

RESEARCH

Open Access



Gaps in the usage and reporting of multiple imputation for incomplete data: findings from a scoping review of observational studies addressing causal questions

Rheanna M. Mainzer^{1,2*}, Margarita Moreno-Betancur^{1,2}, Cattram D. Nguyen^{1,2}, Julie A. Simpson^{3,4}, John B. Carlin^{1,3} and Katherine J. Lee^{1,2}

Abstract

Background Missing data are common in observational studies and often occur in several of the variables required when estimating a causal effect, i.e. the exposure, outcome and/or variables used to control for confounding. Analyses involving multiple incomplete variables are not as straightforward as analyses with a single incomplete variable. For example, in the context of multivariable missingness, the standard missing data assumptions (“missing completely at random”, “missing at random” [MAR], “missing not at random”) are difficult to interpret and assess. It is not clear how the complexities that arise due to multivariable missingness are being addressed in practice. The aim of this study was to review how missing data are managed and reported in observational studies that use multiple imputation (MI) for causal effect estimation, with a particular focus on missing data summaries, missing data assumptions, primary and sensitivity analyses, and MI implementation.

Methods We searched five top general epidemiology journals for observational studies that aimed to answer a causal research question and used MI, published between January 2019 and December 2021. Article screening and data extraction were performed systematically.

Results Of the 130 studies included in this review, 108 (83%) derived an analysis sample by excluding individuals with missing data in specific variables (e.g., outcome) and 114 (88%) had multivariable missingness within the analysis sample. Forty-four (34%) studies provided a statement about missing data assumptions, 35 of which stated the MAR assumption, but only 11/44 (25%) studies provided a justification for these assumptions. The number of imputations, MI method and MI software were generally well-reported (71%, 75% and 88% of studies, respectively), while aspects of the imputation model specification were not clear for more than half of the studies. A secondary analysis that used a different approach to handle the missing data was conducted in 69/130 (53%) studies. Of these 69 studies, 68 (99%) lacked a clear justification for the secondary analysis.

Conclusion Effort is needed to clarify the rationale for and improve the reporting of MI for estimation of causal effects from observational data. We encourage greater transparency in making and reporting analytical decisions related to missing data.

Keywords Missing data, Causal inference, Missingness mechanism

*Correspondence:

Rheanna M. Mainzer
rheanna.mainzer@unimelb.edu.au

Full list of author information is available at the end of the article



© The Author(s) 2024. **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/>.

Background

Observational studies in medical and health-related research often aim to answer a causal question, which we understand as estimation of the average causal effect (ACE) of an exposure on an outcome in a population of interest [1, 2]. Missing data in observational studies often occurs in multiple variables required for the estimation of ACEs, such as the exposure, the outcome and/or the covariates used to control for confounding. Applying standard methods for ACE estimation (e.g., outcome regression with covariate adjustment) using only data from complete records (“complete cases analysis” [CCA]) may lead to selection bias and also overlook precision gains that might be achieved by incorporating information from the incomplete cases [3]. Therefore, missing data need to be carefully considered and addressed to minimise the potential for selection bias and loss of information.

One flexible and widely recommended approach for estimation in the presence of multivariable missingness is multiple imputation (MI) [4–6]. In the first stage of MI, missing data are imputed multiple times with random draws from the predictive distribution of the missing values given the observed data and a specified imputation model. In the second stage, the statistical analysis of interest (e.g., outcome regression with covariate adjustment) is applied to each imputed dataset and the results are combined to obtain a single estimate with associated standard error [4].

To date most reviews of the handling of missing data or the application of MI have been carried out in the context of trials (see [7] and references therein). In contrast, there has been little attention given to how missing data are handled in observational studies, a context in which multivariable missingness is often encountered. Mackinnon (2010) and Hayati Rezvan et al. (2015) reviewed the implementation and documentation of MI in both trials and observational studies [8, 9], and Karahalios et al. (2012) reviewed how missing exposure data are reported in large cohort studies with one or more waves of follow-up [10]. More recently, Carroll et al. (2020) reviewed the handling of missing covariates in observational time-to-event studies in oncology [11], Okpara et al. (2022) reviewed the handling of missing data in longitudinal studies of older adults [12], and Benneville et al. (2023) reviewed the handling of missing covariate data in the field of haematology [13]. However, none of the above reviews focussed on the complexities that arise due to multivariable missingness in exposure, outcome and covariates. The aim of the current study was to review the handling of missing data in observational studies that address causal questions using MI. A scoping review was conducted to systematically benchmark the current state

of practice [14], focussing on four key areas: missing data summaries, missing data assumptions, primary and sensitivity analyses, and MI implementation. In the next section we describe considerations for transparent reporting within each of these four areas to provide context for our review. We then describe our scoping review methodology and present our results. We end with a discussion of our findings and key messages.

Considerations for reporting ACE estimation with MI from incomplete observational data

Several frameworks and guidelines around missing data and the application of MI are available (see [15, 16] and Table 1 of [11]). In this section we outline key considerations when estimating and reporting ACEs from incompletely observed data that are pertinent to the current review from these guidelines.

Missing data summaries

Describing the amount of missing data is an important first step for transparent reporting as the potential for selection bias will generally increase with larger proportions of missing data. When data are missing in a single variable, the number (%) of completely observed values for that variable also summarises the number (%) of complete cases. In contrast, when multiple variables required for analysis are incompletely observed, the number (%) of observed values for each variable may vastly differ from the number (%) with complete cases because of the pattern of missing data, that is, the way in which the variables are jointly missing. In the latter context, a complete description of the missing data would include summaries of the missing data for each variable, as well as summaries of the distinct missing data patterns. Such summaries can be easily obtained in statistical software.

Missing data assumptions

Understanding the process that cause data to be missing, i.e., the “missing data mechanism”, is important because the performance of any estimation method depends critically on this. Sometimes the missing data mechanism will be known (e.g. a machine used for measurement temporarily stopped working), but in most cases it will be unknown and assumptions about the mechanism, along with a justification for these assumptions, are required. Missingness assumptions are often expressed using the classification of missing data patterns as “missing completely at random” (MCAR), “missing at random” (MAR) or “missing not at random” (MNAR) [16, 17]. However, assessing the plausibility of the MCAR/MAR/MNAR assumptions in the context of multivariable missingness is difficult, partly due to the existence of several different, often imprecise, definitions of MCAR, MAR and MNAR

in the literature and the difficulty of interpreting these definitions, [18] and partly because assessment involves making a judgement about the dependence (or lack thereof) of the distribution of the missing data pattern on the observed and missing data [16, 19]. An attractive alternative to using the MCAR/MAR/MNAR framework is to view missing data as a causal problem and to represent assumptions about causes of missingness for each incompletely observed variable using missingness directed acyclic graphs (m-DAGs) [3, 20]. m-DAGs are an extension to standard causal diagrams (DAGs) that include nodes to represent missingness in each incomplete variable, thereby allowing for the clear and transparent specification of assumptions about the causes of missing data, as well as the causal relationships amongst the main variables of interest. Assumptions about the causes of missingness can be justified using expert knowledge, literature or external data (see, e.g., Fig. 3 of [3] and Table 5 of [21]). Although developing a realistic m-DAG can be challenging and time-consuming, m-DAGs lead to assumptions that are more transparent and easier to assess than assumptions expressed using the MCAR/MAR/MNAR framework. Uncertainty about the assumptions depicted in the m-DAG can be assessed using a sensitivity analysis (see next section).

Primary and sensitivity analyses

The next important area for reporting is to justify and describe an appropriate primary method for estimation of the ACE, given the missingness assumptions. It is well known that both a CCA and standard MI (an implementation of MI that does not incorporate an external assumption about a difference between the distribution of the observed and missing data) can provide consistent estimation of the ACE when data are MCAR, that standard MI can provide consistent estimation when data are MAR, and that both approaches may provide biased estimation when data are MNAR. However, in the context of multivariable missingness, a CCA can also provide consistent estimation under missingness mechanisms that could be classified as MAR, and both CCA and MI have been shown in theory and simulations to provide unbiased or approximately unbiased estimation of ACEs across a range of missingness mechanisms that could be classified as MNAR [21]. Therefore, it is not straightforward to justify an estimation approach even if it is believed that data are MAR or MNAR. In contrast, for a given m-DAG, graph theory can be used to establish whether the ACE is *recoverable* (that is, whether it can be estimated unbiasedly from the observed data). If the ACE is recoverable, the process of establishing recoverability can aid in determining whether a CCA and/or

standard MI would be appropriate for estimation (see, e.g., the worked example provided by Lee et. al. [16]). If the ACE is not recoverable, neither standard MI nor a CCA can be used for unbiased estimation, and a more sophisticated approach that incorporates an assumption about a difference in distribution between the missing and observed values is needed. For example, the not-at-random fully conditional specification (NARFCS) procedure extends standard MI to incorporate such assumptions through the inclusion of a sensitivity parameter “delta”, elicited from external information, that represents the difference between the distributions of the observed and missing values [22]. The assumptions made about the missing data and how this justifies the choice of analytic method for the primary analysis should be carefully described.

Sensitivity analyses to reflect uncertainty due to assumptions made about the missing data for the primary analysis are strongly recommended [15, 23]. There are two types of missing data sensitivity analyses to consider; the first is to examine the sensitivity of estimates to the assumptions made about the causes of missing data, e.g. the existence or strength of arrows in the m-DAG. The second type of missing data sensitivity analysis is to examine the sensitivity of estimates to assumptions made for modelling the missing data, such as the form of the imputation model (e.g., linear regression vs predictive mean matching for imputing continuous variables). As with the primary analysis, sensitivity analyses should be justified and described in enough detail that the analysis could be reproduced.

MI implementation

When using standard MI for estimation, quantities that need to be described to ensure that the analysis could be reproduced include, but are not limited to: the imputation method, e.g., multivariate normal imputation or multivariate imputation by chained equations; the imputation model, e.g., which variables are included and in what form; if using multivariate imputation by chained equations, the models/methods that are used to impute each incomplete variable, e.g., linear or logistic regression; the number of imputations conducted; the analysis model that is fitted to obtain estimates within each imputed dataset; and the method for combining estimates across imputed datasets [17]. If using an approach that incorporates an assumption about a difference in distribution between the missing and observed values (e.g., a NARFCS procedure), then, in addition to the above quantities, it is important to describe how the assumption is incorporated in the models used for the estimation procedure.

Methods

The protocol for this scoping review has been published previously [7]. Briefly, we included observational studies that aimed to answer at least one causal research question using MI, published in *International Journal of Epidemiology*, *American Journal of Epidemiology*, *European Journal of Epidemiology*, *Journal of Clinical Epidemiology* and *Epidemiology* between January 2019 and December 2021. These journals were chosen as they are high ranking, general journals in epidemiology that we expected would capture current best practices in the use of MI for estimating ACEs from observational data. This selection of journals has been used previously in a systematic review of epidemiologic practice [24]. A full text search for the term “multiple imputation” was conducted on the journal websites, following the methodology of Hayati Rezvan et al [8]. Causal questions were identified if the study authors explicitly stated that they were estimating an ACE or if the study authors estimated an effect that was given, at least implicitly, a causal interpretation. Studies were excluded from the review if they met any of the following criteria: the study did not aim to answer a causal question, a clear research goal could not be identified, the primary purpose of the article was methodological development, the analysis was based on aggregated data, the article reported qualitative research, the study exposure was assigned to participants by investigators (i.e. a trial), or the study was retracted. The most recent search was performed on 10th June 2022.

A random sample of 10 articles were independently screened and reviewed by two reviewers (RM and KL) to develop the data collection instrument. One reviewer (RM) screened and reviewed all articles. Double data extraction was independently completed for 10% of articles (RM and KL). In addition, a second reviewer (CN or KL) screened articles when there was uncertainty about the inclusion criteria and reviewed articles when there was uncertainty about the information being extracted. Disagreements between reviewers were resolved via discussion with a third reviewer.

A summary of the data extraction items and a copy of the data extraction questionnaire are provided in Table 1 and the Supplementary Material, respectively, of Mainzer et al [7]. Briefly, for each study included in the review, data were extracted on the following: study characteristics; the quantity of missing data; the missing data assumptions made and whether these assumptions were justified; details of the primary analysis and whether or not the primary analysis was justified based on missing data assumptions; details of any secondary/sensitivity analysis conducted that handled the missing data differently from the primary analysis and its justification; and details of the MI implementation. For each study, we

Table 1 Summary of study characteristics for the 130 included papers

Characteristic	n (%)
Publication year	
2019	47 (36%)
2020	45 (35%)
2021	38 (29%)
Journal ^a	
American Journal of Epidemiology	50 (38%)
Epidemiology	24 (18%)
European Journal of Epidemiology	21 (16%)
International Journal of Epidemiology	34 (26%)
Journal of Clinical Epidemiology	1 (1%)
Study design	
Prospective longitudinal study	85 (65%)
Retrospective analysis of routinely collected data	15 (12%)
Pooled cohort analysis	9 (7%)
Case-control study	7 (5%)
Cross-sectional study	5 (4%)
Case-cohort study	2 (2%)
Other ^b	7 (5%)
Type of outcome used for analysis	
Binary	45 (35%)
Categorical (excluding binary)	3 (2%)
Continuous	33 (25%)
Time to event	49 (38%)
Causal question inclusion criteria ^c	
Explicitly stated interest in a causal effect	33 (25%)
Estimate was given a causal interpretation	130 (100%)
Typical signals of a causal analysis ^c	
A directed acyclic graph was used to depict causal assumptions	40 (31%)
A set of variables were identified to control for confounding	106 (82%)
Effect was estimated using a regression model with adjustment for a set of covariates ^d	129 (99%)

^a Number of papers published between January 2019 and December 2021 based using a Pub Med search for (“2019/01/01”[Date - Publication] : “2021/12/31”[Date - Publication]) AND (“Journal name”[Journal]): *American Journal of Epidemiology*, 876; *Epidemiology*, 496; *European Journal of Epidemiology*, 370; *International Journal of Epidemiology*, 814; *Journal of Clinical Epidemiology*, 996

^b Secondary analysis of trial data ($n = 2$); prospective follow-up of cohort recruited for trial ($n = 2$); pooled analysis of data from case-control and cohort studies ($n = 1$); pooled analysis of data from case-control studies ($n = 1$); transportability study using data from 4 clinical trials and 1 observational cohort ($n = 1$)

^c Categories are not mutually exclusive

^d One study used structural equation modelling seemingly without adjustment for covariates, although a causal conclusion was made

defined the “inception sample” as the set of participants who met eligibility criteria for inclusion in the study to answer the research question of interest, where eligibility criteria do not include any requirement for variables

to be complete, and the “analysis sample” as the participants who were included in the analysis to answer the research question of interest. Defining both the inception and analysis sample was necessary as we cannot always establish the size of the inception sample (either because authors neglect to report it, or because they were unable to define it in cases where eligibility data are themselves missing, e.g. for a study using electronic medical records in which patients may not appear in the database unless eligibility-defining measurements have been taken), and therefore needed a way to specify denominators for calculating percentages of missing data (which was the number in the analysis sample). Since our review only included studies with missing data within the analysis sample, the subset with complete cases was always smaller than the analysis sample. Extracted items were summarised using descriptive statistics. Data cleaning and analysis was performed in R [25]. Reporting follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews checklist [26].

Results

Screening process

Figure 1 presents a flow diagram of the article screening process. Of the 304 papers that met the inclusion criteria, 130 papers were included in this review [27–156]. There were 14 articles that were screened by a second reviewer due to uncertainty about inclusion criteria. Double data extraction was completed for a further 14 articles.

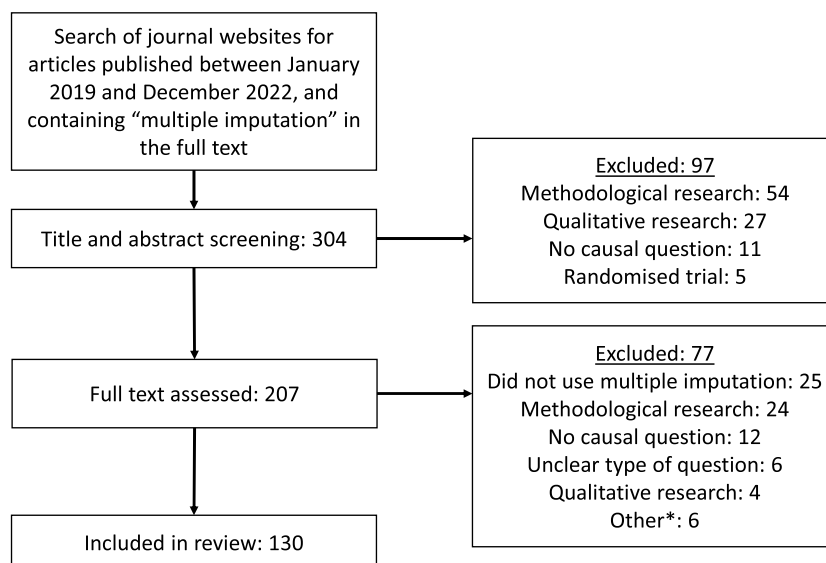
All disagreements were resolved via discussion. Minor changes were made to the review protocol to accommodate unanticipated challenges in data extraction (described in Additional file 1).

Study characteristics

Study characteristics are summarised in Table 1. Most papers included in this review were published in *American Journal of Epidemiology* (38%) or *International Journal of Epidemiology* (26%). The most common study design was a prospective longitudinal study (65%), followed by a retrospective analysis of routinely collected data (12%). The most common outcomes used for analyses were binary (35%) and time to event (38%). Few studies made their causal aim explicit (25%) or presented a DAG to depict causal assumptions (31%). However, most studies identified a set of variables to control for confounding (82%) and almost all studies estimated an effect using a regression model (or a more sophisticated causal effect estimation method such as g-computation) with adjustment for a set of covariates, implicitly or explicitly assumed to be confounders (99%).

Missing data summaries

The reported quantity of missing data is summarised in Table 2 and Fig. 2. The size of the inception sample could not be established in 38% of studies, and 83% of studies derived an analysis sample by excluding individuals with missing data in specific variables. The percentage of complete cases could be established in just



*Aggregate-level data: 2, Trial: 2, Not clear that multiple imputation was used to address the causal question: 1, Retracted: 1

Fig. 1 Article screening process

Table 2 Amount of missing data. Summaries are n (%) unless stated otherwise

Characteristic	Summary
Able to establish the size of the inception sample ^a	
Yes	81 (62%)
No ^b	49 (38%)
Analysis sample was defined by excluding individuals with missing data in specific variables	
Yes	108 (83%)
No	22 (17%)
Complete cases	
Able to establish the % of complete cases	34 (26%)
% of complete cases, median (25 th –75 th percentiles)	85% (75% – 92%)
Only able to establish an upper bound on the % of complete cases	80 (62%)
Upper bound on % of complete cases, median (25 th –75 th percentiles)	84% (72% – 92%)
Not able to establish the percentage of complete cases	16 (12%)
Missing values in the exposure	
Yes, and able to establish the % of missing values	39 (30%)
% of missing values, median (25 th –75 th percentiles)	11% (3% – 16%)
Yes, but only able to establish a lower bound on the % of missing values	4 (3%)
Lower bound on % of missing values, median (25 th –75 th percentiles)	5% (4% – 18%)
Yes, but unable to establish the % or a lower bound on the %	7 (5%)
No	70 (54%)
Unclear	10 (8%)
Missing values in the outcome ^c	
Yes, and able to establish the % of missing values	19 (15%)
% of missing values, median (25 th –75 th percentiles)	9% (5% – 28%)
Yes, but only able to establish a lower bound on the % of missing values	6 (5%)
Lower bound on % of missing values, median (25 th –75 th percentiles)	5% (3% – 17%)
Yes, but unable to establish the % or a lower bound on the %	6 (5%)
No	91 (70%)
Unclear	8 (6%)
Missing values in the covariates	
Yes, in 2 or more	109 (84%)
Yes, in 1 covariate only	7 (5%)
No	7 (5%)
Unable to establish	7 (5%)
Multivariable missingness within analysis sample	
Yes	114 (88%)
No	8 (6%)
Unable to establish	8 (6%)

^a Inception sample defined as the participants who met eligibility criteria for inclusion in the study to answer the research question of interest, where eligibility criteria do not include any requirement for variables to be complete

^b Includes 5 studies where analyses were conducted separately by sub-groups (e.g., sex), but the inception sample for the sub-group could not be identified even though the inception sample for the entire group may have been provided

^c Time-to-event outcomes were not considered to be missing data (we did not treat censored data as missing) except for in two studies where authors explicitly stated that the outcome was imputed

34/130 (26%) studies (median, 25th – 75th percentiles: 85%, 75% – 92%), although an upper bound on the percentage of complete cases that was tighter than 100% (indicating the maximum possible percentage of complete cases given the missing data summaries provided) could be established for another 80/130 (62%) studies (median upper bound, 25th–75th percentiles: 84%, 72% – 92%). Almost all studies (88%) incurred missing data in multiple variables in the analysis sample (despite most studies already arriving at an analysis sample by excluding individuals with missing data in specific variables).

Missing data assumptions

Missing data assumptions are described in Table 3. Most studies (66%) omitted a statement about missing data assumptions entirely. Of the 44 studies that did provide an explicit or indirect statement about missing data assumptions, 35/44 (80%) stated the MAR assumption, 2/44 (5%) stated the MCAR assumption and 6/44 (14%) alluded to data being “not MCAR” but did not distinguish between MAR and MNAR. Eleven of the 44 (25%) studies that provided a statement about missing data assumptions provided a justification for their missing data assumptions (described in Table 3, footnote 3). For example, justifications that were provided for the 6 studies that assumed data were MAR included: describing characteristics associated with missingness and/or conducting formal hypothesis tests ($n = 4$); examining the missingness pattern ($n = 1$), and; because study participants moved homes and/or were impossible to locate ($n = 1$). Of the 130 studies in the review, 31 (24%) linked the justification for the primary analysis to the missing data assumptions.

Primary and sensitivity analyses

Details of the primary and secondary/sensitivity analyses are described in Table 4. Most studies (79%) used MI as the primary analysis method and approximately half (69/130, 53%) of the studies conducted a secondary analysis that handled the missing data differently. Of the 69 studies that conducted a secondary analysis, 70% of studies either provided no justification for conducting the secondary analysis or justified the secondary analysis as a sensitivity analysis without describing to what aspect of their primary analysis they were assessing sensitivity. A further 17/69 (25%) studies provided a vague justification for the secondary analysis, including to examine the influence of missing data (6%), to examine the impact of the missing data method (10%), and to address possible selection bias (9%). 88% of studies that conducted a secondary analysis performed both a CCA and an MI analysis; of these, only

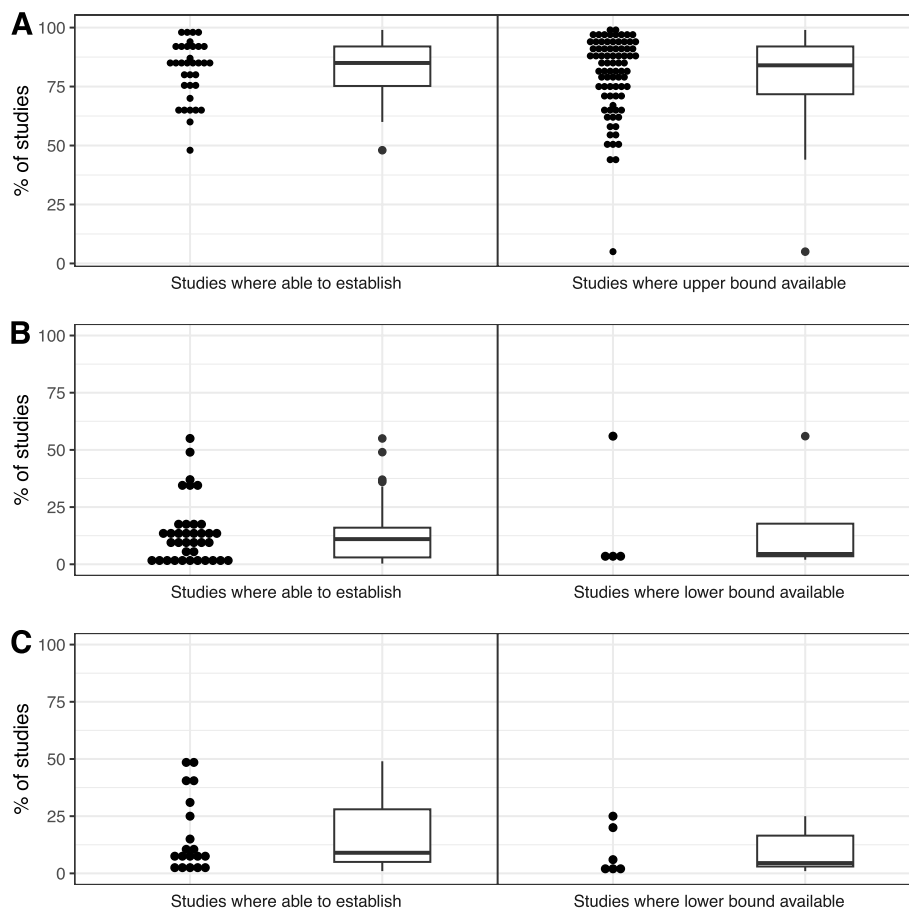


Fig. 2 Dot plots and histograms showing the extent of missing data as a proportion of all participants in the analysis sample (see text for definition). **A** no missing data (complete cases); **B** missing values in the exposure; **C** missing values in the outcome. Left panels: restricted to studies where the percentage could be established; right panels: restricted to studies where the exact percentage could not be established but a conservative bound on the percentage could be established

3 studies (5%) observed a substantial difference between CCA and MI estimates. One study (1%) conducted an “extreme case” analysis that involved single imputation of the outcome under two extreme scenarios, thereby incorporating an external assumption about a difference in distribution between the missing and observed outcome data. However, no studies used a model-based approach such as a NARFCS procedure or elicited external information from subject-matter experts about the difference in distribution between the missing and observed data.

MI implementation

The details of the MI implementation are described in Table 5. Most studies (71%) reported the number of imputations (median, 25th-75th percentiles: 20, 3 – 100). Multivariate imputation by chained equations was the most used imputation method (67% of studies), but the imputation method was unclear for a further 25% of studies. MI was most often conducted in Stata or R. In more than half

of the studies it was unclear whether all analysis variables were included in the imputation procedure (58%), whether auxiliary variables were used in the imputation procedure (55%), and whether interactions were included in the imputation model (57%). Of the 87 studies that reported using multivariate imputation by chained equations, 18 (21%) reported the type of models that were used in the imputation procedure. In approximately two-thirds (65%) of studies, the method that was used to obtain a final MI estimate and its standard error was not stated and could not be deduced from the description in the paper. We assume that most studies would use Rubin’s rules to produce a final estimate and standard error, although alternative approaches are available (see, e.g., [157]).

Discussion

We systematically reviewed the literature to assess the current state of practice in using MI for estimation of causal effects from incompletely observed observational

Table 3 Assumptions about the missing data mechanism

Characteristic	Summary
Missing data assumptions ^a	
No statement of missing data assumptions was provided	86 (66%)
Data were assumed to be MAR	35 (27%)
Data were assumed to be not MCAR	6 (5%)
Data were assumed to be MCAR	2 (2%)
A comprehensive description of missing data assumptions was provided, e.g., using an m-DAG	0 (0%)
Other ^b	1 (1%)
Justification provided for missing data assumptions (as % of papers that made a statement about missing data assumptions, $n = 44$)	
Yes ^c	11 (25%)
No	33 (75%)
Justified the primary analysis using missing data assumptions	
Yes	31 (24%)
No	98 (75%)
Other ^d	1 (1%)

Abbreviations: MAR missing at random, MCAR missing completely at random, MNAR missing not at random, m-DAG missingness directed acyclic graph

^a The assumption may have been stated explicitly or made indirectly. For example, explicit statements of the MAR assumption include: "We assumed the missing at random assumption held and is reasonable", [112] and "We imputed data using multiple imputation by chained equations under the assumption that data were missing at random" [140]. Indirect statements of the MAR assumption include "This multiple imputation approach assumes missing at random", [93] and "We first imputed missing values using multiple imputation by chained equations, which assumes the data are missing at random conditional on the variables in the imputation model" [120]

^b Data assumed to be "MCAR, conditional on age and ethnicity" ($n = 1$)

^c Two studies justified assuming that data were MCAR; justifications included adding the questionnaire to the study after the study began ($n = 1$) and a lack in data registration ($n = 1$). Three studies justified assuming that data were not MCAR; justifications included clinicians ordering tests according to glucose level ($n = 1$), and describing characteristics associated with missingness ($n = 2$). Six studies justified assuming that data were MAR; justifications included describing characteristics associated with missingness and/or conducting formal hypothesis tests ($n = 4$), examining the missingness pattern ($n = 1$) and because children moved homes and/or were impossible to locate ($n = 1$)

^d Justified MI to improve efficiency in the estimators

data. We focussed on four key areas: missing data summaries, missing data assumptions, primary and sensitivity analyses, and MI implementation. Overall, we found that most studies are not reporting missing data, and missing-data-related assumptions, decisions, or analyses with sufficient clarity.

Similarly to other reviews [11–13], we found that the analysis sample is often arrived at by excluding individuals with missing data in certain variables, for example, by using eligibility criteria that require key variables to be completely observed. This is worrying as the preliminary exclusion of individuals may lead to selection bias [158]. It also means that the full extent of missing data is difficult to quantify due to difficulty in identifying the

Table 4 Primary and secondary analyses

Characteristic	n (%)
Method used for the primary analysis	
Standard MI	80 (62%)
Standard MI, combined with weighting ^a	23 (18%)
CCA	21 (16%)
Other ^b	6 (5%)
Secondary analysis conducted that handled the missing data differently	
Yes	69 (53%)
No	61 (47%)
Method used for the secondary analysis (as % of papers that conducted a secondary analysis, $n=69$)	
Standard MI	27 (39%)
Standard MI, combined with weighting ^a	2 (3%)
CCA	26 (38%)
CCA, combined with weighting ^a	6 (9%)
Conducted more than two secondary analyses ^c	8 (12%)
Justification for the secondary analysis, (as % of papers that conducted a secondary analysis, $n=69$)	
Not provided	25 (36%)
As a sensitivity analysis (without further justification)	23 (33%)
To examine the influence of missing data	4 (6%)
To examine the impact of the missing data method	7 (10%)
To address possible selection bias	6 (9%)
To examine robustness to parametric modelling assumptions	1 (1%)
To examine robustness to causal assumptions about the missing data mechanism	0 (0%)
Other ^d	3 (4%)
Conducted both a CCA and an MI analysis, regardless of whether weighting was used or not (as % of papers that conducted a secondary analysis, $n=69$)	
Yes	61 (88%)
No ^e	8 (12%)
Observed a substantial difference between MI and CCA estimates, (as % of papers that conducted both a CCA and a MI analysis, $n=61$)	
Yes	3 (5%)
No	58 (95%)

^a Weights were used to address selection bias due to loss to follow up or censoring. Excludes weights that were used to address confounding bias

^b Treated "missing" as an additional category, with or without weighting ($n = 3$); Single median imputation to obtain exposure ($n = 1$); Single mean imputation for variables with >25% missing values ($n = 1$); MI for one covariate with >25% missing values and single (median/mode) imputation for variables with less than 5% missing values ($n = 1$)

^c Described in Additional file 1, Supplementary Table 1

^d As a sensitivity analysis to examine the robustness of findings to statistical assumptions without stating which statistical assumptions ($n = 1$); as a sensitivity analysis to address possible selection bias and to exploit information in incomplete record participants ($n = 1$); Estimates were presented from both MI and a CCA after fitting two different models (one weighted and one unweighted). No justification was provided for conducting both MI and CCA analyses, but fitting two models was justified by seeing whether the choice of model impacted results and fitting models with and without weights was conducted to see how weighting affected the results ($n = 1$)

^e Standard MI, with and without weighting ($n = 4$); Standard MI, with weighting, with and without imputation of exposure ($n = 1$); Standard MI, with and without inclusion of the outcome in the imputation model ($n = 1$); Three versions of single imputation and standard MI ($n = 1$); Missing treated as an additional category, last value carried forward and standard MI ($n = 1$)

Table 5 Multiple imputation implementation. Summaries are n (%) unless stated otherwise

Characteristic	Summary
Reported number of imputations	
Yes	92 (71%)
No	38 (29%)
Number of imputations, median (25th - 75th percentiles)	20 (3 - 100)
Multiple imputation method	
Multivariate imputation by chained equations	87 (67%)
Multivariate normal imputation	6 (5%)
Other ^a	4 (3%)
Unclear	33 (25%)
Software package used for conducting the multiple imputation analysis	
Stata	40 (31%)
R	33 (25%)
SAS	26 (20%)
SPSS	1 (1%)
Other ^b	14 (11%)
Unclear	16 (12%)
All analysis variables included in imputation model	
Yes	35 (27%)
No	20 (15%)
Unclear	75 (58%)
Auxiliary variables included in imputation model	
Yes	42 (32%)
No	16 (12%)
Unclear	72 (55%)
Interactions included in imputation model	
Yes	2 (2%)
No	54 (42%)
Unclear	74 (57%)
Reported type of models used for imputation (as % of papers that used multivariate imputation by chained equations, $n=87$)	
Yes	18 (21%)
No	69 (79%)
Stated how a final estimate and standard error were obtained	
Either stated, provided code or method could be deduced from software description	45 (35%)
Not stated	85 (65%)

^a Imputation performed using a bootstrapping-based algorithm for panel data in R package Amelia II ($n = 1$), imputation performed in the pan package mitml for multilevel data ($n = 1$), referenced a paper where the MI methods are described rather than providing a description ($n = 1$), used a multiple imputation analysis for exposure and covariates without stating what the analysis was, and used Kaplan-Meier multiple imputation for the outcome as part of a sensitivity analysis ($n = 1$)

^b Study used two software packages for analysis but it was not clear which package was used for MI ($n = 13$), NORM software ($n = 1$)

inception sample. Therefore, for the purposes of reporting the amount of missing data in this review, we considered the amount of missing data within the analysis sample only. However, identifying the exact amount of

missing data within the well-defined analysis sample was also often difficult because summaries were frequently reported per variable without describing missing data patterns.

Details of the assumptions made about the missing data mechanism were often lacking and, when provided, not justified appropriately. A statement of assumptions about the missingness mechanism was provided for just one-third (33%) of studies. This is a marginal improvement over what was found in the reviews conducted by Mackinnon (2010), where 8/50 (16%) observational studies provided a statement that data were MAR, [9] and Rezvan et al. (2015), where 7/30 (23%) observational studies stated or described the assumed missing data mechanism, [8] and illustrates that concerted effort is still needed to improve transparency around missing data assumptions. When a statement about the missing data mechanism was provided, most studies said they assumed data were MAR, but justifications for missingness assumptions were provided in just 11 studies. When they were provided, justifications were generally vague or incomplete. As highlighted in the Introduction, the MCAR/MAR/MNAR assumptions are difficult to interpret and assess in the context of multivariable missingness, so it is not surprising that we found lacking or incomplete justifications for these assumptions. Of note, no study provided a comprehensive description of missing data assumptions, for example, using an m-DAG. Furthermore, the omission of a statement of missing data assumptions entirely from most studies suggests that the critical link between missing data assumptions and estimation methods is not generally appreciated. When missing data assumptions were used to guide the choice of MI as the primary analysis, the most common justification for using MI was because data were assumed to be MAR (without justifying the MAR assumption).

Most studies in this review used standard MI for the primary analysis. Approximately half of the studies conducted a secondary analysis that treated the missing data differently from the primary analysis, but the reason for doing so was almost always omitted or unclear. When studies did carry out two analyses that handled the missing data differently, it was common to conduct both a CCA and MI. Without justification, it is not clear why such an analysis is warranted. It may be to examine the sensitivity of ACE estimates to causal assumptions made about the missing data mechanism for the primary analysis. We speculate another motivation for such an analysis may be the misconception that a CCA is the “normal” approach to dealing with missing data while standard MI provides a more sophisticated analysis that allows you to assess whether the missing data were really an issue or not. However, if under plausible missingness assumptions

neither standard MI nor CCA can provide unbiased estimation, then it would be incorrect to conclude that the missing data “had little impact” on the results. In other words, when there is no unbiased estimate to compare against, the impact of the missing data remains unknown. Of the 61 studies that conducted both a CCA and MI analysis, only 3 (5%) studies observed a substantial difference between MI and CCA estimates. Just one study conducted an analysis that incorporated assumptions about a difference between the missing and observed data distributions. Despite being an area of recent methodological development, our finding that such analyses are not being performed often is similar to findings from previous reviews, see e.g. [8, 159].

MI is increasingly recognised as a method for estimation that needs to be tailored to the target analysis, for example, by careful selection of which variables to include in the imputation model and in what form [6], and by examining the implications of assumptions encoded in an assumed m-DAG [3]. However, the findings from the current review suggest that there is room for improvement in the reporting of MI implementation. For example, certain aspects of the imputation model form were reported just over half of the time despite being needed to judge the appropriateness of the MI model and ensure the analysis can be reproduced.

As noted earlier, several useful frameworks and guidelines around the handling and reporting of missing data are available in the literature. However, the findings from this review suggest that the available guidelines are not being followed. Journals could play more of a role in ensuring appropriate handling and reporting of missing data. Furthermore, although there is growing guidance around using MI in causal inference (see e.g. Chapter 13 of [160]), further guidance is needed when the research question, assumed m-DAG, and/or analysis model are more complex than those considered in the guidance documents, for example, in the context of causal mediation analysis.

The strengths of this review are that it documents the current practices in the use of MI for estimating ACEs from incomplete observational data. Our review followed a clear, pre-specified protocol [7], and by including articles in top general epidemiology journals, we tried to capture current best practice. Furthermore, the analysis conducted for the current study is entirely reproducible as all data and code are available on GitHub: github.com/rheanna-mainzer/MI-scoping-review. This review has several limitations. Authors may have chosen not to provide details on all aspects of handling missing data that we examined, for example, due to strict journal word limits. However, all accompanying supplementary material was also reviewed and used for data extraction. Most of the data extraction was performed by a single reviewer

(RM), with double data extraction performed for 10% of studies, so there may be some extraction errors. Also, it may have been useful to extract additional items or extract items in more detail to better capture the variety of analyses undertaken. However, additional notes on each paper were recorded and are available as part of the complete dataset on GitHub. Lastly, by limiting to five top general epidemiology journals, our results may not reflect papers published in other journals.

Conclusion

The message from our review is clear: there is a need for greater clarity in the conduct and reporting of causal effect estimation using MI with incomplete observational data. Researchers are encouraged to follow the guidance that is available regarding the handling of missing data, to move beyond the MCAR/MAR/MNAR framework and adopt a more transparent approach for outlining missing data assumptions, to use missing data assumptions to justify the estimation method, and to report their assumptions, methods and results systematically.

Abbreviations

ACE	Average causal effect
CCA	Complete case analysis
DAG	Directed acyclic graph
MAR	Missing at random
MCAR	Missing completely at random
m-DAG	Missingness directed acyclic graph
MI	Multiple imputation
MNAR	Missing not at random
NARFCS	Not-at-random fully conditional specification

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12874-024-02302-6>.

Supplementary Material 1

Acknowledgements

Not applicable.

Authors' contributions

RM: Conceptualization, Software, Validation, Formal analysis, Writing – Original Draft, Writing – Review & Editing. MM-B: Conceptualization, Methodology, Writing – Review & Editing. CN: Validation, Writing – Review & Editing. JS: Conceptualization, Methodology, Writing – Review & Editing. JC: Conceptualization, Methodology, Validation, Writing – Review & Editing. KL: Conceptualization, Methodology, Validation, Supervision, Writing – Review & Editing.

Funding

This work was supported by Australian National Health and Medical Research Council (NHMRC) Investigator Grant Leadership Level 1 grants (grant 1196068 awarded to JS and 2017498 to KL), an NHMRC Investigator Grant Emerging Leadership Level 2 (grant 2009572 awarded to MM-B) and an NHMRC Project Grant (grant 1166023). Research at the Murdoch Children's Research Institute is supported by the Victorian Government's Operational Infrastructure Support Program.

Availability of data and materials

The datasets supporting the conclusions of this article are available from RM's GitHub repository: github.com/rheanna-mainzer/MI-scoping-review.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Clinical Epidemiology and Biostatistics Unit, Murdoch Children's Research Institute, Parkville, Victoria 3052, Australia. ²Department of Paediatrics, The University of Melbourne, Parkville, Victoria 3052, Australia. ³Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Parkville, Victoria 3052, Australia. ⁴Nuffield Department of Medicine, University of Oxford, Oxford, UK.

Received: 21 May 2024 Accepted: 2 August 2024

Published online: 04 September 2024

References

- Hernán MA. The C-word: Scientific euphemisms do not improve causal inference from observational data. *Am J Public Health*. 2018;108(5):616–9.
- Lederer DJ, Bell SC, Branson RD, Chalmers JD, Marshall R, Maslove DM, et al. Control of confounding and reporting of results in causal inference studies. Guidance for authors from editors of respiratory, sleep, and critical care journals. *Ann Am Thorac Soc*. 2019;16(1):22–8.
- Moreno-Betancur M, Lee KJ, Leacy FP, White IR, Simpson JA, Carlin JB. Canonical causal diagrams to guide the treatment of missing data in epidemiologic studies. *Am J Epidemiol*. 2018;187(12):2705–15.
- Rubin DB. Multiple imputation for nonresponse in surveys. Hoboken: Wiley; 2004.
- Van Buuren S. Flexible imputation of missing data. Boca Raton: CRC Press; 2018.
- Meng X-L. Multiple-imputation inferences with uncongenial sources of input. *Stat Sci*. 1994;9(4):538–58.
- Mainzer R, Moreno-Betancur M, Nguyen C, Simpson J, Carlin J, Lee K. Handling of missing data with multiple imputation in observational studies that address causal questions: protocol for a scoping review. *BMJ Open*. 2023;13(2):e065576.
- Hayati Rezvan P, Lee KJ, Simpson JA. The rise of multiple imputation: a review of the reporting and implementation of the method in medical research. *BMC Med Res Methodol*. 2015;15:1–14.
- Mackinnon A. The use and reporting of multiple imputation in medical research—a review. *J Intern Med*. 2010;268(6):586–93.
- Karahalios A, Baglietto L, Carlin JB, English DR, Simpson JA. A review of the reporting and handling of missing data in cohort studies with repeated assessment of exposure measures. *BMC Med Res Methodol*. 2012;12(1):96.
- Carroll OU, Morris TP, Keogh RH. How are missing data in covariates handled in observational time-to-event studies in oncology? A systematic review. *BMC Med Res Methodol*. 2020;20(1):134.
- Okpara C, Edokwe C, Ioannidis G, Papaioannou A, Adachi JD, Thabane L. The reporting and handling of missing data in longitudinal studies of older adults is suboptimal: a methodological survey of geriatric journals. *BMC Med Res Methodol*. 2022;22(1):122.
- Bonneville EF, Schetelig J, Putter H, de Wreede LC. Handling missing covariate data in clinical studies in haematology. *Best Pract Res Clin Haematol*. 2023;36(2): 101477.
- Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med Res Methodol*. 2018;18(1):143.
- Lee KJ, Tilling KM, Cornish RP, Little RJ, Bell ML, Goetghebuer E, et al. Framework for the treatment and reporting of missing data in observational studies: The Treatment And Reporting of Missing data in Observational Studies framework. *J Clin Epidemiol*. 2021;134:79–88.
- Lee KJ, Carlin JB, Simpson JA, Moreno-Betancur M. Assumptions and analysis planning in studies with missing data in multiple variables: moving beyond the MCAR/MAR/MNAR classification. *Int J Epidemiol*. 2023;52(4):1268–75.
- Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009;338:b2393.
- Seaman S, Galati J, Jackson D, Carlin J. What is meant by “missing at random”? *Stat Sci*. 2013;28(2):257–68.
- Doretti M, Geneletti S, Stanghellini E. Missing data: a unified taxonomy guided by conditional independence. *Int Stat Rev*. 2018;86(2):189–204.
- Mohan K, Pearl J. Graphical models for processing missing data. *J Am Stat Assoc*. 2021;116(534):1023–37.
- Zhang J, Dashti SG, Carlin JB, Lee KJ, Moreno-Betancur M. Recoverability and estimation of causal effects under typical multivariable missingness mechanisms. *Biomet J*. 2024;66(3):2200326.
- Tompsett DM, Leacy F, Moreno-Betancur M, Heron J, White IR. On the use of the not-at-random fully conditional specification (NARFCS) procedure in practice. *Stat Med*. 2018;37(15):2338–53.
- Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med*. 2012;367(14):1355–60.
- Penning de Vries BBL, van Smeden M, Rosendaal FR, Groenwold RHH. Title, abstract, and keyword searching resulted in poor recovery of articles in systematic reviews of epidemiologic practice. *J Clin Epidemiol*. 2020;121:55–61.
- R Core Team R. R: A language and environment for statistical computing. 2023.
- Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and explanation. *Ann Intern Med*. 2018;169(7):467–73.
- Agier L, Basagaña X, Hernandez-Ferrer C, Maitre L, Tamayo Uria I, Urquiza J, et al. Association between the pregnancy exposome and fetal growth. *Int J Epidemiol*. 2020;49(2):572–86.
- Allison RM, Birken CS, Lebovic G, Howard AW, L'Abbe MR, Morency M-E, et al. Consumption of cow's milk in early childhood and fracture risk: A prospective cohort study. *Am J Epidemiol*. 2020;189(2):146–55.
- Badon SE, Quesenberry CP, Xu F, Avalos LA, Hedderson MM. Gestational weight gain, birthweight and early-childhood obesity: between- and within-family comparisons. *Int J Epidemiol*. 2020;49(5):1682–90.
- Barul C, Richard H, Parent M-E. Night-shift work and risk of prostate cancer: Results from a Canadian case-control study, the Prostate Cancer and Environment Study. *Am J Epidemiol*. 2019;188(10):1801–11.
- Bell GA, Perkins N, Buck Louis GM, Kannan K, Bell EM, Gao C, et al. Exposure to persistent organic pollutants and birth characteristics: The Upstate KIDS Study. *Epidemiol*. 2019;30:S94–100.
- Bell-Gorrod H, Fox MP, Boule A, Prozesky H, Wood R, Tanser F, et al. The impact of delayed switch to second-line antiretroviral therapy on mortality, depending on definition of failure time and CD4 count at failure. *Am J Epidemiol*. 2020;189(8):811–9.
- Bernasconi DP, Antolini L, Rossi E, Blanco-Lopez JG, Galimberti S, Andersen PK, et al. A causal inference approach to compare leukaemia treatment outcome in the absence of randomization and with dependent censoring. *Int J Epidemiol*. 2022;51(1):314–23.
- Bhatta L, Cepelis A, Vikjord SA, Malmo V, Laugsand LE, Dalen H, et al. Bone mineral density and risk of cardiovascular disease in men and women: the HUNT study. *Eur J Epidemiol*. 2021;36(11):1169–77.
- Bijlsma MJ, Wilson B, Tarkiainen L, Myrskylä M, Martikainen P. The impact of unemployment on antidepressant purchasing: adjusting for unobserved time-constant confounding in the g-formula. *Epidemiol*. 2019;30(3):388–95.
- Bjelland EK, Gran JM, Hofvind S, Eskild A. The association of birthweight with age at natural menopause: a population study of women in Norway. *Int J Epidemiol*. 2020;49(2):528–36.
- Blouin B, Casapia M, Kaufman JS, Joseph L, Larson C, Gyorkos TW. Bayesian methods for exposure misclassification adjustment in a mediation analysis: Hemoglobin and malnutrition in the association between: ascariis and IQ. *Epidemiol*. 2019;30(5):e41–2.
- Borch KB, Weiderpass E, Braaten T, Hansen MS, Licaj I. Risk of lung cancer and physical activity by smoking status and body mass index, the Norwegian Women and Cancer Study. *Eur J Epidemiol*. 2019;34(5):489–98.

39. Cepelis A, Brumpton BM, Laugsand LE, Dalen H, Langhammer A, Janszky I, et al. Asthma, asthma control and risk of acute myocardial infarction: HUNT study. *Eur J Epidemiol.* 2019;34(10):967–77.
40. Chasekwa B, Ntozini R, Church JA, Majo FD, Tavengwa N, Mutasa B, et al. Prevalence, risk factors and short-term consequences of adverse birth outcomes in Zimbabwean pregnant women: a secondary analysis of a cluster-randomized trial. *Int J Epidemiol.* 2022;51(6):1785–99.
41. Chen J, van der Duin D, Campos-Obando N, Ikram MA, Nijsten TEC, Uitterlinden AG, et al. Serum 25-hydroxyvitamin D3 is associated with advanced glycation end products (AGEs) measured as skin autofluorescence: The Rotterdam Study. *Eur J Epidemiol.* 2019;34(1):67–77.
42. Chen R, Tedroff K, Villamor E, Lu D, Cnattingius S. Risk of intellectual disability in children born appropriate-for-gestational-age at term or post-term: impact of birth weight for gestational age and gestational age. *Eur J Epidemiol.* 2020;35(3):273–82.
43. Chen Y, Kim ES, VanderWeele TJ. Religious-service attendance and subsequent health and well-being throughout adulthood: evidence from three prospective cohorts. *Int J Epidemiol.* 2020;49(6):2030–40.
44. Chen Z, Glisic M, Song M, Aliahmad HA, Zhang X, Moumjdian AC, et al. Dietary protein intake and all-cause and cause-specific mortality: results from the Rotterdam Study and a meta-analysis of prospective cohort studies. *Eur J Epidemiol.* 2020;35(5):411–29.
45. Chigogora S, Pearce A, Law C, Viner R, Chittleborough C, Griffiths LJ, et al. Could greater physical activity reduce population prevalence and socioeconomic inequalities in children's mental health problems? A policy simulation. *Epidemiol.* 2020;31(1):115–25.
46. Cohen JM, Wood ME, Hernández-Díaz S, Ystrom E, Nordeng H. Paternal antidepressant use as a negative control for maternal use: assessing familial confounding on gestational length and anxiety traits in offspring. *Int J Epidemiol.* 2019;48(5):1665–72.
47. Colen CG, Pinchak NP, Barnett KS. Racial disparities in health among college-educated African Americans: Can attendance at historically black colleges or universities reduce the risk of metabolic syndrome in midlife? *Am J Epidemiol.* 2021;190(4):553–61.
48. Coulombe J, Moodie EEM, Shortreed SM, Renoux C. Can the risk of severe depression-related outcomes be reduced by tailoring the antidepressant therapy to patient characteristics? *Am J Epidemiol.* 2021;190(7):1210–9.
49. Crump C, Friberg D, Li X, Sundquist J, Sundquist K. Preterm birth and risk of sleep-disordered breathing from childhood into mid-adulthood. *Int J Epidemiol.* 2019;48(6):2039–49.
50. Dam V, van der Schouw YT, Onland-Moret NC, Groenwold RHH, Peters SAE, Burgess S, et al. Association of menopausal characteristics and risk of coronary heart disease: a pan-European case-cohort analysis. *Int J Epidemiol.* 2019;48(4):1275–85.
51. Debras C, Chazelas E, Srour B, Julia C, Kesse-Guyot E, Zelek L, et al. Glycaemic index, glycaemic load and cancer risk: results from the prospective NutriNet-Santé cohort. *Int J Epidemiol.* 2022;51(1):250–64.
52. Dekhtyar S, Vetrano DL, Marengoni A, Wang H-X, Pan K-Y, Fratiglioni L, et al. Association between speed of multimorbidity accumulation in old age and life experiences: A cohort study. *Am J Epidemiol.* 2019;188(9):1627–36.
53. Delaney JA, Nance RM, Whitney BM, Crane HM, Williams-Nguyen J, Feinstein MJ, et al. Cumulative human immunodeficiency viremia, antiretroviral therapy, and incident myocardial infarction. *Epidemiol.* 2019;30(1):69–74.
54. Enthoven CA, Tideman JW, Polling JR, Tedja MS, Raat H, Iglesias AI, et al. Interaction between lifestyle and genetic susceptibility in myopia: the Generation R study. *Eur J Epidemiol.* 2019;34(8):777–84.
55. Ferraro AA, Barbieri MA, da Silva AAM, Goldani MZ, Fernandes MTB, Cardoso VC, et al. Cesarean delivery and hypertension in early adulthood. *Am J Epidemiol.* 2019;188(7):1296–303.
56. Flannagan KS, Mumford SL, Sjaarda LA, Radoc JG, Perkins NJ, Andriessen VC, et al. Is opioid use safe in women trying to conceive? *Epidemiol.* 2020;31(6):844–51.
57. Fraser GE, Jaceldo-Siegl K, Orlich M, Mashchak A, Sirirat R, Knutsen S. Dairy, soy, and risk of breast cancer: those confounded milks. *Int J Epidemiol.* 2020;49(5):1526–37.
58. Freedman LS, Agay N, Farmer R, Murad H, Olmer L, Dankner R. Metformin treatment among men with diabetes and the risk of prostate cancer: A population-based historical cohort study. *Am J Epidemiol.* 2022;191(4):626–35.
59. Garcia-Saenz A, de Miguel AS, Espinosa A, Costas L, Aragonés N, Tonne C, et al. Association between outdoor light-at-night exposure and colorectal cancer in Spain. *Epidemiol.* 2020;31(5):718–27.
60. George KM, Lutsey PL, Kucharska-Newton A, Palta P, Heiss G, Osypuk T, et al. Life-course individual and neighborhood socioeconomic status and risk of dementia in the Atherosclerosis Risk in Communities Neurocognitive Study. *Am J Epidemiol.* 2020;189(10):1134–42.
61. Gerlovin H, Posner DC, Ho Y-L, Rentsch CT, Tate JP, King JT Jr, et al. Pharmacoepidemiology, machine learning, and COVID-19: An intent-to-treat analysis of hydroxychloroquine, with or without azithromycin, and COVID-19 outcomes among hospitalized US veterans. *Am J Epidemiol.* 2021;190(11):2405–19.
62. Gero K, Hikichi H, Aida J, Kondo K, Kawachi I. Associations between community social capital and preservation of functional capacity in the aftermath of a major disaster. *Am J Epidemiol.* 2020;189(11):1369–78.
63. Gialamas A, Haag DG, Mittinty MN, Lynch J. Which time investments in the first 5 years of life matter most for children's language and behavioural outcomes at school entry? *Int J Epidemiol.* 2020;49(2):548–58.
64. Giorgianni F, Ernst P, Dell'Aniello S, Suissa S, Renoux C. β -agonists and the incidence of Parkinson disease. *Am J Epidemiol.* 2020;189(8):801–10.
65. Goin DE, Izano MA, Eick SM, Padula AM, DeMicco E, Woodruff TJ, et al. Maternal experience of multiple hardships and fetal growth: extending environmental mixtures methodology to social exposures. *Epidemiol.* 2021;32(1):18–26.
66. Goin DE, Pearson RM, Craske MG, Stein A, Pettifor A, Lippman SA, et al. Depression and incident HIV in adolescent girls and young women in HIV prevention trials network 068: targets for prevention and mediating factors. *Am J Epidemiol.* 2020;189(5):422–32.
67. Gram IT, Park S-Y, Maskarinec G, Wilkens LR, Haiman CA, Le Marchand L. Smoking and breast cancer risk by race/ethnicity and oestrogen and progesterone receptor status: the Multiethnic Cohort (MEC) study. *Int J Epidemiol.* 2019;48(2):501–11.
68. Hamad R, Batra A, Karasek D, LeWinn KZ, Bush NR, Davis RL, et al. The impact of the revised WIC food package on maternal nutrition during pregnancy and postpartum. *Am J Epidemiol.* 2019;188(8):1493–502.
69. Harlow AF, Hatch EE, Wesselink AK, Rothman KJ, Wise LA. Electronic cigarettes and fecundability: Results from a prospective preconception cohort study. *Am J Epidemiol.* 2021;190(3):353–61.
70. Harlow AF, Wesselink AK, Hatch EE, Rothman KJ, Wise LA. Male preconception marijuana use and spontaneous abortion: a prospective cohort study. *Epidemiol.* 2021;32(2):239–47.
71. He J-R, Hirst JE, Tikellis G, Phillips GS, Ramakrishnan R, Paltiel O, et al. Common maternal infections during pregnancy and childhood leukaemia in the offspring: findings from six international birth cohorts. *Int J Epidemiol.* 2021:dyab199.
72. Hillreiner A, Baumeister SE, Sedlmeier AM, Finger JD, Schlitt HJ, Leitzmann MF. Association between cardiorespiratory fitness and colorectal cancer in the UK Biobank. *Eur J Epidemiol.* 2020;35(10):961–73.
73. Hjorth S, Pottegård A, Broe A, Hemmingsen CH, Leinonen MK, Hargreave M, et al. Prenatal exposure to nitrofurantoin and risk of childhood leukaemia: A registry-based cohort study in four Nordic countries. *Int J Epidemiol.* 2021:dyab219.
74. Hlaváčová J, Flegr J, Řežábek K, Calda P, Kaňková Š. Male-to-female presumed transmission of toxoplasmosis between sexual partners. *Am J Epidemiol.* 2021;190(3):386–92.
75. Holzhausen EA, Hagen EW, LeCaire T, Cadmus-Bertram L, Malecki KC, Peppard PE. A comparison of self- and proxy-reported subjective sleep durations with objective actigraphy measurements in a survey of Wisconsin children 6–17 years of age. *Am J Epidemiol.* 2021;190(5):755–65.
76. Hu MD, Lawrence KG, Bodkin MR, Kwok RK, Engel LS, Sandler DP. Neighborhood deprivation, obesity, and diabetes in residents of the US Gulf Coast. *Am J Epidemiol.* 2021;190(2):295–304.
77. India-Aldana S, Rundle AG, Zeleniuch-Jacquotte A, Quinn JW, Kim B, Afanasyeva Y, et al. Neighborhood walkability and mortality in a prospective cohort of women. *Epidemiol.* 2021;32(6):763–72.
78. Inoue K, Mayeda ER, Paul KC, Shih IF, Yan Q, Yu Y, et al. Mediation of the associations of physical activity with cardiovascular events and mortality by diabetes in older Mexican Americans. *Am J Epidemiol.* 2020;189(10):1124–33.

79. Inoue K, Ritz B, Ernst A, Tseng W-L, Yuan Y, Meng Q, et al. Behavioral problems at age 11 years after prenatal and postnatal exposure to acetaminophen: Parent-reported and self-reported outcomes. *Am J Epidemiol.* 2021;190(6):1009–20.
80. Ishii M, Seki T, Kaikita K, Sakamoto K, Nakai M, Sumita Y, et al. Short-term exposure to desert dust and the risk of acute myocardial infarction in Japan: a time-stratified case-crossover study. *Eur J Epidemiol.* 2020;35(5):455–64.
81. Isumi A, Doi S, Ochi M, Kato T, Fujiwara T. Child maltreatment and mental health in middle childhood: A longitudinal study in Japan. *Am J Epidemiol.* 2022;191(4):655–64.
82. Janki S, Dehghan A, van de Wetering J, Steyerberg EW, Klop KWJ, Kimenai HJAN, et al. Long-term prognosis after kidney donation: a propensity score matched comparison of living donors and non-donors from two population cohorts. *Eur J Epidemiol.* 2020;35(7):699–707.
83. Kerschberger B, Boule A, Kuwengwa R, Ciglenecki I, Schomaker M. The impact of same-day antiretroviral therapy initiation under the World Health Organization treat-all policy. *Am J Epidemiol.* 2021;190(8):1519–32.
84. Kim K, Browne RW, Nobles CJ, Radin RG, Holland TL, Omosigbo UR, et al. Associations between preconception plasma fatty acids and pregnancy outcomes. *Epidemiol.* 2019;30:537–46.
85. Lara M, Labrecque JA, van Lenthe FJ, Voortman T. Estimating reductions in ethnic inequalities in child adiposity from hypothetical diet, screen time, and sports participation interventions. *Epidemiol.* 2020;31(5):736–44.
86. Leon ME, Schinasi LH, Lebaillly P, Beane Freeman LE, Nordby K-C, Ferro G, et al. Pesticide use and risk of non-Hodgkin lymphoid malignancies in agricultural cohorts from France, Norway and the USA: a pooled analysis from the AGRICOH consortium. *Int J Epidemiol.* 2019;48(5):1519–35.
87. Lepage B, Colineaux H, Kelly-Irving M, Vineis P, Delpierre C, Lang T. Comparison of smoking reduction with improvement of social conditions in early life: simulation in a British cohort. *Int J Epidemiol.* 2021;50(3):797–808.
88. Lergenmuller S, Ghiasvand R, Robsahm TE, Green AC, Lund E, Rueegg CS, et al. Sunscreens with high versus low sun protection factor and cutaneous squamous cell carcinoma risk: A population-based cohort study. *Am J Epidemiol.* 2022;191(1):75–84.
89. Lerro CC, Hofmann JN, Andreotti G, Koutros S, Parks CG, Blair A, et al. Dicamba use and cancer incidence in the agricultural health study: an updated analysis. *Int J Epidemiol.* 2020;49(4):1326–37.
90. Louie P, Upenieks L, Siddiqi A, Williams DR, Takeuchi DT. Race, flourishing, and all-cause mortality in the United States, 1995–2016. *Am J Epidemiol.* 2021;190(9):1735–43.
91. Love SAM, North KE, Zeng D, Petruski-Ivleva N, Kucharska-Newton A, Palta P, et al. Nine-year ethanol intake trajectories and their association with 15-year cognitive decline among black and white adults: The Atherosclerosis Risk in Communities Neurocognitive Study. *Am J Epidemiol.* 2020;189(8):788–800.
92. Lyall K, Windham GC, Snyder NW, Kuskovsky R, Xu P, Bostwick A, et al. Association between midpregnancy polyunsaturated fatty acid levels and offspring autism spectrum disorder in a California population-based case-control study. *Am J Epidemiol.* 2021;190(2):265–76.
93. Magnus MC, Fraser A, Rich-Edwards JW, Magnus P, Lawlor DA, Häberg SE. Time-to-pregnancy and risk of cardiovascular disease among men and women. *Eur J Epidemiol.* 2021;36(4):383–91.
94. Mårild K, Tapia G, Middtun Ø, Ueland PM, Magnus MC, Rewers M, et al. Smoking in pregnancy, cord blood cotinine and risk of celiac disease diagnosis in offspring. *Eur J Epidemiol.* 2019;34(7):637–49.
95. Mitchell A, Fall T, Melhus H, Wolk A, Michaëlsson K, Byberg L. Is the effect of Mediterranean diet on hip fracture mediated through type 2 diabetes mellitus and body mass index? *Int J Epidemiol.* 2021;50(1):234–44.
96. Mitha A, Chen R, Johansson S, Razav N, Cnattingius S. Maternal body mass index in early pregnancy and severe asphyxia-related complications in preterm infants. *Int J Epidemiol.* 2020;49(5):1647–60.
97. Mollan KR, Pence BW, Xu S, Edwards JK, Mathews WC, O’Cleirigh C, et al. Transportability from randomized trials to clinical care: On initial HIV treatment with efavirenz and suicidal thoughts or behaviors. *Am J Epidemiol.* 2021;190(10):2075–84.
98. Mooldijk SS, Licher S, Vinke EJ, Vernooij MW, Ikram MK, Ikram MA. Season of birth and the risk of dementia in the population-based Rotterdam Study. *Eur J Epidemiol.* 2021;36(5):497–506.
99. Naël V, Pérès K, Dartigues J-F, Letenneur L, Amieva H, Arleo A, et al. Vision loss and 12-year risk of dementia in older adults: the 3C cohort study. *Eur J Epidemiol.* 2019;34(2):141–52.
100. Nøst TH, Alcalá K, Urbarova I, Byrne KS, Guida F, Sandanger TM, et al. Systemic inflammation markers and cancer incidence in the UK Biobank. *Eur J Epidemiol.* 2021;36(8):841–8.
101. O’Brien KM, D’Aloisio AA, Shi M, Murphy JD, Sandler DP, Weinberg CR. Perineal talc use, douching, and the risk of uterine cancer. *Epidemiol.* 2019;30(6):845–52.
102. Ong YY, Sadananthan SA, Aris IM, Tint MT, Yuan WL, Huang JY, et al. Mismatch between poor fetal growth and rapid postnatal weight gain in the first 2 years of life is associated with higher blood pressure and insulin resistance without increased adiposity in childhood: the GUSTO cohort study. *Int J Epidemiol.* 2020;49(5):1591–603.
103. Oude Groeniger J, de Koster W, van der Waal J. Time-varying effects of screen media exposure in the relationship between socioeconomic background and childhood obesity. *Epidemiol.* 2020;31(4):578–86.
104. Pedersen KM, Çolak Y, Vedel-Krogh S, Kobylecki CJ, Bojesen SE, Nordestgaard BG. Risk of ulcerative colitis and Crohn’s disease in smokers lacks causal evidence. *Eur J Epidemiol.* 2021.
105. Pinto Pereira SM, De Stavola BL, Rogers NT, Hardy R, Cooper R, Power C. Adult obesity and mid-life physical functioning in two British birth cohorts: investigating the mediating role of physical inactivity. *Int J Epidemiol.* 2020;49(3):845–56.
106. Pongiglione B, Kern ML, Carpentieri JD, Schwartz HA, Gupta N, Goodman A. Do children’s expectations about future physical activity predict their physical activity in adulthood? *Int J Epidemiol.* 2020;49(5):1749–58.
107. Radojčić MR, Perera RS, Chen L, Spector TD, Hart DJ, Ferreira ML, et al. Specific body mass index trajectories were related to musculoskeletal pain and mortality: 19-year follow-up cohort. *J Clin Epidemiol.* 2022;141:54–63.
108. Ranzani OT, Milà C, Sanchez M, Bhogadi S, Kulkarni B, Balakrishnan K, et al. Association between ambient and household air pollution with carotid intima-media thickness in peri-urban South India: CHAI-Project. *Int J Epidemiol.* 2020;49(1):69–79.
109. Reese H, Routray P, Torondel B, Sinharoy SS, Mishra S, Freeman MC, et al. Assessing longer-term effectiveness of a combined household-level piped water and sanitation intervention on child diarrhoea, acute respiratory infection, soil-transmitted helminth infection and nutritional status: a matched cohort study in rural Odisha. *India International Journal of Epidemiology.* 2019;48(6):1757–67.
110. Reinhard E, Carrino L, Courtin E, van Lenthe FJ, Avendano M. Public transportation use and cognitive function in older age: A quasiexperimental evaluation of the Free Bus Pass Policy in the United Kingdom. *Am J Epidemiol.* 2019;188(10):1774–83.
111. Rhee J, Lofffield E, Freedman ND, Liao LM, Sinha R, Purdue MP. Coffee consumption and risk of renal cell carcinoma in the NIH-AARP Diet and Health Study. *Int J Epidemiol.* 2021;50(5):1473–81.
112. Richardson K, Mattishent K, Loke YK, Steel N, Fox C, Grossi CM, et al. History of benzodiazepine prescriptions and risk of dementia: Possible bias due to prevalent users and covariate measurement timing in a nested case-control study. *Am J Epidemiol.* 2019;188(7):1228–36.
113. Riddell CA, Goin DE, Morello-Frosch R, Apte JS, Glymour MM, Torres JM, et al. Hyper-localized measures of air pollution and risk of preterm birth in Oakland and San Jose. *California International Journal of Epidemiology.* 2021;50(6):1875–85.
114. Rist PM, Buring JE, Rexrode KM, Cook NR, Rost NS. Prospectively collected lifestyle and health information as risk factors for white matter hyperintensity volume in stroke patients. *Eur J Epidemiol.* 2019;34(10):957–65.
115. Robert A, Edmunds WJ, Watson CH, Henao-Restrepo AM, Gsell P-S, Williamson E, et al. Determinants of transmission risk during the late stage of the West African ebola epidemic. *Am J Epidemiol.* 2019;188(7):1319–27.
116. Rogers NT, Blodgett JM, Searle SD, Cooper R, Davis DHJ, Pinto Pereira SM. Early-life socioeconomic position and the accumulation of health-related deficits by midlife in the 1958 British Birth Cohort Study. *Am J Epidemiol.* 2021;190(8):1550–60.

117. Rogers NT, Power C, Pinto Pereira SM. Birthweight, lifetime obesity and physical functioning in mid-adulthood: a nationwide birth cohort study. *Int J Epidemiol.* 2020;49(2):657–65.
118. Rovio SP, Pihlman J, Pahkala K, Juonala M, Magnussen CG, Pitkänen N, et al. Childhood exposure to parental smoking and midlife cognitive function: The Young Finns Study. *Am J Epidemiol.* 2020;189(11):1280–91.
119. Rudolph KE, Gimbrone C, Díaz I. Helped into harm: Mediation of a housing voucher intervention on mental health and substance use in boys. *Epidemiol.* 2021;32(3):336–46.
120. Rudolph KE, Levy J, Schmidt NM, Stuart EA, Ahern J. Using transportability to understand differences in mediation mechanisms across trial sites of a housing voucher experiment. *Epidemiology.* 2020;31(4):523–33.
121. Salmon C, Song L, Muir K, Pashayan N, Dunning AM, Batra J, et al. Marital status and prostate cancer incidence: a pooled analysis of 12 case–control studies from the PRACTICAL consortium. *Eur J Epidemiol.* 2021;36(9):913–25.
122. Sangaramoorthy M, Hines LM, Torres-Mejía G, Phipps AI, Baumgartner KB, Wu AH, et al. A pooled analysis of breastfeeding and breast cancer risk by hormone receptor status in parous Hispanic women. *Epidemiol.* 2019;30(3):449–57.
123. Sato K, Amemiya A, Haseda M, Takagi D, Kanamori M, Kondo K, et al. Postdisaster changes in social capital and mental health: A natural experiment from the 2016 Kumamoto earthquake. *Am J Epidemiol.* 2020;189(9):910–21.
124. Schliep KC, Mumford SL, Silver RM, Wilcox B, Radin RG, Perkins NJ, et al. Preconception perceived stress is associated with reproductive hormone levels and longer time to pregnancy. *Epidemiol.* 2019;30.
125. Schuch HS, Nascimento GG, Peres KG, Mittinty MN, Demarco FF, Correa MB, et al. The controlled direct effect of early-life socioeconomic position on periodontitis in a birth cohort. *Am J Epidemiol.* 2019;188(6):1101–8.
126. Schwartz GL, Leifheit KM, Berkman LF, Chen JT, Arcaya MC. Health selection into eviction: Adverse birth outcomes and children's risk of eviction through age 5 years. *Am J Epidemiol.* 2021;190(7):1260–9.
127. Sellers R, Warne N, Rice F, Langley K, Maughan B, Pickles A, et al. Using a cross-cohort comparison design to test the role of maternal smoking in pregnancy in child mental health and learning: evidence from two UK cohorts born four decades apart. *Int J Epidemiol.* 2020;49(2):390–9.
128. Shiba K, Hanazato M, Aida J, Kondo K, Arcaya M, James P, et al. Cardiometabolic profiles and change in neighborhood food and built environment among older adults: a natural experiment. *Epidemiol.* 2020;31(6):758–67.
129. Shiba K, Hikichi H, Aida J, Kondo K, Kawachi I. Long-term associations between disaster experiences and cardiometabolic risk: A natural experiment from the 2011 great east Japan earthquake and tsunami. *Am J Epidemiol.* 2019;188(6):1109–19.
130. Shiba K, Torres JM, Daoud A, Inoue K, Kanamori S, Tsuji T, et al. Estimating the impact of sustained social participation on depressive symptoms in older adults. *Epidemiol.* 2021;32(6):886–95.
131. Stacy SL, Buchanich JM, Ma ZQ, Mair C, Robertson L, Sharma RK, et al. Maternal obesity, birth size, and risk of childhood cancer development. *Am J Epidemiol.* 2019;188(8):1503–11.
132. Su Y, D'Arcy C, Meng X. Social support and positive coping skills as mediators buffering the impact of childhood maltreatment on psychological distress and positive mental health in adulthood: Analysis of a national population-based sample. *Am J Epidemiol.* 2020;189(5):394–402.
133. Sudharsanan N, Ho JY. Rural–urban differences in adult life expectancy in Indonesia: a parametric g-formula–based decomposition approach. *Epidemiol.* 2020;31(3):393–401.
134. Tefft BC, Arnold LS. Estimating cannabis involvement in fatal crashes in Washington State before and after the legalization of recreational cannabis consumption using multiple imputation of missing values. *Am J Epidemiol.* 2021;190(12):2582–91.
135. Torres JM, Rudolph KE, Sofrygin O, Wong R, Walter LC, Glymour MM. Having an adult child in the United States, physical functioning, and unmet needs for care among older Mexican adults. *Epidemiol.* 2019;30(4):553–60.
136. Torres JM, Sofrygin O, Rudolph KE, Haan MN, Wong R, Glymour MM. US migration status of adult children and cognitive decline among older parents who remain in Mexico. *Am J Epidemiol.* 2020;189(8):761–9.
137. Tsarna E, Reedijk M, Birks LE, Guxens M, Ballester F, Ha M, et al. Associations of maternal cell-phone use during pregnancy with pregnancy duration and fetal growth in 4 birth cohorts. *Am J Epidemiol.* 2019;188(7):1270–80.
138. Vable AM, Duarte CD, Cohen AK, Glymour MM, Ream RK, Yen IH. Does the type and timing of educational attainment influence physical health? A novel application of sequence analysis. *Am J Epidemiol.* 2020;189(11):1389–401.
139. van der Schaft N, Schoufour JD, Nano J, Kiefte-de Jong JC, Muka T, Sijbrands EJG, et al. Dietary antioxidant capacity and risk of type 2 diabetes mellitus, prediabetes and insulin resistance: the Rotterdam Study. *Eur J Epidemiol.* 2019;34(9):853–61.
140. van Gennip ACE, Sedaghat S, Carnethon MR, Allen NB, Klein BEK, Cotch MF, et al. Retinal microvascular caliber and incident depressive symptoms: The Multi-Ethnic Study of Atherosclerosis. *Am J Epidemiol.* 2022;191(5):843–55.
141. van Lee L, Crozier SR, Aris IM, Tint MT, Sadananthan SA, Michael N, et al. Prospective associations of maternal choline status with offspring body composition in the first 5 years of life in two large mother–offspring cohorts: the Southampton Women's Survey cohort and the Growing Up in Singapore Towards healthy Outcomes cohort. *Int J Epidemiol.* 2019;48(2):433–44.
142. Wagner M, Grodstein F, Proust-Lima C, Samieri C. Long-term trajectories of body weight, diet, and physical activity from midlife through late life and subsequent cognitive decline in women. *Am J Epidemiol.* 2020;189(4):305–13.
143. Walsemann KM, Ailshire JA. Early educational experiences and trajectories of cognitive functioning among US adults in midlife and later. *Am J Epidemiol.* 2020;189(5):403–11.
144. Wang C-R, Hu T-Y, Hao F-B, Chen N, Peng Y, Wu J-J, et al. Type 2 diabetes–prevention diet and all-cause and cause-specific mortality: A prospective study. *Am J Epidemiol.* 2022;191(3):472–86.
145. Wang H, László KD, Gissler M, Li F, Zhang J, Yu Y, et al. Maternal hypertensive disorders and neurodevelopmental disorders in offspring: a population-based cohort in two Nordic countries. *Eur J Epidemiol.* 2021;36(5):519–30.
146. Wesselink AK, Bresnick KA, Hatch EE, Rothman KJ, Mikkelsen EM, Wang TR, et al. Association between male use of pain medication and fecundability. *Am J Epidemiol.* 2020;189(11):1348–59.
147. Wesselink AK, Claus Henn B, Fruh V, Orta OR, Weuve J, Hauser R, et al. A prospective ultrasound study of plasma polychlorinated biphenyl concentrations and incidence of uterine leiomyomata. *Epidemiol.* 2021;32(2):259–67.
148. White J, Fluharty M, de Groot R, Bell S, Batty GD. Mortality among rough sleepers, squatters, residents of homeless shelters or hotels and sofa-surfers: a pooled analysis of UK birth cohorts. *Int J Epidemiol.* 2021;51(3):839–46.
149. Williams-Nguyen J, Hawes SE, Nance RM, Lindström S, Heckbert SR, Kim HN, et al. Association between chronic hepatitis C virus infection and myocardial infarction among people living with HIV in the United States. *Am J Epidemiol.* 2020;189(6):554–63.
150. Xiao J, Gao Y, Yu Y, Toft G, Zhang Y, Luo J, et al. Associations of parental birth characteristics with autism spectrum disorder (ASD) risk in their offspring: a population-based multigenerational cohort study in Denmark. *Int J Epidemiol.* 2021;50(2):485–95.
151. Xuan Y, Bobak M, Anusriti A, Jansen EHJM, Paják A, Tamosiunas A, et al. Association of serum markers of oxidative stress with myocardial infarction and stroke: pooled results from four large European cohort studies. *Eur J Epidemiol.* 2019;34(5):471–81.
152. Yisahak SF, Hinkle SN, Mumford SL, Li M, Andriessen VC, Grantz KL, et al. Vegetarian diets during pregnancy, and maternal and neonatal outcomes. *Int J Epidemiol.* 2021;50(1):165–78.
153. Youssim I, Gorfine M, Calderon-Margalit R, Manor O, Paltiel O, Siscovick DS, et al. Holocaust experience and mortality patterns: 4-decade follow-up in a population-based cohort. *Am J Epidemiol.* 2021;190(8):1541–9.
154. Yu D-W, Li Q-J, Cheng L, Yang P-F, Sun W-P, Peng Y, et al. Dietary vitamin K intake and the risk of pancreatic cancer: A prospective study of 101,695 american adults. *Am J Epidemiol.* 2021;190(10):2029–41.
155. Yuan J, Hu YJ, Zheng J, Kim JH, Sumerlin T, Chen Y, et al. Long-term use of antibiotics and risk of type 2 diabetes in women: a prospective cohort study. *Int J Epidemiol.* 2020;49(5):1572–81.

156. Zhou Z, Lin C, Ma J, Towne SD, Han Y, Fang Y. The association of social isolation with the risk of stroke among middle-aged and older adults in china. *Am J Epidemiol*. 2019;188(8):1456–65.
157. Beesley LJ, Taylor JM. A stacked approach for chained equations multiple imputation incorporating the substantive model. *Biometrics*. 2021;77(4):1342–54.
158. Lu H, Cole SR, Howe CJ, Westreich D. Toward a clearer definition of selection bias when estimating causal effects. *Epidemiology*. 2022;33(5):699–706.
159. Bell ML, Fiero M, Horton NJ, Hsu C-H. Handling missing data in RCTs; a review of the top medical journals. *BMC Med Res Methodol*. 2014;14(1):1–8.
160. Carpenter JR, Bartlett JW, Morris TP, Wood AM, Quartagno M, Kenward MG. *Multiple imputation and its application*. Hoboken: Wiley; 2023.

Publisher's Note

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