The cancer « bad luck » hypothesis

Critical review of the ongoing debate and new perspectives

Joint work with H. Omichessan, G. Severi et al.

vittorio.perduca@parisdescartes.fr

Variation in cancer risk among tissues can be explained by the number of stem cell divisions

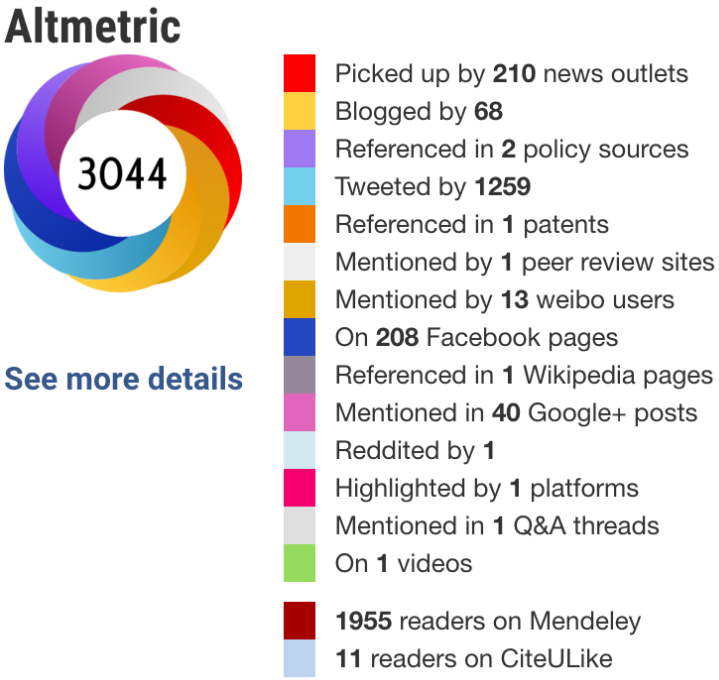
Cristian Tomasetti[1,*], **Bert Vogelstein**[2,*]

+ See all authors and affiliations

Science 02 Jan 2015:
Vol. 347, Issue 6217, pp. 78-81
DOI: 10.1126/science.1260825

Some tissue types give rise to human cancers millions of times more often than other tissue types. Although this has been recognized for more than a century, it has never been explained. Here, we show that the lifetime risk of cancers of many different types is strongly correlated (0.81) with the total number of divisions of the normal self-renewing cells maintaining that tissue's homeostasis. These results suggest that only a third of the variation in cancer risk among tissues is attributable to environmental factors or inherited predispositions. The majority is due to "bad luck," that is, random mutations arising during DNA replication in normal, noncancerous stem cells. This is important not only for understanding the disease but also for designing strategies to limit the mortality it causes.

Impressive coverage by media and social networks, often reporting the wrong message



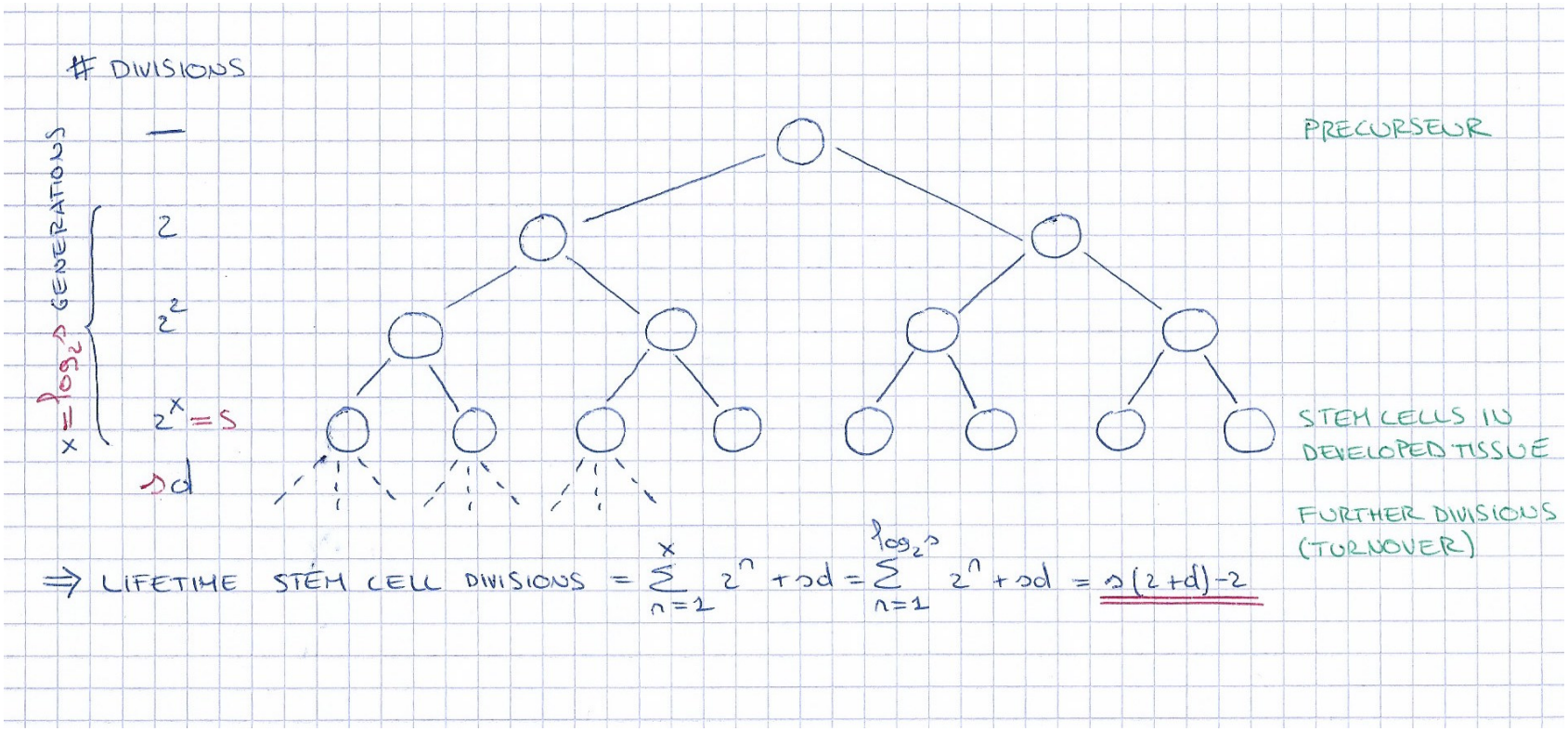
- « The bad luck of cancer » Science
- « Most cancers arise from bad luck » Scientific American
- « Cancer random's assault» The New York Times
- « But 2/3 of the time a person gets #cancer due to bad luck » twit

Contents

- Cancer risk explained by the number of stem cell divisions
- Intrinsic vs. extrinsic tumors?
- Proportion of mutations due to intrinsic and extrinsic factors
- Original result

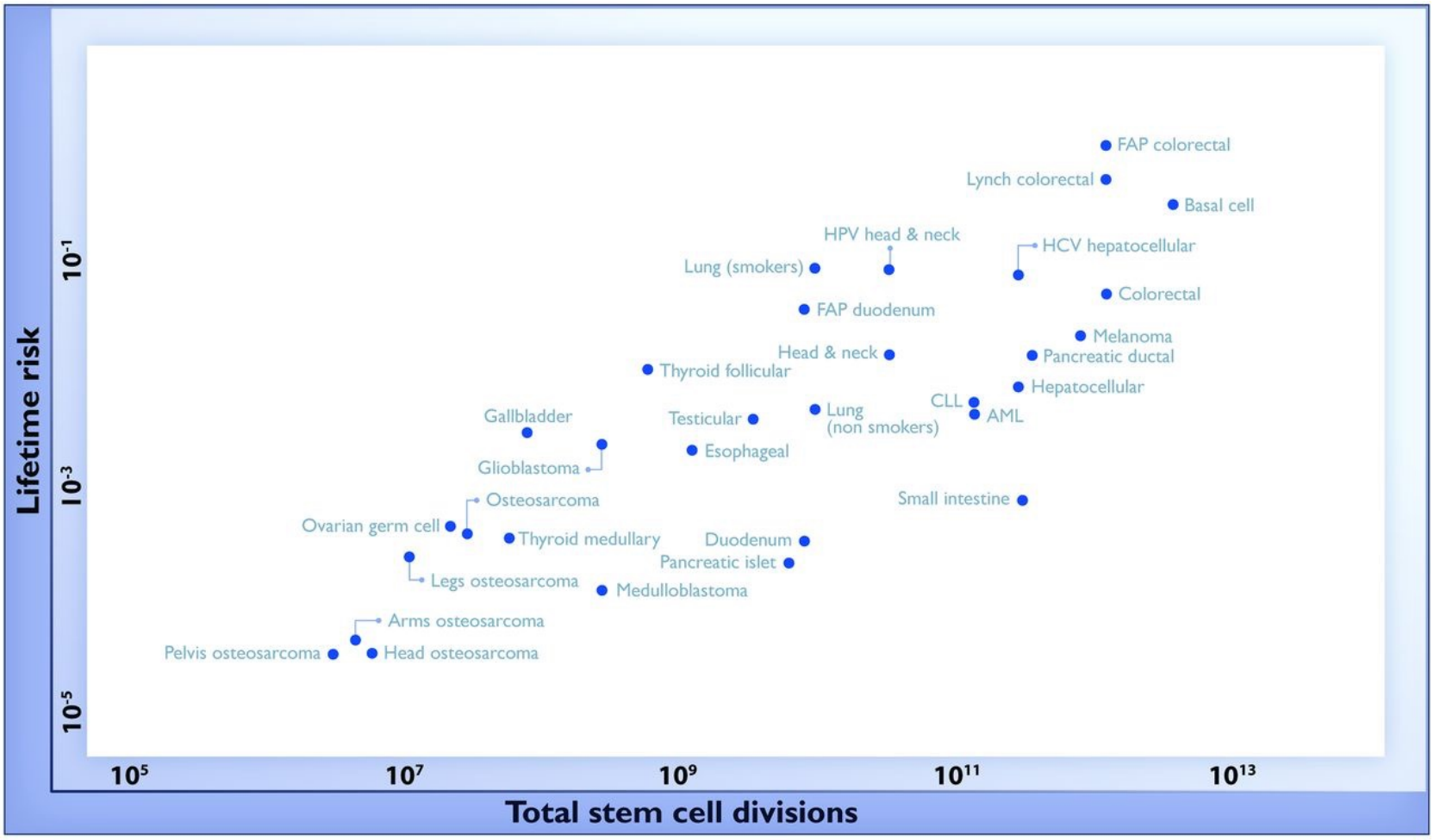
Number of stem cell divisions

- Model:



- Estimation: *s* and *d* found in the literature for 25 different tissues

The famous 2/3



FAP = Familial Adenomatous Polyposis ◇ HCV = Hepatitis C virus ◇ HPV = Human papillomavirus ◇ CLL = Chronic lymphocytic leukemia ◇ AML = Acute myeloid leukemia

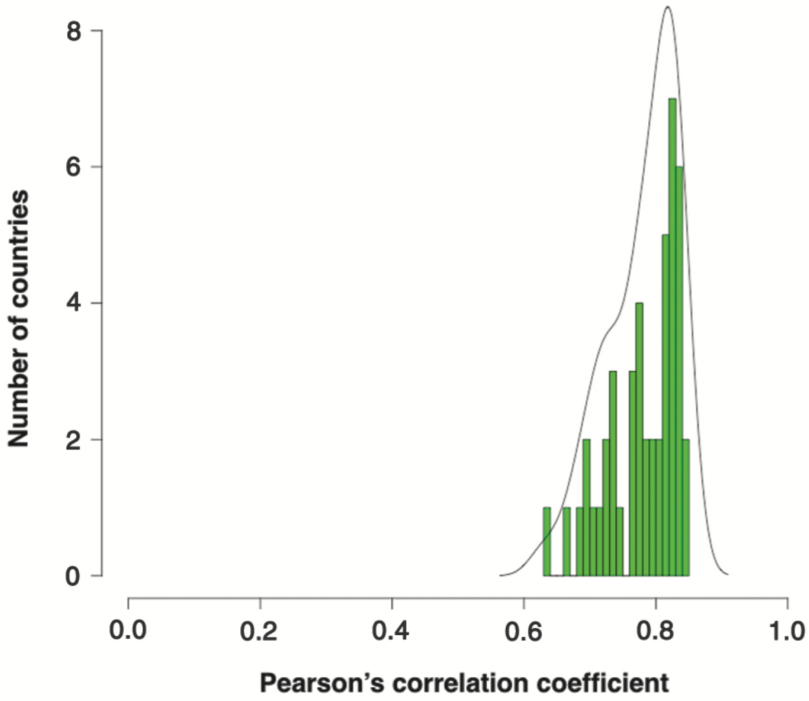
Tomasetti, Science 15

Correlation = 0.81, hence the proportion of the variation in risk across tissues explained by cell divisions is *R2 = 0.66*

Sensitivity analysis

Correlation did not change

- after modifying *s*, *d* and risk values
- considering risk data for 68 countries



Tomasetti, Science 17

Are 2/3 of cancers due to intrinsic unavoidable factors?

No, because

- By comparing the risk and number of cellular divisions of different tissues only an *ecological* correlation can be derived:
  « cancer cases » is not the same as « cancer types »!

- Such correlation cannot be interpreted as the fraction of risk attributable to intrinsic factors.

Primary prevention should not be be questioned:

- the observed correlation is not in contradiction with established findings about the preventability of a many cancers…

Thought experiment

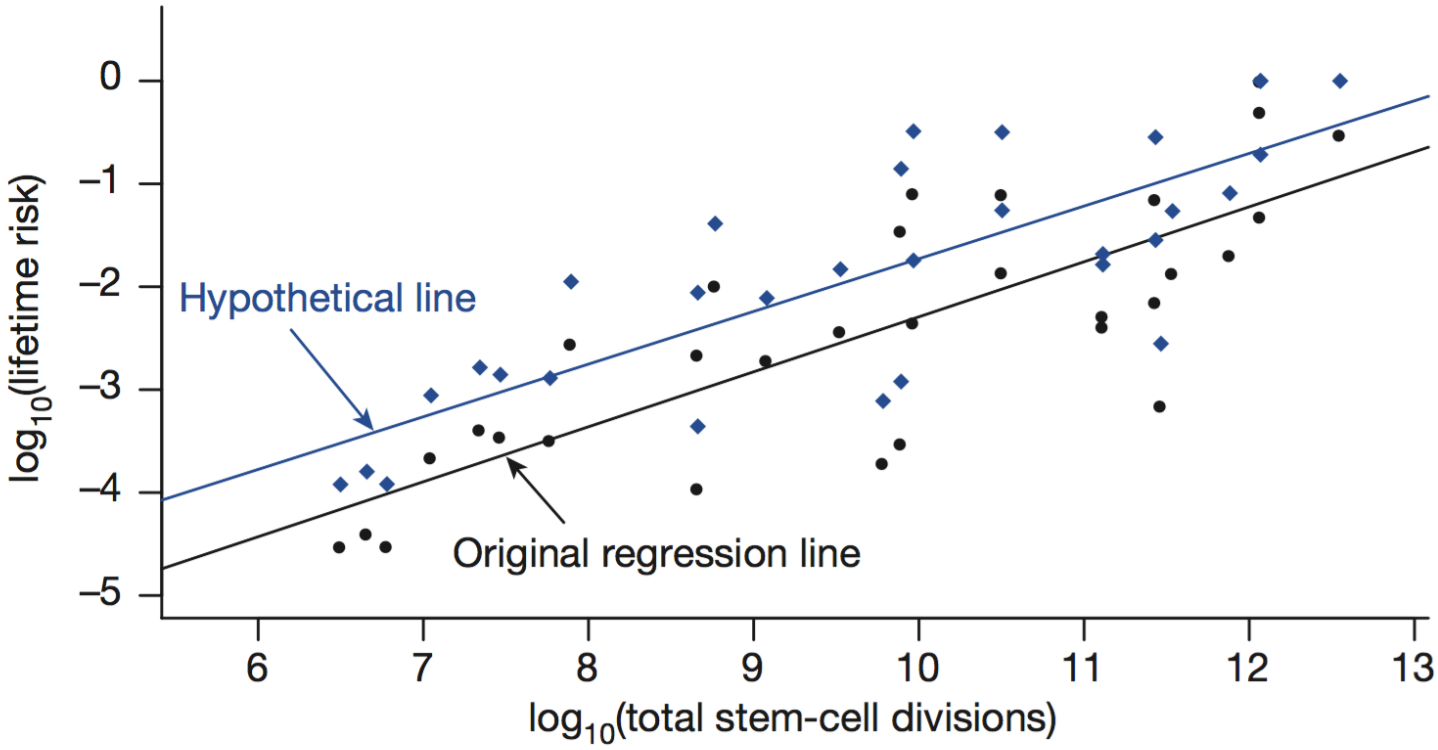


Figure 2 | Correlation analysis of stem-cell division and cancer risk does not distinguish contribution of extrinsic versus intrinsic factors to cancer risk. The black dots are data from figure 1(also shown in supplementary table 1) of Tomasetti & Vogelstein[5], and the black line shows their original regression line. The blue diamonds represent the hypothesized quadrupled cancer risks due to hypothetical exposure to an extrinsic factor such as radiation. The blue regression line for the hypothetical risk data maintains the same correlation as the original black line, albeit reflecting a much higher contribution of extrinsic factors to cancer risk.

Wu, Nature 16

A more realistic scenario

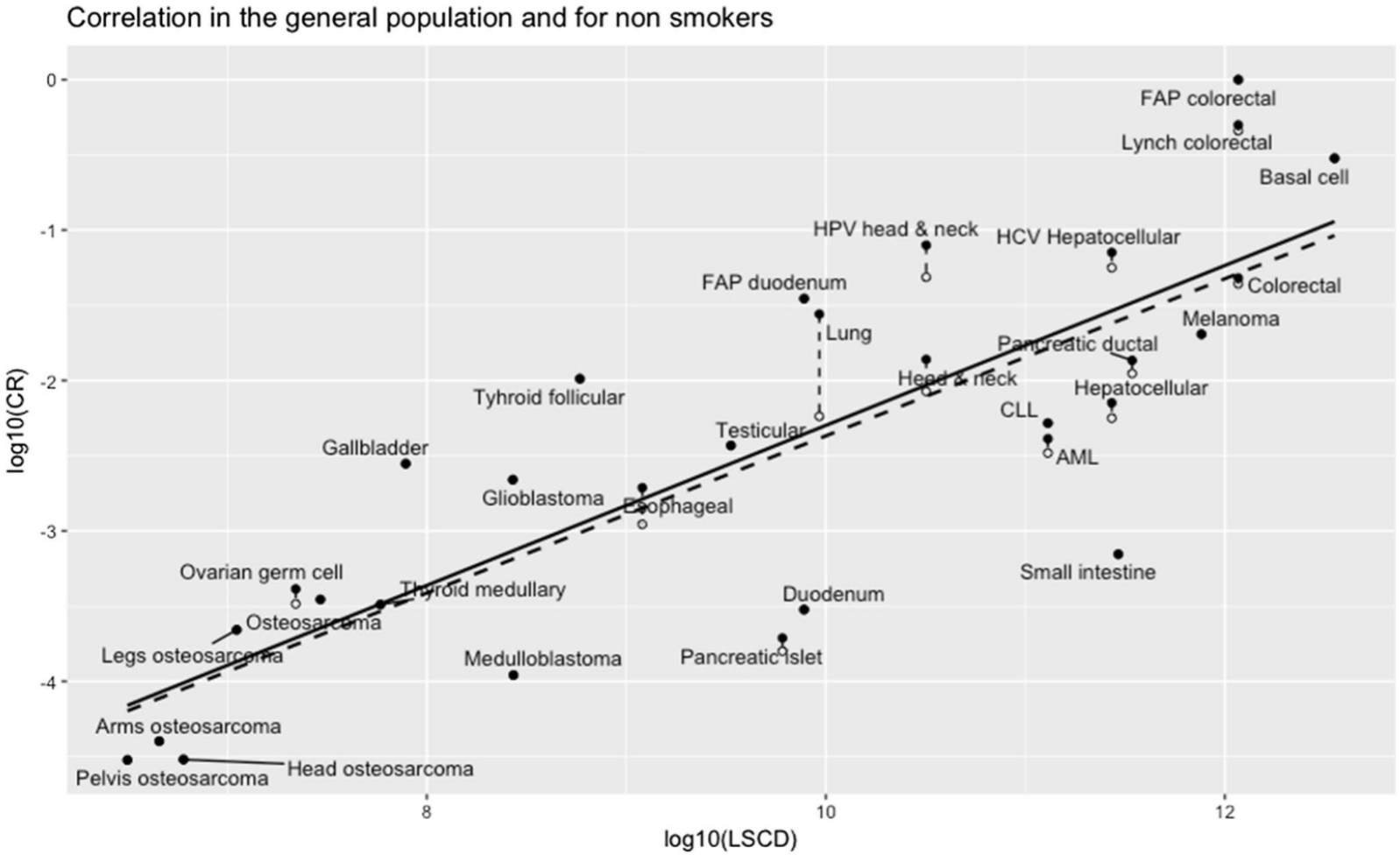


Fig. 2 Linear regression analysis of lifetime stem-cell divisions and cancer risk of tissues in the general population (continuous line) and for non-smokers (dashed line). For tobacco related cancers, a vertical dashed segment connects the dots representing the two risks. Risks for non-smokers were calculated by combining hazard ratio estimates from Agudo et al. [15] and lifetime cancer risk from Tomasetti and Vogelstein [1] and considering a smoking prevalence of 0.3. For both the general population and non-smokers, the correlation is 0.81

Perduca et al, Eur J Epidemiol 19

Partitioning etiological determinants

- Implicit in the interpretation that 2/3 of cancers cases are due to intrinsic factors is that the remaining 1/3 is due to extrinsic (genetic, environmental) factors.

- This is wrong, because relative effects of the determinants of a disease do not add up to 1.

- For instance, environmental factors could have an indirect effect on cancer risk that is mediated by the number of cellular divisions.

Classification of tumors

- Tomasetti's *Extra Risk Score*:

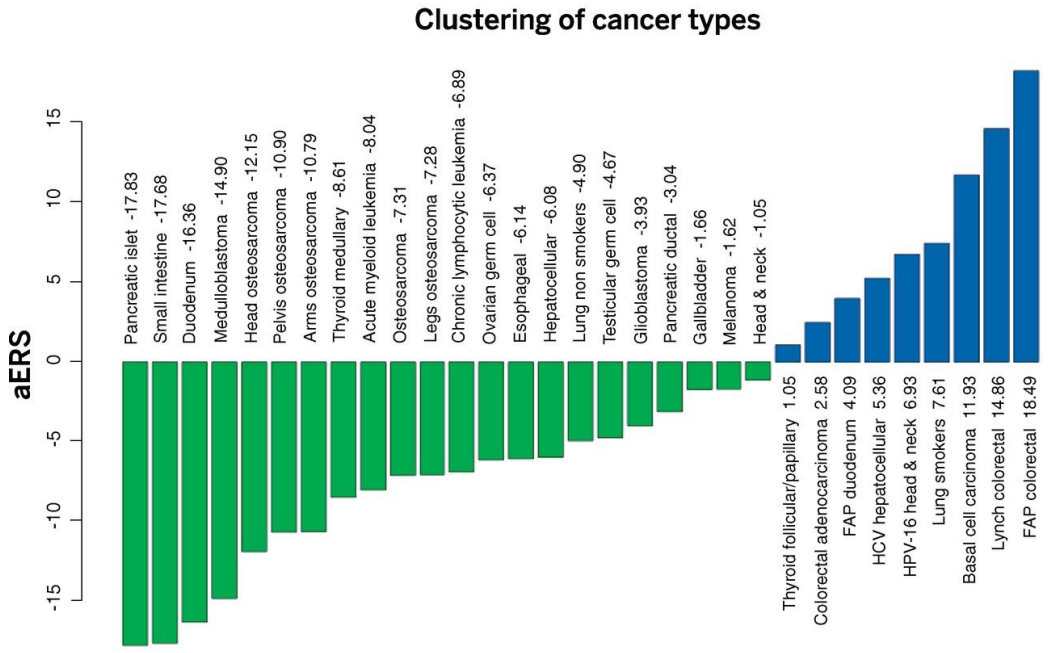
*ERS=log(CR)*log(LSCD)*

- Clustering of cancers based on the ERS:

    1. **E**xtrinsic or **D**eterministic (environment, genetics): *ERS > 0*

    2. **I**ntrinsic or **R**eplicative (cellular divisions): *ERS < 0*

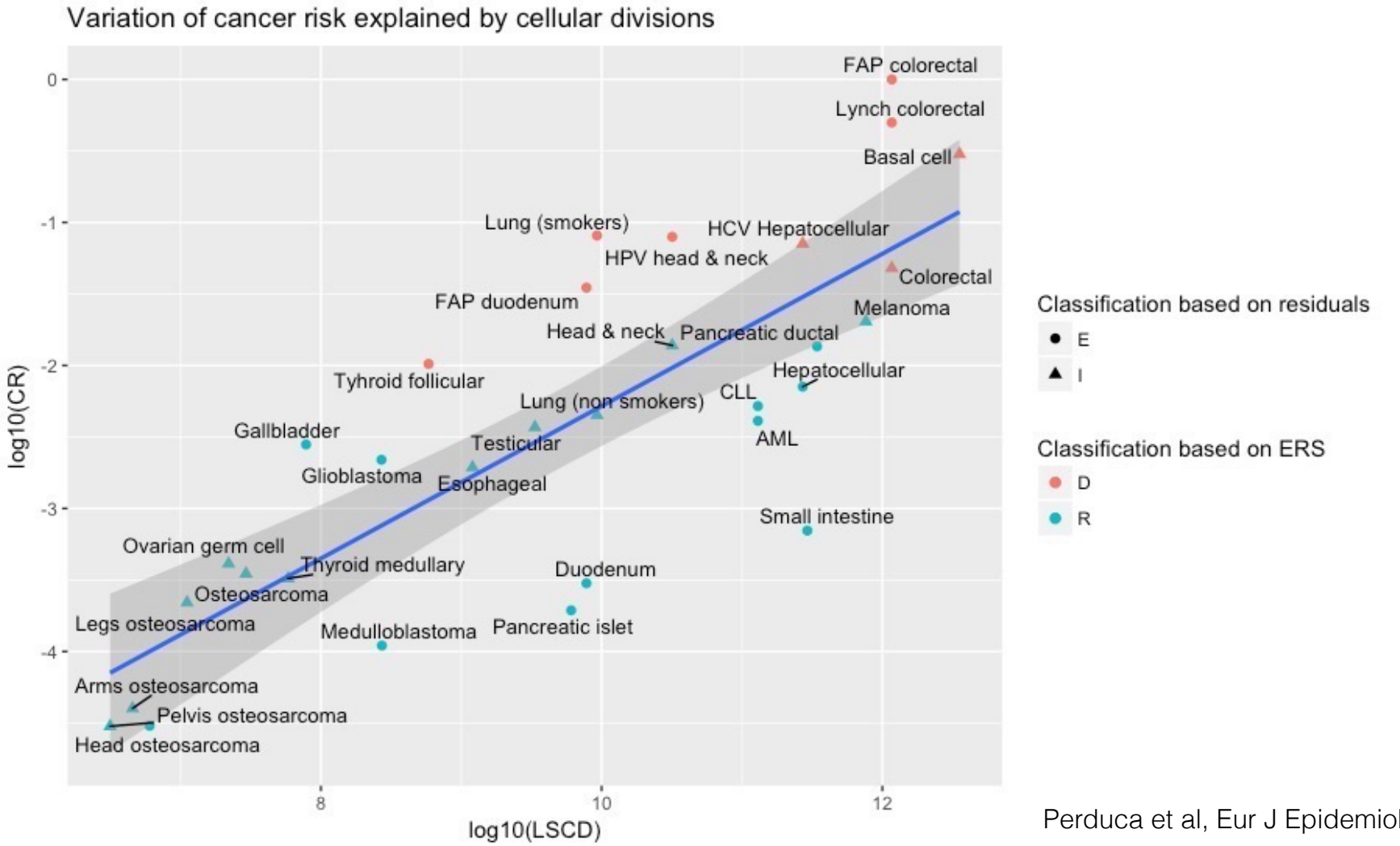
- 2/3 of cancers are classified as intrinsic



Tomasetti, Science 15

A more reasonable proposal

Classify as intrinsic those well predicted by cell divisions:



Perduca et al, Eur J Epidemiol 19

Refining the notion of bad luck in cancer

- We saw that the correlation between number of cell divisions and cancer risk is not a good measure of the importance of intrinsic mechanisms, and therefore bad luck, in cancer.

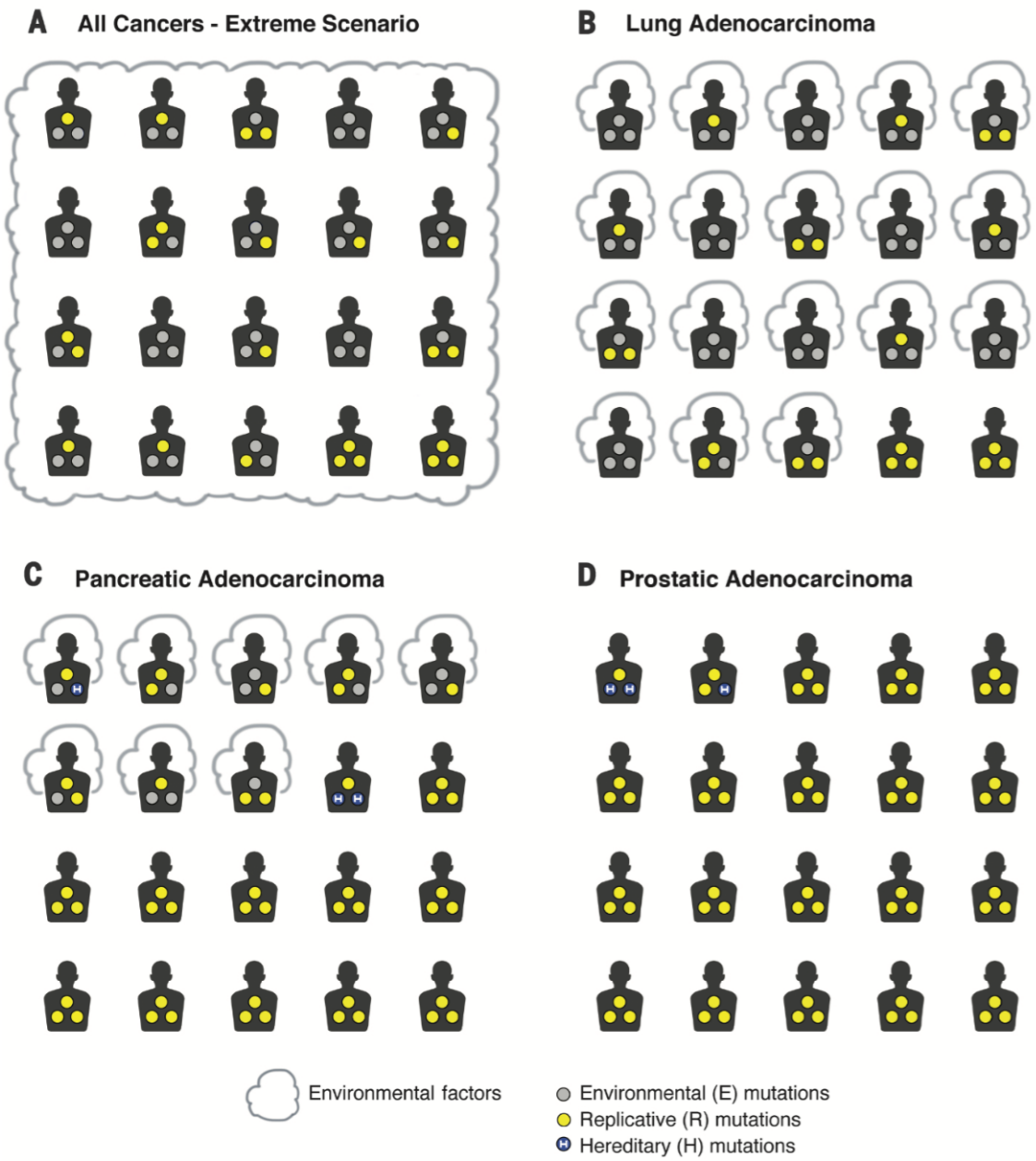
- The proportions of driver mutations due to intrinsic replicative mechanisms (R) and extrinsically factors (genetic H or environmental E) are more promising to investigate the role of chance.

- Warning: again, it could be that driver mutations are largely due to R factors (bad luck) and yet a large fraction of cases is still preventable.

Mutations due to R and preventable cases

Thought experiment A:

- 3 driver mutations required for cancer
- Proportion of mutations due to R = 24/60 = 0.4
- Proportion of mutations due to E = 36/60 = 0.6
- Proportion of preventable cases = 18/20 = 0.9



Tomasetti, Science 17

Proportions of driver mutations due to R, E and H

Tomasetti, Science 17: Mathematical model relating these proportions to

1. Epidemiological data (prevalence and RR)

2. Total number of mutations in unexposed samples (R) and in exposed samples (E and H)

The latter can be estimated from sequencing data.

Importantly, if follows from the model that the proportion of mutations due to E is always less than the fraction of risk attributable to extrinsic factors.

- Mesothelioma: PE = 0.67, PR = 0.33 and PAR = 0.89.

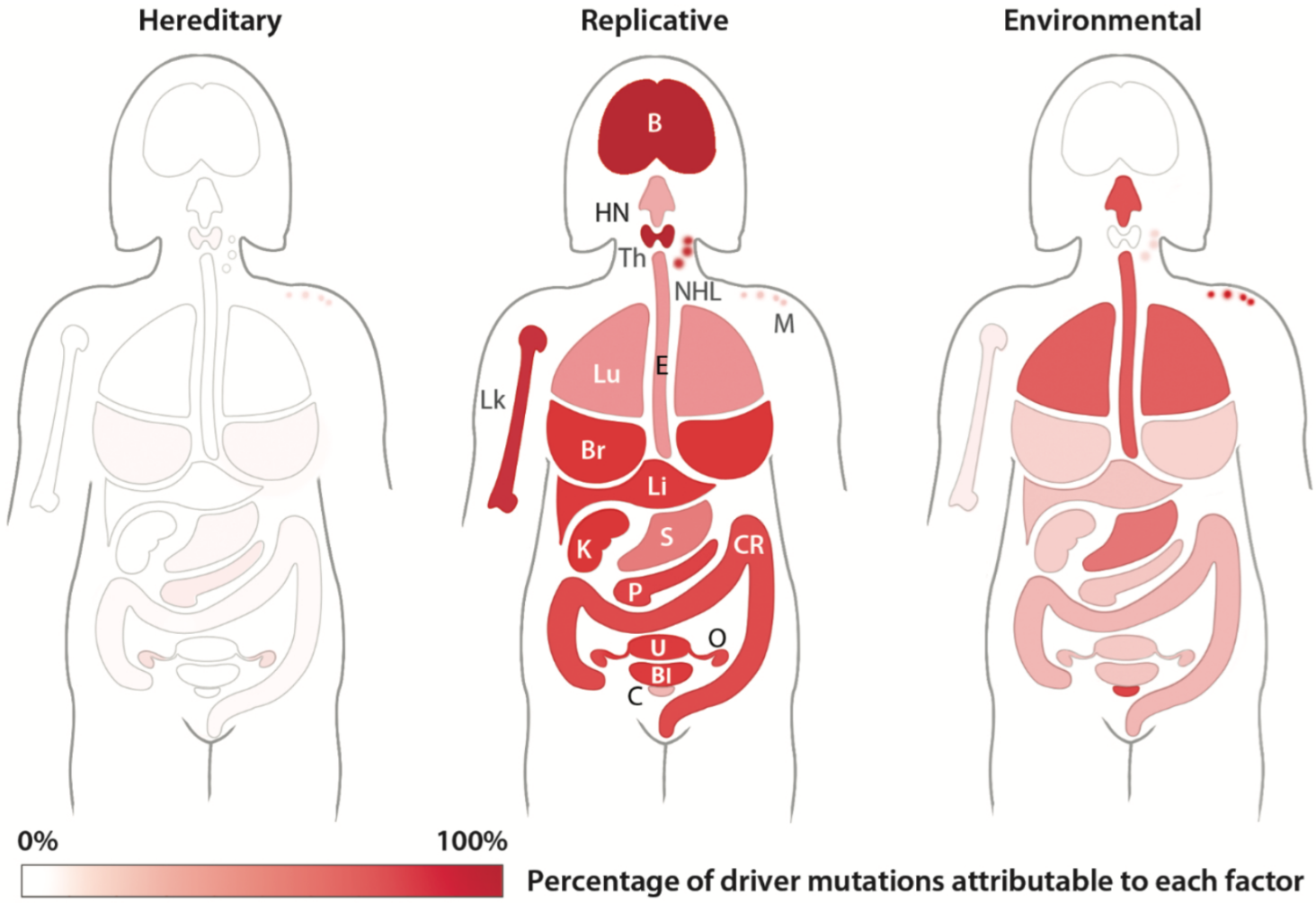


Fig. 3. Etiology of driver gene mutations in women with cancer. For each of 18 representative cancer types, the schematic depicts the proportion of mutations that are inherited, due to environmental factors, or due to errors in DNA replication (i.e., not attributable to either heredity or environment). The sum of these three proportions is 100%. The color codes for hereditary, replicative, and environmental factors are identical and span white (0%) to brightest red (100%). The numerical values used to construct this figure, as well as the values for 14 other cancer types not shown in the figure, are provided in table S6. B, brain; Bl, bladder; Br, breast; C, cervical; CR, colorectal; E, esophagus; HN, head and neck; K, kidney; Li, liver; Lk, leukemia; Lu, lung; M, melanoma; NHL, non-Hodgkin lymphoma; O, ovarian; P, pancreas; S, stomach; Th, thyroid; U, uterus. [Image: The Johns Hopkins University]

Tomasetti, Science 17

Data driven estimates

S. Wu et al. suggested to estimate the proportions of driver mutations using mutational signatures

- these are mutational patterns in cancer genomes left by exposures
- the intensity of mutational signature 1 correlates with age
- this should be an estimate of the proportion of driver mutations due to R
- results are different from Tomasetti's!!

Extended Data Table 3 | Percentages of intrinsic versus extrinsic MS with known and unknown causes in different cancer types

	Intrinsic MS	Extrinsic MS - Known	Extrinsic MS - Unknown	Extrinsic MS - Total
ALL	65.8	34.2	0	34.2
AML	100	0	0	0
Bladder	14.2	71.2	14.6	85.8
Breast	35.5	60.1	4.4	64.5
Cervical	25.3	74.7	0	74.7
CLL	76.7	23.3	0	23.3
Colorectal	17.1	66	16.9	82.9
Esophageal	48	25.3	26.7	52
Glioblastoma	53.8	0	46.2	46.2
Glioma-Low Grade	9.2	2.8	88	90.8
Head & Neck	24.9	75.1	0	75.1
Kidney Chromophobe	17.4	37.5	45.1	82.6
Kidney Clear Cell	66.5	4.1	29.4	33.5
Kidney Papillary	0	15.7	84.3	100
Liver	10.9	21.3	67.8	89.1
Lung Adenocarcinoma	9.1	73.8	17.1	90.9
Lung - Small Cell	0	92.8	7.2	100
Lung-Squamous	0	47	53	100
Lymphoma B-cell	46.3	33.4	20.3	53.7
Medulloblastoma	48.4	0	51.6	51.6
Melanoma	7.2	90.9	1.9	92.8
Myeloma	0	19.9	80.1	100
Neuroblastoma	53.2	0	46.8	46.8
Ovarian	36.6	63.4	0	63.4
Pancreatic	49.9	50.1	0	50.1
Pilocytic Astrocytoma	82.5	0	17.5	17.5
Prostate	32.2	10.2	57.6	67.8
Stomach	22.3	6.1	71.6	77.7
Thyroid	0	39.7	60.3	100
Uterine	10.7	65.5	23.8	89.3

Intrinsic mutational signatures (MS) includes signatures 1A/B, and extrinsic MS includes signatures 2–21, R1–R3, U1 and U2, excluding signature 11 for Temozolomide, an alkylating agent used fo chemotherapy. The blue, yellow and red colours highlight cancers that are have substantial extrinsic risk proportions based on epidemiological data, MS with known causes and MS with unknown causes, respectively. Data from the supplementary figs 59–88 in ref. 31.

Wu et al., Nature 16

Extrinsic mutations are more predictive of cancer risk than cellular divisions

Table 1 Comparison between mutation rates, cumulative stem cell lifetime divisions, hazard ratios for cancer in smokers and mortality rates in smokers and never smokers, for the cancer sites for which information was available in all sources

Cancer site	Mutation rates in smokers[a]	Cumulative stem cell lifetime divisions[b]	Incidence hazard ratio (relative risk) for smoking men[c]	Incidence hazard ratio (relative risk) for former smoking men[c]	Mortality rates smokers with ≥25 cigarettes/day/ non-smokers[d]
Lung adenocarcinoma	150.5	9.272×10^{9} [e]	23.30	5.28	415.2/16.9
Larynx	137.7	3.186×10^{10} [f]	13.24	3.51	17.3/0
Pharynx	38.5	NA	6.67	2.06	19.4/0
Bladder	18.3	NA	3.84	2.15	51.4/13.7
Esophagus (squamous)	N.S.	1.203×10^{9}	3.94	1.26	50.0/5.7
Liver	6.4	2.709×10^{11}	2.92	2.09	31.3/4.4
Pancreas adenocarcinoma	N.S.	3.428×10^{11}	1.62	0.89	52.9/20.6

[a]Statistically significant average number of somatic substitutions per genome per pack-year. From Alexandrov et al. [28]

[b]Cumulative number of divisions of stem cells per lifetime. From Tomasetti and Vogelstein [1]

[c]From Agudo et al. [15]

[d]Cumulative mortality rate per 100,000 persons per year, from Doll et al. [35]

[e]Cumulative number of divisions of stem cells per lifetime. From Tomasetti and Vogelstein [1]

[f]Adenocarcinoma (same rate in smokers and non-smokers)

Perduca et al, Eur J Epidemiol 19

Correlation with risk	Mutation rates	Cellular divisions
Smokers	0.93*	-0.65
Former smokers	0.91*	-0.58

Conclusions

- Extra care is needed when communicating results about variation and probabilities.
- Importance of primary prevention.
- Current models for measuring intrinsic and extrinsic factors are not satisfactory:
    - Mutations are not sufficient for cancer: they must be accompanied by other dysfunctions (eg in the immune system).
    - Epigenetics not taken into account.
    - G, E are most likely not independent of R.

European Journal of Epidemiology
https://doi.org/10.1007/s10654-018-0477-6

ESSAY

Stem cell replication, somatic mutations and role of randomness in the development of cancer

Vittorio Perduca[1,7] · Ludmil B. Alexandrov[2] · Michelle Kelly-Irving[3,4] · Cyrille Delpierre[3,4] · Hanane Omichessan[7] · Mark P. Little[5] · Paolo Vineis[6] · Gianluca Severi[7,8]

Non, le cancer n'est pas principalement dû au hasard

SCIENCE • 04.02.2019

Une étude parue en 2015 a suggéré que le cancer résulterait dans la plupart des cas du hasard, d'une mutation malvenue : un coup de malchance. Gianluca Severi et son équipe du Centre de recherche en épidémiologie et santé des populations (Villejuif) montrent que cette conclusion n'est pas en contradiction avec le fait qu'une forte proportion des cas reste liée aux comportements, et en particulier au tabagisme.