Current Epidemiology Reports (2022) 9:223–232
https://doi.org/10.1007/s40471-022-00306-8

Causal Inference with Case-Only Studies in Injury Epidemiology Research

Andrew G. Rundle[1] · **Michael D. M. Bader**[2] · **Charles C. Branas**[1] · **Gina S. Lovasi**[3] · **Stephen J. Mooney**[4] · **Christopher N. Morrison**[1] · **Kathryn M. Neckerman**[1]

Accepted: 26 August 2022 / Published online: 29 September 2022

Abstract

Purpose of Review We review the application and limitations of two implementations of the "case-only design" in injury epidemiology with example analyses of Fatality Analysis Reporting System data.

Recent Findings The term "case-only design" covers a variety of epidemiologic designs; here, two implementations of the design are reviewed: (1) studies to uncover etiological heterogeneity and (2) studies to measure exposure effect modification. These two designs produce results that require different interpretations and rely upon different assumptions. The key assumption of case-only designs for exposure effect modification, the more commonly used of the two designs, does not commonly hold for injuries and so results from studies using this design cannot be interpreted. Case-only designs to identify etiological heterogeneity in injury risk are interpretable but only when the case-series is conceptualized as arising from an underlying cohort.

Summary The results of studies using case-only designs are commonly misinterpreted in the injury literature.

Keywords Study design · Case-only design · Injury research · Pedestrian injury · Etiologic heterogeneity · Effect modification

Introduction

Retrospective designs that compare cases to comparison groups are well suited to injury epidemiology for two important reasons: injury outcomes are rare at the population level, meaning prospective study is often infeasible; and the latent period between putative causes and injury outcomes is often very short, so it is justifiable to assess exposure and outcome status at the same time. The term "case-only design" has been applied to a wide variety of retrospective study designs, two of which, the case-only design to investigate etiological heterogeneity [1] and the case-only design to measure exposure effect modification [2, 3], have been frequently used in injury epidemiology (for instance, see [4–11]). These two designs, also called case-series, case-case or case^2 studies, share a common design strategy, in which a case-series is stratified into two or more case groups, which are then compared in regard to some putative predictor variable [1, 2, 12]. The variable used to stratify the case-series into sub-groups is the dependent variable in the case-only analyses [1, 2]. For instance, in a case-series of automobile crashes, data on blood alcohol levels of drivers can be used to divide crashes into those with drivers under the influence of alcohol and those with drivers who were not [13]. Logistic regression analyses are then performed to determine whether predictor variables (e.g., age of driver) are associated with one or other of the two automobile crash sub-types. These two case-only designs are attractive for injury epidemiology because of the difficulties in selecting controls and gathering data on

This article is part of the Topical Collection on *Injury Epidemiology.*

Andrew G. Rundle
Agr3@cumc.columbia.edu

1 Department of Epidemiology, Mailman School of Public Health, Columbia University, 722 West 168th Street, Room 727, New York, NY 10032, USA

2 Department of Sociology, Johns Hopkins University, Baltimore, MD, USA

3 Department of Epidemiology and Biostatistics, Drexel University, Philadelphia, PA, USA

4 Department of Epidemiology, University of Washington, Seattle, WA, USA

exposures of interest for injury, which are often of a sensitive nature, such as alcohol and drug use, risk-taking behaviors, mental health issues, and adverse life events [14, 15].

Because the phrase “case-only design” has been used in the literature to refer to a variety of different designs, the reader is cautioned to distinguish these two designs from other designs sometimes referred to as “case-only designs.” Other designs sometimes referred to as “case-only” are case-series that are observed at two time points for exposure status (AKA, case-crossover, self-controlled case series, self-controlled risk interval) or a case-series study in which the proportion of exposed cases is so large, it obviously differs from the general population (e.g., chimney sweeping among scrotal cancer patients) [16]. Furthermore the two designs reviewed here are both referred to in the literature as “case-only designs” and employ similar regression-based analytical approaches, but they generate statistical parameters calling for different interpretations [1, 2, 17]. Furthermore, the odds ratios (ORs) generated from the case-only regression analyses used in these two designs carry very different meanings than the OR parameters estimated by the case-control or cohort analyses [1, 2, 18]. Here, we review the case-only design to investigate etiological heterogeneity [1] and the case-only design to measure exposure effect modification [2, 3]. We describe the value of each design and then describe how each may be used in injury epidemiology, providing example analyses of Fatality Analysis Reporting System (FARS) data on pedestrian fatalities.

How to Interpret Results From These Two Case-Only Designs

To clarify interpretations of case-only parameters, it is worth reviewing why case-control studies offer estimates of causal effects. Case-control studies are valid and interpretable because they can be interpreted as stratified selection from an underlying cohort, and so have the potential to produce unbiased estimates of causal effects of risk factors on outcomes. If done correctly, with samples of cases and controls properly chosen to represent the target population that the study is intended to reflect, case-control studies can be seen as efficient versions of larger, and often time- and resource-prohibitive, cohort studies [3, 19, 20]. It is understood that odds ratios from case-control studies are valid estimates of risk ratios or rate ratios that would have otherwise been estimated from a cohort study that generated the same case-series as analyzed in the case-control study [3]. Estimates identified in case-control studies generalize to the population when the case and control samples represent the exposure history of cases and controls in a broader target population of interest.

To be interpretable, the two case-only designs reviewed here must be understood to utilize the same case-series that otherwise would have been utilized in a case-control study, or if the case-control study were nested in an extant cohort, the case-series that would have been generated from the cohort [1–3, 18]. However, the odds ratio estimated by these case-only designs does not estimate the causal effect of a risk factor on an outcome: to accomplish this, an epidemiologically sound control series is required. The estimand in these two case-only design depends on the type of variable used to stratify or group the case-series. In case-only studies of *etiological heterogeneity*, the stratifying variable describes some inherent characteristic of case-ness that does not have a logical or comparable value for non-cases [1]. In the case-only study of *exposure effect modification*, the stratifying variable can be used to describe both cases and non-cases and could be analyzed as a risk factor for the overall outcome in a full cohort or case-control study [2, 3]. Figure 1 provides a flow chart for conducting and interpreting case-only designs. Discussions of the results of case-only analyses in injury epidemiology rarely state the hypothesis being tested by the design and rarely relate the estimated case-only OR to the OR that would have been estimated in case-control analyses of the case-series under investigation. As such, we argue that the results of these two case-only designs are commonly misinterpreted in injury epidemiology.

Etiological Heterogeneity

The case-only study of etiological heterogeneity tests whether a risk factor has a different causal effect for one case-subtype compared to another case-subtype [1]. It can be used to identify potential mechanisms that may explain how different forms of an outcome come about [1]. It cannot, however, determine population-level risk of first experiencing or contracting those outcomes [1, 12].

Begg and Zhang originally described this design in their study of smoking’s influence on whether a patient had one sub-type of bladder cancer verses another subtype [1, 21]. In this study, cases were classified into two groups based on the presence or absence of a mutation in the p53 gene in the tumor tissue, a classification that had no meaning for controls, as by definition controls have no bladder tumor tissue. Begg and Zhang showed that the odds ratio (OR) for smoking on p53+ status from the case-only design, $\theta_{\text{Case-Only}}$, was equivalent to the ratio of two case-control ORs for smoking estimated from a full case-control study: the OR when p53+ cancer cases were compared to controls, θ_1, and the OR when p53- cancer cases were compared to controls, θ_2. [1] That is, $\theta_{\text{Case-Only}} = \theta_1/\theta_2$.

A logistic regression model predicting tumor p53+ status based *on smoking among cases only is*:

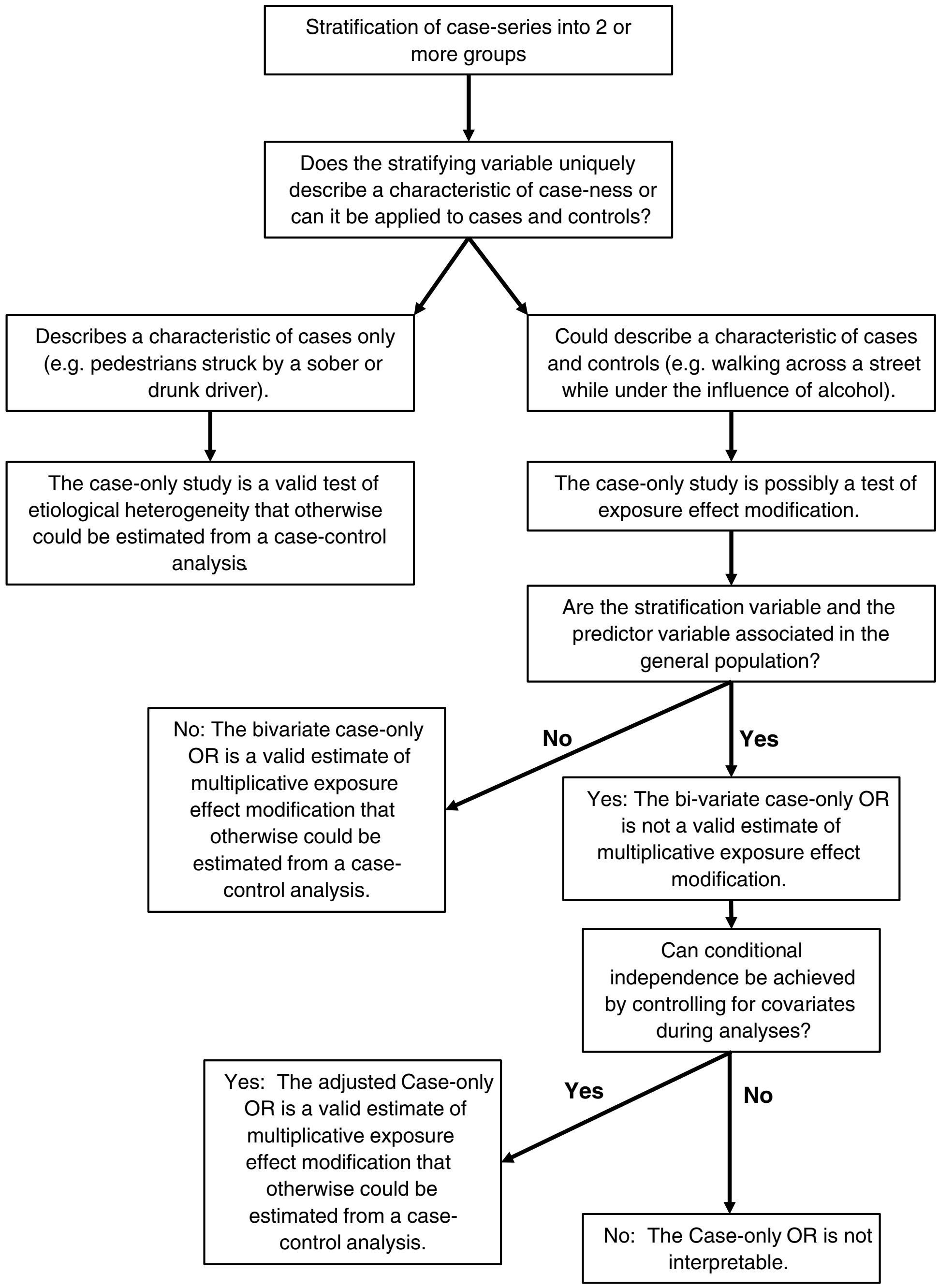

Fig. 1 Flow chart for conducting and interpreting a case-only analysis

$$\log\left[\mathrm{P(Y)}/(1-\mathrm{P(Y)})\right] = \alpha + \beta \mathrm{X_i} \tag{1}$$

where Y represents a case with the p53 mutation and X_i represents whether subject i was a current smoker and $e^{\beta} = \theta_{\text{Case-Only}} = \theta_1/\theta_2$, which can equivalently be estimated from the case-only analyses or case-control analyses of the same case series. The exponentiated coefficient from the case-only model is interpreted as describing the extent to which an exposure differs in its effect on one subtype of cases compared to another, a phenomenon described as etiological heterogeneity [1]. The case-only odds ratio, e^{β}, only reflects the extent to which there is a *difference* in the effect of an exposure between the two subtypes of cases. It does not provide information about the effect of smoking on the risk of getting bladder cancer. An e^{β}*case-only* = 2 could indicate that $\theta_1 = 2$ and $\theta_2 = 1$ or that $\theta_1 = 3$ and $\theta_2 = 1.5$ or even $\theta_1 = 1$ and $\theta_2 = 0.5$; thus, it informs us on the ratio of θ_2 to θ_1 but not the values of θ_2 and θ_1 [1].

The rationale for controlling for covariates in a case-only design of etiologic heterogeneity is similar to that in a traditional case-control or cohort study; there is a class of variables that if uncontrolled for will cause bias in the estimate of the case-only OR for the exposure of interest. Such variables are similar to traditional confounders in that they are associated with the exposure variable of interest, but to "confound" a case-only study of etiologic heterogeneity and create an omitted variable bias, they must also show etiologic heterogeneity for the sub-grouped outcomes in question. If adjustment for a potential confounder causes a similar attenuation for both θ_1 and θ_2, such as with θ_1 being attenuated from 4 to 3 after adjustment and θ_2 being attenuated from 2 to 1.5, their ratio may be unchanged (with θ_1/θ_2 remaining 2 both before and after adjustment in this case).

Exposure Effect Modification

The case-only study of exposure effect modification tests whether one exposure modifies the effect of another exposure (or an intrinsic characteristic of the study participant such as age) [2, 3, 18]. This design measures the extent of multiplicative interaction between two exposures that would otherwise be estimated using a case-control or cohort study, under the assumption that there is no association between the two exposures in the underlying source population (the independence assumption) [2, 3, 18]. If the independence assumption does not hold, the univariate OR from a case-only analysis of exposure effect modification is not interpretable. An advantage of the case-only design is that, if the independence assumption holds, it provides a more statistically efficient estimate of the multiplicative interaction term than would otherwise be generated from case-control analyses of the same case series [2]. This design was first described in the context of gene-by-environment interactions, where the independence assumption was thought to commonly hold [2, 3, 18]. Examples are case-only studies of the interaction between the N-acetyltransferase 2 (NAT-2) genetic polymorphisms and cigarette smoking on bladder cancer risk [17]. Its utility has expanded to include interactions between other pairs of risk factor variables including gene-by-gene and environment-by-environment. Unlike case-only studies of etiological heterogeneity, case-only designs for estimating exposure effect modification use variables that measure exposures and/or characteristics observable in both cases and controls [3, 18].

In the underlying population from which cases arise, two risk factors (e.g., the NAT-2 genetic polymorphism and cigarette smoking) may interact to affect the odds of an outcome (e.g., bladder cancer). If so, the population may be modeled by Eq. (2):

$$\log\left[P(Y)/(1-P(Y))\right] = \alpha + \beta_1 X + \beta_2 Z + \beta_3 X^{*} Z \tag{2}$$

A cohort or case-control study could be devised to estimate such a model, since X and Z can be observed in both cases and controls or in an entire cohort. For example, X may represent a gene (e.g., NAT-2) that modifies the effect of an environmental or behavioral risk factor, Z (e.g., smoking) on disease risk (e.g., bladder cancer) [17].

Using a case-only design, we could estimate the effect modification of X on Z based on predicting the presence of X among cases based on Z under the assumption that X and Z are conditionally uncorrelated among (unobserved) controls. From the example of bladder cancer risk, the OR for the association between NAT-2 genetic polymorphism status (X) and cigarette smoking status (Z) among cases was calculated [17]. In the Eq. (3) below, implemented among cases only, γ_1 is equivalent to β_3 in Eq. (2):

$$\log\left[P(X)/(1-P(X))\right] = \gamma_0 + \gamma_1 Z \tag{3}$$

The rationale for controlling for covariates in a case-only study of exposure effect modification is very different from the rationale for controlling for covariates in a cohort or case-control study [18]. In circumstances where X and Z are associated in the underlying population, covariate control can be used in case-only analyses to establish conditional independence between X and Z, so that a non-biased estimate of the magnitude of effect modification between X and Z can be generated [18]. In practical terms, this means conceptualizing why X and Z are associated in the underlying population and identifying a variable(s), M, such that X and Z are independent, conditional on M [18]. This variable(s) M is then included in the case-only analysis as a covariate, with case-only logistic regression model taking the form of $\log[P(X)/(1\text{-}P(X))] = \gamma_0 + \gamma_1 Z + \gamma_2 M$; however, the OR for M is not of interest itself and is not interpretable.

A multivariate case-only analysis that includes multiple covariates, e.g., $Z_1, Z_2, \ldots, Z_k$, each conceptualized as an exposure, would generate a series of corresponding ORs measuring the effect modification of each Z_k on X. A case-only logistic regression model of the form $\log[P(X)/(1\text{-}P(X))] = \gamma_0 + \gamma_1 Z_1 + \gamma_2 Z_2$, where Z_2 is another exposure, estimates the X^*Z_1 and X^*Z_2 interaction terms from the model $\log[P(Y)/(1\text{-}P(Y))] = \alpha + \beta_1 X + \beta_2 Z_1 + \beta_3 Z_2 + \beta_4 X^*Z_1 + \beta_5 X^*Z_2$ fit in a case-control study, where X and Z_1, and X and Z_2, are assumed to be unassociated in the general population. Practically, however, interpreting multiple effect modifiers of X quickly becomes unwieldy.

Implications for Case-Series Analyses in Injury Epidemiology

When these two case-only designs are implemented in injury epidemiology, we argue, the analyses are commonly misinterpreted, most notably because researchers overlook the independence assumption required for the case-only studies of effect modification. Considering alcohol consumption as a risk factor for pedestrian fatality, a case-only study with the driver's alcohol status as the stratifying variable tests a completely different type of hypothesis than a case-only study with pedestrian alcohol consumption as the stratifying variable [22–28]. In the context of these two options for stratifying a case-series on alcohol involvement, Supplemental Fig. 1 illustrates the relationship between case-only and case-control analyses for etiologic heterogeneity and Supplemental Fig. 2 illustrates the relationship between case-only and case-control analyses for multiplicative interaction. The rest of this section illustrates the application of these two case-only designs using analyses of 2017 and 2018 FARS data. Table 1 shows the results of two case-only analyses of the association between pedestrian age and two groups of pedestrian fatalities, one stratifying based on the driver's alcohol status and the second stratifying on the pedestrian's alcohol status.

In the first analysis, the stratifying variable is whether the *driver* involved in the accident was identified as a "drinking driver" (based either on police reports or a positive alcohol test) [29]. Driver alcohol-involvement can be used to stratify the pedestrian fatality case-series, but cannot logically be used to stratify or describe non-cases, i.e., individuals in the underlying case-control study that were not killed by an automobile. Thus, this case-only analysis is a test of *etiological heterogeneity,* that is, whether the age of the pedestrian is differentially associated with being killed by a drunk versus sober driver. Alternatively, if this case-series was analyzed within a case-control design, two sets of OR for the effect of age group on fatality risk would be calculated. The first would be calculated comparing the age distribution of pedestrians killed by drunk drivers to the age distribution of controls, and the second would be calculated comparing the age distribution of pedestrians killed by sober drivers to the age distribution of controls. The ratio of these two sets of OR for the age groups would equal the OR generated for age from case-only analyses of the case-series alone (see Supplemental Fig. 1).

The results from the first case-only analysis suggest that the effect of age on risk of pedestrian fatality is similar for crashes involving a drunk driver compared to crashes involving sober drivers, up until the age of 60+ years. These analyses suggest that compared to those age 16 to 20 years old, individuals age 60 years or older are less likely to be killed in a crash involving a drunk driver than a crash involving a sober driver. However, these analyses do not provide evidence that this older age group is at lower risk for pedestrian fatality overall. Further adjustment for pedestrian sex and race does not materially alter the ORs for age groups, because, after adjustment for age and each other, sex and race do not show substantial etiologic heterogeneity for the two pedestrian fatality sub-groups.

In the second analysis, the case-series could be conceptualized as part of a case-control study in which the alcohol status of control pedestrians is assessed [14]. Thus, this case-only analysis is a test of *exposure effect modification*. Matched controls could be enrolled, with data collected on whether they were walking outdoors and had recently consumed alcohol at the time when the case was killed [14]. These controls would allow estimation of the prevalence of individuals walking while under the influence of alcohol in the population from which the cases series arose, allowing a valid test of the association between walking while under the influence of alcohol and pedestrian fatality risk [3, 14, 20]. This case-control study could also assess multiplicative interactions between other exposures or study participant characteristics, in our example the age and alcohol consumption status of the pedestrian. The interaction effect would express the extent to which the effect of consuming alcohol and walking on the odds of being fatally struck by a car depends on the pedestrian's age.

When the case-series is analyzed using a case-only approach, the OR estimates the extent of multiplicative exposure effect modification between the pedestrian's alcohol consumption status and the age of the pedestrian. The estimated case-only univariate odds ratios from the FARS data in Table 1 suggest that, in a full case-control analysis, all of the age group × pedestrian alcohol status interaction terms, except for the 60+ age group, would be significantly different from 1. That is, the effect of walking after consuming alcohol on the odds of being fatally struck by a car depends on the age of the pedestrian. Critically, this interpretation of the analysis depends on the assumption that alcohol consumption and pedestrian age are not associated in the general population [2, 3]. However, survey data on alcohol consumption across age groups suggests that this independence assumption does not hold [30]. Thus, it is unlikely that the univariate OR generated in the case-only analyses of the FARS can be interpreted as reflecting valid estimates of exposure effect modification.

Valid interpretation of results from case-only studies of exposure effect modification requires careful attention to this independence assumption—the assumption that the two exposures of interest are unassociated in the source population [2, 3, 18]. The original formulation of case-only studies of exposure effect modification focused on gene-environment interactions, for which the assumption of gene and environment independence in the source population, is often plausible [2, 3, 18, 31]. However, many of the variables used to subset injury case series for case-only analyses of exposure effect modification are socially patterned (e.g., alcohol use) or are associated with social patterning of behaviors (e.g., sex, age, race, mental health). Thus, it seems unlikely that the independence assumption will commonly

Table 1 Results of analyses of the two different case-only designs applied to pedestrian fatalities in the 2017-2018 FARS data[1]

Case-only analysis of etiological heterogeneity for driver's alcohol status: univariate and multivariate analyses							Case-only analysis of exposure effect modification for pedestrian age x pedestrian alcohol status			
	Driver's alcohol status (Dependent variable) Univariate model			Driver's alcohol status (Dependent variable) Multivariate model				Pedestrian's alcohol status (Dependent variable)		
Pedestrian's demographics	Positive (N)	Negative (N)	OR (95% CI)	Positive (N)	Negative (N)	OR (95% CI)	Pedestrian's demographics	Positive (N)	Negative (N)	OR (95% CI)
Age										
16–20	47	516	1	47	516	1	16-20	106	275	1
21–29	157	1465	1.22 (0.85, 1.74)	157	1465	1.23 (0.86, 1.75)	21-29	559	575	2.75 (2.07, 3.66)
30–39	141	1722	0.89 (0.62, 1.28)	141	1722	0.90 (0.62, 1.29)	30-39	640	692	2.64 (1.98, 3.51)
40–49	179	1627	1.22 (0.85, 1.77)	179	1627	1.23 (0.85, 1.78)	40-49	635	652	2.70 (2.05, 3.59)
50–59	175	2204	0.90 (0.62, 1.31)	175	2204	0.90 (0.62, 1.31)	50-59	837	794	3.01 (2.32, 3.89)
60+	180	3456	0.62 (0.44, 0.86)	180	3456	0.61 (0.44, 0.86)	60+	561	1663	0.97 (0.76, 1.31)
Sex										
Male				605	7723	1				
Female				274	3267	1.11 (0.95, 1.28)				
Race										
White				592	7379	1				
Black				174	2173	0.97 (0.80, 1.18)				
Native American				21	243	0.68 (0.40, 1.18)				
Asian				36	427	1.10 (0.75, 1.62)				
Other, Unknown				56	768	0.92 (0.66, 1.28)				

1. For pedestrian fatalities occurring within U.S. Metropolitan areas in 2017 or 2018.

Table 2 Interpretation of ten case-only studies from the literature

Stratification of the case series	Can the stratification variable be applied to a control series?	Type of case-only design	Risk factors tested	Independence assumption required?	Likelihood that independence assumption holds
Pedestrian injuries classified as fatal or not fatal [32]	No	Etiological heterogeneity	Pedestrian age and impairment with alcohol and/or drugs	No	NA
Car crashes with drivers categorized by injury severity [33]	No	Etiological heterogeneity	Driver age and sex, time of day, type of crash partner, intersection type, weather conditions and lighting conditions	No	NA
Fatally injured drivers categorized as positive of negative for cannabinoids [5]	Yes	Exposure effect modification	Urban/rural location, region of the US, state marijuana legalization, time of crash, and driver sex, age and race,	Yes	Unlikely
Fatally injured drivers classified as driving for work or non-work reasons [6]	Yes	Exposure effect modification	Vehicle type, vehicle age, driver sex, driver being alcohol or drug positive and holding a commercial driver's license.	Yes	Unlikely
Car crashes where the drivers were classified as being under the influence of alcohol or not [13]	Yes	Exposure effect modification	Density per square mile of rideshare trips that were in progress at the time of the crash	Yes	Unlikely
Classified suicide decedents as being under the influence of alcohol [34]	Yes	Exposure effect modification	Sociodemographic characteristics of decedent, presence of psychiatric disorders, prior suicide attempts and recent adverse life events.	Yes	Unlikely
Intimate partner violence perpetrators classified as having mental illness or not at the time of the offense [35]	Yes	Exposure effect modification	Perpetrator age, sex, and employment status, and history of previous convictions, alcohol abuse and/or self-harm.	Yes	Unlikely
Categorized deaths as intimate partner violence related or not [11]	Yes	Exposure effect modification	Decedent's age, race, year of death, and weapon used.	Yes	Unlikely
Adolescent suicide decedents classified as younger (10–14) and older (15–17) [36]	Yes	Exposure effect modification	Life stressors (e.g. intimate partner issue, physical health problem, mental health problem) and suicide risk factors (e.g. cutting, recent suicide of loved one).	Yes	Possibly
Suicide decedents classified physician versus non-physician [37]	Yes	Exposure effect modification	Decedent's gender, marital status, age, mental health measures, experience of life stressors.	Yes	Possible for some predictors and unlikely for others

hold in case-only studies of effect modification in injury risk. Moreover, establishing conditional independence in case-only data analyses by controlling for a variable(s) that explains the association between the two exposures is likely to be difficult to do in practice [18]. In our example analyses of FARS data, it is difficult to conceptualize all of the variables that might explain associations between age and alcohol consumption in the general population. Furthermore, like many hospital series or registries used in injury epidemiology, the FARS includes very limited data on the personal or behavioral characteristics of cases that can be used as covariates in a case-only analysis.

Thus, it is likely that the OR generated from many injury case-only studies are uninterpretable due to violations of the independence assumption, even after conditioning on available covariates. Among case-only studies we found that compare pedestrians injured or killed while under the influence of alcohol to pedestrians injured or killed while sober, none interpreted their results within an underlying cohort or case-control framework [22–28]. In addition, none of those studies placed their results within the context of estimating an exposure effect modification or noted the independence assumption [22–28]. These critiques apply more broadly to the use of these two case-only designs in injury epidemiology: these studies rarely state the hypothesis being tested by the design (i.e., exposure effect modification or etiological heterogeneity), nor do they relate the estimated case-only OR to the OR that would have been estimated in case-control analyses of the case-series under investigation. For example, Table 2 relates the designs of ten recent, purposively selected, case-only studies from the literature to the principles discussed here. Two can be classified as studies of etiological heterogeneity. The rest are studies of exposure effect modification; among most of these studies, the independence assumption is unlikely to hold.

Conclusion

The case-only designs reviewed here are commonly used in injury epidemiology research, but in practice these analyses and their interpretation have not been rigorously connected to the epidemiologic study design literature. Discussions of the results of these two types of case-only studies in injury epidemiology rarely state the type of hypothesis being tested by the design (whether effect modification or etiologic heterogeneity) and rarely relate the estimated case-only OR to the OR that would have been estimated in case-control analyses if appropriate controls were available for the case-series under investigation. As such, we argue that these case-only studies in injury epidemiology are commonly misinterpreted, and the underlying assumptions are not stated in a way that supports critical assessment.

When the research goal is to understand distinct causal pathways relevant to injury prevention, case-only studies of etiological heterogeneity may have utility. However, results from such studies should be considered hypothesis-generating and further investigated in full cohort, case-crossover or case-control studies that can estimate causal effects. Conversely, because the independence assumption is unlikely to hold for many of the putative causes of injury, case-only studies of exposure effect modification are unlikely to be interpretable. When such studies are conducted, the researchers should (1) directly address the likely validity of the independence assumption; (2) conceptualize and describe causes of non-independence; and (3) establish conditional independence through inclusion of appropriate covariates in case-only regression analyses. Absent these steps, case-only designs that test for exposure effect modification are unlikely to be useful for understanding the etiology of injuries or for designing interventions.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40471-022-00306-8.

Funding Drs. Rundle, Lovasi, Neckerman, Mooney, Branas, and Bader were supported by a grant from the National Institute on Alcohol Abuse and Alcoholism (R01AA028552); Dr. Morrison was supported by a grant from the National Institute on Alcohol Abuse and Alcoholism (K01AA026327); Dr. Mooney was also supported by a grant from the National Library of Medicine (R00LM012868).

Declarations

Conflict of Interest The authors declare no competing interests.

Human and Animal Rights and Informed Consent This article does not contain any studies with humans or animals performed by any of the authors. Analyzed data were from public sources and include only decedents and as such these analyses are "Not Human Research."

References

1. Begg CB, Zhang ZF. Statistical analysis of molecular epidemiology studies employing case-series. Cancer Epidemiol Biomark Prev. 1994;3(2):173–5.
2. Khoury MJ, Flanders WD. Nontraditional epidemiologic approaches in the analysis of gene-environment interaction: case-control studies with no controls! Am J Epidemiol. 1996;144(3):207–13. https://doi.org/10.1093/oxfordjournals.aje.a008915.
3. Wacholder S. Design issues in case-control studies. Stat Methods Med Res. 1995;4(4):293–309. https://doi.org/10.1177/096228029500400403.
4. Lira MC, Sarda V, Heeren TC, Miller M, Naimi TS. Alcohol Policies and motor vehicle crash deaths involving blood alcohol concentrations below 0.08. Am J Prev Med. 2020;58(5):622–9. https://doi.org/10.1016/j.amepre.2019.12.015.
5. Azagba S, Shan L, Latham K. Rural-urban differences in cannabis detected in fatally injured drivers in the United States. Prev Med. 2020;132:105975. https://doi.org/10.1016/j.ypmed.2019.105975.
6. Rudisill TM, Menon S, Hendricks B, Zhu M, Smith GS. Differences between occupational and non-occupational-related motor vehicle collisions in West Virginia: a cross-sectional and spatial analysis. PLoS One. 2019;14(12):e0227388. https://doi.org/10.1371/journal.pone.0227388.
7. Eichelberger AH, McCartt AT, Cicchino JB. Fatally injured pedestrians and bicyclists in the United States with high blood alcohol concentrations. J Saf Res. 2018;65:1–9. https://doi.org/10.1016/j.jsr.2018.02.004.
8. Boone EM, Rossheim ME, Krall JR, Weiler RM. State helmet laws and helmet use among fatally injured moped riders in the United States, 2011-2015. Traffic Inj Prev. 2018;19(3):270–3. https://doi.org/10.1080/15389588.2017.1383604.
9. Caetano R, Kaplan MS, Kerr W, McFarland BH, Giesbrecht N, Kaplan Z. Suicide, Alcohol Intoxication, and Age Among Whites and American Indians/Alaskan Natives. Alcohol Clin Exp Res. 2020;44(2):492–500. https://doi.org/10.1111/acer.14251.
10. Patten M, Carmichael H, Moore A, Velopulos C. Circumstances of Suicide Among Lesbian, Gay, Bisexual and Transgender Individuals. J Surg Res. 2021;270:522–9. https://doi.org/10.1016/j.jss.2021.08.029.
11. Graham LM, Kafka JM, AbiNader MA, Lawler SM, Gover-Chamlou AN, Messing JT, et al. Intimate Partner Violence-Related Fatalities Among U.S. Youth Aged 0-24 Years, 2014-2018. Am J Prev Med. 2021. https://doi.org/10.1016/j.amepre.2021.09.018.
12. Krumkamp R, Reintjes R, Dirksen-Fischer M. Case-case study of a Salmonella outbreak: an epidemiologic method to analyse surveillance data. Int J Hyg Environ Health. 2008;211(1-2):163–7. https://doi.org/10.1016/j.ijheh.2007.02.006.
13. Morrison CN, D'Ambrosi G, Kamb A, MacManus K, Rundle AG, Humphreys DK. Rideshare Trips and Alcohol-Involved Motor Vehicle Crashes in Chicago. J Stud Alcohol Drugs. 2021;82(6):720–9.
14. Branas CC, Richmond TS, Ten Have TR, Wiebe DJ. Acute alcohol consumption, alcohol outlets, and gun suicide. Subst Use Misuse. 2011;46(13):1592–603. https://doi.org/10.3109/10826084.2011.604371.
15. Li G, Chihuri S. Prescription opioids, alcohol and fatal motor vehicle crashes: a population-based case-control study. Inj Epidemiol. 2019;6:11. https://doi.org/10.1186/s40621-019-0187-x.
16. Cummings P, Weiss NS. Case series and exposure series: the role of studies without controls in providing information about the etiology of injury or disease. Inj Prev. 1998;4(1):54–7. https://doi.org/10.1136/ip.4.1.54.
17. Marcus PM, Hayes RB, Vineis P, Garcia-Closas M, Caporaso NE, Autrup H, et al. Cigarette smoking, N-acetyltransferase 2 acetylation status, and bladder cancer risk: a case-series meta-analysis of a gene-environment interaction. Cancer Epidemiol Biomark Prev. 2000;9(5):461–7.
18. Gatto NM, Campbell UB, Rundle AG, Ahsan H. Further development of the case-only design for assessing gene-environment interaction: evaluation of and adjustment for bias. Int J Epidemiol. 2004;33(5):1014–24. https://doi.org/10.1093/ije/dyh306.
19. Breslow NE. Statistics in epidemiology: the case-control study. J Am Stat Assoc. 1996;91(433):14–28. https://doi.org/10.1080/01621459.1996.10476660.
20. Wacholder S, McLaughlin JK, Silverman DT, Mandel JS. Selection of controls in case-control studies. I Principles Am J Epidemiol. 1992;135(9):1019–28. https://doi.org/10.1093/oxfordjournals.aje.a116396.
21. Zhang ZF, Karpeh MS, Lauwers GY, Marrero AM, Pollack D, Cordon-Cardo C, et al. A case-series study of p53 nuclear overexpression in early-stage stomach cancer. Ann N Y Acad Sci. 1995;768:269–71. https://doi.org/10.1111/j.1749-6632.1995.tb12138.x.
22. Lasota D, Al-Wathinani A, Krajewski P, Goniewicz K, Pawlowski W. Alcohol and Road Accidents Involving Pedestrians as Unprotected Road Users. Int J Environ Res Public Health. 2020;17(23). https://doi.org/10.3390/ijerph17238995.
23. Ortiz N, Ramnarayan M. 1949 - Analysis of Alcohol-Involved Pedestrian Fatalities in the United States, 2003–2015. J Transp Health. 2017;5:S93. https://doi.org/10.1016/j.jth.2017.05.250.
24. Pawlowski W, Lasota D, Goniewicz M, Rzonca P, Goniewicz K, Krajewski P. The Effect of Ethyl Alcohol upon Pedestrian Trauma Sustained in Traffic Crashes. Int J Environ Res Public Health. 2019;16(8). https://doi.org/10.3390/ijerph16081471.
25. Dultz LA, Frangos S, Foltin G, Marr M, Simon R, Bholat O, et al. Alcohol use by pedestrians who are struck by motor vehicles: how drinking influences behaviors, medical management, and outcomes. J Trauma. 2011;71(5):1252–7. https://doi.org/10.1097/TA.0b013e3182327c94.
26. Knight MT, Li T, Dhillon NK, Srour M, Huang R, Margulies DR, et al. Walking Under the influence : association of time of the day with the incidence and outcomes of intoxicated pedestrians struck by vehicles. Am Surg. 2021;87(3):354–63. https://doi.org/10.1177/0003134820947365.
27. Plurad D, Demetriades D, Gruzinski G, Preston C, Chan L, Gaspard D, et al. Pedestrian injuries: the association of alcohol consumption with the type and severity of injuries and outcomes. J Am Coll Surg. 2006;202(6):919–27. https://doi.org/10.1016/j.jamcollsurg.2006.02.024.
28. Ostrom M, Eriksson A. Pedestrian fatalities and alcohol. Accid Anal Prev. 2001;33(2):173–80. https://doi.org/10.1016/s0001-4575(00)00028-2.
29. National Center for Statistics and Analysis. Fatality Analysis Reporting System (FARS) Analytical User's Manual, 1975-2018. Washington, DC: National Highway Traffic Safety Administration2019. Report No.: Report No. DOT HS 812 827.
30. Keyes KM. Age, Period, and Cohort Effects in Alcohol Use in the United States in the 20th and 21st Centuries: Implications for the Coming Decades. Alcohol Res. 2022;42(1):02. https://doi.org/10.35946/arcr.v42.1.02.

31. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol. 2003;32(1):1–22. https://doi.org/10.1093/ije/dyg070.
32. Batouli G, Guo M, Janson B, Marshall W. Analysis of pedestrian-vehicle crash injury severity factors in Colorado 2006-2016. Accid Anal Prev. 2020;148:105782. https://doi.org/10.1016/j.aap.2020.105782.
33. Chen H, Chen Q, Chen L, Zhang G. Analysis of risk factors affecting driver injury and crash injury with drivers under the influence of alcohol (DUI) and non-DUI. Traffic Inj Prev. 2016;17(8):796–802. https://doi.org/10.1080/15389588.2016.1168924.
34. Kolves K, Koo YW, de Leo D. A drink before suicide: analysis of the Queensland Suicide Register in Australia. Epidemiol Psychiatr Sci. 2020;29:e94. https://doi.org/10.1017/S2045796020000062.
35. Oram S, Flynn SM, Shaw J, Appleby L, Howard LM. Mental illness and domestic homicide: a population-based descriptive study. Psychiatr Serv. 2013;64(10):1006–11. https://doi.org/10.1176/appi.ps.201200484.
36. De Silva DA, Diduk-Smith RM. Comparison of Suicides Among Younger and Older Adolescents in Virginia, 2008-2017. Arch Suicide Res. 2021:1–8. https://doi.org/10.1080/13811118.2021.1965929.
37. Gold KJ, Sen A, Schwenk TL. Details on suicide among US physicians: data from the National Violent Death Reporting System. Gen Hosp Psychiatry. 2013;35(1):45–9. https://doi.org/10.1016/j.genhosppsych.2012.08.005.