

PROTOCOL

Open Access



Causal inference in real-world dementia research: a systematic review protocol

Yi Yang^{1*} , Shivani Suresh¹, Jiamin Du^{1,2}, Rachel Brohier¹, Hongen Ma³, Ivan Koychev¹ and Vanessa Raymont^{1*}

Abstract

Background Dementia presents complex challenges for causal inference due to its multifactorial aetiology and slow, heterogeneous progression. Randomized controlled trials are often limited in their potential to fully address these challenges because of ethical and practical constraints. As the field evolves, observational studies incorporating advanced causal inference methods are increasingly used to estimate real-world effects in dementia research. However, the implementation of these methods varies widely and has not been systematically evaluated, with an emerging trend towards integration with techniques such as machine learning. This systematic review will critically examine how causal inference techniques are applied in dementia research, assess their methodological rigor, and identify trends, assumptions, and gaps that may inform future applications and methodological innovation in the field.

Methods Following PRISMA guidelines, searches will be conducted in MEDLINE, EMBASE, Web of Science, PsycINFO, Scopus, and the Cochrane Library for studies published between 1960 and 2024. Eligible studies will include observational designs that use causal inference methods to investigate outcomes such as cognitive decline, disease progression, and quality of life. Data extraction will capture study characteristics, methodological details, and key findings, with risk of bias assessed using ROBINS-I. A narrative synthesis will summarize qualitative results, and meta-analyses will be performed when methodological homogeneity permits.

Discussion This review will address a critical gap in the evaluation of the application of causal inference methods in real-world dementia research. By identifying methodological challenges, underlying assumptions, and emerging analytical techniques, it aims to strengthen research rigor and reproducibility and inform future methodological development, with potential implications for policy and practice in dementia care.

Systematic review registration PROSPERO (CRD42024619228).

Keywords Dementia, Cognitive decline, Causal inference, Observational study, Systematic review

Strengths and limitations of this study

- This review addresses a key gap by evaluating the current landscape and emerging trends in causal inference methodologies applied to dementia research using real-world data.
- By synthesizing evidence on causal inference methods in dementia research, this review aims to strengthen methodological rigor and inform the development of more robust, evidence-based research and practice.

*Correspondence:

Yi Yang

yi.yang@psych.ox.ac.uk

Vanessa Raymont

vanessa.raymont@psych.ox.ac.uk

¹ Department of Psychiatry, University of Oxford, Oxford, UK

² NIHR Applied Research Collaboration Oxford and Thames Valley (NIHR ARC OxTV), Oxford Health NHS Foundation Trust, Oxford, UK

³ Clinical Psychology Department, University of Edinburgh, Edinburgh, UK



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

- The review focuses on English-language, peer-reviewed observational studies, which may limit global representation and exclude insights from non-English publications, grey literature, and experimental research.

Background

Dementia and related neurodegenerative diseases represent a growing global health concern, characterized by complex aetiologies, prolonged disease courses, and a significant societal burden [30, 41]. Dementia is classified under “Major Neurocognitive Disorder” in DSM-5 as a clinical syndrome marked by significant and persistent declines in cognitive functions—such as memory, reasoning, language, and executive functioning—that interfere with daily life and activities [1, 40]. Given that dementia arises from an interplay of genetic, environmental, and lifestyle factors, understanding which factors are causally linked to its onset and progression is essential for designing effective prevention and intervention strategies [37].

Randomized controlled trials (RCTs) are traditionally regarded as the gold standard for establishing causality, but their application in dementia research faces several limitations. The slow progression of the disease, combined with the difficulty of recruiting representative samples of older adults with comorbidities or frailty, often constrains the feasibility of RCTs [18, 29]. Additionally, randomizing participants to potentially harmful exposures or withholding beneficial treatments raises ethical concerns [12]. The prolonged follow-up periods required for dementia studies further increase the cost and complexity of conducting RCTs, making them impractical for many research questions [2, 25]. As a result, observational designs are becoming a critical alternative for investigating real-world exposures and outcomes in dementia research, where RCTs are unfeasible [5, 34].

However, traditional statistical methods used in observational studies, such as correlational or predictive analyses, are limited in their ability to establish causality. These approaches focus on identifying associations but often fail to address confounding, selection bias, and time-varying exposures, which are crucial for causal analysis [14, 33]. Furthermore, they cannot account for counterfactual scenarios, which limits their ability to rigorously estimate causal effects or test causal hypotheses [33]. While there is a growing trend in the use of machine learning techniques, which leverage advanced algorithms and computational power to identify patterns and generating hypotheses from large datasets, they too are limited in addressing biases and confounding without a causal framework [7]. Integrating the exploratory capabilities of machine learning with the rigor of causal inference techniques can provide robust and actionable insights into

complex healthcare datasets, representing a promising direction for future research [7].

Causal inference methods are increasingly recognized as an essential framework for real-world observational studies, addressing the limitations of traditional statistical approaches and RCTs while driving advancements in research methodologies and evidence generation (Listl et al., 2016; [31]). These methods, such as propensity score matching, instrumental variables, and difference-in-differences, aim to mimic the rigor of RCTs by addressing confounding, selection bias, and unmeasured confounders in observational data [14, 21, 33]. For instance, propensity score methods (e.g., matching, weighting, stratification) address confounding in observational studies by summarizing multiple observed covariates into a single score, enabling the creation of comparable groups for estimating treatment effects. These methods offer a flexible and efficient way to reduce bias from measured confounders, but rely on the critical assumption of no unmeasured confounding [3, 4]. In contrast, instrumental variable analysis estimates causal effects in the presence of unmeasured confounding by using external variables—known as instruments—that influence the exposure but affect the outcome only through that exposure [15, 16, 20]. Meanwhile, difference-in-differences evaluates causal effects in observational studies by comparing changes in outcomes between treated and control groups over time. It allows for the estimation of treatment effects while accounting for unobserved, time-invariant confounding, under the assumption that both groups would have followed parallel trends in the absence of the intervention [35, 39]. Together, these methods provide essential tools for strengthening causal inference in observational dementia research, enabling more robust estimation of the effects of genetic, lifestyle, and environmental factors on dementia risk and progression.

Despite their advantages, causal inference methods require rigorous application and careful validation to ensure the reliability and validity of findings. A central challenge is addressing unmeasured confounding—methods like propensity scores can only balance observed covariates, and unmeasured variables cannot be corrected post hoc [13, 27, 43]. Proper model specification is also essential: statistical models must accurately reflect the relationships between exposures, confounders, and outcomes, as misspecification can lead to biased or invalid estimates [33, 38]. Instrumental variable methods, in particular, require strong instruments that are associated with the exposure, independent of unmeasured confounders, and influence the outcome solely through the exposure [11, 20]. Sensitivity analyses are critical for evaluating the

robustness of causal estimates, and high-quality data that comprehensively captures relevant covariates is vital for drawing reliable conclusions [23, 36]. Finally, inconsistency in how these methods are applied and reported in dementia research raises concerns about comparability and generalizability across studies [6, 44].

Methodology-focused systematic reviews play a pivotal role in advancing research by critically evaluating the rigor, consistency, and appropriateness of applied analytical techniques. In dementia research, where unmeasured confounding, population heterogeneity, and the prolonged disease course pose persistent challenges, this review will assess how causal inference methods are being implemented to address these complexities. By identifying best practices, common methodological pitfalls, and underutilized or emerging techniques, it aims to support researchers in adopting more robust, transparent, and reproducible approaches. In doing so, the review contributes not only to strengthening the methodological foundations of future studies, but also to identifying emerging directions and innovations, thereby highlighting methodological developments worth watching in the evolving landscape of causal inference in dementia research. Ultimately, this systematic review aims to enhance the reliability and policy relevance of findings in dementia prevention and care.

Systematic review

Aim

This review aims to evaluate the application of causal inference methods in dementia research to support the generation of robust findings that can inform clinical and policy interventions. The review will address the following questions:

1. Topics: What causal factors have been evaluated for their relationship with dementia and neurodegenerative diseases in observational research?
2. Methods: Which causal inference methods (e.g., instrumental variables, difference-in-differences, target trial emulation, propensity score matching, g-methods) are applied in dementia research?
3. Datasets: How are causal inference methods tailored to different data types (e.g., longitudinal, cross-sectional, pooled datasets), and which approaches are suitable for specific study designs?
4. Causality and Assumptions: How do causal inference studies address challenges such as confounding, bias, and endogeneity, and how consistently are assumptions tested and reported?
5. Rigor and Heterogeneity: How do causal inference studies in dementia research address heterogeneity, such as population diversity and disease characteristics, and how do these variations influence the interpretation and the generalizability of findings?
6. Outcomes: What outcomes are examined using causal inference methods in dementia research, such as cognitive decline, disease progression, or intervention efficacy?
7. Emerging Methods and Gaps: What emerging technologies (e.g., machine learning) and methodologies are advancing causal inference in dementia research, and what gaps remain in their application?
8. Limitations: What practical challenges and systemic barriers influence the application of causal inference methods in observational dementia research, such as those related to implementation, scalability, and data quality?

Methodology

Registration: This study protocol was registered with PROSPERO (CRD42024619228).

Inclusion criteria

- Studies involving individuals with any form of dementia, such as Alzheimer's disease, Parkinson's disease, Lewy body dementia, cognitive decline (including those without an official diagnosis), or mild cognitive impairment (MCI), identified through validated screening tools or clinical diagnoses.
- Human studies only, with no demographic or geographic restrictions.
- Observational studies, including, case-control, cross-sectional, cohort, longitudinal, and pooled multi-cohort analyses, that investigate exposures or interventions related to neurodegenerative diseases or cognitive decline.
- Studies applying causal inference techniques, such as instrumental variables, difference-in-differences, propensity score matching, target trial emulation, G-methods, and regression discontinuity designs.
- Studies published in English language only.
- There is no restriction on exposures or interventions, such as studies examining the following:
 - (1) Lifestyle factors: physical activity, diet, smoking, alcohol consumption, and sleep patterns.
 - (2) Environmental factors: air pollution, occupational exposures, and access to healthcare.
 - (3) Socioeconomic and psychological factors: education level, social isolation, and depression.
 - (4) Biological and genetic factors: biomarkers (e.g., amyloid beta, tau proteins) and genetic predispositions (e.g., APOE ϵ 4 allele).

- (5) Medical and pharmacological interventions: use of medications (e.g., cholinesterase inhibitors), therapies (e.g., cognitive behavioural therapy), or surgical treatments (e.g., deep brain stimulation).
- (6) Digital and technological interventions: use of assistive technologies or telemedicine in managing neurodegenerative conditions.

Exclusion criteria

- Studies focusing on conditions outside the scope of dementia, cognitive decline, mild cognitive impairment, or related neurodegenerative diseases.
- Studies conducted exclusively in experimental settings (e.g., RCTs) or involving purely experimental interventions that lack relevance to real-world observational contexts, unless they include an observational component where causal inference methods are applied.
- Animal or in vitro studies.
- Reviews, descriptive studies, or purely associational studies that do not employ causal inference methods.
- Theoretical models or simulations without empirical data.
- Studies lacking sufficient methodological detail to establish causal relationships between exposures/interventions and outcomes.

Information sources

Peer-reviewed studies will be identified through comprehensive searches in electronic bibliographic databases, including MEDLINE/PubMed, EMBASE, Web of Science, PsycINFO, Scopus, and the Cochrane Library. These databases will be searched systematically to retrieve relevant studies from their inception to the present. Additionally, the reference lists of included studies will be reviewed to identify any further eligible studies that may not have been captured in the initial searches. The search strategy will be periodically updated to ensure the retrieval of the most recent and relevant literature in line with the systematic review's objectives.

Search strategy

The search strategy covers publications from 1960 to 2024 to capture a comprehensive timeline of relevant research. Only English language studies are included to ensure consistency in data extraction and analysis. An initial exploration of Medical Subject Headings (MeSH) and text words related to key concepts are conducted across all selected databases to identify terms relevant to the research questions. Titles, abstracts, and index terms of retrieved articles will be analysed iteratively to refine the strategy and enhance its sensitivity and specificity.

Two independent reviewers will screen titles, abstracts, and full texts for eligibility, with disagreements resolved through discussion or consultation with a third reviewer. The search term framework is provided in Table 1. A detailed list of search terms and results is provided in the table in supplementary material and available on the Open Science Framework: <https://osf.io/kyzvh>. Table 1.

Outcomes

There is no restriction on outcomes, and examples can be the following:

1. Cognitive Decline: Progressive deterioration in cognitive abilities, measured through standardized tests such as the Mini-Mental State Examination (MMSE [10];) and Montreal Cognitive Assessment (MoCA, [28]).
2. Disease Progression: Worsening symptoms or stages tracked through clinical scales like the Clinical Dementia Rating (CDR; [26]) or Hoehn and Yahr Scale [17], or biomarkers such as amyloid-beta and tau proteins.
3. Quality of Life: Well-being and functionality assessed using tools like the Quality of Life in Alzheimer's Disease (QoL-AD; Logsdon et al., [24]) and EuroQol 5 Dimensions (EQ-5D; EuroQol Group, [9]), including caregiver-reported assessments.
4. Mortality and Survival: Documented death due to neurodegenerative diseases or related complications, or time from diagnosis/symptom onset to death or study endpoint, analysed using survival models such as Kaplan–Meier curves.
5. Functional Decline: Loss of independence or ability to perform daily activities, assessed using scales such as Activities of Daily Living (ADL; [19]) and Instrumental Activities of Daily Living (IADL [22];).
6. Behavioural and Psychological Symptoms: Presence of conditions like depression, anxiety, and agitation, assessed with validated scales like the Neuropsychiatric Inventory (NPI [8];).
7. Neurological or Imaging Outcomes: Changes in brain structure or function measured through neuroimaging techniques like magnetic resonance imaging (MRI) and positron emission tomography (PET).
8. Social and Psychosocial Outcomes: Levels of social engagement or support, evaluated through self-reports, caregiver reports, or social network analyses.
9. Treatment Adherence and Adverse Effects: Treatments adherence and reporting of adverse effects,

Table 1 Search term framework for identifying relevant studies in the literature search

Dementia terms	Causal inference terms	Exclusion terms
Dementia	Causal inference	Randomised control trial
Alzheimer*	Causal effect*	RCT
Cognitive decline	Causal estimation*	Review
Cognitive impairment	Counterfactual	
Corticobasal degeneration	Counterfactual framework	
Creutzfeldt Jakob syndrome	Inverse probability of treatment weighting	
Dementia with Lewy bodies	IPTW	
Lewy body dementia	Inverse probability weighting	
Frontotemporal lobar degeneration	IPW	
Frontotemporal dementia	Instrumental variable*	
Late onset dementia	Two-stage least squares	
Memory disorder*	2SLS	
Memory impair*	Generalized method of moments	
Neurodegenerat*	GMM	
Neurocognitive disorder*	Difference-in-differences	
neurodegenerative disease*	DiD	
Parkinson's dementia	Fixed-effects model*	
Pick's disease	Fixed-effects panel data analysis	
Mild cognitive impairment	Propensity score	
	Propensity score matching	
MCI	PSM	
	Inverse probability weighted estimators	
	G-method*	
	G-computation	
	G-formula	
	Marginal structural model*	
	MSM	
	Marginal structural Cox model	
	Mendelian Randomization	
	Target trial emulation	
	Regression discontinuity design*	
	RDD	
	Regression kink design*	
	Structural causal model*	
	SCM	
	Directed acyclic graph*	
	DAG	
	Causal graphical models	
	Causal mediation analysis	
	Natural direct effect	
	NDE	
	Natural indirect effect	
	NIE	
	Synthetic control method*	
	Comparative case study causal inference	
	Causal machine learning	
	Targeted maximum likelihood estimation	
	TMLE	
	Double machine learning	
	DML	
	Interrupted time series analysis	

assessed through patient or caregiver reports and medical records.

10. Other relevant outcomes.

Data management and study selection

Articles identified through the search process will be imported into systematic review software, Rayyan (<https://www.rayyan.ai/>), to ensure the process is rigorous. Duplicate entries will be identified and removed before proceeding with the review. A reference management software, such as Zotero, will be adopted to organize and manage citations. Two independent reviewers will screen the titles and abstracts of the retrieved articles to identify potentially eligible studies based on pre-defined inclusion and exclusion criteria. Full texts of these studies will then be assessed to confirm eligibility. Discrepancies at any stage will be resolved through discussion, and if consensus cannot be reached, a third reviewer will be consulted. The results of the selection process, along with the reasons for exclusion, will be documented and presented in a flow diagram following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 32).

Data extraction and recording

To ensure consistency and minimize errors, a standardized data extraction form will be developed using tools such as Excel or systematic review management software and piloted by the reviewers.

Data extraction will be conducted independently by two reviewers, with discrepancies resolved through discussion or consultation with a third reviewer when necessary. Data extraction will follow a detailed coding framework to ensure consistency, including study characteristics (e.g., author(s), publication year, country, study design and population details such as age, gender, diagnosis, and sample size), exposure/intervention descriptions (e.g., lifestyle factors, medical treatments), causal inference methods (e.g., instrumental variables, difference-in-differences, propensity score matching, target trial emulation), outcomes (e.g., cognitive decline, quality of life), key results (effect estimates, confidence intervals, sensitivity or robustness analyses, heterogeneity), study quality (e.g., confounding adjustments, handling of missing data, assumption tests, and risk of bias), emerging methods and gaps (e.g., machine learning, artificial intelligence), and practical challenges and barriers (e.g., computational complexity, scalability). The extracted findings will be presented using tools such as PRISMA flow diagrams, data extraction tables, and quality assessment forms, ensuring transparency and reproducibility of the review process.

Quality appraisal

The risk of bias and quality of the included studies will be critically evaluated by two independent reviewers, with discrepancies resolved through discussion or, if necessary, consultation with a third reviewer. The ROBINS-I (Risk of Bias in Non-randomized Studies of Interventions) tool will be used to evaluate risk of bias. ROBINS-I assesses bias across seven key domains: confounding, selection of participants, classification of interventions, deviations from intended interventions, missing data, outcome measurement, and selective reporting, and evaluating how effectively studies address these issues. It categorizes risk of bias for each domain and overall as “low risk,” “moderate risk,” “serious risk,” or “critical risk,” based on predefined criteria. By comparing bias levels with those of well-conducted RCTs, ROBINS-I provides a structured framework for assessing the validity of causal estimates in observational research [42].

Additionally, a custom causal inference checklist will be employed to systematically evaluate key aspects of each causal inference method, such as the validity of assumptions (e.g., no unmeasured confounding, proper model specification) and the robustness of methods (e.g., sensitivity analyses, handling of bias), as outlined in Table 2. This checklist will complement the ROBINS-I tool by focusing specifically on the methodological rigor and application of causal inference techniques. The appraisal findings will be integrated into the discussion to provide a comprehensive evaluation of the strengths, limitations, and reliability of the included studies (Table 2).

Quantitative synthesis (Meta-analysis)

Where sufficient data and methodological homogeneity are present, meta-analyses will be conducted to pool effect estimates for specific outcomes and causal inference methods. Statistical heterogeneity will be assessed using the I^2 statistic and Cochran's Q test. Subgroup analyses and meta-regression will explore sources of heterogeneity, such as study design, population characteristics, and data types. Meta-analyses will be conducted using software like R (e.g., *meta* or *metafor* packages) or Stata, and results will be visualized using forest plots.

Discussion

This systematic review will provide a comprehensive synthesis of causal inference methods applied in observational studies of dementia and related neurodegenerative conditions. It will evaluate methods such as instrumental variables, difference-in-differences, G-methods, and propensity score-based approaches,

Table 2 Diagnostics and sensitivity/robustness checks

Causal Inference Method	Key Assumptions/Considerations	Robustness Check
Instrumental variables	Instrument relevance, validity of exclusion restriction, and assessment about weak instruments	Overidentification tests, sensitivity analyses for exclusion restriction, and diagnostics for weak instruments
Difference-in-differences	Parallel trends assumption, adjustment for time-varying confounding, and handling of staggered adoption	Pre-treatment trend visualization, placebo tests, and robustness to alternative model specifications
Propensity score methods (e.g., matching)	Appropriate covariate inclusion, balance diagnostics post-matching, and assessment of overlap and common support	Standardized mean differences, sensitivity analyses for unmeasured confounding, and alternative propensity score models
G-methods (e.g., marginal structural models)	Correct model specification, stabilized weights, and adherence to the positivity assumption	Weight diagnostics, sensitivity to model misspecification, and sensitivity to potential unmeasured confounding
Target Trial Emulation	Specification of a clear trial design (eligibility, treatment, follow-up) and addressing immortal time bias	Alternative trial specifications, validation of key assumptions, and sensitivity to immortal time bias
Regression discontinuity design	Justification for the running variable, balance at the cutoff, and appropriate functional form	Placebo cutoffs, alternative bandwidths, and falsification tests for continuity
General causal estimation / counterfactual approaches	Clear definition of counterfactuals and handling of unmeasured confounding	Sensitivity analyses for confounding, alternative methods, and transparent reporting of limitations

with a focus on their ability to address confounding, bias, and other methodological challenges inherent to real-world data. The review will further examine how these methods are adapted to various data types, study designs, and population characteristics, highlighting their strengths, limitations, and appropriateness for different research contexts. By identifying common methodological challenges and critically assessing the assumptions underlying each approach, the review aims to support more rigorous, transparent, and reproducible causal inference in dementia research. Ultimately, its findings are expected to help bridge the gap between observational evidence and actionable insights, informing future research and methodological development, with potential implications for policy and practice in dementia care and prevention.

Abbreviations

- ADL Activities of Daily Living (Katz et al., 1963)
- APOE Apolipoprotein E
- CDR Clinical Dementia Rating (Morris, 1993)
- DSM-5 Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (American Psychiatric Association, 2013 [1])
- EQ-5D EuroQol 5 Dimensions (EuroQol Group, [9])
- IADL Instrumental Activities of Daily Living (Lawton & Brody, 1969)
- MRI Magnetic Resonance Imaging
- MeSH Medical Subject Headings
- MCI Mild Cognitive Impairment
- MMSE Mini-Mental State Examination (Folstein et al., 1975)
- MoCA Montreal Cognitive Assessment (Nasreddine et al., 2005)
- NPI Neuropsychiatric Inventory (Cummings et al., 1994)
- PET Positron Emission Tomography
- PRISMA Preferred Reporting Items for Systematic Reviews and Meta-Analyses (Page et al., [32])
- QoL-AD Quality of Life in Alzheimer’s Disease (Logsdon et al., [24])
- RCT Randomized Controlled Trial
- ROBINS-I Risk of Bias in Non-randomized Studies of Interventions (Sterne et al., 2016 [42])

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13643-026-03179-w>.

Supplementary Material 1.

Acknowledgements

Funding for the READ-OUT (REAl World Dementia OUTcomes) Project by Alzheimer’s Society, Alzheimer’s Research UK, and the National Institute for Health and Care Research and Gates Ventures is gratefully acknowledged (Grant Number:).

Authors’ contributions

YY conceived the review idea, drafted the protocol manuscript, reviewed and refined the search strategy, and will serve as an additional reviewer during screening. YY will also contribute to results synthesis and drafting of the final manuscript. HM performed and managed the initial literature searches. HM and RB will conduct title and full-text screening. SS and JD will be responsible for data extraction. IK and VR contributed to the development of the research concept. All co-authors will provide input on the interpretation of findings, contribute to the critical revision of the manuscript, and approve the final version of the article.

Funding

This protocol received no specific funding.

Data availability

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable; this study involves no human participants or new data collection.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 27 March 2025 Accepted: 26 March 2026
Published online: 18 April 2026

References

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Arlington, VA: American Psychiatric Publishing; 2013.
- Angeles RC, Berge LI, Gedde MH, Kjerstad E, Vislapuu M, Puauschitz NG, et al. Which factors increase informal care hours and societal costs among caregivers of people with dementia? A systematic review of Resource Utilization in Dementia (RUD). *Health Econ Rev*. 2021;11(1):37. <https://doi.org/10.1186/s13561-021-00333-z>.
- Austin PC. Goodness-of-fit diagnostics for the propensity score model when estimating treatment effects using covariate adjustment with the propensity score. *Pharmacoepidemiol Drug Saf*. 2008;17(12):1202–17. <https://doi.org/10.1002/pds.1673>.
- Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivar Behav Res*. 2011;46(3):399–424. <https://doi.org/10.1080/00273171.2011.568786>.
- Brenowitz WD, Yaffe K. Observational studies in Alzheimer disease: bridging preclinical studies and clinical trials. *Nat Rev Neurol*. 2022;18(12):747–57. <https://doi.org/10.1038/s41582-022-00733-7>.
- Charpignon M-L, Vakulenko-Lagun B, Zheng B, Magdamo C, Su B, Evans K, et al. Causal inference in medical records and complementary systems pharmacology for metformin drug repurposing towards dementia. *Nat Commun*. 2022;13(1):7652. <https://doi.org/10.1038/s41467-022-35157-w>.
- Crown WH. Real-world evidence, causal inference, and machine learning. *Value Health*. 2019;22(5):587–92. <https://doi.org/10.1016/j.jval.2019.03.001>.
- Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The neuropsychiatric inventory. *Neurology*. 1994;44(12):2308–2308. <https://doi.org/10.1212/WNL.44.12.2308>.
- EuroQol Group. EuroQol—A new facility for the measurement of health-related quality of life. *Health Policy*. 1990;16(3):199–208. [https://doi.org/10.1016/0168-8510\(90\)90421-9](https://doi.org/10.1016/0168-8510(90)90421-9).
- Folstein MF, Folstein SE, McHugh PR. Mini-mental state. *J Psychiatr Res*. 1975;12(3):189–98. [https://doi.org/10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6).
- Fu R, Kim SJ. Inferring causality from observational studies: the role of instrumental variable analysis. *Kidney Int*. 2021;99(6):1303–8. <https://doi.org/10.1016/j.kint.2021.03.018>.
- Grill JD, Karlawish J. Addressing the challenges to successful recruitment and retention in Alzheimer's disease clinical trials. *Alzheimers Res Ther*. 2010;2(6):34. <https://doi.org/10.1186/alzrt58>.
- Groenwold RHH, Nelson DB, Nichol KL, Hoes AW, Hak E. Sensitivity analyses to estimate the potential impact of unmeasured confounding in causal research. *Int J Epidemiol*. 2010;39(1):107–17. <https://doi.org/10.1093/ije/dyp332>.
- Guo S, Fraser MW. Propensity score analysis: Statistical methods and applications (2nd ed.). SAGE Publications; 2014.
- Hiu S, Yong T, Hasoon J, Teare MD, Taylor J, Lin N. Instrumental variables in real-world clinical studies of dementia and neurodegenerative disease: systematic review of the subject-matter argumentation, falsification test, and study design strategies to justify a valid instrument. *Brain Behav*. 2024;14(1):e3371. <https://doi.org/10.1002/brb3.3371>.
- Hiu S, Yong T, Teare MD, Taylor J. Systematic review of strategies to justify validity of instrumental variables in observational studies of dementia and neurodegenerative disease. *Alzheimers Dement*. 2023;19(S22):e073290. <https://doi.org/10.1002/alz.073290>.
- Hoehn MM, Yahr MD. Parkinsonism. *Neurology*. 1967;17(5):427–427. <https://doi.org/10.1212/WNL.17.5.427>.
- Horr T, Messinger-Rapport B, Pillai JA. Systematic review of strengths and limitations of Randomized Controlled Trials for non-pharmacological interventions in mild cognitive impairment: focus on Alzheimer's disease. *J Nutr Health Aging*. 2015;19(2):141–53. <https://doi.org/10.1007/s12603-014-0565-6>.
- Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged: the index of ADL: a standardized measure of biological and psychosocial function. *JAMA*. 1963;185(12):914. <https://doi.org/10.1001/jama.1963.03060120024016>.
- Labrecque J, Swanson SA. Understanding the assumptions underlying instrumental variable analyses: a brief review of falsification strategies and related tools. *Curr Epidemiol Rep*. 2018;5(3):214–20. <https://doi.org/10.1007/s40471-018-0152-1>.
- 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):1133–63. <https://doi.org/10.1002/sim.3034>.
- Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist*. 1969;9(3):179–86.
- Lin C, Pena JM, Daoud A. Assessing the unobserved: enhancing causal inference in sociology with sensitivity analysis. *arXiv*. 2024:2311.13410. <https://doi.org/10.48550/arXiv.2311.13410>.
- Logsdon RG, Gibbons LE, McCurry SM, Teri L. Quality of life in Alzheimer's disease: Patient and caregiver reports. *J Ment Health Aging*. 1999;5:21–32.
- Michalowsky B, Xie F, Eichler T, Hertel J, Kaczynski A, Kilimann I, et al. Cost-effectiveness of a collaborative dementia care management—results of a cluster-randomized controlled trial. *Alzheimer's & Dementia*. 2019;15(10):1296–308. <https://doi.org/10.1016/j.jalz.2019.05.008>.
- Morris JC. The clinical dementia rating (CDR). *Neurology*. 1993;43(11):2412–2412-a. <https://doi.org/10.1212/WNL.43.11.2412-a>.
- Musci RJ, Stuart E. Ensuring Causal, Not Casual, Inference. *Prev Sci*. 2019;20(3):452–6. <https://doi.org/10.1007/s11211-018-0971-9>.
- Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53(4):695–9. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>.
- Nicholls SG, Carroll K, Nix HP, Li F, Hey SP, Mitchell SL, et al. Ethical considerations within pragmatic randomized controlled trials in dementia: results from a literature survey. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*. 2022;8(1):e12287. <https://doi.org/10.1002/trc2.12287>.
- Nichols E, Steinmetz JD, Vollset SE, Fukutaki K, Chalek J, Abd-Allah F, et al. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: An analysis for the Global Burden of Disease Study 2019. *The Lancet Public Health*. 2022;7(2):e105–25. [https://doi.org/10.1016/S2468-2667\(21\)00249-8](https://doi.org/10.1016/S2468-2667(21)00249-8).
- Ohlsson H, Kendler KS. Applying causal inference methods in psychiatric epidemiology: a review. *JAMA Psychiatr*. 2020;77(6):637. <https://doi.org/10.1001/jamapsychiatry.2019.3758>.
- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, Moher D. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, n71. 2021. <https://doi.org/10.1136/bmj.n71>.
- Pearl J. Causal inference in statistics: an overview. *Stat Surveys*. 2009;3(none):96–146. <https://doi.org/10.1214/09-SS057>.
- Power MC, Engelman BC, Wei J, Glymour MN. Closing the gap between observational research and randomized controlled trials for prevention of Alzheimer disease and dementia. *Epidemiol Rev*. 2022;44(1):17–28. <https://doi.org/10.1093/epirev/mxoc002>.
- Renson A, Hudgens MG, Keil AP, Zivich PN, Aiello AE. Identifying and estimating effects of sustained interventions under parallel trends assumptions. *Biometrics*. 2023;79(4):2998–3009. <https://doi.org/10.1111/biom.13862>.
- Rippin G, Ballarini N, Sanz H, Largent J, Quinten C, Pignatti F. A review of causal inference for external comparator arm studies. *Drug Saf*. 2022;45(8):815–37. <https://doi.org/10.1007/s40264-022-01206-y>.
- Ritchie CW, Terrera GM, Quinn TJ. Dementia trials and dementia tribulations: methodological and analytical challenges in dementia research. *Alzheimers Res Ther*. 2015;7(1):31. <https://doi.org/10.1186/s13195-015-0113-6>.
- Rubin DB. Causal inference using potential outcomes. *J Am Stat Assoc*. 2005. <https://doi.org/10.1198/016214504000001880>.
- Ryan AM, Kontopantelis E, Linden A, Burgess JF. Now trending: coping with non-parallel trends in difference-in-differences analysis. *Stat Methods Med Res*. 2019;28(12):3697–711. <https://doi.org/10.1177/0962280218814570>.
- Sachdev PS, Blacker D, Blazer DG, Ganguli M, Jeste DV, Paulsen JS, et al. Classifying neurocognitive disorders: the DSM-5 approach. *Nat Rev Neurol*. 2014;10(11):634–42. <https://doi.org/10.1038/nrneurol.2014.181>.

41. Shah H, Albanese E, Duggan C, Rudan I, Langa KM, Carrillo MC, et al. Research priorities to reduce the global burden of dementia by 2025. *Lancet Neurol*. 2016;15(12):1285–94. [https://doi.org/10.1016/S1474-4422\(16\)30235-6](https://doi.org/10.1016/S1474-4422(16)30235-6).
42. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919. <https://doi.org/10.1136/bmj.i4919>.
43. Streeter AJ, Lin NX, Crathorne L, Haasova M, Hyde C, Melzer D, et al. Adjusting for unmeasured confounding in nonrandomized longitudinal studies: a methodological review. *J Clin Epidemiol*. 2017;87:23–34. <https://doi.org/10.1016/j.jclinepi.2017.04.022>.
44. Weuve J, Proust-Lima C, Power MC, Gross AL, Hofer SM, Thiébaud R, et al. Guidelines for reporting methodological challenges and evaluating potential bias in dementia research. *Alzheimers Dement*. 2015;11(9):1098–109. <https://doi.org/10.1016/j.jalz.2015.06.1885>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.