



Sharp, G. C., & Relton, C. L. (2017). Epigenetics and noncommunicable diseases. *Epigenomics*, 9(6), 798-791. <https://doi.org/10.2217/epi-2017-0045>

Peer reviewed version

Link to published version (if available):  
[10.2217/epi-2017-0045](https://doi.org/10.2217/epi-2017-0045)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Future Medicine at <http://www.futuremedicine.com/doi/10.2217/epi-2017-0045>. Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/pure/about/ebr-terms>

# Epigenetics and Non-Communicable Diseases

Gemma C Sharp, Caroline L Relton

Noncommunicable diseases (NCDs), primarily cardiovascular disease, cancer, chronic respiratory disease and diabetes, kill more than 38 million people each year, making them by far the leading cause of death in the world<sup>1</sup>. They develop due to an interplay of genetic, lifestyle and environmental factors, but many NCDs are considered preventable because the most important risk factors, such as tobacco use, alcohol use and obesity, are modifiable<sup>1</sup>. Nevertheless, the prevalence of NCDs is increasing, and ongoing research hopes to inform effective strategies to predict, prevent and treat. This includes research into the role and utility of epigenetic mechanisms, either as an important determinant of NCDs (where causality is important) or as useful markers to help predict their occurrence and/or their consequences (where causality is not important).

## Epigenetics as a determinant of NCDs: where causality is important

Since epigenetic changes can occur in response to external stimuli, can remain stable over time and can lead to long-term changes in gene expression and pathological dysfunctions, it is plausible that epigenetics might contribute to the onset of NCDs. Much of the support for this hypothesis comes from animal studies, which have shown environmentally-induced epigenetic changes within key tissues that are also associated with later-life health consequences. For example, a mouse study published 10 years ago showed that offspring of dams fed a low-protein diet during pregnancy had lower DNA methylation of the angiotensin receptor gene and this was associated with higher blood pressure<sup>2</sup>. More recently, human studies have found epigenome-wide associations between DNA methylation and several NCDs, including a wide range of cancers<sup>3</sup>, asthma<sup>4</sup>, obesity<sup>5</sup> and type 2 diabetes<sup>6</sup>.

While these studies might suggest that epigenetic mechanisms contribute to the onset of NCDs, there are two main potential limitations that should be considered before results from human observational studies are interpreted in this way. Firstly, the association between the NCD and epigenetic mechanism could be confounded by environmental or genetic factors or by cellular composition of the biological sample. Secondly, the association could be explained via a reverse causal pathway, that is, the NCD caused the epigenetic changes rather than the other way around.

The field in general needs better causal evidence to underpin the mechanistic link between epigenetic variation and NCDs<sup>7</sup>. Fortunately, there are a number of strategies that can be applied to help infer causality in studies of epigenetics and NCDs. Some of these are unique to epigenetic epidemiological studies. For example, integrating epigenomic data with genomic and transcriptomic data from the same biological samples to allow a more comprehensive investigation of potential causal pathways to NCDs. Additionally, measuring epigenetics in purified cells, or adjusting for estimated or measured cellular heterogeneity in mixed cell population samples can help to lessen the influence of cell types confounding associations<sup>7</sup>.

Other strategies to improve causal inference are already widely used in non-epigenetic observational studies<sup>8</sup>. These include using a longitudinal study design (rather than a cross-sectional case-control design) where epigenetic data are collected prior to NCD onset, thus

circumventing issues with reverse causation. There are many examples of epigenome-wide studies assessing associations between DNA methylation in neonatal cord blood and emergence of NCDs in childhood, for example, one study found associations between cord blood DNA methylation and cardiovascular risk (increased heart rate) at age 9<sup>9</sup>.

Another strategy is to seek replication of results in an independent sample to provide confidence that the identified associations are reproducible and not spurious. A growing number of consortia are emerging to help facilitate replication and meta-analysis between cohorts with epigenetic data<sup>10</sup>. Stronger evidence for a causal association can be provided by comparing associations across cohorts with different confounding structures. For example, an association between breastfeeding and childhood obesity and high blood pressure that was present in a UK cohort (high-income) was not present in a Brazilian cohort (low- or middle-income), providing evidence for confounding by socioeconomic status<sup>11</sup>. Although we are not aware of this approach having been used in studies of the epigenetics of NCDs, there is certainly scope to do so. In order to control for the large effect of genetics on epigenetics, co-twin designs can be useful. These compare epigenetic modifications in monozygotic twins discordant for an NCD to find associations between epigenetic and disease state while controlling for genetics. Several examples of this approach being applied to study epigenetic causes of NCDs are described elsewhere<sup>12</sup>. One example is a study of 50 twin pairs discordant for major depressive disorder, which found differences in whole blood DNA methylation<sup>13</sup>. Another causal inference technique that can be used in epigenetic epidemiological studies is the negative control design. For example, we compared the association between cord blood DNA methylation and maternal BMI to that of paternal BMI to help infer whether the apparent effect of maternal BMI on offspring DNA methylation is causal or better explained by shared genetic or environmental factors<sup>14,15</sup>. This approach is useful for drawing inferences about factors causing epigenetic changes, which might mediate associations with NCDs. Finally, Mendelian randomization (MR)<sup>16</sup> is a useful causal inference technique that is beginning to be applied to study the role of epigenetics in the development of NCDs. This approach involves using genetic variants robustly associated with exposures to infer causality. The genetic variant serves as a proxy for the exposure, but crucially, any association between the genetic variant and the outcome is not subject to various confounders of the exposure-outcome association. Since the genetic variant cannot be directly influenced by the outcome, MR also helps combat issues with reverse causation<sup>17,18</sup>. The approach can be used to study the epigenetics of NCDs by finding genetic variants near and robustly associated with epigenetic marks and assessing the association between these variants and the NCD. For example, recent studies have used MR to show that DNA methylation at specific loci may be causally associated with obesity, with more evidence of an effect of obesity on DNA methylation than the other way around<sup>5,14,19</sup>. To provide the best evidence for a mechanistic role of epigenetics in causing NCDs, several of the previously discussed strategies should be combined.

### Epigenetics as a predictor of NCDs and their consequences: where causality is not important

Even if epigenetics does not play a causal role in NCD development, there is still great value in studying associations between epigenetics, NCDs, their risk factors and their consequences. Epigenetic information 'captures' information on historical exposure and underlying genetics, and to this end it can be used to improve understanding of the (non-epigenetic) aetiologies of NCDs<sup>20</sup>. For example, recent multi-cohort studies have shown that DNA methylation data

can be a powerful predictor of alcohol consumption<sup>21</sup> and prenatal tobacco exposure<sup>22</sup>. These epigenetic predictors can be used as an alternative to self-report questionnaire data, or where such data are lacking, to assess associations between these risk factors and NCDs. One of the advantages of epigenetic predictors over self-report data is that they are less prone to measurement error introduced by reporting bias. Clinically, such predictors could be useful biomarkers to identify individuals at high risk of developing certain NCDs and targeting preventative strategies.

Additionally, studies such as the recent Mendelian randomization studies of adiposity<sup>5,14,19</sup> demonstrate that NCDs themselves can alter epigenetics and may provide insight into downstream consequences of these diseases. This raises the possibility that epigenetics could be useful in diagnosing NCDs before development of any clinical symptoms, and/or aiding prognosis or stratification of individuals with NCDs for certain treatment options.

## Conclusion

In conclusion, appropriate interpretations of well-designed epigenetic studies offer the opportunity to improve our understanding of the aetiology of NCDs and their outcomes, which will help inform development of better strategies to predict, prevent and treat them.

## References

1. Who, World Health Organization & Who. Global status report on noncommunicable diseases 2014. *World Health* 1–51 (2014). doi:ISBN 9789241564854
2. Bogdarina, I., Welham, S., King, P. J., Burns, S. P. & Clark, A. J. L. Epigenetic modification of the renin-angiotensin system in the fetal programming of hypertension. *Circ. Res.* **100**, 520–526 (2007).
3. Baylin, S. B. & Jones, P. A. A decade of exploring the cancer epigenome — biological and translational implications. *Nat. Rev. Cancer* **11**, 726–734 (2011).
4. DeVries, A. *et al.* Epigenome-wide analysis links SMAD3 methylation at birth to asthma in children of asthmatic mothers. *J. Allergy Clin. Immunol.* (2016). doi:10.1016/j.jaci.2016.10.041
5. Wahl, S. *et al.* Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* **541**, 81–86 (2016).
6. Elliott, H. R. *et al.* The Role of DNA Methylation in Type 2 Diabetes Aetiology – Using Genotype as a Causal Anchor. *Diabetes* db160874 (2017). doi:10.2337/db16-0874
7. Birney, E., Smith, G. D. & Gready, J. M. Epigenome-wide Association Studies and the Interpretation of Disease -Omics. *PLoS Genet.* **12**, e1006105 (2016).
8. Richmond, R. C., Al-Amin, A., Smith, G. D. & Relton, C. L. Approaches for drawing causal inferences from epidemiological birth cohorts: a review. *Early Hum. Dev.* **90**, 769–80 (2014).
9. Murray, R. *et al.* DNA methylation at birth within the promoter of ANRIL predicts markers of cardiovascular risk at 9 years. *Clin. Epigenetics* **8**, 90 (2016).
10. Joubert, B. R. *et al.* DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *Am. J. Hum. Genet.* (2016). doi:10.1016/j.ajhg.2016.02.019
11. Brion, M.-J. A. *et al.* What are the causal effects of breastfeeding on IQ, obesity and blood pressure? Evidence from comparing high-income with middle-income cohorts.

- Int. J. Epidemiol.* **40**, 670–680 (2011).
12. Castillo-Fernandez, J. E., Spector, T. D. & Bell, J. T. Epigenetics of discordant monozygotic twins: implications for disease. *Genome Med.* **6**, 60 (2014).
  13. Davies, M. N. *et al.* Hypermethylation in the ZBTB20 gene is associated with major depressive disorder. *Genome Biol.* **15**, R56 (2014).
  14. Richmond, R. C. *et al.* DNA Methylation and BMI: Investigating Identified Methylation Sites at HIF3A in a Causal Framework. *Diabetes* **65**, 1231–44 (2016).
  15. Sharp, G. C. *et al.* Maternal pre-pregnancy BMI and gestational weight gain, offspring DNA methylation and later offspring adiposity: findings from the Avon Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* (2015). doi:10.1093/ije/dyv042
  16. Davey Smith, G. & Hemani, G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum. Mol. Genet.* **23**, R89–R98 (2014).
  17. Relton, C. L. & Davey Smith, G. Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. *Int. J. Epidemiol.* **41**, 161–76 (2012).
  18. Relton, C. L. & Davey Smith, G. Mendelian randomization: applications and limitations in epigenetic studies. *Epigenomics* **7**, 1239–1243 (2015).
  19. Mendelson, M. M. *et al.* Association of Body Mass Index with DNA Methylation and Gene Expression in Blood Cells and Relations to Cardiometabolic Disease: A Mendelian Randomization Approach. *PLOS Med.* **14**, e1002215 (2017).
  20. Relton, C. L., Hartwig, F. P. & Davey Smith, G. From stem cells to the law courts: DNA methylation, the forensic epigenome and the possibility of a biosocial archive. *Int. J. Epidemiol.* **44**, 1083–1093 (2015).
  21. Liu, C. *et al.* A DNA methylation biomarker of alcohol consumption. *Mol. Psychiatry* **1–12** (2016). doi:10.1038/mp.2016.192
  22. Reese, S. E. *et al.* DNA Methylation Score as a Biomarker in Newborns for Sustained Maternal Smoking during Pregnancy. *Environ. Health Perspect.* (2016). doi:10.1289/EHP333