Establishing causality in epigenetic studies of pre-natal alcohol exposure, alcohol use and alcohol use disorders

N Harlaar*

Abstract

Introduction

The nascent field of epigenetics has attracted growing attention in research on pre-natal alcohol exposure, alcohol use and alcohol use disorders. Animal models and studies in community and clinical human samples have provided tantalising clues that epigenetic patterns may be associated with alcohol use and alcohol-related outcomes. However, the reversible and plastic nature of epigenetic patterns means that questions about causality and direction of effect remain unresolved. This review highlights two particularly refractory challenges to establishing causality in epigenetic studies that use living human subjects: confounding variables and reverse causation. Experimental studies (e.g., using alcohol self-administration tasks) and quasi-experimental designs (e.g., longitudinal studies, sibling comparison designs and instrumental variable approaches) may be used to control for potential confounds and explore the causal impact of epigenetic processes (e.g., DNA methylation) on alcohol-related outcomes.

Conclusion

No single design is a ‘magic bullet’, but the careful use of a combination of designs will help to strengthen causal inferences in epigenetic research in clinical and community samples.

Introduction

One incipient theme in recent research on pre-natal alcohol exposure, alcohol use and alcohol use disorders (AUDs) is epigenetics, the study of mitotically stable but reversible modifications to DNA and nucleosomal histone proteins that have the potential to regulate the transcription of information encoded in the DNA sequence into RNA without altering the DNA sequence itself. There are three main categories of epigenetic mechanisms: histone modifications, DNA methylation, and non-coding RNAs (ncRNAs). These mechanisms contribute to structural changes to chromatin (the complex of DNA and proteins that organise DNA) and affect the ability of transcription factors to access DNA.

Interest in epigenetics stems, in part, from the possibility that epigenetic regulation may be responsive to risk factors (e.g., environmental, genetic, lifestyle, socioeconomic), thus providing researchers with candidate mechanisms for experience-dependent changes in gene transcription. Animal models and emerging research in human samples indicate that repeated exposure to alcohol pre-natally and later in life is associated with potentially long-lasting epigenetic changes that may influence vulnerability to addiction and other health-related consequences of alcohol use. Nonetheless, there are important methodological considerations for this nascent field, especially for research in living human subjects. This brief review considers one issue, that of determining causality. Other issues (e.g., tissue specificity of epigenetic patterns) are addressed elsewhere.

Discussion

The author has referenced some of her own studies in this review. These referenced studies have been conducted in accordance with the Declaration of Helsinki (1964) and the protocols of these studies have been approved by the relevant ethics committees related to the institution in which they were performed. All human subjects, in these referenced studies, gave informed consent to participate in these studies.

Effects of epigenetic factors: conceptual models

The relationships among epigenetic factors, alcohol exposure and alcohol-related outcomes are likely to be complex. In the animal literature, the scenario that has received greatest attention is depicted in Figure 1: alcohol exposure causes epigenetically mediated changes in gene expression, altering brain development or neural plasticity in ways that may ultimately affect behaviour and disease outcomes. For example, studies in rodents have linked pre-natal alcohol exposure to disturbed DNA methylation and ncRNA patterns in offspring. These epigenetic changes lead to long-term alterations in the expression of developmentally relevant genes, thereby disrupting normal neurodevelopmental processes and contributing to symptoms of foetal alcohol syndrome disorders (FASDs). Acute or intensive alcohol exposure in adolescence or adulthood may also induce epigenetic effects. For example, intermittent ethanol administration during the adolescent period leads to histone modifications in reward-related brain regions that impact subsequent ethanol preference and intake. These
findings suggest that epigenetic processes could contribute to the sensitivity of the adolescent brain to the reward-related properties of alcohol. Other work has shown that the anxiolytic effects of alcohol exposure may be due to alcohol-induced histone modifications that lead to chromatin remodelling in the amygdala. Thus, at different life stages, epigenetic processes may act as an intermediate step on the causal pathway to alcohol-related disease outcomes. It is important to note that the model in Figure 1 is likely to be an oversimplification; in addition to direct epigenetic effects linking alcohol exposure to gene expression genes, there may also be indirect pathways. For example, early drug experience may epigenetically prime certain genes by altering their ‘inducibility’, or ability to be expressed in the presence of a substance (an inducer), such that they are more likely to be expressed following subsequent drug use even though they may not initially show any changes in expression.

Causal models of the relationships among epigenetic factors, alcohol exposure and alcohol-related outcomes are far more difficult to address in human samples because ethical and pragmatic considerations mean that experimental designs are often pre-mature, impractical, or impossible. A typical starting point, therefore, is to investigate whether alcohol use or alcohol-related outcomes are simply correlated with variation in epigenetic processes. DNA methylation is the best-characterised and most stable epigenetic modification modulating gene transcription, and because it can be robustly assessed using standard genomic DNA resources, it has been the focus of virtually all human epigenetic research to date. Briefly, DNA methylation entails the transfer of a quartet of atoms, called a methyl group, to the fifth carbon of a cytosine residue, thereby forming 5-methylcytosine. This process typically occurs within the context of cytosine-guanine dinucleotides (CpG sites). Aberrant methylation patterns disrupt transcription factor binding, gene expression, and genomic stability. The precise effects vary depending on genomic location; notably, DNA methylation at promoter CpG sites, which are typically unmethylated, are associated with gene silencing. A growing number of studies in human clinical or community samples have examined associations between alcohol use or AUDs and DNA methylation levels at specific CpG sites or regions (reviewed in ref. 3). Results have varied considerably, reflecting the diversity of methods used and genes targeted. However, a number of positive results have been obtained, including associations at biologically relevant genes or gene pathways.

Although any epigenetic association with alcohol-related phenotypes in a well-designed study is potentially an advance, the studies reported to date are cross-sectional, correlational, and therefore fundamentally descriptive in nature. They offer no way to distinguish causal associations from non-causal associations. Being able to identify causal associations would have both theoretical and applied significance: theoretical significance for the light these findings may shed on the role of epigenetic processes in the aetiology and pathogenesis of alcohol-related diseases such as FASDs and AUDs, and applied significance arising from opportunities to identify and evaluate new, clinically-relevant epigenetic biomarkers.

Non-causal associations in epigenetic research in human samples can occur for several reasons, in addition to chance, including confounding and reverse causation. Confounding factors are important because they may obscure observed associations among alcohol use, alcohol-related outcomes and DNA methylation. This type of model has been widely used or implied in published empirical studies of alcohol-related DNA methylation patterns in human samples, which are generally observational and cross-sectional. Specifically, alcohol exposure induces stable epigenetic changes that alter the expression of genes that influence brain development or neural plasticity in ways that may ultimately affect behaviour and disease outcomes.

Figure 1: The role of epigenetic processes in the causal path from environmental risk factors (e.g., chronic or acute alcohol use; prenatal alcohol exposure) to disease or behavioural outcomes (e.g., susceptibility to addiction). This type of model, which is a simplified abstraction, is often used or implied in animal models of epigenetic responses to prenatal alcohol exposure or acute or chronic alcohol treatment. Specifically, alcohol exposure induces stable epigenetic changes that alter the expression of genes that influence brain development or neural plasticity in ways that may ultimately affect behaviour and disease outcomes.

Figure 2: Correlation model that asks whether epigenetic variation (e.g., DNA methylation) is associated with individual differences in alcohol exposure or alcohol-related disease outcomes. This type of model has been widely used or implied in published empirical studies of alcohol-related DNA methylation patterns in human samples, which are generally observational and cross-sectional.
pertinent are age, gender, ethnicity, smoking and genetic variation. Even though it may be possible to statistically adjust for known (measured) confounds as well as some unknown confounds (with newer modelling techniques\textsuperscript{10}), adjustment of confounding factors in observational studies is not a panacea\textsuperscript{11}. Reverse causation is also problematic. Consider the hypothesis that chronic heavy alcohol use leads to aberrant methylation at certain genes. An experiment is devised in which methylation in DNA from individuals who meet a pre-established criterion for heavy drinking is compared with methylation from controls, matched on potential confounding factors. Even if findings consistent with the hypothesis were obtained, such that cases show significantly higher methylation levels at certain CpG sites (above chance levels) compared with controls, it is not possible to say whether chronic heavy alcohol use alters DNA methylation patterns (i.e., a causal effect) or whether altered DNA methylation patterns are simply epiphenomena that are downstream from other changes that really drive chronic heavy drinking (i.e., reverse causation). These feedback loops are depicted in Figure 4.

**Beyond observation studies: alternative research designs**

The prospects for epigenetic research on alcohol use and alcohol-related outcomes in human samples are not entirely gloomy. For some research questions, it may be possible to implement an experimental design. For example, a laboratory alcohol self-administration task in which alcohol is administrated intravenously might be employed to examine DNA methylation profiles in individuals with a positive history of pre-natal alcohol exposure or a family history of AUDs, compared with controls\textsuperscript{12}. By tightly controlling alcohol consumption, this method may shed light on possible epigenetic priming effects or alcohol-induced changes in methylation in groups at high risk for AUDs. Self-administration tasks may also be combined with pharmacotherapy. Medications that effectively reduce alcohol consumption in clinical settings (e.g., the non-selective opioid antagonist naltrexone) can attenuate alcohol self-administration for some individuals in laboratory experiments\textsuperscript{13} accordingly, it might be interesting to examine drug effects on alcohol consumption and DNA methylation levels at relevant genes, or to study epigenetic predictors of drug response. Importantly, control over extraneous variables is usually greater in experimental designs than in other methods. For example, alcohol administration tasks can be designed to ensure precise control.

**Figure 3:** Confounding factors suggest an alternative non-epigenetic pathway in models linking environmental risk factors to disease or behavioural outcomes. Known or potential confounding factors in epigenetic studies of alcohol use and alcohol use disorders include age, gender, genetic variation and smoking.

**Figure 4:** In addition to confounding, various feedback loops may lead to reverse causation in models linking environmental risk factors to disease or behavioural outcomes. The red arrows highlight that disease states and phenotypic outcomes may impact directly upon epigenetic processes as well as risk factors themselves (e.g., as a result of negative reinforcement during the transition from alcohol use to dependence).
over an individual’s alcohol exposure. Nonetheless, challenges remain. For example, it is unclear whether a single session of acute alcohol exposure is sufficient to induce methylation changes; it may be the case that reliable, persistent epigenetic modifications occur only after prolonged alcohol exposure. Additionally, adjusting for all potential confounding factors is difficult, especially in experiments that include group comparisons (e.g., AUD cases vs. controls).

When it is not possible to conduct experimental studies, we may still be able to shed light on causality and direction of effect by using quasi-experimental designs; that is, methods with design elements that rule out plausible explanations for an association, thus providing some kind of approximation to random assignment to experimental conditions. Three types of quasi-experiments are of note. First, longitudinal designs are important for increasing our understanding of epigenetic changes over time. Specifically, with well-designed longitudinal studies, we may be able to gain purchase on the temporal direction of observed associations between methylation changes and AUDs while taking into account possible confounding factors, including age, which is of special relevance to AUDs given that the severity of AUD symptoms tend to increase with age. Longitudinal studies also permit the analysis of potential mediating or interactive effects of lifetime biological, psychological and social risk. The number of epigenetic studies in longitudinal cohorts is increasing, but remains relatively limited, in part because few studies have routinely collected DNA samples from the same individuals at multiple points in the lifespan (reviewed in ref. 15). There have been no published longitudinal studies examining the effects of pre-natal alcohol exposure, alcohol use or alcohol-related outcomes across time.

A second natural experiment is the sibling comparison design, where cases are compared with unaffected sibling controls. This design may be considered a simple extension of the matched case–control design. However, instead of explicitly matching on a set of measured variables, the use of siblings as controls will automatically match on many unmeasured, even unknown, variables that may otherwise confound observed associations among methylation and alcohol use or alcohol-related outcomes, such as social background and other factors that affect siblings similarly.

An especially popular version of the sibling comparison design is the use of discordant monozygotic (MZ) twins. MZ twins share their genetic sequence, birth date and sex, along with other factors shared by regular siblings, which affords a substantial degree of control for both genetic and environmental liabilities that are shared by twins. Disease-associated epigenetic variation has been identified in MZ twins discordant for a range of behavioural and medical disorders. Typically, each study has reported modest, but consistent, differential methylation in moderate to large numbers of genes relevant to the phenotype. To date, no discordant MZ twin study of epigenetic effects on alcohol use or alcohol-related outcomes has been reported.

A weaker but more practical alternative to the discordant MZ twin design is to examine siblings discordant for AUDs, or, if the effects of pre-natal alcohol exposure are of interest, sibling pairs in which exposure to alcohol use in utero differed (i.e., because the mother drank heavily in one pregnancy but less heavily in another). This study design controls for differences in familial influences that may contribute to methylation differences, although it is not possible to determine whether these familial influences are genetic, environmental or age related. To date, two published epigenetic studies on AUDs have used sibling designs. The first study compared DNA methylation patterns at cancer-related genes in individuals with chronic alcoholism and matched sibling controls. There were no significant differences in the average methylation score (average methylation level over 1287 Cpg sites) between cases and siblings, although there was a small decrease in overall methylation scores in cases who smoked compared with their non-alcohol dependent siblings who did not. In contrast, a more recent study identified 1381 differentially methylated Cpg sites in AUD cases and sibling controls.

A third natural experiment that is germane to epigenetic studies in humans is instrumental variable (IV) analysis, which may be used to estimate the average causal effect of an exposure on an outcome in the presence of unmeasured confounding. The best-characterised IV approach for epigenetic studies is two-step Mendelian randomisation, which first seeks to examine the causal impact of a risk factor (e.g., alcohol use) on DNA methylation patterns, using a specific genetic variant as a proxy for the risk factor, and then tests the causal impact of DNA methylation on the outcome of interest (e.g., AUD symptoms). This approach mitigates the problem of reverse causation because genetic variants are randomly assigned within a family at conception and are not reversely modified by one’s behaviours or environment. Mendelian randomisation has been used quite widely in research on alcohol consumption and pre-natal alcohol exposure, but has yet to be employed in epigenetic studies. A related IV design is the causal inference test (CIT), which infers causal status for a potential mediator (e.g., DNA methylation) between a genetic locus and a phenotype based on a set of mathematical conditions. Proof-of-principle for the value of CIT in epigenetic studies has been demonstrated for rheumatoid arthritis, but like Mendelian randomisation, CIT has not yet been used in epigenetic studies of alcohol use or alcohol-related outcomes.

Like experimental designs, natural experiments are not without challenges. For example, it is unclear whether a simple extension of the matched case–control design is sufficient to induce methylation changes; it may be the case that reliable, persistent epigenetic modifications occur only after prolonged alcohol exposure. Additionally, adjusting for all potential confounding factors is difficult, especially in experiments that include group comparisons (e.g., AUD cases vs. controls).
limitations. Longitudinal studies are vulnerable to non-random attrition (e.g., individuals with the greatest alcohol use problems may be the most likely to drop out), and are slow and difficult to establish. It may be possible to harness existing longitudinal cohorts, but cohorts that are optimal for epigenetic studies (e.g., that follow initially disease-free individuals over the course of many years and obtain extensive phenotypic data and biological samples at multiple time points) are exceptional and relatively scarce. Sibling comparison studies cannot rule out the effects of environmental factors that vary among siblings (e.g., perceived psychosocial stress), or are correlated with an environmental risk factor (e.g., drinking) or with the outcome (e.g., AUD symptoms). Additionally, biases arise if siblings are less similar with regard to confounders than to the exposure under study, which will mean that they are likely to differ more from each other than two randomly selected persons from the same population having the same exposure levels. Discordant MZ twin studies are challenging due to the relatively high heritability of AUDs (~50% [ref. 29]), which may make it difficult to identify and recruit enough sufficiently discordant twins. Additionally, they cannot be used to study pre-natal factors (e.g., pre-natal alcohol exposure) and other factors shared by co-twins. Finally, all studies require large samples for adequate statistical power. Power may be especially important for Mendelian randomisation and other IV studies that use genetic variants, because genetic variants (especially when considered individually) typically account for only a tiny proportion of phenotypic variance.

Conclusion

It is difficult to over-state the challenges of establishing causal effects in epigenetic studies of pre-natal alcohol exposure, alcohol use and AUDs in human samples. Addressing these challenges is crucial because it will help to elucidate the functional role of the AUD-associated epigenetic variation and its potential utility in terms of diagnostics or therapeutics. This review has described experimental and quasi-experimental designs that may be employed to explore the causal impact of DNA methylation. Each design has its own particular strengths and limitations, and none is free of the latter. As such, they may be more effective for identifying probable causal effects than for determining precisely what form and how many causal pathways are involved in the relationships among alcohol use, alcohol-related outcomes and epigenetic processes. Nevertheless, especially when used in combination (e.g., embedding a sibling design with a longitudinal study), they can do much to strengthen causal inferences, and thus they play a unique position in bridging gaps between basic and pre-clinical research in animal models and exploratory, observational studies in humans.

Acknowledgements

N.H. is supported by K99AA020536-0141. The funders had no role in the decision to publish or in the conception and preparation of the manuscript.

References