Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Psychology

Hidden hypotheses in 'hypothesis-free' genome-wide epigenetic associations

Edward D Barker[1], Susanna Roberts[1] and Esther Walton[2]

The recent interest in epigenetics within mental health research, from a developmental perspective, stems from the potential of DNA methylation to index both exposure to adversity and vulnerability for mental health problems. Genome-wide technology has facilitated epigenome-wide association studies (EWAS), permitting 'hypothesis-free' examinations in relation to adversity and/or mental health problems. In EWAS, rather than focusing on *a priori* established candidate genes, the genome is screened for DNA methylation, thereby enabling a more comprehensive representation of variation associated with complex disease. Despite their 'hypothesis-free' label, however, results of EWAS are in fact conditional on several *a priori* hypotheses, dictated by the design of EWAS platforms as well as assumptions regarding the relevance of the biological tissue for mental health phenotypes. In this short report, we review three hidden hypotheses — and provide recommendations — that combined will be useful in designing and interpreting EWAS projects.

Addresses
[1] Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, UK
[2] Medical Research Council Integrative Epidemiology Unit and Bristol Medical School, Population Health Sciences, University of Bristol, UK

Corresponding author: Barker, Edward D (ted.barker@kcl.ac.uk)

Current Opinion in Psychology 2019, **27**:13–17

This review comes from a themed issue on **Genetics**

Edited by **Brian Boutwell** and **Michael A White**

https://doi.org/10.1016/j.copsyc.2018.07.009

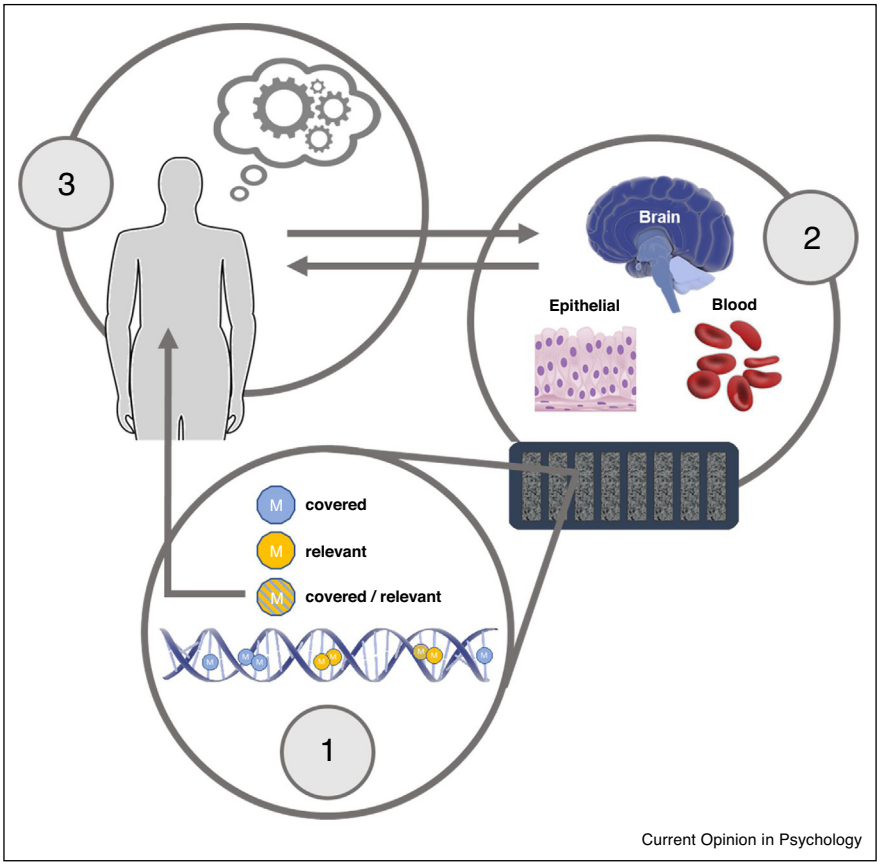
Understanding the biological mechanisms by which early psychosocial adversity associates with long-term mental health problems may have the potential to facilitate the development of effective screening, intervention strategies and health policy decisions [1]. Recent research has focused on the degree to which adversity disrupt gene regulation through epigenetic processes, thereby providing a mechanism by which the environment can have lasting effects on measurable mental health phenotypes [2••]. High profile studies suggest that epigenetic changes associated with early adversities [3,4] and even lifestyle choices [5•] can be observed across the life span, and that these long-term epigenetic modifications are associated with risk for a range of health outcomes [6]. These studies have generally focused on DNA methylation (DNAm) for two reasons: it is currently the best understood epigenetic mechanism and array-based technologies are readily available, which provides coverage of hundreds of thousands of methylation sites across the genome [7]. This combination of basic science and genome-wide technology has facilitated numerous epigenome-wide association studies (EWAS), permitting 'hypothesis-free' examinations in relation to adversity and/or mental health problems.

The logic underlying EWAS is comparable to genome-wide association studies (GWAS [8•]). Rather than focusing on DNAm in proximity to candidate genes, the genome is screened for DNAm, thus enabling a more comprehensive representation of variation associated with complex disease. As with GWAS (e.g. [9,10]), despite their 'hypothesis-free' label, results of EWAS are in fact conditional on several *a priori* hypotheses, dictated by the design of EWAS platforms as well as assumptions regarding the relevance of the biological tissue for the mental health phenotypes under investigation. In this short report, we review three hidden hypotheses (see Figure 1) — and provide recommendations — that combined will be useful in designing and interpreting EWAS projects.

Hidden hypothesis 1: EWAS coverage is sufficient for complex psychiatric problems

Array-based platforms have become widespread in psychology research, largely due to their ease of use, relatively high through-put, and well standardised and validated pipelines for processing, quality control, and analysis techniques. In particular, the Illumina 450k and EPIC arrays feature 480 000–850 000 probes targeting nearly 99% of RefSeq genes, as well as a range of other genomic categories, such as CpG islands, shores and shelves, miRNA promoters and enhancers, where DNAm can be influenced by and/or impact transcription in distal genomic regions [11••]. Compared with the Ilumina 450k, the newer Illumina EPIC 850k array provides much greater coverage of ENCODE and FANTOM5 enhancers [12••], and shows higher genetic influence underlying DNAm probes [13]. Nevertheless, these microarrays are limited in the number of sites they can assess, and thus lack true genome-wide measurements [14].

Figure 1



Hidden hypotheses in epigenome-wide approaches. *Note*: (1) = Hypothesis 1: EWAS coverage is sufficient for complex psychiatric problems; (2) = Hypothesis 2: peripheral tissue is meaningful for mental health problem(s); and (3) = Hypothesis 3: biology can be meaningful to phenotype of interest.

Furthermore, during the design process of the 450k and EPIC arrays, CpG sites were chosen as potentially biologically informative based on consultation with a consortium of DNA methylation experts [15]. Whilst the coverage of genes and CpG islands on these microarrays are comprehensive, it does not represent a complete picture of methylated cytosines across the genome. Selection was, in part, based on data from a number of phenotypes (some medical in nature such as cancer), and thus is not specifically targeted to brain-based, stress-related complex mental health phenotypes. This is an important point: if a sizeable proportion of the CpG sites tested are not relevant to the phenotype of interest, the likelihood of detecting relevant results is reduced.

Hidden hypothesis 2: peripheral tissue is meaningful for mental health problem(s)

The second hidden hypothesis relates to the tissue that is used to quantify DNAm. The majority of mental health research is based on DNAm profiles obtained from peripheral tissues from living persons, such as blood and saliva. When investigating outcomes such as conduct disorder or depression, however, the brain is often the main tissue of interest when it comes to mechanistic interpretations of results [16••]. To this end, research suggests that the correspondence of methylation profiles from blood and saliva to the brain is in fact quite limited, but can be higher with cross-tissue genetic influence [13,17]. This presents a critical disadvantage if the investigator would like to use the peripheral tissue as a surrogate of the central nervous system (CNS; the brain).

One promising avenue is to establish DNAm as a biomarker for mental illness. A biomarker does not have to be mechanistic (i.e. CNS surrogate). Indeed, blood-based biomarkers have been used for diagnostics, predictive risk, disease monitoring and/or treatment response in cancer, cardiovascular and infectious disease [18,19]. However, even within a biomarker framework, the assumption is often that distinct peripheral tissues are interchangeable and equally suited for biomarker detection, when in fact it is highly probable that peripheral tissues themselves correspond differently to environmental adversity and/or disease state [14]. For instance,

biomarkers for mental health traits (e.g. depression) may be more detected in blood than saliva, as blood is more central to inflammatory processes related to stress and disease [16••].

Hidden hypothesis 3: biological relevance for the phenotype of interest

The last hypothesis relates to the assumption that biology can be informative to the phenotype itself. Focal phenotypes (e.g. oppositional defiant disorder, anxiety) in mental health research are often complex and multiply determined [20]. The lack of established robust biomarkers for mental health problems (e.g. [19]) may suggest that some of these traits might not strongly associate with detectable biological processes. Furthermore, effect size associations in EWASes are often very small suggesting that — while significant — distinguishing the importance of DNAm in the aetiology of the mental health phenotype may prove difficult [21••]. Perhaps unsurprisingly then, most EWAS in mental health include some form of gene ontology analysis, which queries the role of larger biological systems based on existing databases [22]. These analyses result in general statements such as 'neurodevelopment' or the 'immune system' being involved in the aetiology of a given phenotype. Whether these broad categories play indeed a substantial role in the aetiology of the mental health problem is often hard to determine given the post hoc nature of the interpretation. Relatedly, many EWASes have tried to infer downstream effects of observed variation in DNAm such as differences in gene expression. Many of these studies find very little in terms of functional relationships, but a small number do report downstream biological associations (e.g. [21••]). In general, it has proven difficult to pinpoint EWAS-related biological relevance of observed DNAm changes, even if they are in genes which seem 'plausible' based on reported functionality and previous literature.

Recommendation for hidden hypothesis 1: EWAS coverage

An alternative to using arrays with limited coverage is to use next-generation sequencing-based approaches to interrogate the whole methylome [21••]. However, these methodologies are high in cost and time intensive. Despite the limitations described above, pragmatic and strategic study design can maximise utility and interpretation of results of the Illumina 450k and EPIC arrays. For example, for researchers interested in targeting CpGs likely to associate with 'brain-based' mental illnesses, an *a priori* set of CpGs (e.g. a 'systems approach') could be isolated from the array data, which could still span thousands of loci. The suggestion is to prioritize CpGs within biological systems that are known to associate with variation in post-mortem brain samples [23••] or even structural or functioning brain imaging [24] if this is of primary interest to the investigator.

The second recommendation for optimising the use of EWAS CpGs is to target those probes with underlying genetic influence — methylation quantitative trait loci (mQTLs). This approach may have the advantage that cross-tissue concordance (e.g. blood, saliva, post-mortem brain) appears higher for CpGs that show cross-tissue genetic influence [13]. Another advantage of mQTLs is that CpGs under considerable genetic influence are less affected by confounds [11••,13,25]. However, while mQTLs are a worthwhile approach, it is a relatively new area and at present, there is a small proportion of methylation sites with consistently reported mQTLs [11••,13]. Furthermore, large-scale and detailed information on tissue-specific mQTLs is still sparse.

Recommendation for hidden hypothesis 2: peripheral tissue and phenotype

One strategy to maximise the interpretability of EWAS projects is to examine DNAm as a biomarker for mental health problems that have mechanistic underpinning in tissues other than the brain, such as blood. A wide-range of psychiatric disorders have been associated with immune function as measured by peripheral inflammation [26]. Furthermore, there is good evidence from animal studies, and increasing evidence in humans, that peripheral inflammatory markers can affect brain areas implicated in certain psychiatric disorders [27]. Consequently, adversity-related immune processes and DNAm may be well measured in blood samples (see [28••]). For biomarkers to be useful, they must be cost effective, drawn from accessible tissue and predictive of future risk [29]. Biomarkers for brain-based disorders (e.g. depression) have proven more difficult to establish [19]. Liu *et al.* [30] performed an EWAS on blood tissues across 13 population-based cohorts and reported that a composite biomarker (consisting of 144 CpGs) discriminated drinkers from non-drinkers. It was thus suggested that a blood-based DNAm diagnostic test could be developed. It is important to note, however, that in addition to methodological considerations [31], the Liu *et al.* study was cross-sectional, thus it may prove difficult to use this specific biomarker as a predictor of future alcohol use, as the variation in DNAm may be the result of chronic drinking (i.e. reverse causality [32]). Importantly, large-scale meta-analyses based on new and growing consortia (e.g. PACE [33•]) are beginning to report consistent epigenetic effects on traits such as schizophrenia or smoking behaviour (e.g. [34,35]) which suggests that we may begin to be able to utilise this information to further optimize DNAm biomarker approaches.

Recommendation for hidden hypothesis 3: phenotype and biology

Several suggestions have been put forward to address the complex nature of the biology that may underlie mental health problems. Most notably, the Research Domain Criteria (RDoC) initiative has proposed alternative

approaches to study mental illness by integrating many different levels of information including genetics, neurocircuits and behaviour [36]. Methylation-based research can be integrated into an RDoC perspective. Here, researchers could employ a two-stage analysis, first investigating epigenetic effects on intermediate dimensions of mental health and then, using the results as biomarkers to query the more complex phenotypes. For, example, if externalising difficulties (e.g. ADHD, aggression) are the focal phenotype, rather than performing an EWAS directly on the disorder(s), the researchers could instead, as the first step, perform an EWAS on brain imaging endophenotypes of the externalising phenotype (e.g. [24]). In the second step, the results of the EWAS could be used to create poly-epigenetic genetic biomarker score (e.g. [28••]) to be (potentially) associated with the disorder. This type of two stage of EWAS may examine the epigenetic changes associated with antecedents of diagnosable mental health conditions, which would be could be more useful as a risk biomarker than a biomarker of the actual diagnosis.

Conclusion

The recent interest in epigenetics, from a developmental perspective, stems from the potential of DNA methylation to index both exposure to adversity and vulnerability for mental health problems [2••]. To this end, there has been substantial activity in examining EWASes of adversity-related disorders, such as conduct disorder [37] and psychosis [38]. Of interest, from these EWAS, DNAm in genes that underlie stress response, neurotransmitter activity and immune regulation have been identified. These preliminary findings may provide a useful framework for more in-depth investigations — potentially as CNS surrogates or biomarkers — of the biological pathogenesis of a mental health problem. However, we argue that understanding hidden hypotheses within the EWAS is an important first step in interpreting the results in relation to mental health phenotypes.

Conflict of interest statement

Nothing declared.

Acknowledgement

This manuscript was supported by a grant from the Economic and Social Research Council (ES/R005516/2) to EDB.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Shonkoff JP, Boyce WT, McEwen BS: **Neuroscience, molecular biology, and the childhood roots of health disparities: building a new framework for health promotion and disease prevention**. *J Am Med Assoc* 2009, **301**:2252-2259 http://dx.doi.org/10.1001/jama.2009.754.

2. •• Barker ED, Walton E, Cecil CAM: **Annual research review: DNA methylation as a mediator in the association between risk exposure and child and adolescent psychopathology**. *J Child Psychol Psychiatry* 2018, **59**:303-322 http://dx.doi.org/10.1111/jcpp.12782.
Very good review on the degree to which DNA methylation links risk exposure and psychopathology.

3. Klengel T, Mehta D, Anacker C, Rex-Haffner M, Pruessner JC, Pariante CM, Pace TW, Mercer KB, Mayberg HS, Bradley B *et al.*: **Allele-specific FKBP5 DNA demethylation mediates gene–childhood trauma interactions**. *Nat Neurosci* 2013, **16**:33.

4. Rijlaarsdam J, Cecil CA, Walton E, Mesirow MS, Relton CL, Gaunt TR, McArdle W, Barker ED: **Prenatal unhealthy diet, insulin-like growth factor 2 gene (IGF2) methylation, and attention deficit hyperactivity disorder symptoms in youth with early-onset conduct problems**. *J Child Psychol Psychiatry* 2017, **58**:19-27.

5. • Richmond RC, Simpkin AJ, Woodward G, Gaunt TR, Lyttleton O, McArdle WL, Ring SM, Smith AD, Timpson NJ, Tilling K: **Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC)**. *Hum Mol Genet* 2015, **24**:2201-2217.
Longitudinal study describing the transient, but also partially persistent effects of maternal smoking on offpsring DNA methylation.

6. Tobi EW, Slieker RC, Luijk R, Dekkers KF, Stein AD, Xu KM, Slagboom PE, van Zwet EW, Lumey LH, Heijmans BT: **DNA methylation as a mediator of the association between prenatal adversity and risk factors for metabolic disease in adulthood**. *Sci Adv* 2018, **4** eaao4364.

7. Feinberg AP: **The key role of epigenetics in human disease prevention and mitigation**. *N Engl J Med* 2018, **378**:1323-1334.

8. • Rakyan VK, Down TA, Balding DJ, Beck S: **Epigenome-wide association studies for common human diseases**. *Nat Rev Genet* 2011, **12**:529.
Overview of the strengths and, 2018 overview of the strengths and challenges of epigenetic research for complex diseases, with recommendations for study design and other considerations.

9. Kitsios GD, Zintzaras E: **Genome-wide association studies: hypothesis-"free" or "engaged"?** *Transl Res* 2009, **154**:161-164.

10. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, Hirschhorn JN: **Genome-wide association studies for complex traits: consensus, uncertainty and challenges**. *Nat Rev Genet* 2008, **9**:356.

11. •• Gaunt TR, Shihab HA, Hemani G, Min JL, Woodward G, Lyttleton O, Zheng J, Duggirala A, McArdle WL, Ho K *et al.*: **Systematic identification of genetic influences on methylation across the human life course**. *Genome Biol* 2016, **17**:1.
Extremely well-done research studying genetic influences on DNA methylation.

12. •• Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, Van Djik S, Muhlhausler B, Stirzaker C, Clark SJ: **Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling**. *Genome Biol* 2016, **17**:208.
Rigorous comparison of the EPIC BeadChip to the HM450 BeadChip and Whole Genome Sequencing, to evaluate agreement, coverage and content.

13. Lin D, Chen J, Perrone-Bizzozero N, Bustillo JR, Du Y, Calhoun VD, Liu J: **Characterization of cross-tissue genetic-epigenetic effects and their patterns in schizophrenia**. *Genome Med* 2018, **10**:13.

14. Jones MJ, Moore SR, Kobor MS: **Principles and challenges of applying epigenetic epidemiology to psychology**. *Annu Rev Psychol* 2018, **69**:459-485.

15. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, Delano D, Zhang L, Schroth GP, Gunderson KL *et al.*: **High density DNA methylation array with single CpG site resolution**. *Genomics* 2011, **98**:288-295.

16. •• Bakulski KM, Halladay A, Hu VW, Mill J, Fallin MD: **Epigenetic research in neuropsychiatric disorders: the "tissue issue"**. *Curr Behav Neurosci Rep* 2016, **3**:264-274.

Very good review on the challenges but also potential avenues of using blood tissue in neuropsychiatric research.

17. Walton E, Hass J, Liu J, Roffman JL, Bernardoni F, Roessner V, Kirsch M, Schackert G, Calhoun V, Ehrlich S: **Correspondence of DNA methylation between blood and brain tissue and its application to schizophrenia research**. *Schizophr Bull* 2015. sbv074.

18. Ladd-Acosta C, Fallin MD: **The role of epigenetics in genetic and environmental epidemiology**. *Epigenomics* 2015, **8**:271-283 http://dx.doi.org/10.2217/epi.15.102.

19. O'Bryant SE, Mielke MM, Rissman RA, Lista S, Vanderstichele H, Zetterberg H, Lewczuk P, Posner H, Hall J, Johnson L: **Blood-based biomarkers in Alzheimer disease: current state of the science and a novel collaborative paradigm for advancing from discovery to clinic**. *Alzheimer's Dement* 2017, **13**:45-58.

20. Barker ED, Cecil CAM, Walton E, Meehan AJ: **Genetic influences on childhood disruptive behaviors**. In *The Wiley Handbook of Disruptive and Impulse-Control Disorders.* Edited by Lochman J, Matthys W. USA: Wiley-Blackwell; 2017.

21. •• Breton CV, Marsit CJ, Faustman E, Nadeau K, Goodrich JM, Dolinoy DC, Herbstman J, Holland N, LaSalle JM, Schmidt R *et al.*: **Small-magnitude effect sizes in epigenetic end points are important in children's environmental health studies: the Children's Environmental Health and Disease Prevention Research Center's Epigenetics Working Group**. *Environ Health Perspect* 2017, **125**:511.

Thorough review discussing methodologies in epigenetic research, robust findings (even of small effect sizes) in longitudinal work, and implications for the future.

22. Michels KB, Binder AM, Dedeurwaerder S, Epstein CB, Greally JM, Gut I, Houseman EA, Izzi B, Kelsey KT, Meissner A, Milosavljevic A, Siegmund KD, Bock C, Irizarry RA: **Recommendations for the design and analysis of epigenome-wide association studies**. *Nat Methods* 2013, **10**:949-955 http://dx.doi.org/10.1038/nmeth.2632.

23. •• Hannon E, Lunnon K, Schalkwyk L, Mill J: **Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes**. *Epigenetics* 2015, **10**:1024-1032.

This is an important paper, one of the first to show concordance in DNAm across tissue types, particularly blood and brain.

24. Walton E, Cecil CAM, Suderman M, Liu J, Turner JA, Calhoun VD, Ehrlich S, Relton CL, Barker ED: **Longitudinal epigenetic predictors of amygdala:hippocampus volume ratio**. *J Child Psychol Psychiatry* 2017, **58**:1341-1350.

25. Hannon E, Weedon M, Bray N, O'Donovan M, Mill J: **Pleiotropic effects of trait-associated genetic variation on DNA methylation: utility for refining GWAS loci**. *Am J Hum Genet* 2017, **100**:954-959.

26. Nusslock R, Miller GE: **Early-life adversity and physical and emotional health across the lifespan: a neuroimmune network hypothesis**. *Biol Psychiatry* 2016, **80**:23-32.

27. Felger JC, Li Z, Haroon E, Woolwine BJ, Jung MY, Hu X, Miller AH: **Inflammation is associated with decreased functional connectivity within corticostriatal reward circuitry in depression**. *Mol Psychiatry* 2016, **21**:1358.

28. •• Barker ED, Cecil CAM, Walton E, Houtepen LC, O'Connor TG, Danese A, Jaffee SR, Jensen SKG, Pariante C, Roberts S, McArdle W, Gaunt TR, Relton CL: **Inflammation-related epigenetic risk and child and adolescent mental health: a prospective study from pregnancy to mid-adolescence**. *Dev Psychopathol* 2018. (in press).

First longitudinal study to employ inflammation-related poly epigenetic genetic risk scores in relation to pre- and post-natal adversities, neuro-cognitive function and mental health problems.

29. Ladd-Acosta C: **Epigenetic signatures as biomarkers of exposure**. *Curr Environ Health Rep* 2015, **2**:117-125.

30. Liu C, Marioni RE, Hedman AK, Pfeiffer L, Tsai PC, Reynolds LM, Just AC, Duan Q, Boer CG, Tanaka T *et al.*: **A DNA methylation biomarker of alcohol consumption**. *Mol Psychiatry* 2016 http://dx.doi.org/10.1038/mp.2016.192.

31. Hattab M, Clark S, van den Oord E: **Overestimation of the classification accuracy of a biomarker for assessing heavy alcohol use**. *Mol Psychiatry* 2017 http://dx.doi.org/10.1038/mp.2017.181.

32. Barker ED: **Epigenetics, early adversity and child and adolescent mental health**. *Psychopathology* 2018, **51**:71-75.

33. • Felix JF, Joubert BR, Baccarelli AA, Sharp GC, Almqvist C, Annesi-Maesano I, Arshad H, Baïz N, Bakermans-Kranenburg MJ, Bakulski KM *et al.*: **Cohort profile: Pregnancy And Childhood Epigenetics (PACE) Consortium**. *Int J Epidemiol* 2017.

Important example of consortia type efforts.

34. Hannon E, Dempster E, Viana J, Burrage J, Smith AR, Macdonald R, St Clair D, Mustard C, Breen G, Therman S *et al.*: **An integrated genetic-epigenetic analysis of schizophrenia: evidence for co-localization of genetic associations and differential DNA methylation**. *Genome Biol* 2016, **17**:176.

35. Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, Reese SE, Markunas CA, Richmond RC, Xu C-J: **DNA methylation in newborns and maternal smoking in pregnancy: genome-wide consortium meta-analysis**. *Am J Hum Genet* 2016, **98**:680-696.

36. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, Sanislow C, Wang P: **Research domain criteria (RDoC): toward a new classification framework for research on mental disorders**. *Am Psychiatric Assoc* 2010, **167**:748-751.

37. Cecil CAM, Walton E, Jaffee SR, O'Connor T, Maughan B, Relton CL, Smith RG, McArdle W, Gaunt TR, Ouellet-Morin I, Barker ED: **Neonatal DNA methylation and early-onset conduct problems: a genome-wide, prospective study**. *Dev Psychopathol* 2018, **30**:383-397.

38. Fisher HL, Murphy TM, Arseneault L, Caspi A, Moffitt TE, Viana J, Hannon E, Pidsley R, Burrage J, Dempster EL *et al.*: **Methylomic analysis of monozygotic twins discordant for childhood psychotic symptoms**. *Epigenetics* 2015, **10**:1014-1023.