2%) out of the 162 studies utilised the recommended methods (*category A*), all of which aimed at examining several specific risk factors.

Among the studies that widely explored potential risk factors, the majority ($n=77$, 87.5%) mutually adjusted for all studied risk factors, i.e. including all risk factors into a multivariable regression model (*category C-a*). In the remainder, two (2.3%) adjusted for the same basic confounders for all risk factors separately (*category B-a*), one (1.1%) adjusted for the same confounders with some of the risk factors being mutually adjusted (*category E*), and eight (9.1%) did not report sufficient information (*category F*). Notably, no study has adjusted for confounders for each risk factor separately (*category A*).

Table 3 Basic characteristics of included studies by objectives ($N = 162$)

Characteristics	Widely explored potential risk factors ($n = 88$)	Examined several specific risk factors ($n = 74$)
Study design		
Cohort study	66 (75.0)	57 (77.0)
Case–control study	22 (25.0)	17 (23.0)
Sample size		
< 200	9 (10.2)	2 (2.7)
200 ~	15 (17.1)	5 (6.8)
500 ~	17 (19.3)	4 (5.4)
1000 ~	30 (34.1)	27 (36.5)
10,000 ~	17 (19.3)	36 (48.7)
Journal Impact Factor quartiles in publication year		
Q1	34 (38.6)	47 (63.5)
Top 5%	4 (4.5)	12 (16.2)
Q2	28 (31.8)	22 (29.7)
Q3 and Q4	26 (29.5)	5 (6.8)
Region of the first author		
Europe	19 (21.6)	23 (31.1)
North America	18 (20.5)	19 (25.7)
Asia	36 (40.9)	28 (37.8)
Others	15 (17.1)	4 (5.4)
Degree of the first author		
Doctor's degree	29 (33.0)	33 (44.6)
Master's or bachelor's degree	8 (9.1)	3 (4.1)
No information	51 (58.0)	38 (51.4)
Affiliation type of the last corresponding author		
University	51 (58.0)	48 (64.9)
Hospital	11 (12.5)	2 (2.7)
Research institute	2 (2.3)	1 (1.4)
Both university and other institute	23 (26.1)	20 (27.0)
Others	1 (1.1)	3 (4.1)
Including authors from epidemiology or biostatistics		
Yes	16 (18.2)	39 (52.7)
No	51 (58.0)	23 (31.1)
Unknown	21 (23.9)	12 (16.2)
Is there a unidirectional relationship between any two of the studied risk factors?		
Yes	88 (100)	51 (68.9)
Probably yes	0 (0)	14 (18.9)
No or unsure	0 (0)	9 (12.2)
Selected outcome		
Cardiovascular diseases	57 (64.8)	40 (54.1)
Diabetes	26 (29.6)	11 (14.9)
Dementia	5 (5.7)	23 (31.1)

Values are No. (%).

Among the studies that examined several specific risk factors, more than half of them made mutual adjustment for all risk factors, either by adjusting for basic confounders (*category C-b*, 29.7%) or additional potential confounders (*category C–c*, 28.4%). Ten (13.5%) studies separately adjusted for potencial confounders for each risk factor (*category A*). The remaining studies adjusted for the same confounders for all risk factors, with all risk

Table 4 Methods for confounder selection in the included studies ($n=162$)

Methods for confounder selection	Studies widely explored potential risk factors ($n=88$)	Studies examined several specific risk factors ($n=74$)	Total
Prior knowledge only	24 (27.3)	60 (81.1)	84 (51.9)
Using causal graphs	1 (1.1)	1 (1.4)	2 (1.2)
Data-driven approaches only	28 (31.8)	2 (2.7)	30 (18.5)
Univariate analyses only	15 (17.0)	1 (1.4)	16 (9.9)
Stepwise only	9 (10.2)	0 (0)	9 (5.6)
Both univariate analysis and stepwise	3 (3.4)	1 (1.4)	4 (2.5)
Other (e.g. change in estimate)	1 (1.1)	0 (0)	1 (0.6)
Both prior knowledge and data-driven approaches	20 (22.7)	10 (13.5)	30 (18.5)
Using causal graphs	0 (0)	0 (0)	0 (0)
Insufficient information to judge	16 (18.2)	2 (2.7)	18 (11.1)

Table 5 Methods for confounder adjustment in the included studies

Methods for confounder adjustment	Studies widely explored potential risk factors ($n=88$)	Studies examined several specific risk factors ($n=74$)
A. Each risk factor was adjusted for potential confounders separately	0 (0)	10 (13.5)
B. All risk factors were separately adjusted for the same confounders		
B-a. All risk factors were separately adjusted for the same basic confounders	2 (2.3)	5 (6.8)
B-b. All risk factors were separately adjusted for the same potential confounders	0 (0)	3 (4.1)
C. All risk factors were mutually adjusted		
C-a. For studies that widely explored potential risk factors, all risk factors were mutually adjusted	77 (87.5)	NA
C-b. For studies that examined several specific risk factors, all risk factors were mutually adjusted on the basis of adjusting for basic confounders	NA	22 (29.7)
C–c. For studies that examined several specific risk factors, all risk factors were mutually adjusted on the basis of adjusting for potential confounders	NA	21 (28.4)[a]
D. For studies that examined several specific risk factors, all risk factors were adjusted for the same confounders, with mutual adjustment being unclear	NA	12 (16.2)[a]
E. All risk factors adjusted for same variables, with some risk factors being mutually adjusted	1 (1.1)	1 (1.4)
F. Unable to judge	8 (9.1)	0 (0)

NA not applicable.

[a] Among the 21 studies in the category C-c and the 12 studies in the category D, six and three studies showed no/uncertain unidirectional relationships between any two risk factors, respectively

factors not being mutually adjusted (*category B-a*, 6.8%; *category B-b*, 4.1%), some risk factors being mutually adjusted (*category E*, 1.4%), or it was unclear if risk factors were mutually adjusted (*category D*, 16.2%). Notably, among the 21 studies in *category C–c* and the 12 studies in *category D*, six and three studies showed no/uncertain unidirectional relationships between any two risk factors, respectively.

Figure 3 illustrates the relationship between the methods for confounder selection and adjustment. All studies that adjusted for confounders for each risk factor separately (*category A*) are based on prior knowledge, whereas variable selection methods that rely solely on data-driven approaches directly lead to mutual adjustment.

Table 6 outlines the characteristics of comparisons between the recommended method and other categories. In general, studies that used the recommended method to adjust for confounders were more likely to have a large sample size (> 10,000), be published in Q1 journals, be conducted in Europe and include authors from the fields of epidemiology or biostatistics.

Additional file 1: Tables S4 and S5 detail the methods for confounder adjustment in studies published in the top 5% of journals and high-impact medical and

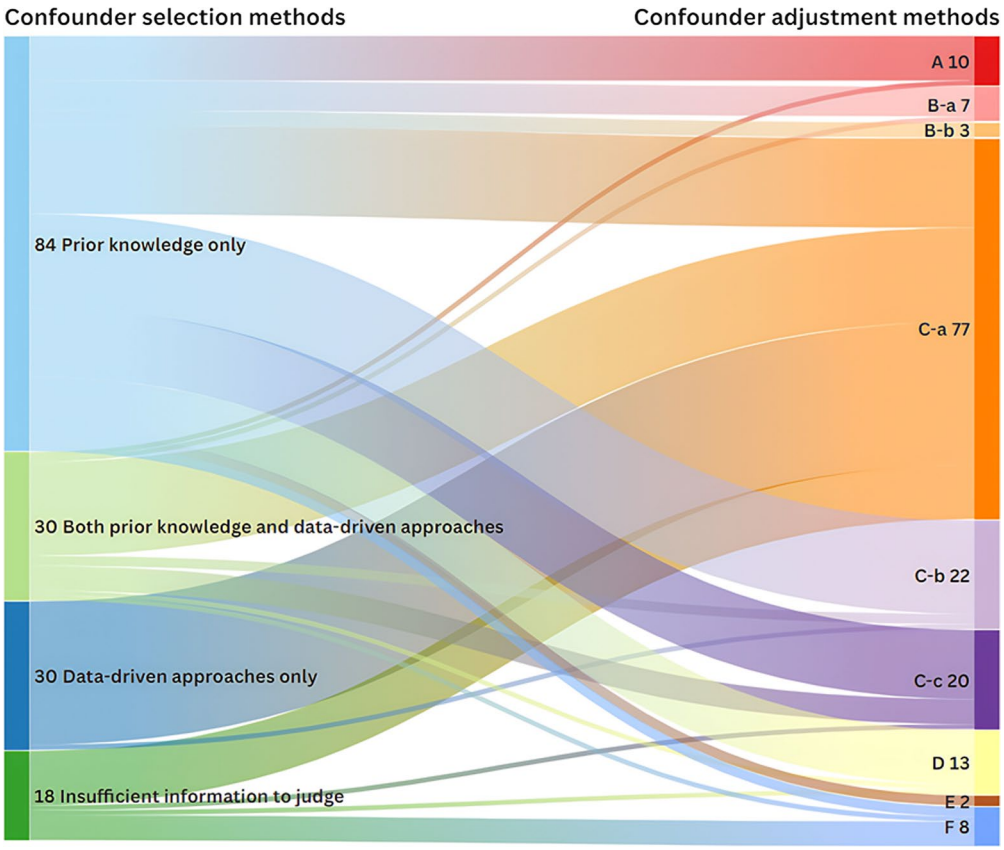


Fig. 3 Sankey diagram of the relationship between the methods for confounder selection and the methods for confounder adjustment. Notes: A, each risk factor was adjusted for potential confounders separately; B-a, all risk factors were separately adjusted for the same basic confounders; B-b, all risk factors were separately adjusted for the same potential confounders; C-a, for studies that widely explored potential risk factors, all risk factors were mutually adjusted; C-b, for studies that examined several specific risk factors, all risk factors were mutually adjusted on the basis of adjusting for basic confounders; C–c, for studies that examined several specific risk factors, all risk factors were mutually adjusted on the basis of adjusting for potential confounders; D, for studies that examined several specific risk factors, all risk factors were adjusted for the same confounders, with mutual adjustment among them being unclear; E, all risk factors adjusted for same variables, with some risk factors being mutually adjusted; F, unable to judge

epidemiology journals. Only two (published in *BMJ* and *International Journal of Epidemiology*) of the 16 studies used the recommended method.

Discussion

This methodological study provided insights concerning the methods for confounder adjustment in observational studies investigating multiple risk factors. Although the concept of confounder and the principle of confounder adjustment have been increasingly recognised in recent decades, substantial variation existed in the methods for confounder adjustment across studies investigating multiple risk factors. Among the 162 studies focused on CVD, diabetes and dementia published between January 2018 and March 2023, only ten implemented the recommended method, specifically, adjusting for potencial confounders separately for each risk factor; notably, all of these aimed at examining several specific risk factors. The most prevalent category was mutual adjustment for all risk factors. Regarding methods used for confounder selection, studies examining several specific risk factors relied on prior knowledge more frequently and data-driven approaches less frequently, compared to studies that widely explored potential risk factors.

Observations and remarks

Among the included studies, only one of the identified six categories of methods for confounder adjustment was recommended. The methods classified in the remaining categories—either separately adjusting for the same confounders or conducting mutual adjustment for all risk factors—may lead to various adjustment issues, including insufficient adjustment [22], overadjustment and unnecessary adjustment [12]. Of which, insufficient adjustment

Table 6 Comparison of characteristics between the recommended method and other categories

Characteristics	Studies that adopted the recommended method or not	
	Adopted ($n=10$)	Not adopted ($n=152$)
Study design		
Cohort study	10 (100)	114 (75.0)
Case–control study	0 (0)	38 (25.0)
Sample size		
< 200	0 (0)	11 (7.2)
200 ~	0 (0)	20 (13.1)
500 ~	0 (0)	21 (13.8)
1000 ~	2 (20.0)	55 (36.2)
10,000 ~	8 (80.0)	45 (29.6)
Journal Impact Factor quartiles		
Q1	9 (90.0)	72 (47.4)
Top 5%	2 (20.0)	14 (9.2)
Q2	1 (10.0)	49 (32.2)
Q3 and Q4	0 (0)	31 (20.4)
Region of the first author		
Europe	5 (50.0)	37 (24.3)
North America	2 (20.0)	35 (23.0)
Asia	2 (20.0)	62 (40.8)
Others	1 (10.0)	18 (11.8)
Degree of the first author		
Doctor's degree	4 (40.0)	58 (38.2)
Master's or bachelor's degree	1 (40.0)	10 (6.6)
No information	5 (40.0)	84 (55.3)
Affiliation type of the corresponding author		
University	8 (80.0)	86 (56.6)
Hospital	1 (10.0)	17 (11.2)
Research institute	0 (0)	3 (2.0)
Both university and other institute	1 (10.0)	42 (27.6)
Others	0 (0)	4 (2.6)
Including authors from epidemiology or biostatistics		
Yes	7 (70.0)	48 (31.6)
No	2 (20.0)	72 (47.4)
Unknown	1 (10.0)	32 (21.1)

Values are No. (%)

and overadjustment may mask the true effect of interest, while unnecessary adjustment does not affect the true effect (in expectation) but may reduce its statistical precision, which is acceptable to some extent [21].

Most of the included studies adjusted for confounders by entering all risk factors into a multivariable regression model, indicating that all risk factors were mutually adjusted. This method is appropriate only when all the studied risk factors are confounded with each other, which depends on the relationships among them. If there is a unidirectional relationship between any two of the studied risk factors (e.g. age and hypertension), the reported associations can be puzzling (called "Table 2 fallacy"), as the association for hypertension is considered a "total effect" while the association for age is "direct effect" [8]. Therefore, this method would cause overadjustment bias for the upstream risk factors (e.g. sex, age, education, family history), because the adjustment of relatively

downstream risk factors (e.g. behavioural factors, diseases) might adjust for mediators. In contrast, if there is no unidirectional relationship between the studied risk factors, but rather a correlation and shared the same set of confounders, as seen among lifestyle factors (such as smoking, diet, alcohol consumption and physical activity), the mutual adjustment is reasonable [25]. However, the vast majority (more than 95%) of the included studies investigating multiple risk factors exhibited unidirectional relationships between studied risk factors.

Some included studies have separately adjusted for the same confounders for all risk factors. Most of them only adjusted for basic confounders, such as age and sex [26–28], which might lead to insufficient adjustment when these were not the minimally sufficient adjustment set for some risk factors. While the remaining small fraction of studies adjusted for variables in addition to basic confounders, this was only appropriate when the studied risk factors were independent of each other and shared the same set of confounders. Otherwise, this may lead to both insufficient adjustment and unnecessary adjustment. In addition, due to the poor reporting on confounder adjustment, in studies that adjusted for the same confounders, mutual adjustment among them was unclear [29–31]. Therefore, these studies would introduce the aforementioned biases based on their actual adjustment practices.

In comparison to studies exploring potential risk factors (hypothesis-generating), those examining several specific risk factors (hypothesis-driven) more frequently utilised a prior knowledge approach for confounder selection, employed the recommended method for confounder adjustment and were more often published in Q1 journals. This could be attributed to several reasons. First, studies examining several specific risk factors were more likely to involve authors specialised in epidemiology or biostatistics, who could enhance the quality of confounder adjustment. Second, the selected outcomes were common diseases (CVD, diabetes and dementia), which have already been extensively studied for potential risk factors. Thus, studies exploring potential risk factors might not be prioritised for publication in high-impact journals. Finally, the high-impact journals were more likely to enforce compliance with reporting checklists and have submissions reviewed by epidemiologists or biostatisticians.

In addition, all ten studies that used the recommended method for confounder adjustment were studies examining several specific risk factors (hypothesis-generating). Although studies exploring potential risk factors aim to generating hypotheses, often using statistical significance as the criterion for identifying potential risk factors, it is still necessary to adjust for confounders following the confounder adjustment rules (e.g. cannot adjust for mediators). Otherwise, some potential factors may be obscured, as shown in our previously mentioned example (Additional file 1: Tables S1). Furthermore, in hypothesis-generating studies, while relationships between various risk factors and the outcome may not be explicitly recognised, the relationships across many of these risk factors are often well-established (e.g. gender, age, education level, smoking and alcohol use); therefore, inappropriate adjustment should be avoided.

Recommendations and implications

In studies investigating multiple risk factors, either widely exploring potential risk factors (hypothesis-generating) or examining several specific risk factors (hypothesis-driven), it is recommended to adjust for confounders for each risk factor-outcome relationship separately, rather than simply including all factors in a multivariable model. The causal graph is strongly suggested to identify confounder set for each risk factor-outcome relationship. When employing the directed acyclic graph (DAG), it is necessary to create multiple DAGs for all interested risk factor-outcome relationships, as a single DAG can only specify one exposure [19, 20]. Since complete knowledge of the relationships among all covariates is often unavailable, constructing a DAG can be challenging. VanderWeele proposed a more practical method, the modified disjunctive cause criterion, is also recommended, which includes: controlling for variables that cause the risk factor, the outcome or both; excluding known instrumental variables; and including covariates that act as proxies for unmeasured variables that are common causes of both the risk factor and outcome [21]. This criterion is simpler and requires less time than DAG, and similarly, the set of confounders for each risk factor-outcome relationship should be established separately.

Data-driven approaches, such as univariate analysis and stepwise methods, were commonly used to select variables. These methods cannot distinguish between confounders, which require adjustment, and mediators, which do not; therefore, these methods lead to mutual adjustment among the selected variables, increasing the risk of introducing adjustment bias and producing invalid causal inferences [17]. When sample sizes are large enough, we recommend relying exclusively on knowledge-based approaches. While data-driven methods are essential for small datasets and rare events, they should be integrated with knowledge-based approaches before application.

There is considerable room for improving the transparency of reporting on confounder adjustment. We strongly recommend authors to follow the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement. Journals should take action to improve the reporting quality. The strategy for advancing transparency in the reporting of observational studies,

implemented by the *PLOS Medicine* editors, provided a possible solution [32]. It mandates authors to not only complete the STROBE checklist with page references but also incorporate relevant text excerpted from the manuscript to elucidate their compliance with each item [32].

Strengths and limitations

This is the first study that investigated the methods for confounder adjustment and related issues in studies investigating multiple risk factors, shedding light on future research for confounder selection and adjustment.

There were several limitations. First, instead of conducting a systematic review, which is out of our scope, we only included studies published within the last 5 years and restricted the outcomes to three common chronic diseases, which may limit the generalisability of our findings. Nonetheless, we included 162 articles from 31 journal categories, which we believe can provide insight into the current confounder adjustment status of studies investigating multiple risk factors, to some extent. Second, due to the included studies spanning the years covering the period of the COVID-19 pandemic, therefore the methodological quality of the included studies might be relatively lower. Third, due to the lack of standardised guidelines or recommendations for confounder adjustment in studies investigating multiple risk factors, we identified potential methods from the included studies and defined the recommended approach without formal validation (e.g. external expert consultation), which may inevitably introduce subjectivity. However, we classified the methods through group discussions and consultation with a senior epidemiologist and defined the recommended method based on general causal inference principles, which could reduce misclassification. Additionally, we provided detailed justifications for confounder adjustment in each included study to enhance transparency and objectivity. Fourth, in assessing the methods for confounder adjustment, our judgement was based solely on whether confounders were adjusted separately for each risk factor, without evaluating the appropriateness of these adjusted confounders. Lastly, we did not assess the direction and magnitude of the bias introduced by the inappropriate methods for confounder adjustment on the effect estimates. While this was beyond the scope of our study, it is an area that warrants further exploration in future research.

Conclusions

The methods for confounder adjustment in observational studies investigating multiple risk factors have not received adequate attention. There was substantial variation across the included studies. Only a few studies employed the recommended method, specifically, adjusting for confounders separately for each risk factor, whereas mutual adjustment was the most commonly adopted method. We recommended researchers consider confounders for each risk factor-outcome relationship separately, rather than simply including all factors in a multivariable model.

Abbreviations

CVD	Cardiovascular diseases
DAG	Directed acyclic graph
JCR	Journal Citation Reports
SCIE	Science Citation Index Expanded
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12916-025-03957-8.

Additional file 1. sMethods. Classification rules on methods for confounder adjustment. Fig. S1 Flowchart of literature screening and selection process. Table S1 Example of comparing effect estimates between two confounder adjustment methods for exploring potential factors associated with medication intake in essential tremor patients. Table S2 Search strategy in PubMed (searched in 27 Mar 2023). Table S3 Data extraction form. Table S4 Confounder adjustment methods of studies published in top 5% journals and high impact medical and epidemiology journals. Table S5 Details of confounder adjustment methods in the top 5% journals and high impact medical and epidemiology journals.

Additional file 2. Detailed extracted information from the included studies.

Acknowledgements

Thanks to Prof. Hongzhuan Tan from Central South University for his valuable suggestions on the methodology of the manuscript. Thanks to Dr. Ting Gan from Queensland University of Technology for her kind assistance with English grammar checking. Thanks to the peer reviewers and editors for their valuable comments that helped improve the manuscript. Thanks to ChatGPT for English editing.

Authors' contributions

IXYW and YYG conceptualised and designed the study. YYG conducted the literature search. YYG, LHX and HY performed the literature selection, data extraction and analysis. YYG and LHX wrote the draft. IXYW, GCZ, SJL, DKS and BYX critically revised the manuscript. All authors read and approved the final manuscript.

Funding

The work was supported by the National Natural Science Foundation of China (82203154); Fundamental Research Funds for the Central Universities (No. 3332024205), the Natural Science Foundation of Hunan Province (No. 2023JJ30734) and the Graduate Student Innovative Scientific Research Project of Central South University, China (2024ZZTS0558). The funding source was not involved in the study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Data availability

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

Declarations

Ethics approval and consent to participate

Not applicable.

Consent to publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Epidemiology and Biostatistics, Xiangya School of Public Health, Central South University, Changsha, China. [2]Thoracic Surgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China. [3]Hunan Provincial Key Laboratory of Clinical Epidemiology, Central South University, Changsha, China.

Received: 9 October 2024 Accepted: 17 February 2025
Published online: 05 March 2025

References

1. Hemkens LG, Ewald H, Naudet F, Ladanie A, Shaw JG, Sajeev G, Ioannidis JPA. Interpretation of epidemiologic studies very often lacked adequate consideration of confounding. J Clin Epidemiol. 2018;93:94–102.
2. VanderWeele TJ, Shpitser I. On the definition of a confounder. Ann Stat. 2013;41(1):196–220.
3. Lash TL, VanderWeele TJ, Haneuse S, Rothman KJ. Modern epidemiology, 4th edition. Lippincott Williams & Wilkins (LWW). 2021.
4. Too G, Wen T, Boehme AK, Miller EC, Leffert LR, Attenello FJ, Mack WJ. D'Alton ME, Friedman AM: Timing and risk factors of postpartum stroke. Obstet Gynecol. 2018;131(1):70–8.
5. Ismail M, Mok VC, Wong A, Au L, Yiu B, Wang Z, Chu WC, Chan AY, Fan FS, Ma SH, et al. Risk factors for delayed-onset dementia after stroke or transient ischemic attack-a five-year longitudinal cohort study. Int J Stroke. 2022;17(5):517–25.
6. Sarfo FS, Ovbiagele B, Akpa O, Akpalu A, Wahab K, Obiako R, Komolafe M, Owolabi L, Ogbole G, Calys-Tagoe B, et al. Risk factor characterization of ischemic stroke subtypes among West Africans. Stroke. 2022;53(1):134–44.
7. Lopez-Jaramillo P, Joseph P, Lopez-Lopez JP, Lanas F, Avezum A, Diaz R, Camacho PA, Seron P, Oliveira G, Orlandini A, et al. Risk factors, cardiovascular disease, and mortality in South America: a PURE substudy. Eur Heart J. 2022;43(30):2841–51.
8. Westreich D, Greenland S. The table 2 fallacy: presenting and interpreting confounder and modifier coefficients. Am J Epidemiol. 2013;177(4):292–8.
9. Gardner RC, Barnes DE, Li Y, Boscardin J, Peltz C, Yaffe K. Medical and psychiatric risk factors for dementia in veterans with and without traumatic brain injury (TBI): a nationwide cohort study. J Prev Alzheimers Dis. 2023;10(2):244–50.
10. Mosing MA, Lundholm C, Cnattingius S, Gatz M, Pedersen NL. Associations between birth characteristics and age-related cognitive impairment and dementia: a registry-based cohort study. PLoS Med. 2018;15(7): e1002609.
11. Tomata Y, Li X, Karlsson IK, Mosing MA, Pedersen NL, Hägg S. Joint impact of common risk factors on incident dementia: a cohort study of the Swedish Twin Registry. J Intern Med. 2020;288(2):234–47.
12. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. Epidemiology. 2009;20(4):488–95.
13. Peters SAE, Carcel C, Millett ERC, Woodward M. Sex differences in the association between major risk factors and the risk of stroke in the UK Biobank cohort study. Neurology. 2020;95(20):e2715–26.
14. Millett ERC, Peters SAE, Woodward M. Sex differences in risk factors for myocardial infarction: cohort study of UK Biobank participants. BMJ. 2018;363: k4247.
15. Xiang L, Zhou X, He R, Gao Y, Li M, Zeng S, Cao H, Wang X, Xu Y, Zhao G, et al. Medication status and related factors in essential tremor patients: a cross-sectional study in China. Neuroepidemiology. 2023;57(4):260–70.
16. Green MJ, Popham F. Interpreting mutual adjustment for multiple indicators of socioeconomic position without committing mutual adjustment fallacies. BMC Public Health. 2019;19(1):10.
17. Talbot D, Massamba VK. A descriptive review of variable selection methods in four epidemiologic journals: there is still room for improvement. Eur J Epidemiol. 2019;34(8):725–30.
18. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. Epidemiology. 1992;3(2):143–55.
19. Tennant PWG, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, Harrison WJ, Keeble C, Ranker LR, Textor J, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. Int J Epidemiol. 2021;50(2):620–32.
20. Ferguson KD, McCann M, Katikireddi SV, Thomson H, Green MJ, Smith DJ, Lewsey JD. Evidence synthesis for constructing directed acyclic graphs (ESC-DAGs): a novel and systematic method for building directed acyclic graphs. Int J Epidemiol. 2020;49(1):322–9.
21. VanderWeele TJ. Principles of confounder selection. Eur J Epidemiol. 2019;34(3):211–9.
22. Evans D, Chaix B, Lobbedez T, Verger C, Flahault A. Combining directed acyclic graphs and the change-in-estimate procedure as a novel approach to adjustment-variable selection in epidemiology. BMC Med Res Methodol. 2012;12:156.
23. van Zwieten A, Tennant PWG, Kelly-Irving M, Blyth FM, Teixeira-Pinto A, Khalatbari-Soltani S. Avoiding overadjustment bias in social epidemiology through appropriate covariate selection: a primer. J Clin Epidemiol. 2022;149:127–36.
24. Staerk C, Byrd A, Mayr A. Recent methodological trends in epidemiology: no need for data-driven variable selection? Am J Epidemiol. 2024;193(2):370–6.
25. Yang J, Qian F, Chavarro JE, Ley SH, Tobias DK, Yeung E, Hinkle SN, Bao W, Li M, Liu A, et al. Modifiable risk factors and long term risk of type 2 diabetes among individuals with a history of gestational diabetes mellitus: prospective cohort study. BMJ. 2022;378: e070312.
26. Aradine EM, Ryan KA, Cronin CA, Wozniak MA, Cole JW, Chaturvedi S, Dutta TLM, Hou Y, Mehndiratta P, Motta M, et al. Black-White differences in ischemic stroke risk factor burden in young adults. Stroke. 2022;53(3):e66–9.
27. Bahall M, Seemungal T, Legall G. Risk factors for first-time acute myocardial infarction patients in Trinidad. BMC Public Health. 2018;18(1):161.
28. Ezquerra-Lázaro I, Cea-Soriano L, Giraldez-García C, Ruiz A, Franch-Nadal J, Diez-Espino J, Nogales P, Carramiñana F, Javier Sangros F, Regidor E. Lifestyle factors do not explain the difference on diabetes progression according to type of prediabetes: results from a Spanish prospective cohort of prediabetic patients. Diabetes Res Clin Pract. 2019;153:66–75.
29. Pantell MS, Prather AA, Downing JM, Gordon NP, Adler NE. Association of social and behavioral risk factors with earlier onset of adult hypertension and diabetes. JAMA Netw Open. 2019;2(5): e193933.
30. Kozela M, Doryńska A, Bobak M, Pająk A. Accumulation of psychosocial risk factors and incidence of cardiovascular disease: a prospective observation of the Polish HAPIEE cohort. Kardiol Pol. 2019;77(5):535–40.
31. Yong HY, Mohd Shariff Z, Mohd Yusof BN, Rejali Z, Tee YYS, Bindels J, van der Beek EM. Independent and combined effects of age, body mass index and gestational weight gain on the risk of gestational diabetes mellitus. Sci Rep. 2020;10(1):8486.
32. PLOS Medicine Editors. Observational studies: getting clear about transparency. PLoS Med. 2014:11(8):e1001711.

Publisher's Note