Author’s response to reviews

Title: A cross-sectional study of global DNA methylation and risk of colorectal adenoma

Authors:

Will D King (kingw@queensu.ca)
Janet E Ashbury (ashburyj@queensu.ca)
Sherryl A Taylor (sherryl1@ualberta.ca)
M Yat Tse (yat.tse@queensu.ca)
Stephen C Pang (pangsc@queensu.ca)
Jacob A Louw (louwi@hdh.kari.net)
Stephen J Vanner (vanners@hdh.kari.net)

Version: 2 Date: 26 May 2014

Author’s response to reviews: see over
May 23, 2014

Dafne Solera, Executive Editor  
BMC Cancer  
BioMed Central  
236 Gray’s Inn Road  
London WC1X 8HB  
United Kingdom

Dear Dr. Solera,

We thank the reviewers for their thoughtful and thorough review of this manuscript ‘A cross-sectional study of global DNA methylation and risk of colorectal adenoma’. In response to these reviews we have revised the manuscript. We think that this has resulted in an improvement – in particular with regard to the overall objective and message of the manuscript.

A point-by-point response to each of the reviewer comments and description of resulting changes to the manuscript are provided below. (Reviewer comments are highlighted in bold.)

______________________________

**Reviewer #1**

1a) Authors state throughout the article that there is an association between lower levels of LINE-1 DNA methylation in normal colon tissue and increased adenoma risk. However, according to the results they show, this could be considered as an overstatement. Despite authors find a significant relationship between the lowest methylation levels quartile and adenoma risk in men and both sexes together (not in women), the significance is too marginal to be considered clinically relevant.

We feel that this comment and several that follow from Reviewer #1 stem from our lack of clarity in terms of the objective and contribution of this research. In response to this comment we have clarified throughout the text that the objective of this research was to contribute to a better understanding of colorectal cancer etiology and not to focus on a clinical or diagnostic marker.

**Revisions:**

Abstract – Introduction – a minor modification to the last sentence to remove the words “in a screening population” which may have implied that the objective was a diagnostic tool. Rather, this is the study population and it is identified in the methods.
Abstract – Conclusion section. We removed statement on implications for screening and intervention programs and limited the conclusion to findings supporting a role for global DNA methylation in colon mucosa as an early event in colorectal etiology.

Introduction – 1st paragraph. Reference to the identification of asymptomatic individuals at increased risk of colorectal cancer has been removed. The focus of this introduction is solely on the potential etiologic role of global DNA methylation in tissue and blood on CRC.

Discussion section, last paragraph – the implication that findings support global hypomethylation as a potential biomarker for identifying individuals at higher risk of developing colorectal cancer who could be targeted for screening and intervention programs has been removed. This revised conclusion focuses on the contribution of findings to providing some support for a relationship between LINE-1 DNA methylation in colon mucosal tissue and adenoma risk

1b) Authors categorize methylation levels in quartiles but show no evident reason to do that. Due to the marginal significance levels, it could be possible that other categorization of the variable may produce non-significant results. In order to avoid the need to define arbitrary groups, authors should consider using methylation levels as a quantitative variable in their logistic regression model. Moreover, this could also increase their statistical power.

We agree with the reviewer that given the marginal significance levels of the main findings that other categorizations could result in findings which are not statistically significant at 0.05. We chose to create a categorical representation in order to obtain effect measures which were straightforward to present and interpret and in order to avoid the necessary assumption of a (log) linear relationship between a continuous measure of methylation level and adenoma risk. That said, we agree that an analysis of a continuous representation of methylation would be informative (see revisions below).

We created a categorization based on quartiles for a number of reasons:
1. Basing category boundaries on a percentile distribution ensures adequate sample size within categories.
2. Quartiles are commonly used in the literature and appropriate given our total sample size (in comparison to quintiles).
3. A percentile approach avoids “gerrymandered” categories which may maximize effects and over fit the data.

Revisions: We have considered the reviewer suggestion of using a quantitative (continuous) representation of methylation level. We agree that this representation adds to a balanced presentation of effects and have added this to the main analysis tables. This change is reflected in the description of study methods, in tables 2 and 3, and in the text describing the results.

2) Even if authors could show a stronger and more reliable association between low methylation levels and adenoma risk, they should prove that this global methylation marker is useful in terms of sensitivity and specificity compared to currently used early-diagnosis methods. Authors should also show ROC curves and AUC scores so that they could be compared to FOBT tests.

As indicated in response to comment #1, we feel that this comment stems from our lack of clarity in presentation of study objectives and have revised several sections of the text to address this.
We acknowledge that this is not a marker that would be useful in terms of early diagnosis. We feel that the contribution of this research is in terms of understanding CRC etiology. Therefore, there is no utility in analyzing in terms of sensitivity and specificity or ROC curves.

Revisions: see Comment #1 above

3) Authors find moderate correlation (r=0.36) between colon tissue methylation and leukocyte methylation. This is completely expected, since it is well known that methylation status is highly cell type dependent. In any case, the lack of association between leukocyte methylation status and adenoma presence raises concerns in terms of the usefulness of this marker as a non-invasive screening tool.

Again, we think that this comment stems from our lack of clarity in presentation of study objectives and we have revised several sections of the text to address this. We have removed implicit and explicit references to the utility of DNA methylation as utilized in this study as a screening tool.

In the etiologic context, the lack of a relationship between blood leukocytes measures and adenoma risk is relevant for epidemiologic studies which may focus on methylation measures and adenoma risk, or on the relationship between environment and lifestyle measures and global DNA methylation level. Reviewwe #2 (comments #8 and #9) also comment on this point.

Revisions: We have commented on this in the discussion

- Minor Essential Revisions

4) Authors should state whether they have followed the STARD guidelines (STAndards for the Reporting of Diagnostic accuracy studies), which is a requirement for any diagnostic paper published by BMC Cancer.

This paper is not aimed at creating a diagnostic marker and we have made revisions in response to previous items raised above in order to clarify this issue. As a result we have not addressed STARD - as this is not a diagnostic paper.

5) Authors should show p-values in their result tables.

Revisions: We have added p-values to the results tables (table 2 and 3).

Reviewer #2

1. Although very well written, the introduction is too long and could be shortened by at least one page.

Revisions: We have shortened the introduction in response to reviewer #1 (Comment #1) by removing background related to the identification of asymptomatic individuals at increased risk of colorectal cancer. In addition we have removed the paragraph in the introduction which summarized evidence from tumour cell lines and mouse models.

2. As the occurrence or prevalence of adenomas in this study population is not a rare disease, the OR may not provide the best estimate and may overestimate the association. Previous studies
showed that for this reason it is preferable to estimate a prevalence risk ratio (see for instance Behrens T et al. Different methods to calculate effect estimates in cross-sectional studies. A comparison between prevalence odds ratio and prevalence ratio Methods Inf Med. 2004;43(5):505-9.), by for instance using a COX-model with time=0.

We agree that the Odds Ratio as applied to this data would overestimate the Prevalence Ratio and we have been careful in the text not to interpret the odds ratios as prevalence ratios. We selected the Odds Ratio and logistic regression over the alternatives of Prevalence Ratios via log-binomial or Poisson regression for a number of reasons.
1. The odds ratio provides a robust measure of direction, strength and statistical significance of the relationships of interest.
2. We do not have the ability in this dataset to estimate prevalence ratios as our study population is not representative of a population or patient cohort. Specifically, the subjects include those with adenoma and those with normal colonoscopies. Patients diagnosed with serrated adenomas or only hyperplastic polyps were not included. As a result, a model estimating Prevalence Ratio would be problematic.
3. The alternative models are potentially problematic with respect to statistical significance tests – where log-binomial regression may underestimate p-values (smaller) and Poisson regression tends to result in overestimation of p-values (McNutt L et al. Estimating the Relative Risk in Cohort Studies and Clinical Trials of Common Outcomes, AJE 2003, 157(10) pp 940-943.)

Revisions: We have added a comment in the methods which addresses this issue and the interpretation of Odds Ratios in this study.

3a. Why do the authors think differences are observed between men and women? This should be discussed in the Discussion section. Were the analyses planned to be conducted for men and women separately before data-analysis or was this decided based on the findings? Are sex-specific distributions necessary?

There was a priori rationale for sex-specific analysis. Specifically there are sex-specific differences in the incidence of CRC and the incidence of adenomas. There are also observed sex specific difference in terms of risk factors for CRC and adenoma. Most importantly sex-specific differences in the relationship between DNA methylation and bladder cancer outcomes have been observed.

With respect to the need for sex-specific quartiles in the analysis, we observed higher methylation levels in men compared to women – consistent with the literature. We chose to create sex-specific quartiles (and sex-specific standardized continuous values) as a means of controlling for sex in the all subjects combined analysis and in order to have consistent categorization in the sex-specific analysis.

Revisions: A paragraph has been added to the discussion which addresses the rationale for sex-specific analysis and the interpretation of sex-specific results.

3.b. The numbers for women, especially in the adenoma group are relatively low. This could also explain the non-significant findings. Maybe it would have been better to use tertiles instead of quartiles.

We used quartiles in all analysis for consistency. We agree that sample size/statistical power could impact the sex-specific analysis.
Revisions: A comment on limited statistical power in sex-specific analysis was been added to the discussion.

4. In the genders combined analysis, would it be possible to stratify for number of adenomas observed and size of adenomas?

The sample size, in particular number of adenoma subjects, does not facilitate this analysis.

5. Is global hypomethylation indeed a useful marker of increased adenoma risk when it should be measured in the colon and the adenoma is already there?

We feel that this comment is consistent with the concerns of reviewer #1 with respect to investigating clinical markers.

Revisions: In response to this comment and those of reviewer #1 we have clarified throughout the text that the objective of this research is to contribute to the understanding of colorectal cancer etiology.

6. It is not clearly explained why two biopsies are used for the analyses. Is for those biopsies that are taken closer to the adenomas a stronger association expected? DNA methylation appears quite heterogeneous with a correlation coefficient of 0.66 between the two biopsies. What does that mean for the reliability of this potential biomarker? This should be discussed in more detail in the Discussion section.

Two biopsies are used in an attempt to represent the overall methylation level in descending colon mucosa. The correlation of 0.66 is a useful finding on its own – as it suggests some heterogeneity of methylation level within the colon and indicates that one measure is not sufficient.

Revisions: A sentence emphasizing the purpose of the two biopsies has been added to the methods section. A sentence commenting on the heterogeneity in the two samples was added to the discussion.

7. The response rate is relatively low (61%). Could this have implications for the results as those who participated may be more health conscious?

We feel that response rates do not likely impact the study results given that the study objectives are oriented towards understanding a biologic relationship postulated to be consistent irrespective of the population studied. Furthermore, given that the objective measures of methylation and adenoma are obtained after consent to participate, there is little opportunity to introduce systematic bias.

Revisions: This issue was addressed to some degree in the discussion and this paragraph has been revised to specifically address response rates.

8. The fact that no association between blood LINE-1 methylation and colorectal adenomas is observed is very important for large epidemiological studies in which it is not possible to do endoscopies. Given the wide confidence intervals was this study large enough to evaluate this?

We agree that the smaller sample size for this analysis could contribute to null results.

Revisions: A comment on the smaller sample size and reduced statistical power for the blood leukocyte analysis has been included in the discussion.
9. Given the fact that in etiological studies in which suspected environmental and lifestyle factors are assessed the participants usually do not undergo a colonoscopy, the question is whether this is indeed a very useful biomarker in large epidemiological studies.

We agree that this study points to the importance of colon tissue methylation over blood leukocyte methylation in future epidemiologic studies. This comment is similar to that raised by reviewer #1 (comment #3). A comment on the implication of this finding has been added to the discussion.

**Revisions:** See response to reviewer #1 (Comment #3)

10. I do agree that several potential confounding variables could be in the causal pathway and should not be adjusted for. However, why would age be different?

We agree with this interesting point. Age is a strong risk factor for colorectal cancer – it has the potential to act through different biologic mechanisms and as a result it is difficult to determine whether it should be controlled for in analysis (i.e. whether it is upstream of DNA methylation in a pathway or an independent risk factor). We presented crude and age-adjusted analysis as we felt that readers would prefer to see each analysis. Interestingly, age did not confound the relationship between DNA methylation and adenoma risk in this data.

11. Why is the LINE-1 methylation not assessed in the adenomas as well?

Previous research has demonstrated aberrant methylation patterns in adenoma and colorectal cancer and this was not a study objective.

11. Why was the normal appearing colon mucosa not histologically reviewed? I guess because no tissue was available anymore.

During the colonoscopy pinch biopsies were obtained from macroscopically healthy, normal appearing mucosa. The study did not have the opportunity to histologically review these pinch biopsies.

**Minor essential revisions**

**M1. Abstract: include results blood LINE-1 and adenoma risk**

The null results between blood leukocyte LINE-1 DNA methylation and adenoma risk have been identified in the abstract.

**M2 Introduction and Discussion: How small were the previous small observational studies?**

The sample size of the two previous studies of this relationship have been identified in the text in the introduction and discussion.

**M3. Explain why this age-group is used.**

The target age range was aged 40 to 65. This age group was targeted to represent an etiologically relevant age period where DNA methylation patterns, which might be considered precursors to the appearance of
colorectal adenoma, could be investigated. Higher rates of epigenetic deregulation have been observed with increasing age, and this is a phenomena which would inhibit the utility of this study if older patients (>65 years of age) were included.

**M4. Could a sensitivity analysis be conducted by including only those with a positive family history?**

For the majority of patients (n=241) the indication for colonoscopy was a positive family history. Results are similar when the analysis is restricted to those with a family history indication for colonoscopy.

**Revisions**: A statement describing this result was added to the results section.

**M5. Could a figure be added on the correlation between DNA-methylation in mucosa versus blood? Or between biopsy 1 and 2?**

The relationship between measures of DNA methylation in blood leukocyte, buccal cells, and colon tissue samples is the focus of a recently accepted paper from this study. (Ashbury JA, et al. Biomarkers measured in buccal and blood leukocyte DNA as proxies for colon tissue global methylation. International Journal of Molecular Epidemiology and Genetics, *in press*). We have included a reference to this paper in the Discussion.

**M6. LINE-1 is a proxy for global hypomethylation. It is a proxy for genome-wide DNA methylation.**

We have utilized LINE-1 methylation as a proxy for global DNA methylation in this study. In this context global DNA methylation refers to the overall genome-wide content of methylated cytosines within CpG (cytosine-phosphate-guanine) sites. We realize that we had used the terms global methylation and genome-wide DNA methylation interchangeably in the text. In the literature the term genome-wide methylation is used in this context, but is also used in reference to measurement of gene-specific methylation across the genome.

**Revision**: We have consistently used the term global DNA methylation and have removed all occurrences of genome-wide methylation.

**M7. The number of words about the bisulfite conversion could be limited by adding a reference.**

**Revisions**: The text has been reduced and an appropriate reference added.

**M8. reference 4: correct the website**

**Revisions**: Corrected
We hope that our manuscript is now acceptable for publication in BMC Cancer. As reviewer #2 states, we think that this is ‘an article of importance in the field’ of epidemiology.

Thank you for your consideration. We look forward to hearing from you at your earliest convenience.

Yours sincerely,

Will King, PhD,
Department of Public Health Sciences,
Queen’s University, Kingston, ON, Canada
(613) 533-6000, ext. 74735
Fax: (613) 533-6686; Email: kingw@queensu.ca