# Our cancer risk is not written in the stars

If Tomasetti and Vogelstein had not used the words "bad luck" in their paper on how 'variation in cancer risk among tissues can be explained by the number of stem cell divisions', the media might not have covered the story the way they did. In this interview with *The Cancer Letter*, Bertram Kramer, head of Cancer Prevention at the US National Cancer Institute, tries to clear up some of the confusion.

**he Cancer Letter (TCL):** What was your overall impression of the Tomasetti and Vogelstein paper?

**Bertram Kramer (BK):** I found the paper interesting. What they did was they didn't generate any new experimental evidence, obviously. They searched the literature for reports on numbers of stem cells and number of divisions of the stem cells.

They used well-accepted concepts that the risk of mutations or number of mutations are relatively constant for a given cell division – in statistical terms, a stochastic process – that is, any given division, you don't know which gene is going to mutate, but for every given division, you can predict, relatively accurately, how many mutations are going to occur in the division. You just don't know which cell it's going to happen to. But if you have enough cells, then a statistical analysis of this stochastic process gives you, generally, a pretty good idea of how many mutations there are, and the number of mutations to be a risk factor for cancer.

## **TCL:** What were the authors trying to achieve in their analysis?

**BK:** They took well-known concepts, went to the literature, looked for the number of stem cells in any given class of tumors or tissue type, and looked for reports of the number of divisions.

The innovation they added – actually directly plotting the number of anticipated mutations or divisions with the cancer risk – and what I found interesting was that, relative to most biological processes, they got a pretty tight correlation between the number of stem cell divisions and the risk of cancer.

The variation in cancer risk across the tumor types for which they had any data was about 65%, and that's a pretty tight correlation, in biological terms. So it fits with the existing notions of the association between mutations and cancer. I found that interesting. I think they took existing literature and results and, for the first time to my knowledge, plotted them looking for variation across cancers using that information and got a tight correlation.

So it's not conceptually different from what was, in essence, accepted, in terms of the association, but what they did was plot it graphically, and as it often happens, you get some biological input by taking existing data and graphing them.

That's what I took as particularly

THE CANCER LETTER This interview was first publ

This interview was first published in *The Cancer Letter* (www.cancerletter.com) on 16 January 2015, and is republished with permission. © The Cancer Letter 2015. The interview was conducted by Matthew Ong

### FOCUS



was widely – but often inaccurately – covered in media across the world

interesting in the paper. I wouldn't have predicted that the correlation would be quite that high, and so I found it intriguing that it was. That's the good part.

TCL: What have news reports missed in their coverage of the paper's findings? **BK:** On the parts that I think may have either been misinterpreted or picked up in the press and took an extra step too far, was going beyond the actual data to some of the implications. I don't think that, given those observations, you can conclude with any confidence what would be the best strategy to decrease mortality for a given cancer.

I don't think that tells you *a priori* whether the best strategy will be screening; or the best strategy

will be primary prevention; or the best strategy will be treatment. Unfortunately, you're left with the hard grunt work of testing various strategies to see which is the most effective amongst the three for decreasing mortality.

A case in point would be that they unfortunately didn't have reported evidence on stem cells or stem cell divisions from two very common cancers – prostate cancer and breast cancer – and for both of those cancers we at least have some evidence about whether or not screening works, or how effective it is, and it would have added to the paper if they had some stem cell division data on those. There have been randomized trials at least to test the inference that screening would or wouldn't work.

The next important thing, which I think was sort of missed in the press –

even the paper itself says something that appears to equate that stochastic process with bad luck. I personally think that the use of the phrase 'bad luck' can be easily misinterpreted. Stochastic processes have a crisp scientific definition, but 'bad luck' doesn't. The lay public may interpret incorrectly in this case, in my opinion, that 'bad luck' simply means "it's in the stars, it's your fate, there's nothing you can do about it." And 'bad luck' is not equivalent to random mutations in a stochastic process.

**TCL:** *What would be a good analogy?* **BK:** Let's say you're dealing with traffic patterns. The heavier the traffic, the more accidents there are going to be. There is a tight correlation between the

# 'Bad luck' means to most people, 'nothing you can do about it, you are meant to have cancer'

number of cars on the roads and the number of accidents, but that doesn't mean that it's pure bad luck if you have an accident.

Statisticians can predict that, for a given road at a given time and given road conditions, there's going to be a certain risk and a certain number of accidents. And the correlation almost certainly is going to be very tight, but that doesn't mean that the individual car driver has no control, and might as well give up because whether they have an accident is purely bad luck. They can choose to drive differently.

So aggressive drivers are at a higher risk than slower or safer drivers. And the same is true for speed limits. It's well known and it has been well described that for every mile per hour that you raise the speed limit, or every five or 10 miles per hour, the rate of mortalities or fatalities can go up.

But that doesn't mean for an individual driver, it's just pure bad luck. Because individual drivers and individual cars have a different risk of traffic fatality depending on how they drive, even if they're driving at the same speed in the same speed zone.

The other thing which was not picked up by most of the press was that the correlation they were even looking at, leaving aside the issue of cause and effect, because this isn't even designed to determine cause and effect – they were looking at classes of tumors.

They lined up 31 classes of tumors, and they found out that the correlation was surprisingly high, and I found that interesting. But they were not looking at risk of individual tumors. Even if it were true that two-thirds of the variability among tumor types is associated with the number of stem cell divisions, it doesn't mean that two-thirds of all cancers are predetermined.

Let's say you have an extremely common tumor and 10 extremely rare tumors, and you plot the number of stem cell divisions for those 11 tumors. The 11 tumors may line up very nicely along that diagonal line, that is, they fit a pattern that, across tumor types, there is a pretty tight association between stem cell divisions and cancer risk.

But remember, the most common tumor accounts for most of the cancers. And if that most common tumor is attributable in large measure to a known environmental carcinogen, then the overwhelming majority of cancers, individual cancers, will be preventable. And so a clear case in point would be lung cancer, which we know that 90% of lung cancers are probably attributable to smoking and preventable if people don't smoke at all.

And yet there are many, many rare tumors for which we don't have any known environmental cause, and even in the aggregate, if you add them all up, they don't come anywhere close to the number of lung cancers.

So just one simple preventive intervention would prevent the overwhelming majority of all those cancers even if the association tells you that, across cancer types, two-thirds are due to the stochastic process of mutation.

Let's say there were only five cases of every other cancer type there is, and they added up to a total of 200 cases a year, and there were 150,000 cases a year of lung cancer, 90% of which were attributable to smoking, then the overwhelming majority of individual cancers would be preventable, even if a regression curve tells you that across cancer classes, there is a pretty tight correlation with stochastic processes.

And in this case, let's take lung cancer, which we know 90% are preventable by no smoking, and skin cancer, especially non-melanoma skin cancer – which is more common than all the other cancers combined, including lung cancer – and we know that non-melanoma skin cancers are largely preventable by avoiding intensive sun overexposure, the biggest risk factor for non-melanoma skin cancer.

The number of non-melanoma skin cancers just completely outweighs all other cancers combined. And so, even though skin cancer fits on that regression line, and is part of the pattern of cancer types, sun avoidance would still prevent an inordinately large number of total cancers in the country.

Unfortunately, the term 'bad luck' got picked in a number of news outlets. Just the term 'bad luck' can be misleading. 'Bad luck' just means, to most people, "nothing you can do about it, you are meant to have cancer."

And since the term was – for the sake of simplicity or I would say, over-simplicity – equated with a more precise statistical phenomenon, stochastic risk, that led to the sense that, "Gee, there's not much you can do about cancer, it's just all in the stars." That has an unfortunate connotation, and I think that was the biggest error of translation of the results.

Lawmakers, and physicians, by the way, and health professionals and the lay

public often respond to news articles, and if they are misinterpreted, then it can lead to policy decisions, which are obviously made on behalf of the lay public.

## **TCL:** Do you have any other observations that you'd like to highlight?

**BK:** Another thing I wanted to point out that I found interesting in Figure 1 of the paper – the correlation seems good relative to many biological phenomenon. One thing I took from it, and it wasn't emphasized in the article, is that you can sort of visually look at the vertical distance between any given individual cancers on that regression line.

The further it is away from the regression line, the more that one could suspect that there is something going on, if it is cause and effect, there's something additional going on that explains the higher incidences for the curves that are well above the line. And sure enough, that fits the pattern very nicely, so it's interesting to look at.

The best example is lung cancer. When you look at lung cancer (smokers) and lung cancer (nonsmokers), there is a very large vertical difference between those. So lung cancer (smokers) as you'd expect, the point is way above that regression line.

And the same is true, for example, for HPV head and neck cancer and other cancers, and hepatitis B liver cancer is way above the line relative to the rest of liver cancer. It fits that one would say, "Gee, the further vertically the point is from the line, especially if it's north of the line, the more may be going on, over and above the stochastic random process."

That is one indicator that something else might be going on: how far above, vertically, the regression line, a given point is. That's not pure, it's very rough, but nevertheless, if you look at some of the points, they fit that pattern.

General colorectal cancer is right on



#### FIGURE 1: THE DATA AT THE HEART OF THE DEBATE

The relationship between the number of stem cell divisions in the lifetime of a given tissue and the lifetime risk of cancer in that tissue

FAP – familial adenomatous polyposis, HCV – hepatitis C virus, HPV – human papillomavirus, CLL, chronic lymphocytic leukaemia, AML – acute myeloid leukaemia

*Source:* C Tomasetti and B Vogelstein. (2015) Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science* 347:78–81, reprinted by permission from AAAS

the regression line, but those with a genetic predisposition (FAP) for colorectal cancer are way above that regression line vertically. Each of those points that are very far away from the line seems to fit that pattern.

Now, always, an environmental carcinogen, you have to be very cautious before you say, it must be an environmental carcinogen. A case in point is thyroid follicular cancer – the incidence may be driven by screening for thyroid cancer and screening tests are much better at picking up thyroid follicular than other forms of thyroid cancer. So all it means is that the incidence is considerably higher than you have expected simply based on the formula of stem cells and number of divisions.

I think that we can be pretty confident that there are some causative reasons

for the vertical difference. Certainly, we can be confident in the case of smoking and lung cancer. That's a well-established causative factor. I think we can be confident in the case of HPV infections for head and neck cancer. We're pretty confident that that's causative.

In the case of thyroid follicular cancer, I think the weight of evidence is that screening increases the risk of thyroid cancer even if there are no known new carcinogens. And I think there is a large body of evidence that some of the incidence, and sometimes a large measure of incidence in some cancers, is attributable to screening and overdiagnosis, such as picking up very indolent, non-life-threatening cancers just by simply dipping into a reservoir of silent, non-progressive tumors with a screening test.