Review

Is there a link between genome-wide hypomethylation in blood and cancer risk?

Kevin Brennan and James M. Flanagan*

Epigenetics Unit, Department of Surgery and Cancer, Hammersmith Hospital, Imperial College, London W12 0NN

*Author for correspondence: Dr. James M. Flanagan

Epigenetics Unit, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London

4th Floor IRDB, Hammersmith Campus

Du Cane Road, London W12 0NN

Tel: 02075941804 Fax: 02075942129

Email: j.flanagan@imperial.ac.uk

Running title: Genome-wide hypomethylation and cancer risk

Conflicts of Interest: None
Abstract

Cancer cells display widespread genetic and epigenetic abnormalities, but the contribution to disease risk, particularly in normal tissue prior to disease, is not yet established. Genome-wide hypomethylation occurs frequently in tumours and may facilitate chromosome instability, aberrant transcription and transposable elements reactivation. Several epidemiological case-control studies have reported genomic hypomethylation in peripheral blood of cancer patients, suggesting a systemic effect of hypomethylation on disease predisposition, which may be exploited for biomarker development. However, more recent studies have failed to reproduce this. Here we report a meta-analysis, indicating a consistent inverse association between genomic 5-methylcytosine levels and cancer risk (95% CI 1.2-6.1), but no overall risk association for studies using surrogates for genomic methylation, including methylation at the LINE-1 repetitive element (95%CI 0.8-1.7). However, studies have been highly heterogeneous in terms of experimental design, assay type and analytical methods. We discuss the limitations of the current approaches, including the low inter-individual variability of surrogate assays such as LINE1 and the importance of using prospective studies to investigate DNA methylation in disease risk. Insights into genomic location of hypomethylation, from recent whole genome, high-resolution methylome maps, will help address this interesting and clinically important question.
Introduction

Epigenetics is the mitotically heritable control of gene expression and chromatin structure through covalent modification by DNA methylation, post-translational modifications of histone proteins, and control of gene expression by non-coding RNAs [1, 2]. DNA methylation is the addition of methyl groups to cytosines within CpG dinucleotides, to form 5-methylcytosine (5meC) and is catalysed by the DNA methyltransferases (DNMTs) [2, 3]. DNA methylation is dependent on the one carbon metabolism pathway, which requires dietary supply of folate, homocysteine and choline, and other micronutrients [4, 5]. The role of DNA methylation as a negative regulator of transcription initiation at CpG islands, regions of high CpG density found in approximately 60% of genes, is well established, and is required for silencing of developmental and tissue specific genes [2, 6, 7]. DNA Methylation facilitates transcriptional repression independently through preventing binding of DNA polymerases and transcription factors, or through interaction with methyl binding domain proteins (MBDs) and other epigenetic mechanisms such as repressive histone marks [8, 9]. The functional roles of DNA methylation elsewhere in the genome is less well understood, however, DNA methylation is thought to confer genomic stability and integrity, and high methylation at repetitive elements is proposed to protect against expression of transposable elements (TEs) and endogenous retroviruses [10, 11]. Epigenetic regulation is often studied in the context of environmental and population health, as widespread DNA methylation patterns are known to be affected by environmental, lifestyle and demographic factors that affect complex disease risk, such as diet, carcinogen exposure, reproductive factors and age [12, 13]. Furthermore, this ‘plasticity’ or susceptibility to chemical alteration of the epigenome provides promise for therapeutic intervention where abnormal epigenetic patterns occur in disease or in relation to disease risk [14]. A current research topic is the potential of DNA methylation variability in normal tissues such as peripheral blood, for development cancer predisposition biomarkers [15, 16]. This review and meta-analysis aims to synthesise all of the current data and to answer the
question of whether there is a link between genome-wide DNA hypomethylation and cancer risk.

DNA methylation and cancer

DNA methylation dysregulation occurs in almost all cancers, and is becoming established as a hallmark of cancer [1, 11, 17]. Most research has been concerned with local hypermethylation of promoter CpG islands, as transcriptional repression due to promoter hypermethylation of tumour suppressor genes such as MLH1 and BRCA1 is implicated in driving several cancers [7, 8]. However, the first described cancer-associated epigenetic phenomenon was the lower net percentage of 5meC in tumour tissue compared with equivalent normal tissue, known as ‘genome-wide’ or ‘global’ hypomethylation, that occurs frequently in all cancer types [18, 19]. Despite considerable research, it remains unclear whether promoter hypermethylation and genome-wide hypomethylation are mechanistically linked, or are independent processes. [20, 21]. Whole-genome DNA methylation (methylome) analysis has so far revealed than an overall increase in methylation variability (hypervariability) may be more prevalent than discreet changes in DNA methylation levels in cancer, suggesting a widespread loss of epigenetic control [9, 17, 22]. Recent research has shown that most cancer-associated DNA hypermethylation occurs at ‘CpG shores’, regions flanking CpG islands, rather than within CpG islands [17, 23]. Furthermore, genome-wide hypomethylation is largely confined to large genomic ‘hypomethylation blocks’ [9, 17], that tend to occur in regions that display intermediate methylation levels in normal tissue, termed partially methylated domains (PMDs)[17, 22], and regions of low CpG density [23]. That hypomethylation is specifically associated with repetitive elements has been called into question, with only a modest enrichment of repeats in hypomethylation blocks [17, 24]. Such apparent compartmentalisation of hypomethylation may reflect higher-order changes in chromatin structure [9]. Whereas genome-wide hypomethylation has long been implicated in loss of transcriptional repression, recent evidence suggests that hypomethylation at PMDs or
gene-bodies may be often associated with gene repression through formation of repressive chromatin [24]. Therefore, understanding of cancer epigenetics is rapidly advancing with the advent of whole-methylome sequencing, revealing a greater complexity and organisation to the epigenome than was previously appreciated.

**Causes and consequences of genome-wide hypomethylation**

The cause of genome-wide hypomethylation remains poorly-understood, and several factors have been implicated [11, 21]. Dietary insufficiency of folate and other one-carbon metabolism-dependent micronutrients, or germline or somatic mutations in members of the pathway can induce hypomethylation in tumour, and normal tissues including blood. [4, 14, 25, 26] Exogenous exposures, such as carcinogen exposure may influence methylation by inducing DNA damage or by affecting DNMT enzyme activity [4, 11, 27, 28]. Furthermore, the efficiency of DNA methylation may be affected by age [12, 29]. Lastly, neoplasia may induce ‘field-effect’ hypomethylation in cancer and surrounding histologically normal tissues, through sequestration of DNA methylation enzymes and substrates by rapidly proliferating, dividing cancer cells, resulting in failure of DNA methylation maintenance in surrounding tissue [11, 30-35].

Several tumorigenic events may result from genome-wide hypomethylation. These include chromosomal instability, which may facilitate gene-dosage alterations, mutations, genetic recombination, large deletions, or translocations [2, 11, 36]. Hypomethylation may facilitate altered expression of oncogenes and/or tumour suppressors, as well as aberrant transcription of non-coding RNAs via transcriptional read-through subsequent to a loss of repression at repetitive DNA [37-39]. Expression of repetitive elements is a feature of many cancers, with largely unknown consequences [37, 40]. Whether genome-wide hypomethylation represents an early or causative tumorigenic event or a passive consequence of cancer remains unknown (reviewed elsewhere [21]). The detection of hypomethylation in cancer precursor lesions and normal adjacent tissue, and the induction of
cancers in animal models with experimentally induced hypomethylation suggest a causative role [30, 41-43]. However, the occurrence of many cancers without apparent hypomethylation, and the progression of hypomethylation with cancer stage suggest a passive role [20, 21].

**Transposable element hypomethylation**

Due to the conventional view that genome-wide hypomethylation is associated with repetitive elements, surrogate assays for genome-wide methylation have been developed specifically targeting consensus sequences within repetitive elements, including LINEs (LINE1), SINEs (Alu) and Satellite repeats (Sat2)[44, 45]. LINE1 (Long Interspersed Nucleotide Element 1, hereafter referred to as ‘L1’) is the only autonomous (capable of independent retrotransposition (transposition via an RNA intermediate)), and most highly expressed TE in the human genome, comprising ~17% of the human genome, with over 500,000 copies [11, 46, 47]. A full length L1 element is ~6kb long with a bi-directional, non-canonical promoter and two open reading frames (ORFs) coding for an endonuclease and retrotransposition machinery proteins[11, 46, 47]. L1 transcription is largely regulated by DNA methylation of the 5’ promoter; however, most L1 elements are truncated and cannot be transcribed [37, 38, 40]. Approximately 100 L1 elements are functionally capable of retrotransposition, however, only a few contribute to the vast majority of retrotransposition events [46]. Alu elements, of which there are multiple families, are the most common TE in the human genome, with ~1.1 million copies, comprising roughly 11% of the genome [11, 47]. Alu elements are non-autonomous and require the L1 transposition machinery for transposition[11]. Satellite repeats (including Sat2) are short tandemly repeated non-coding DNA, frequently in centromeric and heterochromatic regions of chromosome 1 [48]. Both L1 and Alu elements are heavily methylated in normal somatic tissue, however, hypomethylation of both, especially L1, is often detectable in tumours [20, 30, 49]. Whereas transcription of TEs are required for their transposition, transposition-independent consequences of transcription may have functional consequences, as an estimated 7% of
the human transcriptome is derived from transcription start sites (TSS) within L1 elements [37] and hypomethylation induced transcription of L1 elements within host gene introns can effect host gene transcription through RNA interference [39] and through driving host gene ectopic expression from the L1 antisense promoter [50].

DNA methylation variability and cancer risk

We and others have hypothesised that the epigenetic variability may contribute to risk of cancer development [14, 15, 51-53]. “Epigenetic epidemiology” refers to the investigation of epigenetic patterns associated with disease risk, and such research holds great potential for cancer prevention and cancer risk-biomarker development, especially where abnormal epigenetic patterns may be detectable in easily accessible tissues such as blood, which is suitable for population screening [15, 16].

Methylation variability may contribute to, or be predictive of, risk of tumorigenesis in other tissues for several reasons. Many innate cancer risk factors, such as age, anthropometric factors, and genetic factors affect epigenetic patterns in blood [54, 55, 56]. For example, rare epimutations in MLH1 and MSH2, which consist of local DNA hypermethylation events detectable in all tissues and conferring high risk of colorectal cancer, have been shown to occur due to genetic polymorphisms [57, 58]. Blood DNA methylation is also affected by exposure to environmental and lifestyle cancer risk factors, such as smoking, alcohol and other carcinogens [59, 60], therefore, blood methylation may provide useful biomarkers for acquired or environmentally induced cancer risk [12, 13]. Lastly, inflammation predisposes to risk of many cancers [61, 62], and affects DNA methylation [63], suggesting that blood methylation may reflect immune effects on cancer risk.

Promising for this avenue of research is the identification of a variable methylated region within the ATM gene associated with breast cancer risk [64]. Furthermore, hypermethylation
of BRCA1 is reportedly associated with increased prevalence of BRCA1-methylated breast tumours [65, 66], suggesting a functional link between blood detectable DNA methylation and disease histology [67-69]. High-throughput discovery studies have identified cancer-associated DNA methylation signatures in blood of patients with breast [69], ovarian [68] bladder [67] and head and neck (HCC) cancers [70], providing potential cancer diagnostic biomarkers.

The most widely studied putative epigenetic risk marker for cancer is genome-wide or ‘global’ DNA methylation in blood. The studies investigating this, their findings, and the research approaches employed, will be discussed further.

Current methods for investigation of genome-wide hypomethylation

A challenge in investigating genome-wide DNA methylation in population studies is the lack of cost-effective, high-throughput assays that have both wide genome coverage and high resolution [45]. Early studies used methods that measure genomic 5meC content using 5meC-specific antibodies [71], methyl-acceptance assays [72] and HPLC [73]. Such methods however, give no information about the spatial arrangement or genomic location of DNA methylation and require large amounts of input DNA, making them unsuitable for population studies using precious patient samples. Widespread DNA methylation patterns can be measured using methylation sensitive restriction enzymes with relatively high resolution; however, methylation analysis is biased towards regions of high CpG-density, as the enzyme restriction sites occur at sequences such as CCGG, and CGCG for the HpaII and Hha1 restriction enzymes, respectively [45, 74]. Most DNA methylation assays are based on bisulphite conversion of DNA, where incubation of DNA with sodium bisulfite causes deamination of unmethylated, but not methylated, cytosines to uracil [45]. Sequencing of bisulphite converted DNA represents the ‘gold standard’ for DNA methylation analysis [16, 75], however, real-time PCR, restriction enzyme based (combined bisulphite restriction
analysis (COBRA)[76]), and microarray-based methods[75] are also used. Most popular and convenient for population studies is the measurement of methylation at the repetitive elements such as L1, Alu and Sat2, which are distributed at high frequency throughout the human genome [11, 44]. These are considered ‘surrogate’ measures of genome-wide DNA methylation as their methylation is thought to reflect genome-wide methylation levels [77]. However, the efficacy of these surrogate assays for genome-wide methylation has been questioned [14, 78] (discussed below), and a re-interpretation of the results reported using these assays is warranted.

Comparison of published reports of genome-wide methylation in blood and cancer risk.

Twenty three publications have reported population based cancer case-control studies investigating blood genomic DNA methylation in relation to cancer risk (Table 1). It is important to note that the comparability of these studies is limited by several differences between each, including cancer type, assay, study design (prospective/retrospective), length of time between blood-draw and diagnosis, sample size, sex (male/female/mixed), ethnicity, cancer treatment exposure and analytical/statistical methods. Furthermore, populations at different cancer risk are included, for instance one report included elderly men at high cancer risk [79], whereas another included Asian women at lower risk of breast cancer [80]. Some reports included more than one “study”, due to use of different assays, different patient populations or different study designs, so altogether there were 34 individual studies of genome-wide methylation and cancer incidence/prevalence (Table 1). The greatest limitation to comparison between studies was reporting of data analyses. Twenty six studies reported an odds ratio (OR) for cancer associated with DNA methylation, generated by categorical analysis. Categorical analysis compares the OR for cancer for individuals displaying methylation within the lowest category (tertile, quartile or decile) of methylation, compared
with individuals within the highest (reference) category. The choice of categorical 'split' for methylation affects the OR, with fewer categories providing more conservative analysis, and this varied between studies (Table 1). OR was the most consistently reported, comparable, and representative (of overall results) factor between studies, and was therefore used for meta-analysis. Therefore, eight studies reporting risk analysis at the mean/median level only (none of which showed significant case-control differences) could not be included, resulting in an over-estimate of any positive effects [69, 73, 79, 81-83] (Table 1). We have recently reported methylation analysis of blood L1 methylation in two population-based prospective case-control studies for breast cancer risk, but did not report categorical analysis [78]. However, for the purpose of comparison with other studies, we have included categorical analyses of these in this review. Additionally, as four studies (three reports) [31, 84, 85] used the lowest, rather than the highest methylation category as the reference category, inverse ORs were calculated for these studies. Importantly, use of different methylation 'split' (tertile, quartile, quintile or decile) for categorical analysis between reports is a potential confounding factor; however, this cannot be easily corrected without obtaining the raw data for each study. A meta-analysis of many of the relevant studies was recently reported by Woo et al [86], which did not include five relevant and recent reports [64, 84, 87-89]. Therefore, we report a revised meta-analysis including these studies (Figure 1). A thorough literary search was carried out to identify all case-control studies relationship between genomic DNA methylation in blood, and either cancer incidence or prevalence (Table 1). Included were studies investigating any cancer type and using any quantitative measurement of genome-wide DNA methylation, including 5meC measures and surrogate assays. Excluded were studies measuring methylation at single-locus sites, or measuring methylation is tissues other than blood. Summary estimates were weighted by sample size, using random-effects models due to significant inter-study heterogeneity. (Full details of search strategy and meta-analysis are provided in supplementary methods). The overall summary estimate OR for all using a random effects model was OR=1.4 (0.9-1.9), suggesting no overall association with
cancer risk. Summary ORs for studies using L1, Alu and Sat2 repetitive elements were not significantly associated with cancer risk. It appears that total 5meC genomic content is the most consistent association with cancer prevalence, as all five studies identified significant associations between hypomethylation and cancer prevalence in categorical analysis, four of which also found significant association at the mean level. Though the effect sizes are variable, with smaller studies showing the largest ORs, it appears evident that genomic 5meC content is reduced in blood of cancer patients, suggesting potential for biomarker development. Interestingly, these studies represent the earliest published reports investigating blood genomic methylation in relation to cancer risk, and no similar study has attempted to replicate these findings since. Given that most studies using surrogate assays are null for cancer risk, it may be worthwhile returning to 5meC measures in an effort to reproduce this association. Interestingly, the only study [84] investigating widespread genomic methylation, using the restriction-enzyme and bisulphite-pyrosequencing-based LUMA assay, identified a strong protective effect of genomic DNA hypomethylation on breast cancer risk. This inconsistent finding was likely influenced by the use of restriction enzymes, which measure methylation mainly at CGIs, which tend to be unmethylated, meaning that the only detectable change would be an increase in methylation [84].

There is no overall association between methylation at L1 and cancer risk according to this analysis. Whereas early studies suggested an association of blood L1 hypomethylation with cancer, later studies have failed to reproduce this. Furthermore, the direction of effect on cancer prevalence of L1 hypomethylation was inconsistent between significant studies [85], suggesting that this association may have occurred due to chance. One consideration is that heterogeneity between cancer types may affect the overall findings from this meta-analysis. However, in three breast cancer studies there is no association of L1 methylation with either cancer incidence [64] or prevalence [73, 84]. Of five prospective studies, only one [79] identified an association between low L1 methylation and cancer incidence, however, small sample size, combining of multiple cancer types, and selection of a very high-risk population
(elderly males) may be confounding factors for this study. Four more recent, larger studies have failed to identify any such association in gastric [90], breast [64] or hepatocellular [89] cancer, suggesting that L1 methylation is not a cancer risk factor.

Factors affecting epigenetic epidemiological studies

Study design

Many retrospective studies investigating blood DNA methylation in cancer patients inappropriately use the term ‘risk’ to describe associations between methylation and disease. Association with disease risk of blood DNA methylation variability can only be determined using prospective studies, with blood samples collected several years prior to disease development, due to the possibility of ‘reverse causality’, i.e. the alteration of blood DNA methylation by presence of active disease [15, 68, 89]. Whereas the effects of active cancer on blood DNA methylation are unknown, proliferation of lymphocytes, or depletion of required substrates or enzymes may affect blood DNA methylation [91]. Additionally, blood methylation may be affected by the presence of circulating tumour cells [92], though the contribution to overall blood methylation of this small fraction of tumour cells may not be detectable. An important factor for prospective studies is latency, i.e. the duration between sample collection and cancer diagnosis [15], as methylation may be affected by as yet undetected cancer in samples collected shortly before diagnosis. Furthermore, the degree to which latency affects methylation-risk association may help determine whether methylation variability confers long-term or transient risk, or whether methylation is likely to be an early tumorigenic event. Retrospective studies are useful for investigating the relationship between DNA methylation variability and cancer prevalence, and may be used for development of diagnostic cancer biomarkers [16, 68]. A potential confounding factor in
some retrospective studies is the possible effect of cancer treatment, as several commonly
used neo-adjuvant chemotherapeutic drugs are known affect inhibit activity of the folate-
mediated one-carbon metabolism pathway [93, 94]. Whereas this treatment-induced
inhibition of DNA methylation could potentially explain the occurrence of genome-wide
hypomethylation in post-treatment blood samples, four studies identifying significant
associations between genomic hypomethylation and cancer prevalence, including one L1
study [87], and three 5meC studies [31, 72, 73] used pre-treatment sample only.
Nonetheless, the exposure of cases, but not controls to cancer treatment may represent a
confounding factor. Sample size is an important factor for case-control studies, and a
confounding factor when comparing different population-based case control studies. Studies
of small sample size (n<40 cases) [72, 79], may have little power to detect subtle
methylation variability, perhaps leading to a bias towards non-significance, or to chance
detection [16]. Most of the studies relevant to this report were of small sample size.
Conversely, larger sample sizes may lead to detection of statistically significant results with a
very small overall effect size [95], which can be misleading. The largest study to date
included 1000 cases/control pairs, at alpha=0.05 this study would have 80% power to detect
a difference of 0.12 standard deviations. For the L1 assay, with a standard deviation of 1.8%
across the population, this would equal a difference in means of 0.06%, which is well below
the technical variability in this assay (0.5-5% [78, 88, 89, 96-99]) and may be biologically
meaningless.

Sample selection

Various patient characteristics are potential modifiers of both cancer risk and DNA
methylation. Such factors may represent intermediate factors in the link between DNA
methylation and disease risk, but may also represent potential confounding factors within the
study. For example, where DNA methylation is affected by smoking, an association between
DNA methylation and cancer risk may be apparent due to uneven numbers of smokers in the case and control groups. In order to reduce the effect of such potential confounders, case-control pairs should be matched on potential confounding patient characteristics, as well as technical factors such as experimental batch [16]. Furthermore, statistical adjustment for these factors should be applied during analysis. Age is a strong modifier of DNA methylation for some genes [12, 100], and is the biggest risk factor for most cancers, therefore, adjustment for age is essential [12, 100]. L1 methylation is significantly lower in females than males [80, 101, 102], probably due to differences in the X and Y chromosomes [103], therefore, population studies investigating genome-wide DNA methylation must be stratified by gender [79, 80, 101, 104]. Whereas gender may represent a confounding factor for some of the included studies [72, 85], statistical adjustment may have helped to reduce this bias. Both L1 methylation and cancer risk are modified by ethnicity and environmental carcinogen exposures [55, 80, 99, 105]. Hospital-based studies [106] may include controls confounded by conditions unrelated to the disease under investigation. Ideally, controls samples should represent healthy individuals, matched on all potential confounding factors, both technical and biological. Case-control studies nested within prospective cohorts remain the gold standard for this type of analysis [16, 107].

Assay measurement

Repetitive element surrogate assays provide a practical and cost-effective indicator of genomic hypomethylation in tumour tissues and cell lines, where L1 methylation is highly variable [108]. However, in blood, L1 methylation displays little variability [109], and it remains unclear how sensitive such assays are to detect subtle DNA methylation variability. Methylation of TE, particularly L1, is often reported as genome-wide or 'global' DNA methylation. However, L1 pyrosequencing measures methylation at only three to four CpG sites within a pool of L1 elements based on a consensus sequence [44], and is, therefore,
not representative of genome-wide methylation [14]. The detection of genome-wide hypomethylation by repetitive-element assays in tumour DNA is likely due to the occurrence of a proportion of repetitive elements within hypomethylated domains, rather than to a specific enrichment at repetitive elements [17, 24]. Key to the utility of repetitive element methylation as a surrogate for genome-wide methylation is the reported correlation between 5meC measured by HPLC, and methylation of L1 and Alu measured by MethylLight, a methylation-sensitive PCR-based system, in a panel of cell lines and tissues. However, these correlations were moderate (r=0.66 and r=0.70 for L1 and Alu respectively), and the authors recommended using a composite measure of two repetitive elements, ALU-M2 and Sat2-M1(r=0.85), rather than individual elements, as a surrogate for genome-wide methylation. The only study to measure the correlation between L1 pyrosequencing and 5meC in blood, or indeed any tissue, failed to find any correlation (r=-0.204), though the sample size was small (n=27) [73]. Whereas L1 methylation measured by pyrosequencing is reduced in cell lines exposed to DNA methylation inhibitors [44], and in many tumours, it is unclear whether L1 methylation reflects subtle genomic methylation variability in blood. This discrepancy may explain the stronger association with cancer risk of blood 5meC than blood L1 methylation.

Two prospective cancer risk studies detected blood hypomethylation at Alu [97] and Sat2 [89], but not at L1 in the same samples, and methylation of L1 and Alu are not correlated [99]. This is inconsistent with these assays detecting “genome-wide methylation”, but suggests that hypomethylation may be restricted to specific genomic sequences. Widely apparent is the greater variability of repetitive element methylation between different studies than between cases and controls within individual studies [14]. L1 methylation varies with ethnicity [55, 105], likely due to L1 genetic variability [14, 90]. L1 sequence heterogeneity may also cause cellular and allelic heterogeneity in L1 methylation [110], and may lead to underestimation of methylation levels [77]. L1 pyrosequencing however, displays little technical variation [108], and it will be important to determine whether measures of 5meC are as technically reproducible. L1 elements are differentially methylated at different
genomic loci [50, 103, 111], and at different CpG loci within the L1 consensus sequence [98], therefore assays preferentially amplifying different L1 elements or measuring methylation at different CpG sites may produce inconsistent results.

Finally, methylation variability at individual ‘functional’ repetitive elements may provide greater biomarker potential than methylation of ‘global’ repetitive elements, as hypomethylation of an aberrantly transcribed L1 within the MET oncogene, but not global L1 methylation, was detected in normal bladder in patients with bladder cancer [50].

Statistical Analysis

Categorical analysis of DNA methylation using odds ratio is the most frequently used measure of disease risk for case control epigenetic epidemiology studies [16]. Furthermore, categorical analysis is useful for investigating potentially non-linear relationships between methylation and disease risk and retains more power to detect significant differences compared with linear regression. However, different approaches may introduce bias, for instance a prominent inconsistency between studies included in this report is the use of different levels of methylation categories, including tertile, quartile, quintiles and deciles. When a pre-specified analysis plan and statistical power should have been described and used for the study. The narrow ranges of DNA methylation reported, especially at L1, would mean that the difference in percentage methylation between categories would be far below the detection sensitivity of the assay, for instance the technical variation for pyrosequencing, one of the most quantitative assays currently available, is around 2-3% [64]. In our study we observed an intraclass correlation coefficient (ICC) for LINE1 in blinded duplicate samples of 0 (95% CI 0-0.61) which suggests higher within individual variability than between individual variability. A major problem with developing small methylation differences as biomarkers is the overlap between methylation of cases and controls or lack of specificity [14].
Publication factors

According to the STROBE-ME (Strengthening the reporting of observational studies in epidemiology-molecular epidemiology) guidelines, all basic statistics of a biomarker measure distribution (mean, median, range and variance) and details of all other analyses should be reported in studies investigating molecular biomarkers for disease risk [112]. However, many reports do not include these, making direct comparison between studies difficult. Publication bias frequently affects meta-analyses of observational studies [113], and there is some evidence that it affects the current analysis.

All of the included reports showing significant association between L1 methylation and cancer included investigations of L1 methylation only, however all reports showing negative results for L1 also included studies showing significant associations between another methylation marker and cancer incidence/prevalence, consistent with a bias towards publication of significant associations. There appears to be a trend towards lower effect size with later publication date, as OR is significantly correlated with publication year among L1 methylation studies (Spearman r=0.49, p=0.05). This appears to be independent of sample size, as the correlations between OR and sample size (r=-0.19), and sample size and year of publication (r=0.2), are not significant. A funnel plot including L1 studies only (Supplementary Figure 1) does not appear to show asymmetry or publication bias, however the ability of this plot to indicate publication bias may be limited by the small number of studies (n=14), and variable direction of effect in studies showing significant associations.

A reporting bias exists whereby categorical analysis of methylation-risk associations tends only to be reported if it reveals statistically significant results, and if the authors deem categorical analysis appropriate. The inclusion of categorical analysis of our previously reported L1 studies represents an attempt to address this bias; however, this could only be estimated for other studies that reported no association with cancer risk (Supplementary Figure 2).
expected, inclusion of our studies, and other recent null studies, attenuated the significant association between L1 methylation and cancer risk reported by a recent meta-analysis [86]. Inclusion of all relevant reports is critical to the accuracy of meta-analyses, [95], therefore publication and reporting biases may affect our ability to identify the true effect size.

Other considerations for epigenetic risk studies

A concern for all studies investigating blood DNA methylation in relation to cancer in other tissues is the tissue-specificity of DNA methylation [91, 100]. Several studies have indicated that methylation of repetitive elements is tissue specific, most variable in tumour tissue, and not correlated between tumour and blood [81, 98, 114]. Consistently, evidence suggests that genomic hypomethylation in tumour and normal adjacent tissue of bladder and colon cancer was not detectable in blood [30, 50], suggesting that hypomethylation is restricted to the disease affected tissue. Furthermore, ~36% of L1 elements display tissue-specific expression [37]. Methylation variability between blood cell-types is an important consideration for retrospective studies, as differential methylation may reflect cellular proliferation. A recent microarray based study found that ~39% of CpG sites analysed were differentially methylated between leukocyte subgroups [91]. However, a recent report did not find any association between L1 methylation and blood cell count [115]. Another potential caveat for all cancer biomarker studies is tumour subtype heterogeneity, where the biomarker under investigation is a feature of only one disease subtype. For example, in tumour tissue, L1 hypomethylation frequently occurs in cancers displaying chromosome instability, but rarely in cancers with microsatellite instability [20, 36, 116]. Consistently, L1 hypomethylation is detectable in some colon tumour samples, but not others [20]. Therefore, if blood L1 hypomethylation is associated with colorectal cancer risk, it may predict risk of some, but not all cancers.
Conclusions and future studies

Epigenetic epidemiology holds great promise for identifying biomarkers of cancer risk and understanding cancer etiology. It is important to investigate methylation variability at individuals loci in the context of widespread methylation changes associated with cancer; therefore, understanding the contribution of genome-wide DNA methylation to cancer susceptibility and development is needed. The challenges of investigating a genome-wide epigenetic phenomenon in population studies are formidable, due to the requirements of high genome coverage, highly quantitative DNA methylation measurement and large sample size [16, 45]. We conclude from our meta-analysis that genome-wide DNA methylation, as measured by the surrogate L1 methylation assay, is not associated with cancer risk, demonstrated most appropriately by several prospective cohort studies. However, there appears evidence for an association of genomic hypomethylation, measured using more representative assays, such as HPLC or methyl-acceptance assays, with cancer prevalence, though this remains to be validated in prospective studies.

A thorough investigation of the relationship between genome-wide DNA methylation and cancer risk would require whole-genome bisulphite sequencing of large cohorts of prospectively collected blood samples, with careful sample selection, case-control matching, and deep sequencing to achieve the sensitivity required to detect subtle DNA methylation variability [16, 45]. Such an approach would provide information about genome-scale methylation changes as well as local methylation variability, thus providing candidate regions for risk biomarker development. An equivalent study using retrospectively collected blood samples may provide candidate diagnostic biomarkers. Whole-genome bisulphite sequencing is extremely costly, however, technological improvement will make this feasible in the coming years [45, 75]. Whereas current research using whole-genome sequencing of small numbers of samples cannot provide information about methylation variability across populations, further classification of the regions that consistently undergo methylation changes in cancer may provide candidate regions for a sequence capture approach in
population studies. New primer technology will enable the investigation of epigenetic regulation at specific ‘functional’ repetitive elements, such as those L1 elements capable of retrotransposition [47, 50], or driving ectopic expression of neighbouring genes [37, 50]. Furthermore, availability of the first microarrays specifically covering repetitive elements will provide high-throughput methods of investigating repetitive element methylation on a genome scale [117]. The most popular tool for current investigation of DNA methylation variability in population studies is the Illumina 450K methylation beadchip, which measures methylation at individual CpG sites across the entire human genome. These arrays are highly quantitative, high-throughput, and largely unbiased in terms of genomic coverage, making them suitable for biomarker discovery and epigenome-wide association studies [16]. Use of these assays will likely provide interesting evidence for the implications of DNA methylation variability in disease susceptibility and development in coming years.

Acknowledgements

This work was funded by Breast Cancer Campaign fellowship to JMF. KB and JMF are funded by Breast Cancer Campaign. The authors would like to thank Professor Robert Brown for his critical review of this manuscript.

References


103. Singer, H., et al., Methylation of L1Hs promoters is lower on the inactive X, has a tendency of being higher on autosomes in smaller genomes and shows inter-individual variability at some loci. Hum Mol Genet, 2012. 21(1): p. 219-35.


Figure Legends:

Figure 1. Meta Analysis of Odds Ratios reported in studies investigating Genome-wide DNA methylation in peripheral blood DNA for cancer risk. Test for heterogeneity showed highly significant heterogeneity across all studies (p<0.001), and specifically in the analysis of 5meC (p<0.001) and LINE1 (p<0.001), but not significant for Alu (p=0.121) or Sat2 (p=0.827). Squares represent the size of the study and are centred on the OR with whiskers representing the 95% CI's. Random Effects (RE) model was used for all summary analyses and the width of the diamond represents the confidence intervals. CRC=colorectal cancer, CRA=colorectal adenoma, BGS=Breakthrough Generations Study, EPIC=European Prospective Investigation into Cancer and Nutrition.

Table 1. Study details abstracted from all reports included in meta-analysis and all relevant studies excluded. Categorical split represents the level of categorisation of DNA methylation used in the study. Country/continent refers to the place of sample collection. Prospective/retrospective and pre/post chemotherapy indicate whether blood samples were collected specifies whether blood samples were collected before or after diagnosis and neo-adjuvant chemotherapy, respectively. 'Newly diagnosed' indicates whether retrospectively collected samples were collected from patients reported as 'newly' or 'recently' diagnosed. Cases and controls give a brief indication of the way in which case and control subjects were recruited and defined. 'Reported restricted ethnicity' indicates whether publications for each study reported exclusion of study participants based on race. ‘no’ indicates that no race exclusion was reported. Other values indicate the race to which the sample group was restricted.
Figure 1

5meC
- 2003 Pufulete (CRC)
- 2003 Pufulete (CRA)
- 2008 Lim
- 2008 Moore
- 2009 Choi

Alu
- 2010 Hou
- 2011 Zhu
- 2012 Gao
- 2012 Wu

LUMA
- 2010 Wilhelm
- 2010 Mirabello
- 2011 Zhu
- 2011 Zhu
- 2011 Rusiecki
- 2011 Liao
- 2011 Di
- 2012 Gao
- 2012 Xu
- 2012 Cash
- 2012 Brennan (BGS)
- 2012 Brennan (EPIC)
- 2012 Wu

LINE1
- 2007 Hsiung
- 2012 Wu
- 2012 Wu

LRE
- 2007 Hsiung
- 2012 Wu
- 2012 Wu

Sat2
- 2007 Hsiung
- 2012 Wu
- 2012 Wu

Summary (RE) 5meC
OR= 3.65 (1.20, 6.09)

Summary (RE) Alu
OR=1.31 (0.93, 1.68)

Summary (RE) LINE1
OR=1.24 (0.76, 1.72)

Summary (RE) Sat2
OR= 1.55 (0.99, 2.10)

Summary (RE) All
OR=1.39 (0.91, 1.87)
<table>
<thead>
<tr>
<th>Reference</th>
<th>Author</th>
<th>Publication date</th>
<th>Journal</th>
<th>Cancer type</th>
<th>measurment</th>
<th>assay</th>
<th>Categorical split</th>
<th>sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>72</td>
<td>Pufulete</td>
<td>5/2/2003</td>
<td>GASTROENTEROLOGY</td>
<td>colorectal</td>
<td>5meC</td>
<td>Methyl acceptance</td>
<td>tertiles</td>
<td>mixed</td>
</tr>
<tr>
<td>72</td>
<td>Pufulete</td>
<td>5/2/2003</td>
<td>GASTROENTEROLOGY</td>
<td>colorectal adenoma</td>
<td>5meC</td>
<td>Methyl acceptance</td>
<td>tertiles</td>
<td>mixed</td>
</tr>
<tr>
<td>31</td>
<td>Lim</td>
<td>1/1/2008</td>
<td>GASTROENTEROLOGY</td>
<td>colorectal</td>
<td>5meC</td>
<td>HPLC and MS</td>
<td>tertiles</td>
<td>female</td>
</tr>
<tr>
<td>106</td>
<td>Moore</td>
<td>4/9/2008</td>
<td>LANCET ONCOL</td>
<td>bladder</td>
<td>5meC</td>
<td>HPCE and hpaII*</td>
<td>quartiles</td>
<td>mixed</td>
</tr>
<tr>
<td>73</td>
<td>Choi</td>
<td>7/7/2009</td>
<td>CARCINOGENESIS</td>
<td>breast</td>
<td>5meC</td>
<td>Mass spectrometry</td>
<td>tertiles</td>
<td>female</td>
</tr>
<tr>
<td>99</td>
<td>Hou</td>
<td>10/15/2010</td>
<td>INT J CANCER</td>
<td>gastric</td>
<td>Alu</td>
<td>pyrosequencing</td>
<td>tertiles</td>
<td>female</td>
</tr>
<tr>
<td>79</td>
<td>Zhu</td>
<td>3/22/2011</td>
<td>CANCER CAUSE CONTROL</td>
<td>all (combined)</td>
<td>Alu</td>
<td>pyrosequencing</td>
<td>quartiles</td>
<td>male</td>
</tr>
<tr>
<td>97</td>
<td>Gao</td>
<td>1/31/2012</td>
<td>BRIT J CANCER</td>
<td>gastric</td>
<td>Alu</td>
<td>pyrosequencing</td>
<td>quartiles</td>
<td>female</td>
</tr>
<tr>
<td>48</td>
<td>Wu</td>
<td>6/7/2012</td>
<td>CARCINOGENESIS</td>
<td>breast</td>
<td>Alu</td>
<td>methylight</td>
<td>quartiles</td>
<td>female</td>
</tr>
<tr>
<td>74</td>
<td>Xu</td>
<td>2/27/2012</td>
<td>FASEB J</td>
<td>breast</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>deciles</td>
<td>mixed</td>
</tr>
<tr>
<td>101</td>
<td>Wilhelm</td>
<td>3/1/2010</td>
<td>CLIN CANCER RES</td>
<td>bladder</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>tertiles</td>
<td>male</td>
</tr>
<tr>
<td>88</td>
<td>Mirabello</td>
<td>3/17/2010</td>
<td>BMC MED GENET</td>
<td>testicular</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>tertiles</td>
<td>female</td>
</tr>
<tr>
<td>99</td>
<td>Hou</td>
<td>10/15/2010</td>
<td>INT J CANCER</td>
<td>gastric</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>tertiles</td>
<td>mixed</td>
</tr>
<tr>
<td>79</td>
<td>Zhu</td>
<td>3/22/2011</td>
<td>CANCER CAUSE CONTROL</td>
<td>all (combined)</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>quartiles</td>
<td>male</td>
</tr>
<tr>
<td>79</td>
<td>Zhu</td>
<td>3/22/2011</td>
<td>CANCER CAUSE CONTROL</td>
<td>all (combined)</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>quartiles</td>
<td>male</td>
</tr>
<tr>
<td>104</td>
<td>Rusiecki</td>
<td>4/3/2011</td>
<td>EPIGENOMICS</td>
<td>gastric</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>tertiles</td>
<td>mixed</td>
</tr>
<tr>
<td>85</td>
<td>Liao</td>
<td>6/11/2011</td>
<td>PLOS ONE</td>
<td>renal cell</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>quartiles</td>
<td>mixed</td>
</tr>
<tr>
<td>87</td>
<td>Di</td>
<td>10/12/2011</td>
<td>J ZHEJIANG UNIV-SC B</td>
<td>hepatocellular</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>quartiles</td>
<td>mixed</td>
</tr>
<tr>
<td>97</td>
<td>Gao</td>
<td>1/31/2012</td>
<td>BRIT J CANCER</td>
<td>gastric</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>quartiles</td>
<td>female</td>
</tr>
<tr>
<td>74</td>
<td>Xu</td>
<td>2/27/2012</td>
<td>FASEB J</td>
<td>breast</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>quintiles</td>
<td>female</td>
</tr>
<tr>
<td>80</td>
<td>Cash</td>
<td>3/1/2012</td>
<td>INT J CANCER</td>
<td>bladder</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>tertiles</td>
<td>mixed</td>
</tr>
<tr>
<td>15</td>
<td>Brennan</td>
<td>5/1/2012</td>
<td>CANCER RES</td>
<td>breast</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>quartiles</td>
<td>female</td>
</tr>
<tr>
<td>15</td>
<td>Brennan</td>
<td>5/1/2012</td>
<td>CANCER RES</td>
<td>breast</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>quartiles</td>
<td>female</td>
</tr>
<tr>
<td>89</td>
<td>Wu</td>
<td>5/12/2012</td>
<td>CARCINOGENESIS</td>
<td>hepatocellular</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>quartiles</td>
<td>mixed</td>
</tr>
<tr>
<td>47</td>
<td>Wu</td>
<td>6/7/2012</td>
<td>CARCINOGENESIS</td>
<td>breast</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>quartiles</td>
<td>female</td>
</tr>
<tr>
<td>76</td>
<td>Hsiung</td>
<td>1/16/2007</td>
<td>CANCER EPIDEM BIOMAR</td>
<td>head and neck</td>
<td>LRE1</td>
<td>COBRA*</td>
<td>tertiles</td>
<td>mixed</td>
</tr>
<tr>
<td>89</td>
<td>Wu</td>
<td>5/12/2012</td>
<td>CARCINOGENESIS</td>
<td>hepatocellular</td>
<td>Sat2</td>
<td>methylight</td>
<td>quartiles</td>
<td>mixed</td>
</tr>
<tr>
<td>47</td>
<td>Wu</td>
<td>6/7/2012</td>
<td>CARCINOGENESIS</td>
<td>breast</td>
<td>Sat2</td>
<td>methylight</td>
<td>quartiles</td>
<td>female</td>
</tr>
<tr>
<td>Study ID</td>
<td>Name</td>
<td>Date</td>
<td>Journal</td>
<td>Tissue</td>
<td>Method</td>
<td>Categorical Split</td>
<td>Race</td>
<td>Sex</td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>------------</td>
<td>------------------------</td>
<td>--------</td>
<td>--------------</td>
<td>-------------------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>26</td>
<td>Cho</td>
<td>12/30/2010</td>
<td>ANTICANCER RES</td>
<td>breast</td>
<td>Alu</td>
<td>MethyLight</td>
<td>na</td>
<td>female</td>
</tr>
<tr>
<td>69</td>
<td>Widschwendter</td>
<td>7/16/2008</td>
<td>PLOS ONE</td>
<td>breast</td>
<td>Alu</td>
<td>MethyLight</td>
<td>na</td>
<td>female</td>
</tr>
<tr>
<td>73</td>
<td>Choi</td>
<td>7/7/2009</td>
<td>CARCINOGENESIS</td>
<td>breast</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>na</td>
<td>female</td>
</tr>
<tr>
<td>26</td>
<td>Cho</td>
<td>12/30/2010</td>
<td>ANTICANCER RES</td>
<td>breast</td>
<td>LINE1</td>
<td>MethyLight</td>
<td>na</td>
<td>female</td>
</tr>
<tr>
<td>98</td>
<td>Piyathilake</td>
<td>5/27/2011</td>
<td>NUTRITION</td>
<td>CIN*</td>
<td>LINE1</td>
<td>pyrosequencing</td>
<td>na</td>
<td>female</td>
</tr>
<tr>
<td>83</td>
<td>Patchsung</td>
<td>5/15/2012</td>
<td>PLOS ONE</td>
<td>bladder</td>
<td>LINE1</td>
<td>COBRA</td>
<td>na</td>
<td>mixed</td>
</tr>
<tr>
<td>26</td>
<td>Cho</td>
<td>12/30/2010</td>
<td>ANTICANCER RES</td>
<td>breast</td>
<td>Sat2</td>
<td>MethyLight</td>
<td>na</td>
<td>female</td>
</tr>
</tbody>
</table>

**Table 1.** Study details abstracted from all reports included in meta-analysis and all relevant studies excluded. Categorical split represents the level of categorisation of DNA methylation. Country/continent refers to the place of sample collection. Prospective/retrospective and pre/post chemotherapy indicate whether blood samples were collected before or after diagnosis and neoadjuvant chemotherapy, respectively. 'Newly diagnosed' indicates whether retrospectively collected samples were collected from patients reported as 'newly' or 'recently' diagnosed. Cases and controls give a brief indication of the way in which case and control subjects were recruited and defined. 'Reported restricted ethnicity' indicates whether publications for each study reported exclusion of study participants based on race. 'No' indicates that no race exclusion was reported. Other values indicate the race to which the sample group was restricted.
<table>
<thead>
<tr>
<th>First Author</th>
<th>reported restricted ethnicity</th>
<th>Country/continent</th>
<th>n (case/ctrl)</th>
<th>Prospective/retrospective</th>
<th>newly diagnosed</th>
<th>pre/post chemotherapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pufulete</td>
<td>no</td>
<td>UK</td>
<td>28/76</td>
<td>retrospective</td>
<td>yes</td>
<td>pre-treatment</td>
</tr>
<tr>
<td>Pufulete</td>
<td>no</td>
<td>UK</td>
<td>35/76</td>
<td>retrospective</td>
<td>yes</td>
<td>pre-treatment</td>
</tr>
<tr>
<td>Lim</td>
<td>no</td>
<td>U.S.</td>
<td>115/115</td>
<td>retrospective</td>
<td>prospective</td>
<td>pre-treatment</td>
</tr>
<tr>
<td>Moore</td>
<td>no</td>
<td>Spain</td>
<td>775/397</td>
<td>retrospective</td>
<td>yes</td>
<td>unreported</td>
</tr>
<tr>
<td>Choi</td>
<td>no</td>
<td>U.S.</td>
<td>176/173</td>
<td>retrospective</td>
<td>yes</td>
<td>pre-treatment</td>
</tr>
<tr>
<td>Hou</td>
<td>caucasian</td>
<td>Poland</td>
<td>302/421</td>
<td>retrospective</td>
<td>yes</td>
<td>mixed</td>
</tr>
<tr>
<td>Zhu</td>
<td>no</td>
<td>Boston</td>
<td>205/486</td>
<td>retrospective</td>
<td>no</td>
<td>post-treatment</td>
</tr>
<tr>
<td>Gao</td>
<td>no</td>
<td>Shanghai</td>
<td>192/381</td>
<td>prospective</td>
<td>prospective</td>
<td>pre-treatment</td>
</tr>
<tr>
<td>Wu</td>
<td>no</td>
<td>US/Australia</td>
<td>266/334</td>
<td>retrospective</td>
<td>no</td>
<td>post-treatment</td>
</tr>
<tr>
<td>Xu</td>
<td>no</td>
<td>US</td>
<td>1055/1101</td>
<td>retrospective</td>
<td>yes</td>
<td>post-treatment</td>
</tr>
<tr>
<td>Wilhelm</td>
<td>no</td>
<td>U.S.</td>
<td>285/465</td>
<td>retrospective</td>
<td>yes</td>
<td>unreported</td>
</tr>
<tr>
<td>Mirabello</td>
<td>no</td>
<td>U.S.</td>
<td>114/151</td>
<td>retrospective</td>
<td>no</td>
<td>post-treatment</td>
</tr>
<tr>
<td>Hou</td>
<td>caucasian</td>
<td>Poland</td>
<td>300/419</td>
<td>retrospective</td>
<td>yes</td>
<td>mixed</td>
</tr>
<tr>
<td>Zhu</td>
<td>no</td>
<td>Boston</td>
<td>30/487</td>
<td>prospective</td>
<td>prospective</td>
<td>pre-treatment</td>
</tr>
<tr>
<td>Zhu</td>
<td>no</td>
<td>Boston</td>
<td>204/478</td>
<td>retrospective</td>
<td>no</td>
<td>post-treatment</td>
</tr>
<tr>
<td>Rusiecki</td>
<td>no</td>
<td>Oman</td>
<td>110/104</td>
<td>retrospective</td>
<td>yes</td>
<td>unreported</td>
</tr>
<tr>
<td>Liao</td>
<td>caucasian</td>
<td>Europe</td>
<td>328/654</td>
<td>retrospective</td>
<td>yes</td>
<td>unreported</td>
</tr>
<tr>
<td>Di</td>
<td>no</td>
<td>China</td>
<td>315/356</td>
<td>retrospective</td>
<td>yes</td>
<td>pre-treatment</td>
</tr>
<tr>
<td>Gao</td>
<td>no</td>
<td>Shanghai</td>
<td>192/382</td>
<td>prospective</td>
<td>prospective</td>
<td>pre-treatment</td>
</tr>
<tr>
<td>Xu</td>
<td>no</td>
<td>US</td>
<td>1064/1100</td>
<td>retrospective</td>
<td>yes</td>
<td>post-treatment</td>
</tr>
<tr>
<td>Cash</td>
<td>Han Chinese</td>
<td>China</td>
<td>510/528</td>
<td>retrospective</td>
<td>no</td>
<td>post-treatment</td>
</tr>
<tr>
<td>Brennan</td>
<td>white</td>
<td>UK</td>
<td>252/255</td>
<td>prospective</td>
<td>prospective</td>
<td>pre-treatment</td>
</tr>
<tr>
<td>Brennan</td>
<td>no</td>
<td>Europe</td>
<td>263/257</td>
<td>prospective</td>
<td>prospective</td>
<td>pre-treatment</td>
</tr>
<tr>
<td>Wu</td>
<td>no</td>
<td>Taiwan</td>
<td>305/1254</td>
<td>prospective</td>
<td>yes</td>
<td>pre-treatment</td>
</tr>
<tr>
<td>Wu</td>
<td>no</td>
<td>US/Australia</td>
<td>279/340</td>
<td>retrospective</td>
<td>no</td>
<td>post-treatment</td>
</tr>
<tr>
<td>Hsiung</td>
<td>no</td>
<td>Boston</td>
<td>278/526</td>
<td>retrospective</td>
<td>yes</td>
<td>post-treatment</td>
</tr>
<tr>
<td>Wu</td>
<td>no</td>
<td>Taiwan</td>
<td>305/1254</td>
<td>prospective</td>
<td>yes</td>
<td>pre-treatment</td>
</tr>
<tr>
<td>Wu</td>
<td>no</td>
<td>US/Australia</td>
<td>266/333</td>
<td>retrospective</td>
<td>no</td>
<td>post-treatment</td>
</tr>
<tr>
<td>Name</td>
<td>Country</td>
<td>Sample Size</td>
<td>Study Design</td>
<td>Treatment</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
<td>-------------</td>
<td>---------------</td>
<td>-----------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>Cho</td>
<td>Turkey</td>
<td>40/40</td>
<td>Retrospective</td>
<td>Yes</td>
<td>Pre-treatment</td>
<td></td>
</tr>
<tr>
<td>Widschwender</td>
<td>Germany</td>
<td>169/180</td>
<td>Retrospective</td>
<td>Prof.</td>
<td>Pre-treatment</td>
<td></td>
</tr>
<tr>
<td>Choi</td>
<td>U.S.</td>
<td>19/18</td>
<td>Retrospective</td>
<td>Yes</td>
<td>Pre-treatment</td>
<td></td>
</tr>
<tr>
<td>Choi</td>
<td>Turkey</td>
<td>40/40</td>
<td>Retrospective</td>
<td>Yes</td>
<td>Pre-treatment</td>
<td></td>
</tr>
<tr>
<td>Piyathilake</td>
<td>U.S.</td>
<td>103/273</td>
<td>Prospective</td>
<td>Pros.</td>
<td>Pre-treatment</td>
<td></td>
</tr>
<tr>
<td>Patchung</td>
<td>Thailand</td>
<td>61/45</td>
<td>Retrospective</td>
<td>Yes</td>
<td>Unreported</td>
<td></td>
</tr>
<tr>
<td>Pobsook</td>
<td>Thailand</td>
<td>34/43</td>
<td>Retrospective</td>
<td>Yes</td>
<td>Pre-treatment</td>
<td></td>
</tr>
<tr>
<td>Cho</td>
<td>Turkey</td>
<td>40/40</td>
<td>Retrospective</td>
<td>Yes</td>
<td>Pre-treatment</td>
<td></td>
</tr>
<tr>
<td>Cases (description)</td>
<td>Controls (description)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------------------</td>
<td>------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cancer screening</td>
<td>Unaffected on screening</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cancer screening</td>
<td>Unaffected on screening</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cancer screening</td>
<td>Unaffected on screening</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hospital-based</td>
<td>hospital-based</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hospital-based</td>
<td>patients (unrelated conditions)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hospital-based</td>
<td>cancer registry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>baseline screening of prospective population cohort</td>
<td>unaffected at baseline and follow up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incident cases from prospective cohort</td>
<td>unaffected on follow up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high-risk families</td>
<td>unaffected sisters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cancer registry, first primary breast cancer</td>
<td>randomly selected (digit dialing)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cancer registry (hospitals)</td>
<td>From previous study</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high-risk families</td>
<td>unaffected relatives</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cancer registry (hospitals)</td>
<td>randomly selected (registry)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incident cases from prospective elderly male cohort</td>
<td>unaffected at case-selection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>baseline screening of prospective population cohort</td>
<td>unaffected at case-selection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hospital-based</td>
<td>donors/government workers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hospital-based</td>
<td>patients (unrelated conditions)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hospital-based</td>
<td>unaffected at case-selection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incident cases from prospective female cohort</td>
<td>unaffected on follow up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cancer registry, first primary breast cancer</td>
<td>randomly selected (digit dialing)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cancer registry</td>
<td>recruited from registry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incident cases from prospective female cohort</td>
<td>unaffected on follow up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incident cases from prospective female cohort</td>
<td>unaffected on follow up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high-risk families</td>
<td>unaffected on follow up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hospital-based</td>
<td>unaffected sisters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incident cases from prospective cohort</td>
<td>recruited from registry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high-risk families</td>
<td>unaffected sisters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hospital-based</td>
<td>healthy volunteers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>-------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hospital-based</td>
<td>unaffected on routine health check</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hospital-based</td>
<td>Unaffected relatives/visitors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hospital-based</td>
<td>healthy volunteers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hospital-based</td>
<td>women with one or less CIN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hospital-based</td>
<td>healthy elderly volunteers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hospital-based</td>
<td>volunteers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hospital-based</td>
<td>healthy volunteers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Is there a link between genome-wide hypomethylation in blood and cancer risk?

Kevin Brennan and James M. Flanagan


Updated version
Access the most recent version of this article at:
doi:10.1158/1940-6207.CAPR-12-0316

Supplementary Material
Access the most recent supplemental material at:
http://cancerpreventionresearch.aacrjournals.org/content/suppl/2012/11/07/1940-6207.CAPR-12-0316.DC1

Author Manuscript
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.