



RESEARCH NOTE

REVISED A research note regarding "Variation in cancer risk among tissues can be explained by the number of stem cell divisions" [version 2; referees: 2 approved]

Maxime Tarabichi, Vincent Detours

IRIBHM, Université Libre de Bruxelles (ULB), Brussels, B-1070, Belgium

v2 First published: 22 Aug 2016, 5:2044 (doi: 10.12688/f1000research.9448.1)
 Latest published: 10 Nov 2016, 5:2044 (doi: 10.12688/f1000research.9448.2)

Abstract

Tomasetti and Vogelstein argued that 2/3 of human cancers are due to 'bad luck' and that "primary prevention measures [against cancer] are not likely to be very effective". We demonstrate that their calculations for hepatocellular carcinomas overlooked a major subset of these cancers proven to be preventable through vaccination. The problem, which is not limited to hepatocellular carcinoma, arises from the general reliance of their analysis on average incidences in the United States and the omission of incidences in specific risk groups.



This article is included in the **Preclinical Reproducibility and Robustness** channel.

Open Peer Review

Referee Status:

Invited Referees

1 2

REVISED

version 2

published
10 Nov 2016

version 1

published
22 Aug 2016



report



report

- 1 **Benjamin Roche**, Centre for Ecological and Evolutionary Research on Cancer France, **Cindy Gidoin**, Université Montpellier France
- 2 **Maddalena Fratelli**, IRCCS Mario Negri Institute for Pharmacological Research Italy, **Marco Bolis**, IRCCS Mario Negri Institute for Pharmacological Research Italy

Discuss this article

Comments (0)

Corresponding author: Vincent Detours (vdetours@ulb.ac.be)

How to cite this article: Tarabichi M and Detours V. **A research note regarding "Variation in cancer risk among tissues can be explained by the number of stem cell divisions" [version 2; referees: 2 approved]** *F1000Research* 2016, 5:2044 (doi: [10.12688/f1000research.9448.2](https://doi.org/10.12688/f1000research.9448.2))

Copyright: © 2016 Tarabichi M and Detours V. This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](#) (CC0 1.0 Public domain dedication).

Grant information: MT was supported by an FNRS/FRIA grant. VD was funded by FNRS, grant J009714F.

Competing interests: No competing interests were disclosed.

First published: 22 Aug 2016, 5:2044 (doi: [10.12688/f1000research.9448.1](https://doi.org/10.12688/f1000research.9448.1))

REVISED Amendments from Version 1

We thank the reviewers for their comments. We did not follow the suggestions for supplemental remarks because we wanted to maintain the specific focus of our research note. There have been many commentaries questioning the 'bad luck' theory that interested readers can easily find. We included references to the Boyd-Orr cohort.

See referee reports

Introduction

Tomasetti and Vogelstein¹ claimed that for tumors of relatively low incidence arising in organs undergoing many stem cell divisions “primary prevention measures are not likely to be very effective” because they arise mostly from random mutations fixed during stem cell division, independently of specific genetic or environmental factors. This conclusion—which has received much press coverage²⁻⁶—has important implications for public health and environmental research and policies. Here we re-interpret the results in light of additional data.

Results

The authors argued that 2/3 of the variation of cancer incidence among human organs could be explained by the total number of

lifetime stem cell divisions (lscd) which, according to them, drives the stochastic accumulation of random mutations. Yet, the incidence variation among organs is not informative about incidence variation among different risk groups. For example, worldwide cancer incidence variations and their association with regional risk factors are well documented⁷. But the study by Tomasetti and Vogelstein rests mostly on current average USA incidence statistics and is therefore blind to population-specific risk factors.

Tomasetti and Vogelstein did, however, consider risk-group specific incidences for a few cancers. For example, they calculated the excess risk score (ERS) for hepatocellular cancer (HCC) for the USA subpopulation infected by the hepatitis C virus (HCV) and the non HCV-infected subpopulation. The risk was 5.36 for HCV and -6.08 in non HCV cancers, which corresponded to the D-tumor (deterministic) and R-tumors (replicative) classes, respectively. This seems to support the validity of the ERS. But what would have been the classification of HCC if, as for most other cancers in the study, only the USA average incidence would have been taken into account? The ERS would be -5.65, well within the range of R-tumors, leading to the conclusion that HCC is a less preventable cancer (Figure 1). This would be a dangerous distraction from the fact that 10 to 33% of them, depending on world regions, are caused by HCV infections that are both preventable and treatable when responsible health policies are implemented. Furthermore, is the non-HCV HCC not preventable, as its classification suggests? Fifty

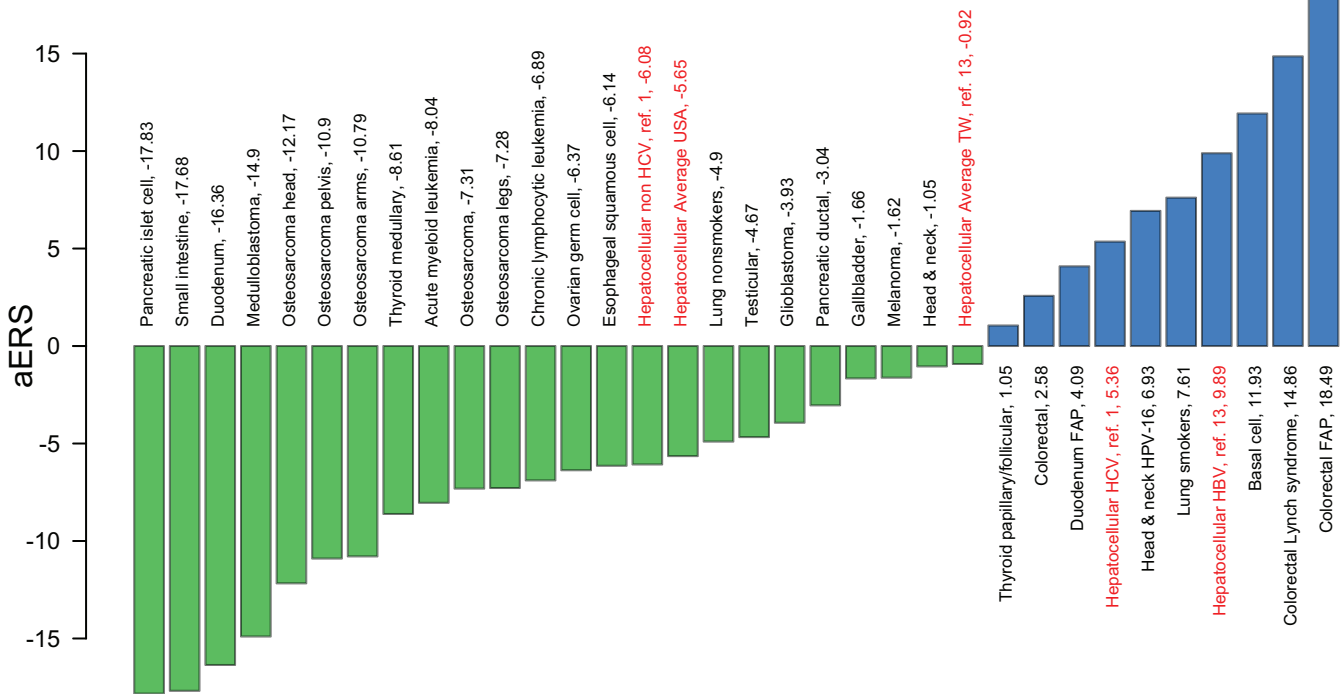


Figure 1. Effect of risk stratification of hepatocellular carcinoma incidence. This figure is an enhanced version of Figure 2 in ref. 1 showing the adjusted ERS (aERS=ERS + 18.49) for human cancers. 'Hepatocellular Average USA' and 'Hepatocellular Average TW' denotes the entire population of HCC patients, including both HCV and non HCV cases, in the USA and in Taiwan, respectively. Incidence was taken from the SEER database. 'Hepatocellular HBV' denotes HCC patients who are also HBV carriers. HCC lifetime risk for HBsAg-positive patients was taken from ref. 13.

nine percent of HCC cases in the developing world are associated with hepatitis virus B (HBV) infection⁷, which greatly increases the probability of developing the disease (Figure 1). Universal HBV vaccination has resulted in a 65–75% reduction of HCC incidence in 6–14 years children from Taiwan⁸. Other overlooked preventable risk factors for HCC include obesity, alcoholic cirrhosis, exposure to aflatoxin B and schistosomiasis. We focused on HCC due to space constraints, but similar arguments could be made for most cancers analyzed in ref. 1.

We also included in Figure 1 the ERS for the overall Taiwanese population. It is between the D- and R-tumors and higher than for the USA population. This is consistent with the fact that HCC is more preventable in Taiwan where HCV and HBV are more prevalent and supports, it seems, the potential usefulness of the ERS. Importantly, however, the ERS for all HCC rests on the same lscd estimate, thus the incidence data alone would produce the same ranking of the HCC groups [$ERS = \log_{10}(\text{lscd}) \times \log_{10}(\text{incidence})$]. On a more fundamental level, the ERS does not provide an absolute quantification of determinism because we do not know the baseline ERS for cancers occurring in the *proven* absence of any risk factor. Is this baseline universal or is it organ-specific? If the latter is correct, then the ERS will not be comparable among organs and will not be more informative than incidence data alone, as we have noted for HCC. If not, the ERS scale will be universal and the lscd will add information useful for the comparison of cancer determinism between organs. The modalities of DNA repair varies across the stem cell compartments of different organs⁹, suggesting an organ-specific baseline.

To our knowledge, a substantial variation of the lscd in the general population cannot be excluded. Hence the stratification problem encountered for incidence data may also arise because of lscd variation. The authors wrote that factors “such as those that affect height and weight” could play a role. These factors are corroborated by epidemiology^{10,11}. In addition, we consider highly plausible that tissue repair following chronic and possibly preventable damage may also significantly affect the lscd. Similarly, the relation between mutation rates and lscd is modulated by a range of factors, including DNA repair efficiency and activation of APOBEC DNA mutators¹². All of these were averaged out as were most known cancer risk factors.

In order to demonstrate the robustness of the correlation between the lscd and cancer lifetime risk, Tomasetti and Vogelstein varied randomly their lscd estimates over four orders of magnitude. We reproduced this calculation except that both lscd and incidences were varied randomly by two orders of magnitude (N=10,000). This calculation confirmed the robustness of the correlation (median $\rho=0.54$, 95% CI: 0.32–0.72; median $p=0.002$, 95% CI: 0.000005–0.08). We also collapsed to a single data point cancers sharing

the same lscd estimate to address statistical independence concerns. Again, the correlation remained strong ($\rho=0.67$, $p=0.0009$). The R script and the data to derive these results are given in supplementary material.

Conclusion

The remarkable relation between cancer incidence and lscd uncovered by Tomasetti and Vogelstein is statistically robust. The ERS is typically high for known deterministic cancers. But we demonstrated that a cancer with a low ERS can include a sizable fraction of preventable diseases. This proves that their classification scheme, in its current form, is not suitable to gauge the likely effectiveness of prevention measures and to direct funding for research on cancer etiology. Many more risk factors for cancers will likely be discovered in the future. Hence, cancers ascribed to ‘bad luck’ today due to lack of proper risk stratification may someday become explainable and, hopefully, preventable.

Dataset 1. R script and data to reproduce the analysis

<http://dx.doi.org/10.5256/f1000research.9448.d133564>

A description of each entry is provided in the readme file.

Methods

Data were retrieved online from Table S1 of the supplementary material of ref 1.

Additional data were retrieved from ref 13. Incidence for HCC with HBV was computed as the sum of incidences in women and men, divided by 2 (HCC with HBV incidence=0.17685%). Lscd for all HCC was set to 2.709×10^{11} , taken from ref 1.

Then data were analyzed in R v3.1.3¹⁴. Like in ref 1, both Spearman’s and Pearson’s correlation coefficients and p -values were computed. Here, we report values for Spearman’s correlations.

First, we reproduced Figure 2 of ref 1 after including data of HCC in Taiwan and with HBV. aERS were computed as described in ref 1, *i.e.* $aERS = \log_{10}(\text{lscd}) \times \log_{10}(\text{incidence}) + 18.49$.

Second, to assess the stability of the correlation between lscd and the incidence upon measurement errors, both lscd and incidence data were randomly multiplied or divided by 10, spanning two orders of magnitude. Incidences values were capped at 1. This was repeated 100 times for each variable, amounting to $100 \times 100 = 10,000$ pairs of randomly shifted lscd and incidences. From these 10,000 comparisons, distributions of correlation coefficients and p -values were obtained, from which confidence intervals were derived at percentiles 0.025 and 0.975.

Finally, we recomputed the correlation coefficient and p -value between *lscd* and incidence data after removing osteosarcomas and doublons, *i.e.* cancer originating from the same tissue type, thus having the same *lscd*. These doublons included *Colorectal adenocarcinoma with FAP*, *Colorectal adenocarcinoma with Lynch syndrome*, *Hepatocellular carcinoma with HCV*, *Head & neck squamous cell carcinoma with HPV-16*, and *Lung adenocarcinoma (smokers)*.

Data availability

F1000Research: Dataset 1. R script and data to reproduce the analysis, [10.5256/f1000research.9448.d133564](https://doi.org/10.5256/f1000research.9448.d133564)¹⁵

Author contributions

MT ran the analyses. MT and VD conceived the analyses and wrote the paper.

Competing interests

No competing interests were disclosed.

Grant information

MT was supported by an FNRS/FRIA grant. VD was funded by FNRS, grant J009714F.

References

- Tomasetti C, Vogelstein B: **Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions.** *Science*. 2015; **347**(6217): 78–81.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Boseley S: **Two-thirds of adult cancers largely “down to bad luck” rather than genes.** *The Guardian*. 2015.
[Reference Source](#)
- Knapton S: **Most cancers are caused by bad luck not genes or lifestyle, say scientists.** 2015; [Accessed: 15-Aug-2016].
[Reference Source](#)
- New Study: Two-Thirds of Cancer Cases Can Be Attributed to Bad Luck.** 2016; [Accessed: 15-Aug-2016].
[Reference Source](#)
- Dunham W: **Biological bad luck blamed in two-thirds of cancer cases, researchers say.** *Washington Post*. 2015; [Accessed: 15-Aug-2016].
[Reference Source](#)
- Grady D: **Cancer’s Random Assault.** *The New York Times*. 2015.
[Reference Source](#)
- American Cancer Society: **Global Cancer, Facts and Figures.** 2nd edition, Atlanta, 2011.
[Reference Source](#)
- Chang MH, Chen CJ, Lai MS, *et al.*: **Universal Hepatitis B Vaccination in Taiwan and the Incidence of Hepatocellular Carcinoma in Children. Taiwan Childhood Hepatoma Study Group.** *N Engl J Med*. 1997; **336**(26): 1855–1859.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Blanpain C, Mohrin M, Sotiropoulou PA, *et al.*: **DNA-damage response in tissue-specific and cancer stem cells.** *Cell Stem Cell*. 2011; **8**(1): 16–29.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Albanes D: **Height, early energy intake, and cancer. Evidence mounts for the relation of energy intake to adult malignancies.** *BMJ*. 1998; **317**(7169): 1331–1332.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Altman AJ, Schwartz AD: **Malignant diseases of infancy, childhood and adolescence.** *Major Probl Clin Pediatr*. 1978; **18**: 1–515.
[PubMed Abstract](#)
- Alexandrov LB, Nik-Zainal S, Wedge DC, *et al.*: **Signatures of mutational processes in human cancer.** *Nature*. 2013; **500**(7463): 415–421.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Huang YT, Jen CL, Yang HI, *et al.*: **Lifetime risk and sex difference of hepatocellular carcinoma among patients with chronic hepatitis B and C.** *J Clin Oncol*. 2011; **29**(27): 3643–3650.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- R Development Core Team: **R: A language and environment for statistical computing.** Vienna, Austria: R Foundation for Statistical Computing, 2009.
[Reference Source](#)
- Tarabichi M, Detours V: **Dataset 1 in: Comment on “Variation in cancer risk among tissues can be explained by the number of stem cell divisions”.** *F1000Research*. 2016.
[Data Source](#)

Open Peer Review

Current Referee Status:  

Version 1

Referee Report 25 October 2016

doi:10.5256/f1000research.10176.r16419



Maddalena Fratelli, Marco Bolis

Pharmacogenomics Unit, Department of Molecular Biochemistry and Pharmacology, IRCCS Mario Negri Institute for Pharmacological Research, Milano, Italy

This research note adds on the current discussion on the results of Tomasetti and Vogelstein, demonstrating that stochastic effects associated with stem cell division contribute in a substantial way to cancer incidence. As this conclusion has important and possibly harmful implications in the public health decisions, this deserves attention.

On the one hand, this note confirms robustness of the T&V analysis of the correlation between lscd and cancer risk. On the other hand, it points out that cancers predicted to be replicative may include a proportion of preventable cases. This is demonstrated by the higher ERS index for hepatocellular carcinoma in Taiwan, where HCV and HBV prevalence, and therefore its preventability, is higher than in the USA.

Minor points:

1. It would be interesting to show variation in ERS for other tumors such as, for example, esophageal and thyroid cancer. As discussed in Garattini & Tavani (2016), the incidence of these tumors show a great geographical variability. Even if it is not possible to associate these differences to preventable factors, this would add information on the global impact of geographical variation on ERS variability.
2. The discussion on lscd is also interesting. Perhaps the authors should make it explicit that more research is needed on the determination of the number of stem cell divisions and its dependence on factors such as different cell types in the same organ, damage, inflammation, individual height or weight and even ethnicity.
3. Finally, the authors should better explain why they deemed it important to demonstrate the robustness of T&V analysis, as they reproduced it with minor changes.

We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Referee Report 23 September 2016

doi:10.5256/f1000research.10176.r16205

✓ **Benjamin Roche¹, Cindy Gidoin²**

¹ Centre for Ecological and Evolutionary Research on Cancer, Montpellier, France

² Centre for Ecological and Evolutionary Research on Cancer, Université Montpellier, Montpellier Cedex, France

We think that this article could be an interesting addition to the current debate about the 'bad luck' theory. Briefly, this paper shows that the Adjusted Extra Risk Score (aERS) is dependent on the risk group that has been considered and that calculating this score at a country-scale level, without considering precisely all risk-groups, could give biased estimates. This suggests that prevention measures cannot be discarded (at least based only on this aERS), especially when a pathogen could be involved in cancer pathogenesis.

We just have few minor remarks:

- In the results section, it is said that there is no relation between cancer and height or weight. Some references, especially with the Boyd-Orr cohort, exist; ^{1 2}
- We think it could be also worth mentioning that many cancers are due to infections, which therefore calls for a prudent interpretation before withdrawing prevention strategies against some cancers.

References

1. Albanes D: Height, early energy intake, and cancer. Evidence mounts for the relation of energy intake to adult malignancies. *BMJ*. 1998; **317** (7169): 1331-2 [PubMed Abstract](#)
2. Altman AJ, Schwartz AD: Malignant diseases of infancy, childhood and adolescence. *Major Probl Clin Pediatr*. 1978; **18**: 1-515 [PubMed Abstract](#)

We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.
