A Causal (Regression) Analysis for Cancer Formationunder a Corresponding Virus Attack

Mr. (Carson) Lam Kai Shun

B.Sc., M.Sc.(Eng), M.Sc.(ITE), HKU, FSASS, India

email: h9361977@connect.hku.hk

Abstract:

Cancer or COVID-19 virus can be serious harmful diseases to our human beings. However, by understanding how these disease will be formed, we may construct the corresponding statistical causality regression model or even the artificial intelligence one. This may help us develop the corresponding drugs or vaccine. Indeed by applying the techniques as stated in one of my paper [11], one may even control the mutation or the evolution of the cancer or the COVID-19 virus whenever the Lorenz attractor or the Butterfly effects existing during the virus mutation process. This may be a break through in the history of fighting against our human beings' nature enemy — virus.

Background

Cancer is the name given to a group of serious diseases that humans have long tried to cure. Indeed, the latest and most advanced research shows that different viruses play a significant role in the formation of various cancers. In the following, I shall outline a causal (regression) analysis [1] for cancer formation under a corresponding virus attack. According to Joyce et al., [2], alcohol can induce hepatitis C virus (HCV)-hepatocellular carcinoma (HCC) that causes the DNA methylation of repetitive elements. These elements may include long interspersed nuclear element-1 (LINE-1) and All element (Alu). Joyce et al. [2] further concludes that HCV infection is highly connected with the loss of DNA methylation in specific REs, implicating molecular mechanisms in liver cancer development.

Investigating DNA methylation in more depth, Cho et al. [3] finds that tobacco smoking may change the transcription and methylation states of extracellular matrix organisation-related genes. It is worth noting here that tuberculosis (TB) may increase the risk of lung cancer, which is highly related to smoking. In such a case, DNA methylation patterns will change, altering the transcription states of genes. To be more precise, proteins will bind to methylated DNA. This DNA will then form complexes with the proteins that formulated during the process of deacetylated histones. Therefore, when DNA has undergone the process of methylation, the nearby histones will also deacetylate, resulting in compounded inhibitory effects on transcription. Similarly, demethylated DNA will not cause the attachment of deacetylated enzymes to the histones, allowing DNA to be acetylated and more mobile, thus promoting transcription.

Next, when transcription indirectly caused by the virus takes place, cellular lipidomics may occur (Chakrabarti et al.,) [4]. As a result, this could lead to lipid metabolism and cause metabolic health issues. Snaebjornsson et al. [5] explains that by altering lipid metabolism, one may even develop a possible therapeutic window for cancer treatment. Thus, with reference to my paper in causal (regression) analysis, as shown in [1], one can deduce the following causality for the possible formation of cancer:

Virus infection → DNA methylation → transcription occurs → lipidomics → metabolic problem → cancer formation

Using suitable bio-informatics cancer data, one can follow the procedure as stated in [1] to establish a Hayes process model [6] for just such a virus cancer formation mechanism. Specifically, one can perform heterogeneous analysis, together with spatial analysis, to determine the cancer formation mechanism. Indeed, other virus infection mechanisms, such as COVID-19, can also be

modelled using the same or similar methods, as long as the details of the infection can be ascertained from cellular experiments.

According to Stephen et al. [7], one can employ the following phenomenological model for the hepatitis B virus:

$x'(t) = \lambda - dx - \beta\nu x$ _____ equation (A)

$y'(t) = \beta\nu x - ay$ _____ equation (B)

$\nu'(t) = ky - \mu\nu$ _____ equation (C)

where x, y, and ν are the numbers of uninfected liver cells, infected cells, and free virions, respectively; λ represents the (constant) rate where the uninfected liver cells are produced; and dx represents the linear term of maintaining tissue homeostasis in the face of hepatocyte turnover.
During infection, healthy liver cells are assumed to become infected at a rate βνx, where β is the mass action rate constant describing the infection process. Infected liver cells are killed by immune cells at rate ay and produce free virions at rate ky, where k is the so-called 'burst' constant. Free virions are cleared by lymphatic and other mechanisms at rate μν, where μ is a constant.

Major Results

Mathematically, such a system of differential equations can be made to interact with my HK-Lam theory [8] and, in terms of the Lorenz differential system (or other differential equation models such as SERI), can be expressed in the form of a matrix. In addition, it is feasible that these matrices are capable of being written using HKLam's (net-seizing) theorem. The following is an example:

$$\begin{pmatrix} -\mathbf{6} & \mathbf{6} & 0 \\ -z+\beta & -1 & 0 \\ 0 & x & \beta \end{pmatrix} \begin{pmatrix} \mathrm{L} \\ \mathrm{T} \\ 1 \end{pmatrix}$$ = Linear regression of causal domino (LRA 1) ——— (equation 1)

$$\begin{pmatrix} -\mathbf{6} & \mathbf{6} & 0 \\ -z+\beta & -1 & 0 \\ 0 & x & \beta \end{pmatrix}$$ = Linear regression expression (LRE 2)——————————— (equation 2)

(LRE 2)$\begin{pmatrix} \mathrm{L} \\ \mathrm{T} \\ 2 \end{pmatrix}$ = (LRA 1) ——————————————————————————— (equation 3)

Substitute $\begin{pmatrix} \mathrm{L} \\ \mathrm{T} \\ 2 \end{pmatrix}$ in equation (3) back into equation (1) and continues this (in fact which is initially an infinite and) recursive process (such that this is a kind of mathematical formalism) until the linear transformation $\begin{pmatrix} \mathrm{L} \\ \mathrm{T} \\ 2 \end{pmatrix}$ is found in the pre-calculated optimal approximation for a given set of values (σ1, β1, β1) in the Lorenz attractor. In this case, we have got the wanted real values of the linear transformation $\begin{pmatrix} \mathrm{L} \\ \mathrm{T} \\ 2 \end{pmatrix}$ at that a particular given set of valued point, say (σ1, β1, β1). Conversely, if there are sufficiently large amount of $\begin{pmatrix} \mathrm{L} \\ \mathrm{T} \\ 2 \end{pmatrix}$s, we can estimate back the corresponding true val-

ues of (σ2, β2, β2) and get the optimal values in the Lorenz attractor. Hence, in terms of weather management, we can determine the consequence risks behind. Thus, we may associate with the most feasible warnings and give in advance by applying some suitable decision theories. In such a case, we are actually using the HKLam Theory to net-seize those changes in our earth-weathering butterfly effect (or the Lorenz attractor). Similar cases happen in other differential equation models such as in the viruses mutation in micro-biology (cancer research) together with the spread of other viruses (influenza and COVID-19 for example) in our public health.

Take for some case studies, lest's consider a set of values for (σ = 10, β, β = 8 / 3, z = 1, x = 1) and interact with one of the regression model equation in Lam [1], we may get:

$$\begin{pmatrix} -10 & 10 & 0 \\ -1+\beta & -1 & 0 \\ 0 & 1 & 8/3 \end{pmatrix} \begin{pmatrix} L \\ T \\ 1 \end{pmatrix} = (4.3685.53 + 26.71*\text{wind} - 2.185*\text{wettest} - 2054.05 *$$

temperature

———————- equation(1')

Approximate the matrix by a linear regression, we also have:

$$\begin{pmatrix} -10 & 10 & 0 \\ -1+\beta & -1 & 0 \\ 0 & 1 & 8/3 \end{pmatrix} = AX + \varepsilon_0;$$

———————- equation(2')

Thus, by equating the above two equations, we get:

$$(AX + \varepsilon_0) \begin{pmatrix} L \\ T \\ 2 \end{pmatrix} = (4.3685.53 + 26.71*\text{wind} - 2.185*\text{wettest} - 2054.05 * \text{temperature};$$ or

$$\begin{pmatrix} L \\ T \\ 2 \end{pmatrix} = (AX + \varepsilon_0)^{-1} * [(4.3685.53 + 26.71*\text{wind} - 2.185*\text{wettest} - 2054.05 * \text{temperature}]$$

——————— equation(3')

Let B = (4.3685.53 + 26.71*wind - 2.185 *wettest -2054.05 * temperature

$$\begin{pmatrix} -10 & 10 & 0 \\ -1+\beta & -1 & 0 \\ 0 & 1 & 8/3 \end{pmatrix} [(AX + \varepsilon_0)^{-1} * B] = B$$

$$\begin{pmatrix} -10 & 10 & 0 \\ -1+\beta & -1 & 0 \\ 0 & 1 & 8/3 \end{pmatrix} = B * [(AX + \varepsilon_0)^{-1} * B]^{-1}$$

Continue the above process for the second time, we get:

$$B * [(AX + \varepsilon_0)^{-1} * B]^{-1} \begin{pmatrix} L \\ T \\ 3 \end{pmatrix} = B$$

$$B * [(AX + \varepsilon_0)^{-1} * B]^{-1} = A'X + \varepsilon'_0$$

$$(A'X + \varepsilon'_0) \begin{pmatrix} L \\ T \\ 4 \end{pmatrix} = B$$

$$\begin{pmatrix} L \\ T \\ 4 \end{pmatrix} = (A'X + \varepsilon'_0)^{-1} * B$$

$$B * [(AX + \varepsilon_0)^{-1} * B]^{-1} * [(A'X + \varepsilon'_0)^{-1} * B] = B$$

$$B * [(AX + \varepsilon_0)^{-1} * B]^{-1} = B * [(A'X + \varepsilon'_0)^{-1} * B]$$

$$\begin{pmatrix} -10 & 10 & 0 \\ -1+\beta & -1 & 0 \\ 0 & 1 & 8/3 \end{pmatrix} = B* [B * [(AX + \varepsilon_0)^{-1} * B]^{-1} *B]$$

$= B^2 [(AX + \varepsilon_0)^{-1}]^{-1}$ if we assume $(A*B)^{-1} = B^{-1}* A^{-1}$ and the communicative, distributive properties of matrix

multiplication.

If furthermore, the Lorenz matrix can be QR decomposed, then we may have:

$$\begin{pmatrix} -10 & 10 & 0 \\ -1+\beta & -1 & 0 \\ 0 & 1 & 8/3 \end{pmatrix} = B^2\,[(QR)^{-1}]^{-1}$$

The above sample Lorenz matrix (with assumed parameters) shows that we can always express it in the formative and recursive format using linear regression approximation.

In general, formative recursion computing can be solved using a five-step process:

1. Find out the simplest possible input
2. Visualise and play around with examples
3. Relate hard cases to simpler cases
4. Generalise the pattern
5. Write computer code combining the recursive pattern with the base case

Next, we apply the mathematical/statistical model to simulate the virus interaction that appears in the biological animal model for the corresponding type of cancer. Doing so allows one to further develop the corresponding hepatitis anti-cancer vaccine through reconstruction of the T-cells using CRISPR technology, and train them to attack those cancer cells and finally eliminate them. One can evaluate and improve the efficacy of the vaccine using big data analysis.

Conclusions

Finally, it is important to note that one can mathematically model the virus infection processes that lead to liver cancer using Wnt signalling (e.g., corresponding to patterns in metabolism

in colon cancer) [9]. The result being that one might be able to find a suitable virus, such as an oncolytic virus, for the treatment of colon cancer, for example [10]. Moreover, it might be possible to "reconstruct" the selected virus using CRISPR (according to the mathematical model computed previously) and train it to attack those cancer cells whenever we humans cannot find a suitable virus that is suitable used as a vaccine for our liver cancer like the case in . Hence, we might be able to develop the necessary drugs (that contain the suitable virus) to attack the respective liver cancer virus and balance the micro-organisms that co-existed in the surrounding cancer infected area. Theoretically, the tumour size will gradually be diminished, and we may even cure the liver cancer caused by the hepatitis C virus completely. Similarly, we can apply the same mathematical and statistical modelling method for other diseases such as HIV AIDS, and develop corresponding drugs using CRISPR re-construction for suitable viruses. This author wants to remark that there is an international study done by the South African professor that when the
COVID-19 virus meets the AIDS one in the human body, COVID-19 can be mutated and further evolved into a harmful one while on the other hand, COVID-19 can be mutated and evolved in the Netherland minks to become less harmful just like the case in small pot. This is indeed an interesting result obtained and may be used in the control of virus mutation for future. Actually, HKLam Theory can thus be applied as stated in the above section if the Butterfly effects or the Lorenz attractor exists in such kind of COVID-19 mutation in order to find the saddle or the equilibrium point [11]. Thus, this can then be acted as some form of control to harmful virus evolution. Last but not least, the SARS-CoV-2 virus mutates within HIV patients because of enzymes that copy RNA are prone to make errors [12]. This event implies the needs of the development in corresponding anti-enzymes of SARS-CoV-2 drugs to prevent continuous mutation in that virus among HIV patients. (**Remarks**: One may indirectly verify my idea about the SARS-CoV-2 becomes less harmful by interviewing those SARS-CoV-2 patients qualitatively (or even quantitatively) from the first generation to the latest sixth generation one in Hong Kong to determine their level of sickness etc.)

References

1. Lam Kai Shun (Dec, 2020) Scholar Journal of Physics, Mathematics and Statistics; Vol 7, Issue 12; p.398–412
https://www.saspublishers.com/media/articles/SJPMS_712_398–412.pdf

2. Zheng, Yinan, Hlady, Ryan A, Joyce, Brian T, Robertson, Keith D, He, Chunyan, Nannini, Drew R, Kibbe, Warren A, Achenbach, Chad J, Murphy, Robert L, Roberts, Lewis R, and Hou, Lifang. "DNA Methylation of Individual Repetitive Elements in Hepatitis C Virus Infection-induced Hepatocellular Carcinoma." *Clinical Epigenetics* 11.1 (2019): 145. Web.

3. Cho, Young-Dan, Kim, Pil-Jong, Kim, Hong-Gee, Seol, Yang-Jo, Lee, Yong-Moo, Ku, Young, Rhyu, In-Chul, and Ryoo, Hyun-Mo. "Transcriptomics and Methylomics in Chronic Periodontitiswith Tobacco Use: A Pilot Study." *Clinical Epigenetics* 9.1 (2017): 81. Web.

4. Chakrabarti, A., Membrez, M., Morin-Rivron, D. *et al.* Transcriptomics-driven lipidomics (TDL) identifies the microbiome-regulated targets of ileal lipid metabolism. *npj Syst Biol Appl* **3,** 33 (2017). https://doi.org/10.1038/s41540-017-0033-0

5. Snaebjornsson, Marteinn Thor, Janaki-Raman, Sudha, and Schulze, Almut. "Greasing the Wheels of the Cancer Machine: The Role of Lipid Metabolism in Cancer." *Cell Metabolism* 31.1(2020): 62-76. Web.

6. Hayes, Andrew F. *Introduction to Mediation, Moderation, and Conditional Process Analysis*. New York: Guilford Publications, 2014. Methodology in the Social Sciences. Web.

7. Gourley, Stephen A, Kuang, Yang, and Nagy, John D. "Dynamics of a Delay Differential Equation Model of Hepatitis B Virus Infection." *Journal of Biological Dynamics* 2.2 (2008): 140-53. Web.

8. Lam Kai Shun (Dec, 2020) Scholar Journal of Physics, Mathematics and Statistics; Vol 7, Issue 10; p.272-294 https://www.saspublishers.com/media/articles/SJPMS_710_272-294.pdf

9. Jung, Youn-Sang, and Park, Jae-Il. "Wnt Signaling in Cancer: Therapeutic Targeting of Wnt Signaling beyond β-catenin and the Destruction Complex." *Experimental & Molecular Medicine* 52.2 (2020): 183-91. Web.

10. WEIN, Lawrence M, WU, Joseph T, and KIRN, David H. "Validation and Analysis of a Mathematical Model of a Replication-competent Oncolytic Virus for Cancer Treatment: Implications forVirus Design and Delivery." *Cancer Research (Chicago, Ill.)* 63.6 (2003): 1317-324. Web.

11. Shun CL, A Saddle point finding method for lorenz Attractor through business machine learning algorithm, Journal of Physical Mathematics, Vol 11:4, 2020, https://www.hilarispublisher.com/open-access/a-saddle-point-finding-method-for-lorenz-attractor-through-business-machine-learning-algorithm.pdf

12. Ewen Callaway, The coronavirus is mutating — does it matter?, Sept 2020, Nature publishing https://www.nature.com/articles/d41586-020-02544-6