Editorial

From stem cells to the law courts: DNA methylation, the forensic epigenome and the possibility of a biosocial archive

The growth in epigenetics continues to attract considerable cross-disciplinary interest, apparently representing an opportunity to move beyond genomics towards the goal of understanding phenotypic variability from molecular through organismal to the societal level. The epigenome may also harbour useful information about life-time exposures (measured or unmeasured) irrespective of their influence on health or disease, creating the potential for a person-specific biosocial archive. Furthermore, such data may prove of use in providing identifying information, providing the possibility of a future forensic epigenome. The mechanisms involved in ensuring that environmentally induced epigenetic changes perpetuate across the life course remain unclear. Here we propose that a potential role of adult stem cells in maintaining epigenetic states provides a useful basis for formulating such epidemiologically-relevant concepts.

Epigenetics encompasses different mechanisms of gene expression regulation, the most commonly discussed ones being DNA methylation, histone modifications and non-coding RNAs. The epigenetic mechanism most studied to date is DNA methylation—addition of a methyl group (CH₃) at the 5’ position of a cytosine base, typically at CpG dinucleotides which are often clustered in CpG-rich DNA segments called CpG islands. This chemical addition is made by a covalent bond and is stable over time.

Since epigenetic processes are believed to be modifiable by environmental (i.e. non-genetic) factors, much of the current interest in epigenetics lies in understanding how the environment influences gene expression—even though, as currently analysable, epigenetics is unlikely to hold all of the answers for this broad and rather complex question. Technological advances now allow epigenome-wide investigations to be performed in large human populations at affordable costs—especially with respect to DNA methylation. Although other measures are technically possible, they remain prohibitively expensive on a population scale. Epidemiological studies involving DNA methylation pertain to one of the most recent branches of epidemiology—epigenetic epidemiology. Large-scale profiling of DNA methylation levels has been applied to case-control studies, densely-phenotyped cohorts and in some instances serial samples from the same longitudinal cohort. This allows investigation of the determinants of variation in DNA methylation and their importance in the context of different health outcomes and traits, as evidenced by the steady rise in epigenetic epidemiology publications (Figure 1) and the numerous studies reported in this issue of the IJE.

Long-term effects of environmental exposures on methylation

Methylation modifications are relatively stable. Indeed, such stability is important in the maintenance of methylation modifications that regulate developmental processes, such as cell differentiation. On the other hand, methylation modifications are sometimes reversible, and there is evidence that demethylation plays an important role in biological processes. Stochastic processes also contribute to some of the plasticity observed in the methylome.

The short-term responsiveness of the epigenome to environmental challenges is an area that has not been widely studied. Experimental studies of short-term exposure to particulate air matter have demonstrated that DNA methylation changes occur within hours of exposure. Inference can also be made with respect to responsiveness to exposures in utero detected at birth, given the specific time window of exposure. Critical windows of exposure can however be difficult to determine in many instances, as
the epigenetic measure inevitably captures a cross-sectional measure of epigenetic differences as opposed to being contemporaneous with the timing of the exposure itself.

Long-term effects of an early-life exposure on methylation levels might mediate at least some of the associations between exposures and phenotypes at later life stages, although scepticism needs to be maintained in the breadth of claims regarding DNA cytosine methylation, given it plays no important role in the development of such important model organisms as Caenorhabditis elegans,\textsuperscript{33} Drosophila melanogaster and Saccharomyces cerevisiae\textsuperscript{34} (a conclusion not altered by detection of very low levels of cytosine methylation in Drosophila\textsuperscript{35} and adenine methylation in both Drosophila and C. elegans\textsuperscript{36}). It is commonplace in epidemiological studies to see claims that epigenetic processes may provide biological plausibility to potentially causal associations, which then represent potential targets for intervention.\textsuperscript{2,5,8,37–39}

Epigenetic mediation through persistent changes induced in early life has indeed become a widely accepted hypothetical model for developmental programming,\textsuperscript{38–40} particularly in generic reviews of this field.\textsuperscript{39} This has motivated many studies, including examples published in this issue of the IJE, involving: early-life determinants of methylation levels shortly after birth\textsuperscript{16,17,20} and in adulthood\textsuperscript{14}; longitudinal associations of methylation levels with neurocognitive function and behaviour in children\textsuperscript{18} and with physical and cognitive fitness in the elderly;\textsuperscript{23} and the potential mediating role of DNA methylation in the association between maternal smoking and birth weight.\textsuperscript{15} It should be noted, however, that epigenetic persistence is not a prerequisite for a programmed effect, since a transient change in DNA methylation could, in theory, set other biological processes in train which then precipitate long-term effects.

Outside of the developmental programming literature there are also examples of long-term health effects of an exposure potentially mediated by DNA methylation. Lung cancer risk is higher among past smokers compared with never smokers, and the relative risk is maintained over time after quitting smoking.\textsuperscript{41} It is possible that long-term epigenetic modifications are mediators of this association, since past smoking has recently been associated with DNA methylation levels even decades after cessation\textsuperscript{42} (and could be considered a very useful biomarker of exposure at the tissue level). Nevertheless, disentangling mediation from other association-driving mechanisms—such as confounding and reverse causation—requires careful consideration,\textsuperscript{2,5,8,37} and statistical approaches to such mediation analysis are severely compromised by measurement error.\textsuperscript{43–46} Furthermore, the tissue from which methylation data are generated—blood, in the case of the above-referenced paper on dynamic methylation changes in relation to quitting smoking\textsuperscript{41}—is not a plausible candidate for a mediator between smoking and long-term, post-cessation, risk of lung cancer. These serious problems are unlikely to restrain over-confident claims of causality and mediation in the reporting of studies incorporating methylation measures, however.

Perusal of existing literature indicates that evidence of long-term effects of exposures on DNA methylation is largely limited to assumptions of persistence made in studies where methylation is measured at a single time point and related to historical exposure data. This approach is, of course, limited in the inferences that can be made, although access to serial samples from the same individuals with prospectively collected exposure data can help to improve this.

**Modelling methylation change over time**

Longitudinal data sets are valuable in many epidemiological contexts. Repeated measures of epigenetic signatures allow the modelling of change in methylation over time in tandem with single or repeated measures of an exposure that occurs prior to the methylation measurement.\textsuperscript{47} One can assess the persistence of differential methylation observed at birth across the life course, whether any changes are reversible and what factors may explain reversibility. For example, an epigenome-wide association study identified that methylation levels in seven gene regions were associated with maternal smoking during pregnancy. Four of these remained associated with maternal smoking throughout childhood and adolescence.\textsuperscript{48} Furthermore, models can be developed and tested to evaluate the intensity, duration and timing of an exposure on DNA methylation.\textsuperscript{48} The extent to which methylation data can indicate the particular timing of exposures—for example, during the intrauterine period or during puberty—is currently poorly understood.
A DNA methylation score derived from smoking-responsive DNA methylation sites has previously been used as an indicator of smoking status, i.e. to categorize current, former and never smokers and the widely used ‘epigenetic clock’ has been used to predict age from methylation patterns. The use of the ‘epigenetic clock’ as an indicator of biological rather than chronological age is increasingly being mooted, based on the assumption that DNA methylation signatures provide an index of cellular ageing (as in the case of telomere length).

The forensic epigenome

The potential utility of DNA methylation in forensics is beginning to generate interest, to the extent that by 2009 a popular prime-time TV programme, *Law and Order: Special Victims Unit* ran an episode entitled ‘Perverted’, in which DNA methylation analysis was utilized to demonstrate that DNA at a crime scene had been generated *in vitro* (and was unmethylated rather than the *in vivo* methylated copy) and thus appeared to have been planted. Although it was a far-fetched storyline, the methods alluded to in the programme had been reported in *Forensic Science International: Genetics*. The body fluid or tissue source of DNA obtained at crime scenes can also be of forensic importance, and DNA methylation can help in this identification.

Other potential forensic uses include distinguishing between a monozygotic (MZ) twin pair to establish which twin left a DNA sample at a crime scene. Methylation patterns, unlike the germ-line genome, can be identifying due to the phenotypic information that they reflect (although potentially, somatic mutations detected in complete high-coverage genome sequencing could also identify a particular MZ twin). This is not entirely of purely theoretical interest, as identical twins have indeed gone unprosecuted in such situations. Between-twin methylation differences appear stable enough to be useful even when there is a substantial time interval between when a sample from one twin was recovered from a crime scene and the twin pair had samples collected and examined. Indeed, if it transpires that there is any non-germ-line genetic variation-based paternal-to-offspring transmission of methylation, as some epigenetic enthusiasts claim, this could even be used in paternity tests involving an offspring of one of a pair of male MZ twins. Transgenerational epigenetic inheritance has been reviewed in detail but remains a contentious area of epigenetic research, in particular with respect to the public health importance of epigenetic variations transmitted across generations.

Despite overblown claims regarding potential forensic uses for DNA-based face-shape prediction, common genetic variation cannot provide much useful information about individual characteristics beyond sex and (probabilistically and problematically) ethnicity and related characteristics including eye, hair and skin colour, known collectively as ‘externally visible characteristics’, or EVCs.

DNA methylation offers the possibility of moving beyond conventional EVCs and adding identification of other aspects of the bodily habitus of the (generally unwitting) source of forensic blood (or other tissue) samples. The epigenetic clock, mentioned above, is producing mean absolute differences of chronological and estimated age of only 3–4 years in adults, which would certainly be useful for narrowing the range of those who could have been the source of recovered samples. Approaches using fewer markers have been formally tested within a forensics framework with demonstration of reasonable robustness to long-term room temperature blood sample storage and application to non-blood DNA sources.

As outlined above, smoking behaviour is a characteristic that can be reasonably reliably predicted using DNA methylation data, with separation of current from ex-smokers being possible. Alcohol consumption has also been investigated and useful indicators based on DNA methylation data may be developed, but evidence is too limited to generate these at present. For other aspects of habitus, such as body mass index, the combination of genetic and epigenetic data can improve prediction over the use of genetic data only, but this does not approach the level of being useful for identification.

Knowledge of smoking and alcohol-drinking behaviours could certainly help in the identification of recovered samples. In other situations, DNA methylation data could in principle point to both identifying characteristics and to illegal activity. For example, there has been considerable interest in potential epigenetic effects of cocaine, methamphetamine and other substance use, although this is mainly limited to animal studies at present.

Speculating further, maternal behaviours during pregnancy could leave marks on the offspring epigenome that have forensic implications as well as the epidemiological implications discussed earlier. As alluded to above, maternal smoking during pregnancy leaves such identifying and in some cases persistent offspring DNA methylation changes, providing retrospective indicators of this behaviour. There is also mounting evidence that maternal obesity, underweight or other dietary perturbation during pregnancy is associated with changes to the offspring’s epigenome. This has more than theoretical interest, since in the USA various states have introduced legislation (in one case referred to as ‘the cocaine mom act’) through which mothers who have participated in behaviours that could damage their fetuses during pregnancy can be subject
to criminal prosecution.79 The role of epigenetic understanding in formulating evidence as to whether, in general, maternal behaviours during pregnancy can influence fetal development and outcomes has been discussed in the legal context80 but, beyond this, DNA methylation data could be used retrospectively to establish that such behaviours have been practised during a particular pregnancy. In addition to smoking, maternal alcohol use during pregnancy has also been suggested to produce identifiable methylation changes in offspring,81 although this is not well established as yet. Offspring methylation indicators of other maternal behaviours, largely in the context of animal model studies, are currently subject to intense research activity.82–84 Indeed, continuing the speculation beyond the realms to which it should probably go, paternal behaviours that lead to detrimental effects on the offspring through transmitted epigenetic changes could also become a topic for investigation, taking the exclusive focus away from mothers (with its potential of compromising the constitutional rights of one societal group).85

The biosocial archive

Clearly the very same persistent epigenetic markers that allow for a potential forensic epigenome can also usefully serve in the context of epidemiological studies, contributing to the delineation of the full range of exposures to which individuals are subject (sometimes referred to as the exposome).86 The potential synthesis and integration of exposure history across the life course, during which social and biological processes are inscribed on the epigenome, suggests the possibility of a biosocial archive. The ability to better characterize exposures, including prenatal exposures using samples collected postnatally, offers a tool of considerable value to epidemiologists.

Just as researching the health effects of smoking acted as an impetus for the birth of chronic disease epidemiology,87 smoking has probably been the most widely investigated exposure in epigenetic epidemiology, as discussed above. Smoking-related methylation of the F2RL3 locus predicts mortality from cancer, cardiovascular disease and other causes, even after statistical adjustment for reported smoking behaviour.88 This probably indicates that the methylation measure provides a better indicator of long-term exposure to smoking than do reports of the behaviour, a conclusion strengthened by the finding that the associations of reported smoking behaviour with mortality were virtually abolished following adjustment for F2RL3 methylation. It is obvious that smoking is an underlying cause of these categories of mortality, but adjustment for a better indicator of the exposure attenuates the apparent effect of reported smoking behaviour. One important implication of this is that use of such methylation markers could reduce the residual confounding that remains after adjustment for reported smoking in observational epidemiological studies and, given that such residual confounding can—89 and does90—generate misleading evidence from such studies, this is potentially of considerable value.

Investigating the effects of maternal smoking on offspring health is an area of considerable research activity, and some studies rely on retrospective reports of whether mothers smoked during pregnancy. Persistent methylation markers assayed on offspring blood samples could clearly add to exposure classification in such situations.48

There are many studies of environmental exposures and DNA methylation, but differences in technologies used for methylation assessment and statistical analysis make drawing firm conclusions, of the sort available for smoking and age, difficult.91,92 Potentially exciting findings, for example an apparent methylation marker of prenatal lead exposure on umbilical cord blood samples,93 require robust replication. The very considerable value of such exposure indicators makes this an exciting area of research, which is likely to yield much that is worthwhile and can be implemented in ongoing epidemiological studies with banked blood samples.

In some cases, markers of particular exposures that have been searched for have not been found. For example, it would be very valuable to have a persistent indicator of having been born pre-term, given the association of pre-term birth with many later-life health outcomes. Whereas many gestational age-related methylation differences were observed in umbilical cord blood samples, these were not found to persist at later ages.47

A final issue relating to the epidemiological context is that the potential to identify previously unrevealed characteristics of individuals from methylation array data has led to suggestions that this produces particular ethical concerns with respect to the sharing of such data.94 and has led to debate about this.95 Epigenetic data can and are considered in the same ethical and governance framework as other personal data in epidemiological studies, although their potential to reflect a more detailed exposure history than is revealed by questionnaire or other data sources may warrant further consideration by consenting study participants.

Mechanisms to explain the stability of environmentally-induced variation in DNA methylation

The concepts of a biosocial archive and forensic epigenome both rely upon the perpetuation (or measurable attenuation over time) of environmentally-induced
epigenetic changes. However, the mechanisms explaining such persistence remain unclear. The majority of evidence available from epigenetic epidemiology studies to date is based on blood cells or cells from the oral mucosa. Both blood and oral epithelium are tissues with a rapid turnover of cells and therefore are dependent on adult stem cells (ASCs) for their maintenance, or tissue homeostasis as this has been termed, involving both tissue repair following injury and physiological tissue renewal. Briefly, ASCs are a cell sub-type that promotes tissue renewal by replenishing more differentiated cells (Figure 2). There are many differences between ASCs and more differentiated cells within the same tissue, including both gene expression and cellular environment.

A biological model where ASCs can act as reservoirs of environmental stresses, by continued tissue replenishment with cells harbouring a somatic event that is mitotically transmissible (e.g. a genetic mutation or an epigenetic modification), has been proposed in the context of cancer. The general idea is that, if an exposure causes an oncogenic somatic event in non-ASCs, any increased cancer risk associated with such an exposure would be short lived, especially in high-turnover tissues. However, this is not consistent with the notion that cancer (and many other non-communicable diseases) result from long-term exposure to different risk factors including past exposures—such as a history of tobacco smoking in lung cancer. Such increased disease risk that persists after cessation of exposure to a particular environmental factor must involve modifications occurring during the exposure period, which are maintained across cell division.

Evidence to date on epigenetic responsiveness to environmental, lifestyle, behavioural and other factors is largely limited to peripheral tissues, due to ease of sampling. The question arises as to whether other tissues (with lower turnover) provide an equal or better archive of exposure. Brain tissue from the prefrontal cortex has been linked to methylation changes induced by chronic pain, and other such examples can be found. However, the presence of persistent epigenetic changes in high turnover tissues, such as blood, may represent a more robust and responsive measure of dynamic change than a tissue with low turnover.

Recent literature, reviewed above, has provided evidence of sustained methylation differences according to past exposures. However, relatively little attention has been devoted to the cellular mechanisms underlying such maintenance. An obvious interpretation of sustained differential methylation is that there are still cells remaining in which the methylation modification arose. Given the high turnover of haematopoietic and epithelial tissues, this hypothesis is unlikely to apply in such cases. It is possible that some methylation modifications favour cell survival, thus increasing the lifespan beyond that of other cells. However, this explanation is ad hoc and is unlikely to apply in many cases—although this could occur if such modifications affect the function of genes associated with cell senescence or apoptosis. Furthermore, the increase in lifespan would have to be substantial to account for long-term effects after several years of exposure cessation. The counter argument has also been posited, which is that DNA methylation changes promote genomic instability and may therefore promote cell death rather than survival.

A second possibility is that the signal maintenance arises from the daughter cells of the parent cells where the methylation modification took place and was transmitted during mitosis (mitotic stability). Although this is a more plausible hypothesis, it does not explain why some exposure-associated methylation patterns are maintained over time, whereas others are not. One possibility is that a methylation change will only be maintained across (long) time periods—especially in high turnover tissues—if it arises in one or more ASCs. Although exposure continuation would likely increase the methylation signals, the epigenetic consequences of an exposure in the past (such as past smoking) could be maintained in the tissue due to cell replenishment by epigenetically-modified ASCs. This model would be a plausible biological mechanism for the maintenance of the absolute risk over time among individuals no longer exposed to the risk factor—as in the elevated lung cancer risk in former smokers. Importantly in the epidemiological context, although it is understandable that the development of such models has focused on epigenetic changes which influence disease risk, methylation and other epigenetic responses to exposures which do not have any consequences for disease, but can serve as exposure indicators, may be maintained in the same way.

This model also provides a plausible explanation for short-term associations between an exposure and methylation levels. According to this model, such associations can result from methylation modifications not having occurred in ASCs (although other processes could lead to the same
outcome). This could happen either by chance (which types of cells happen to get most exposure, or which respond with a methylation change, a process that is likely to be probabilistic) or through biological mechanisms. For example, it is possible that modifiability of some CpG regions depends on the cell sub-type within a tissue hierarchy.\textsuperscript{106} It is also well known that ASCs have protective mechanisms, including protection against external stresses in their microenvironment.\textsuperscript{101} Therefore, it would be expected that the proportion of cells epigenetically affected by a given exposure in the ASC compartment would be influenced by the proportion of cell sub-types in a given tissue.

An additional implication of this model is that even a methylation modification that occurred in an ASC might not persist over time. Methylation modifications in these cells could be reversed before cell division or after a few cell divisions\textsuperscript{106}, and additional mechanisms might also be involved. When an ASC divides, there are three potential outcomes: one committed cell (i.e. a cell that will generate differentiated cells) and one ASC (which is referred to as an asymmetrical division), two committed cells or two ASCs (symmetrical divisions). Such flexibility is important for prioritizing between ASC niche maintenance and replenishment of differentiated cells in response to specific stimuli.\textsuperscript{107} If the ASC in which the epigenetic event occurred undergoes a symmetrical division producing two committed cells, the epigenetic event will not be maintained in the ASC compartment\textsuperscript{102}, and, therefore, will be detectable for a shorter period of time after the end of the exposure. A summary of the models of proliferation and maintenance of ASCs is provided in Figure 3.

Mitotic heritability of cellular phenotypes is one of the two historical origins of the notion of ‘epigenetics’ (attributed to David Nanney, and contrasting to the developmental notion of how genes result in phenotypes, associated with Conrad Waddington)\textsuperscript{108,109} and methylation is, of course, only one of several mechanisms that could lead to such stability across mitosis.\textsuperscript{110,111} The stage of development (from embryonic stem cells through subsequent somatic stem cells and the niches they enter) at which a mitotically stable change occurs could influence the range of tissues in which this change is observed. The understanding of these processes of cellular differentiation and the evolution and dynamic nature of methylation signatures, both in general\textsuperscript{112,113} and in particular lineages (such as B cells),\textsuperscript{114} is rapidly developing. The influence of particular exposures (e.g. diet, smoking, lead) at particular stages and levels of chronicity on DNA methylation within cellular lineages is being explored\textsuperscript{115} and greater knowledge of these processes will help to inform theories of epigenetic persistence in different tissue types.

The ASC model, Mendelian randomization, ageing and interventions

Implications of adopting an ASC model to explain methylation stability over time extend to conceptual issues regarding Mendelian randomization.\textsuperscript{116} This technique has been applied in different fields of epidemiology, including epigenetic epidemiology. In the latter, Mendelian randomization can be used to obtain more robust evidence regarding: (i) the effects of an exposure on methylation levels; (ii) the effect of methylation levels on a disease outcome; and (iii) combining both in order to evaluate whether methylation levels are mediators of the association between the exposure and the outcome (a strategy referred to as two-step Mendelian randomization).\textsuperscript{121}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{ASC_model.png}
\caption{Symmetric and asymmetric stem cell division and maintenance of environmentally induced epigenetic change.}
\end{figure}
Because a genetic instrument for methylation levels has the same genotype regardless of the cell type and life-course stage, it approximates to the situation of long-term exposure-associated methylation modifications better than the case of short-term associations. This is particularly important if two-step Mendelian randomization is used to assess the mediating role of methylation modifications in relation to the effects of a time-restricted exposure, such as maternal smoking in pregnancy. If maternal smoking in pregnancy (in a hypothetical situation) only has short-term effects on methylation, then genetic instruments for such methylation differences would not necessarily provide valid evidence regarding the consequences of maternal smoking. This implies that using Mendelian randomization to investigate methylation modifications as mediators would be more reasonable in cases where there are long-term associations between the exposure and methylation. Understanding the (possible) role of ASCs in long-term epigenetic modifications may help in producing a broader understanding of conceptual considerations necessary for optimal study design and causal inference.

Another implication of considering the importance of cell lineages relates to physiological processes, such as aging. As discussed above, there are considerable age-related methylation modifications, useful in age-prediction at the population level. If such modifications occurred in committed cells, they would not accumulate— in a single cell—over time. Therefore, it would be more plausible to think that they occur in ASCs, thus re-emphasizing (assuming that the proposed model is true) the importance of this cell sub-type to the ageing process.

In the context of epigenetic markers as potential targets for intervention, the proposed role of ASCs in long-term maintenance of methylation modification would also be important. Referring back to the example of the association of past smoking with methylation and lung cancer risk, if long-term methylation modifications mediate sustained increase in risk of lung cancer risk in former smokers, then intervening on such modifications towards a ‘never smoker pattern’ could contribute to reduce cancer risk. If such long-term modifications result from tissue renewal by epigenetically modified ASCs, then a drug (or any other intervention) would have to be able to reach the ASC compartment of the relevant tissue to reverse the methylation modifications in these cells. In this regard, the protection mechanisms of ASCs would likely influence the efficacy of methylation-targeted interventions.

**Relevance to the epidemiologist**

In considering the concepts of a biosocial archive, a forensic epigenome and an ASC model for epigenetic persistence, one can speculate upon their relevance to the epidemiologist. The epigenome, or more specifically DNA methylation patterns, could enhance (and in some instances replace) exposure measurement or be considered as a biomarker in a diagnostic, predictive or prognostic context, adding a molecular dimension to a conventional observational study.

Epigenetic measures could also be considered as a surrogate end point in a clinical trial setting. Given the current state of knowledge, setting up randomized controlled trials to evaluate long-term effects of an exposure on methylation levels might be unjustifiable in most cases. However, such an investigation could be incorporated in trials with other primary goals—for example, in a randomized controlled trial evaluating the efficacy of an intervention to quit smoking. If the intervention is successful (e.g. increases smoking quitting rates) and the time from the first to the last measurement of the trial is sufficient to evaluate long-term effects on methylation, then an intention-to-treat analysis with methylation levels as the outcome could be performed. Moreover, strategies to improve causal inference in observational studies—such as Mendelian randomization—can be applied to observational data to investigate whether an exposure produces (at an aggregate level) a tendency towards a particular methylation change.

It is unlikely that epidemiological approaches will be able to resolve the outstanding questions regarding the hypothetical role of ASCs in epigenetic persistence of environmentally responsive DNA methylation changes; molecular biology will perhaps play a more important role here. Nevertheless, the ASC hypothesis may serve to help in the interpretation of epidemiological observations where data are generated from the longitudinal analysis of exposure-induced epigenetic changes.

**Conclusions**

The field of epigenetic epidemiology continues to provide evidence that the epigenome may serve as a very useful exposure indicator. There are broad applications in epidemiological studies, as a refined exposure indicator and biosocial archive but also in more specialized instances where prediction of phenotype can be gleaned from molecular signatures captured in the epigenome. Initial evidence that some epigenetic marks persist over decades whereas others are short lived raises the question as to the possible mechanisms underlying the stability of environmentally induced epigenetic changes. ASCs provide a plausible biological mechanism (unlikely to be the only one) underlying epigenetic persistence, although considerably more empirical data are required on this. ASCs can be considered as potential biological archives of events that
have occurred throughout the life course (which also contributes to the broader notion of the forensic epigenome). Such considerations also imply that in addition to the well-recognized between-tissue specificity complication in epigenetic epidemiological studies, within-tissue (in this case, cell differentiation stage) specificity is also of importance.

**Funding**

G.D.S. and C.L.R. are partially supported by the ESRC (RES-060-23-0011) ‘The biosocial archive: transforming lifecourse social research through the incorporation of epigenetic measures’. This work was supported by the UK Medical Research Council Integrative Epidemiology Unit, Cancer Research UK [C18281/A19169] and the University of Bristol (MC_UU_12013_1, MC_UU_12013_2).

**Conflict of interest:** G.D.S.’ mother smoked and drank alcohol during her pregnancy with him, and (jokingly) suggested he sue her for the damage done.

Caroline L Relton*

MRC Integrative Epidemiology Unit, School of Social & Community Medicine, University of Bristol, Bristol, UK

Fernando Pires Hartwig

Epidemiology Division, Federal University of Pelotas, Pelotas, Brazil

George Davey Smith

MRC Integrative Epidemiology Unit, School of Social & Community Medicine, University of Bristol, Bristol, UK

*Corresponding author. MRC Integrative Epidemiology Unit, School of Social & Community Medicine, University of Bristol, Bristol BS8 2BN, UK. E-mail: caroline.relton@bristol.ac.uk

**References**


44. Vanderweele TJ, Valeri L, Ogburn EL. The role of measurement error and misclassification in mediation analysis. Epidemiology 2012;23:561–64.
86. Wild CP. The exposome: from concept to utility. *Int J Epidemiol* 2012;41:24–32.