Scholars Journal of Physics, Mathematics and Statistics
Abbreviated Key Title: Sch J Phys Math Stat
ISSN 2393-8056 (Print) | ISSN 2393-8064 (Online)
Journal homepage: https://saspublishers.com/sjpms/

OPEN ACCESS

Big Data Investigation into the Causes and Treatment of Caries in Kindergarteners

Lam Kai Shun*

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

DOI: 10.36347/sjpms.2020.v07i10.005 | **Received:** 28.09.2020 | **Accepted:** 08.10.2020 | **Published:** 11.10.2020

***Corresponding author:** Lam Kai Shun

Abstract **Original Research Article**

It has been thousands of years that tooth decay is a health problem among human beings (Chu, 2000). The disease is like our common daily influenza. The aim of this paper is to use the heated topic big data analysis and its related statistical mathematics to predict the possible behavior behind kindergarten children tooth-care response — a predictive medicine for the prevention. Moreover, the paper also develops a thought experiment from the Bayes' Decision tree. The aim is to determine some suitable strategies in the case of kindergarten tooth-caries — for regenerative medicine. In the research, I have found that Butterfly Effect can form a predictive philosophy. I rationale it with Bayes Theory and map each outcome with the corresponding Domino effects (Heinrich Theorem). While in the middle part, I insert random variables respectively as the connection. This event forms a completely new theory which can catch the chaos and dominos of the Butterfly Effects (philosophy) or the so-called Lorenz system. I propose the name should be the (HKLam's) Net-Seizing Theory. When my theory is expressed in terms of linear transformation, random matrix and regression, one may use it in the prediction model (for the posterior distribution) of human behaviour etc. If we construct a reversed Bayesian tree with all necessary posterior predicted distributions (models), we may get the wanted forecasting prior (distribution) model and is recursive formalism or the Bayes filter. We may establish the corresponding regression tree etc. Furthermore, with the Savage theory, one can apply the machine learning technique to generate the necessary policy for handling the social problem which is just like the child's tooth care shown in this research paper — there is a need to subsidy our kindergarten as early as possible for the result of best social return.

Keywords: Kindergarteners Data Causes, Butterfly Effects.

INTRODUCTION

Tooth decay is a common health issue as prevalent as influenza. Dental caries can often occur in children and young adults before the appearance of permanent teeth. Research has shown that fluoride can prevent tooth decay effectively. Recent studies have proven that the chemical compound, silver diamine fluoride when it is using in the tooth decay, it can treat dental carious lesions. Hence, the primary focus of this paper is to investigate the causes of cavities in kindergarteners (as they have milk teeth) in Hong Kong [3-6 years or students from nursery (K1) to upper kindergarten (K3)] and the, efficky of fluoride treatment. Then, one may set up the wanted predictive model for cavity's development and hence develop a plan for the strategies of regenerative medicine in our teeth.

Background — the Decision making that comes from the Butterfly

The forward part of my decision making body (like the Butterfly Effect) is first illustrated from the divergent part of a Bayesian tree diagram that follows by a series of domino events using for particularly in the generation of a suitable policy. (See Fig. 1)

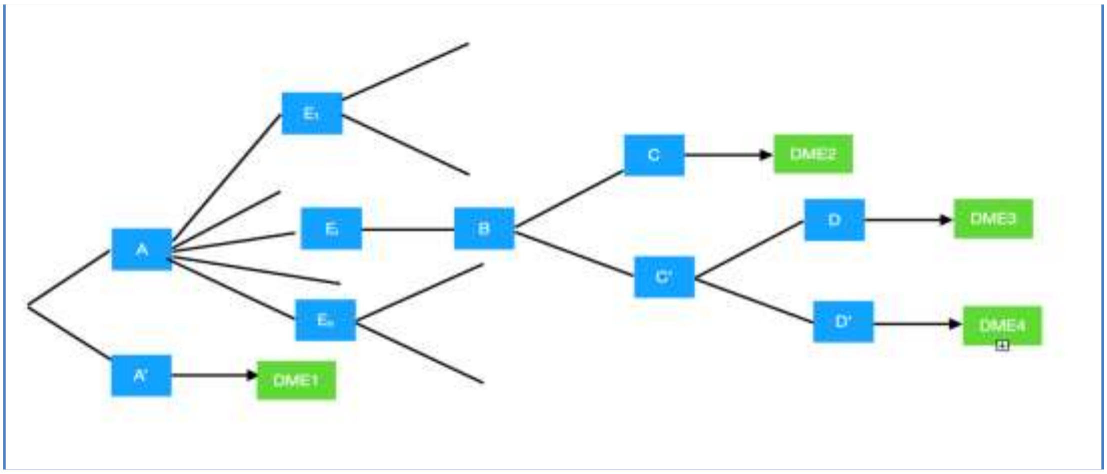


Fig-1: Divergent part of my proposed decision making body

E_1, E_2, E_n are independent events with different probabilities Pr_1, Pr_2, Pr_n acts as the outcome probabilities. A and E_1, E_2, E_n have their respective causal dependencies.

For the convergent part of Bayes theory plus Heinrich one (like the second part of the mirrored image of the Bayes Theorem) is mapped to those domino consequences (Dom_1, Dom_2, Dom_3,…Dom_n) corresponds to the different possibilities event E_1, E_2, …, E_n and the probabilities. Hence, one can use the Bayesian tree with the domino events (e.g., Dom_n,…,Dom_2, and Dom_1) to evaluate the efficiency of the hidden Markov Model.(SEM).

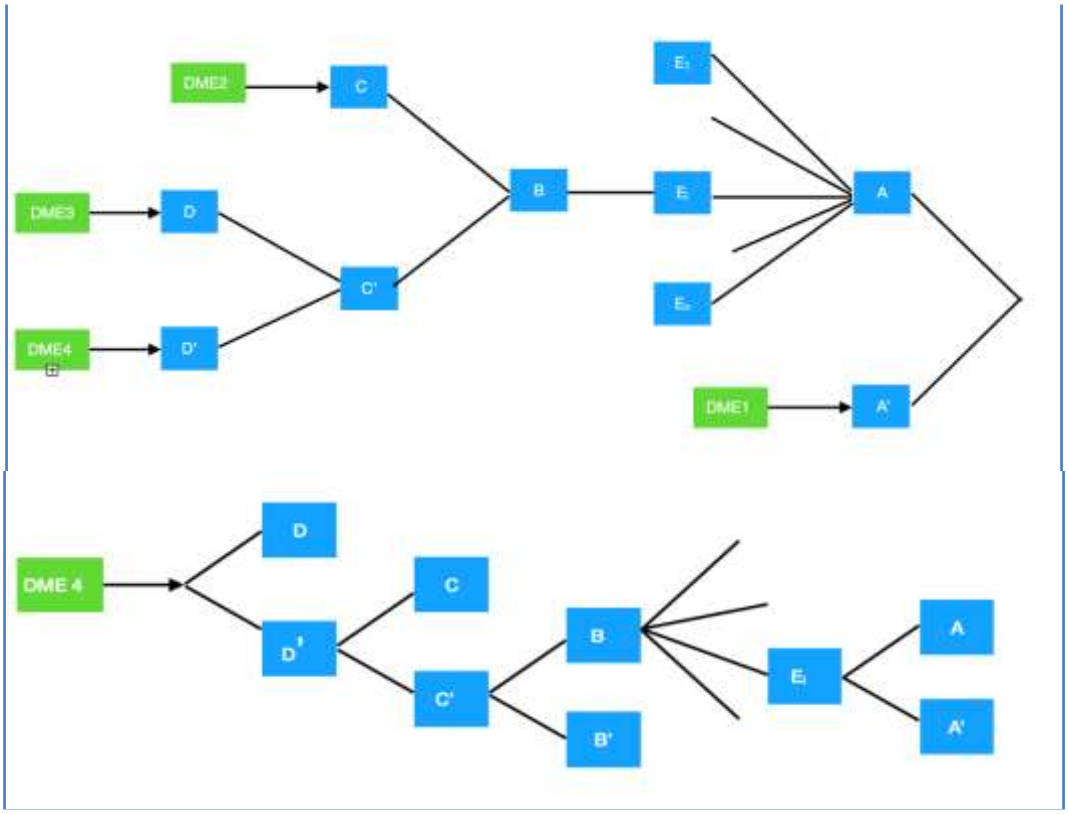


Fig-2: The convergent part of my proposed decision making body

Thirdly, for the series of domino effect and the converse of the Bayes Theorem, they constitute the backward part of my proposed decision making body.

Indeed, DME 4 represents the corresponding domino consequence event 4 which is mapped to the inverse Bayes Tree with the outcome event D and D'. One may then move upward the tree until event A and A'. I note that both the forward and the backward parts of the proposed philosophy come from the Bayes theory and the Heinrich Theorem. Thus I summarise my proposed one in catching the Butterfly Effects (maybe I name it as the (HKLam's) Net-Seizing theory) like the following:

1. A Bayes' probability tree diagram [1] which lists all the feasible outcomes of an event (Forward and Backward section from Bayes, Convergent and Divergent Bayesian Tree from Decision Making — Savage Theory [1]);
2. A series of domino effect consequences or the Heinrich Theory in risk management and engineering [2];
3. The immediate part (random variables) of linear mapping [3] or transformation is added such that one may rationale [4] the philosophy. Hence, things can be approximated through the regression model.

My Original Discovery (Contribution) in the Linear Mapping of the Bayes Theorem

Consider the following imaginary statistical experiment that is often used when teaching statistics:

One should first toss three (or more) coins together in front of the participants, such that heads are on the front and tails are on back. The participants are then required to list all possible outcomes after tossing the coins.

In general, the set of all feasible results (outcomes) would be:

S = {HHH, HHT, THH, HTT, THT, TTH, TTT} (suppose one will obtain such expected result). There are also other combinations and permutations of heads and tails when each of them are assigned with a number. Hence, the outcome space is not just unique.)

Finally, the participants must count the number of tails for each possible outcome and then list them on a table, as seen below:

Possible Outcomes (s_i)	**No. of tails obtained t_i**
HHH	**0**
HHT	**1**
HTH	**1**
THH	**1**
HTT	**2**
HTT	**2**
TTH	**2**
TTT	**3**

(Clearly, the above table does NOT list all possible outcomes for tails.)

A function f (which should be called the random variable) must also be considered. Intuitively, f maps all the possible outcomes (s_i) to the number of tails T_i

f_i: $s_i \longmapsto t_i$ (i.e. no. of tails obtained as it takes values t = 0,1,2,3 … by HKU, Dr. K.T. Leung's lecture note 1993-1994)
or strictly speaking

f: $S \longrightarrow T$ (i.e. **R – Real Number**)

Actually, through the linear mapping, the non-linear chaos can be converted into the linear one through a suitable choice of random variables. One of the cases is the mapping between Bayes outcome space, random variables and the domino. Hence one can find the matching linear regression model approximations for different research areas. Obviously, in such sense, we can "catch the chaos & the dominos". This event is known as capturing the "Butterfly Effects (or the philosophy)" behind. Thus, I propose such kind of methodology should be termed as the HKLam's Net-Seizing theory. When my theory is expressed in terms of both mathematics and statistics, it becomes

[Random Matrix] [LT] = Linear regression of the causal domino effects
And
[Linear regression of the causal domino effects] [LT] = [Random Matrix]

For the first part of Bayes plus Heinrich, one may further get a diagram (figure 3) that rationales it and gets the structure of my Net-Seizing theory. Actually, one should insert the probability that multiplied with weight Pri*Wgi between the random variables RVi and the last branch of event Ei. For the converse part of my Net-Seizing theory, (figure 4) each domino causal effect is corresponding to a suitable linear transformation (random variable), and further mapped to a Bayesian tree. Hence, one can then express my net-seizing theory in terms of mathematics and statistics and get the following diagram (figure 5) in matrix notation. The convergent (or the backward but NOT the convergent — mirror part) of the matrix notation diagram (figure 6) is also true.

1 HTTPS://PLATO.STANFORD.EDU/ENTRIES/DECISION-THEORY/#JEFTHE

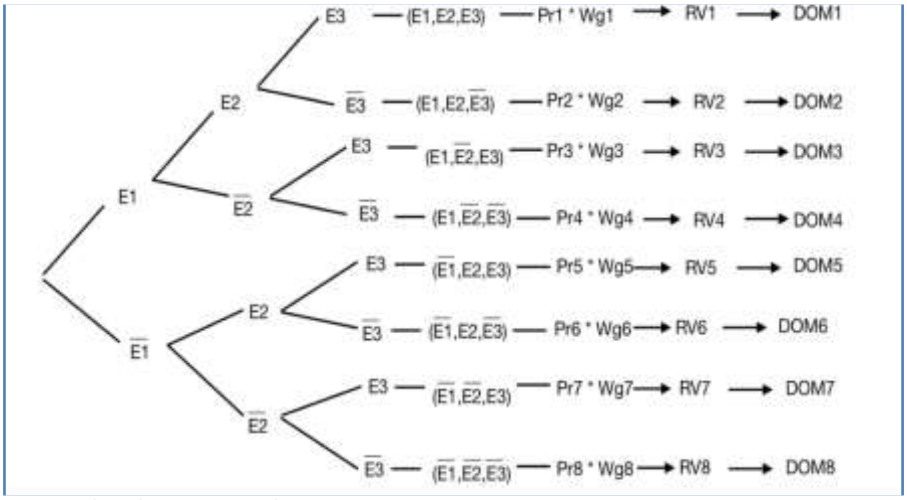


Fig-3: The rationalisation of the forward part of my proposed "Net-seizing" theory with random variables as the immediate connection between events and dominos.

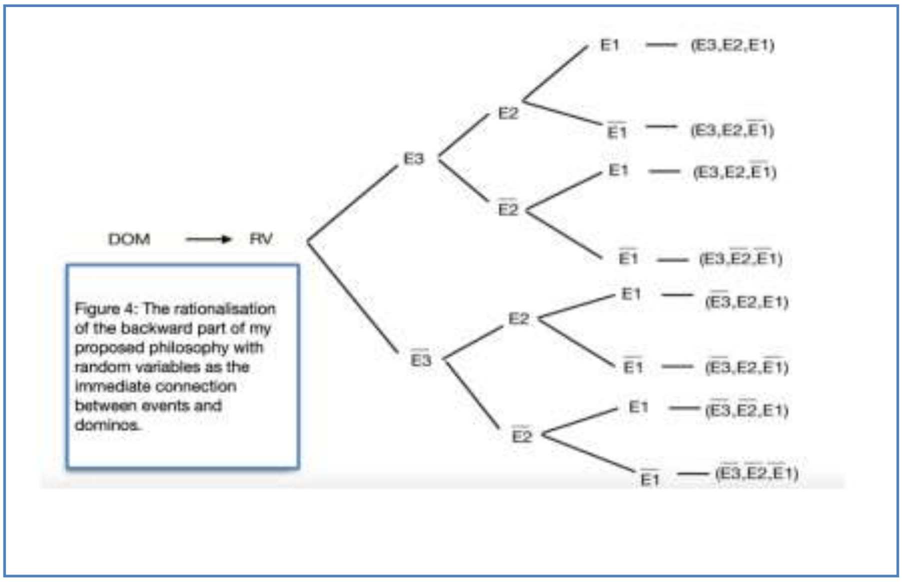


Fig-4: The rationalisation of the converse part of my proposed "Net-Seizing" theory with random variables as the immediate connection between events and dominos.

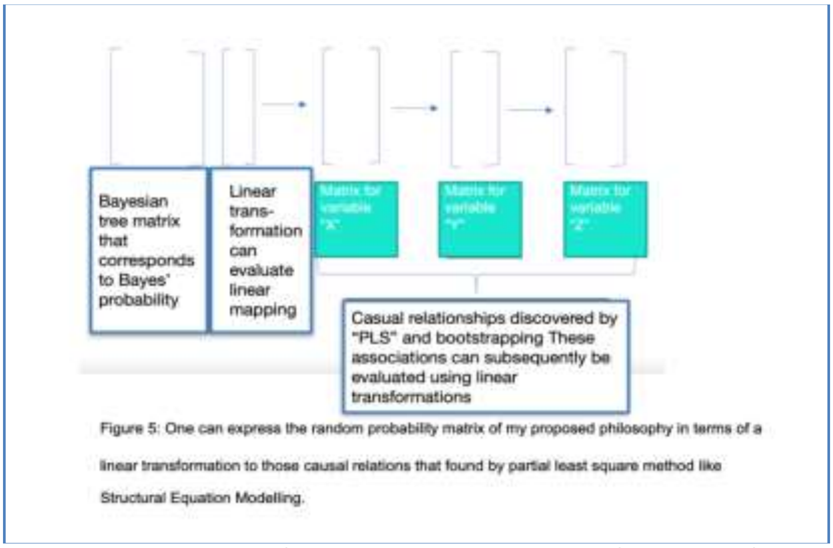


Fig-5: One can express the random probability matrix of my "Net-Seizing" theory in terms of a linear transformation to those causal relations that found by the partial least square method like the Structural Equation Modelling.

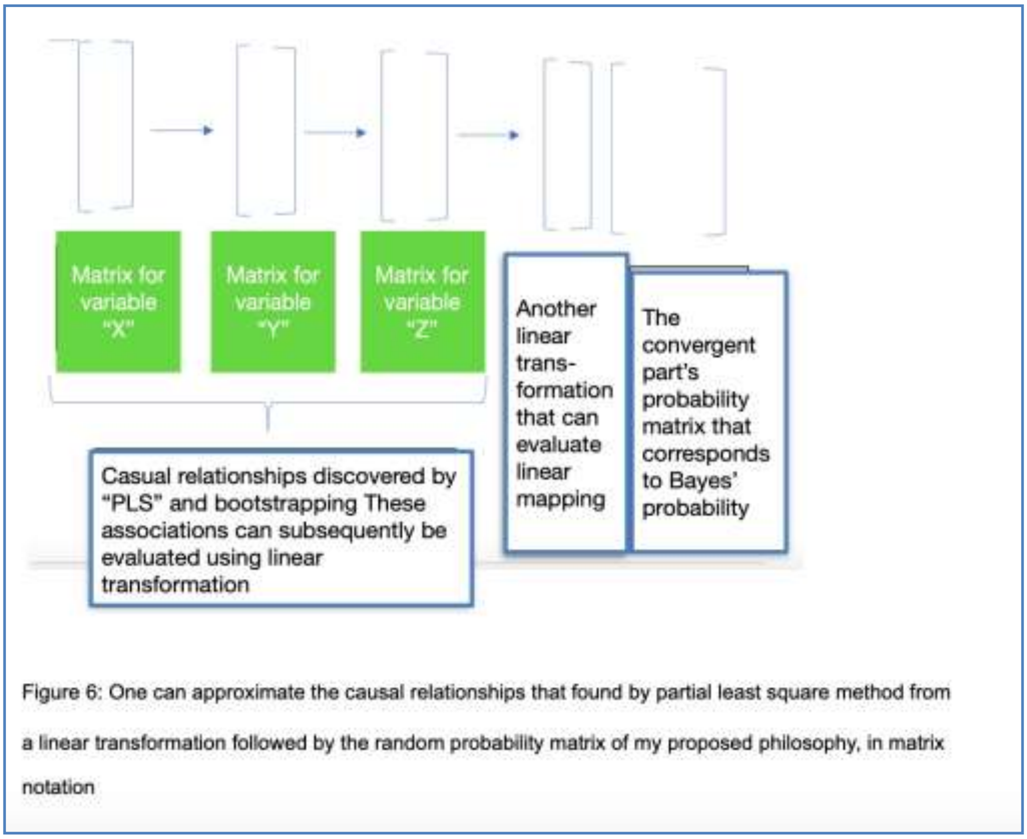


Fig-6: One can approximate the causal relationships that found by the partial least square method from a linear transformation following by the random probability matrix of my "Net-Seizing" theory, in matrix notation

The study of the students reading behaviour in Lam Nov, 2018 can be acted as my proposed philosophy's example. The following diagrams outline the corresponding inverseBayesian tree diagram and domino consequences such that one may find the respective conditional probabilities for the convergent part of my proposed philosophy:

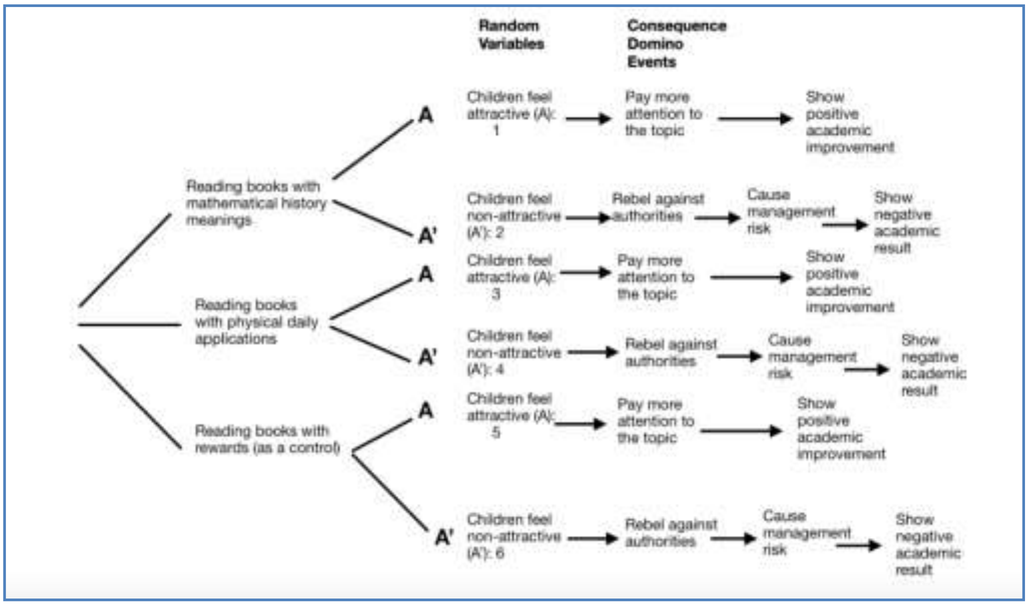


Fig-7: The forward pat of the students' reading behaviour, random variables and the domino consequences (Similar to Savage Theory in decision making)

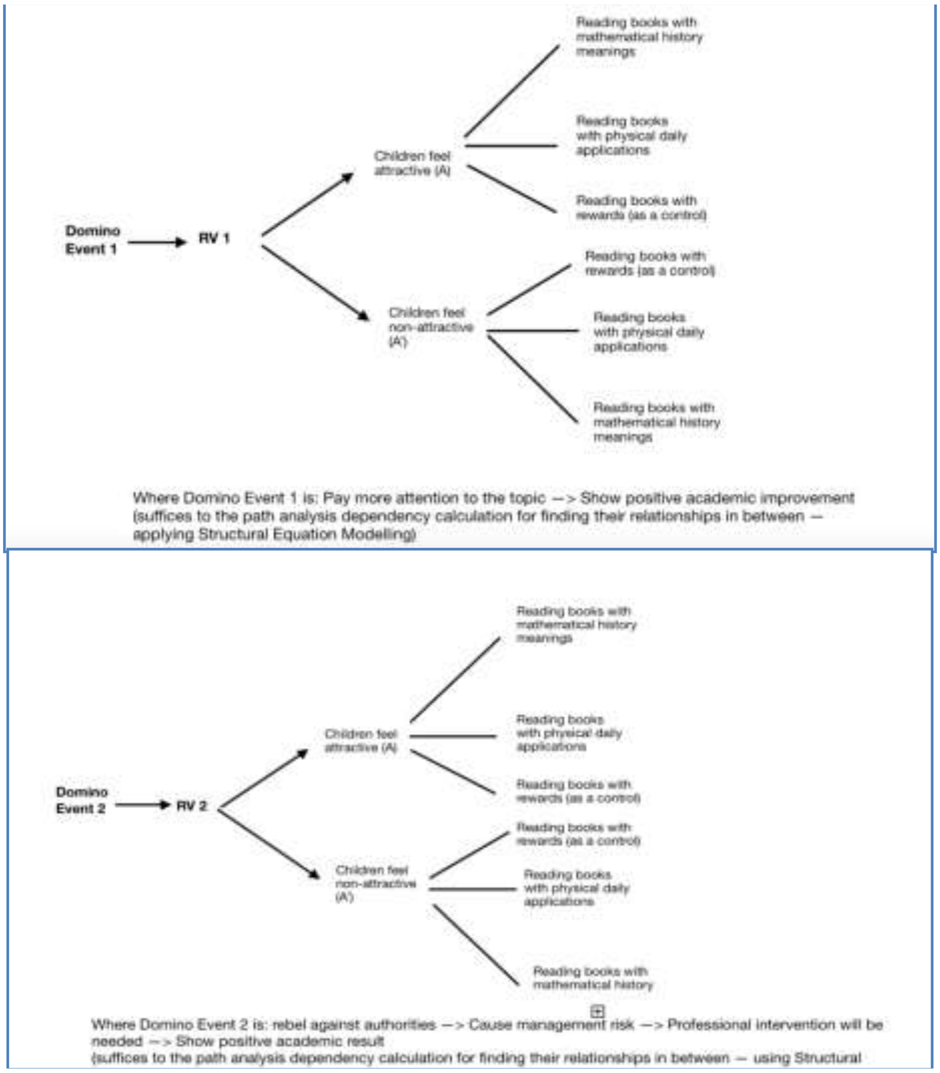


Fig-8: The backward parts of the students' reading behaviour and the domino consequences

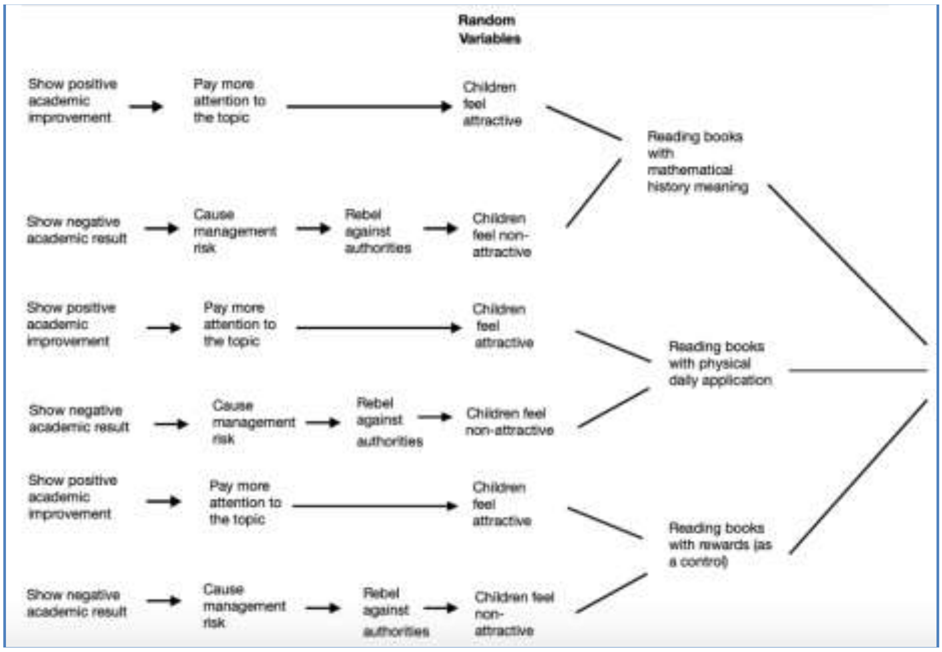


Fig-9: The convergent parts (mirrored image of Bayes' Theory) of the students' reading behaviour and the domino consequences

Practically, we can calculate the corresponding conditional probabilities of the convergent part of Bayes plus Heinrich from the inverse of the Bayesian tree diagram like the following:

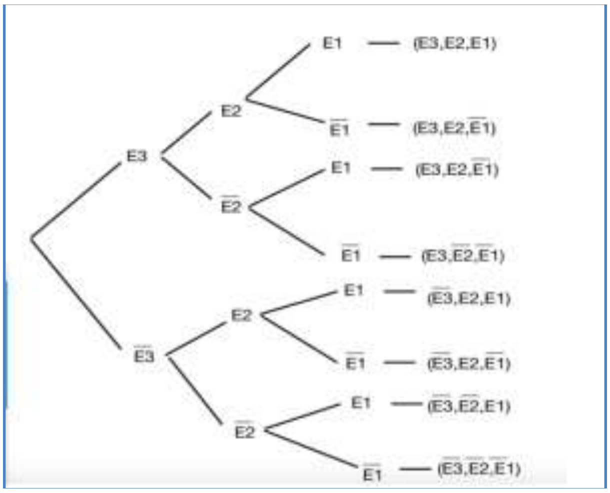


Fig-10: The inverse part of the Bayes theorem for computing the required conditional probabilities to the convergent part of my proposed philosophy

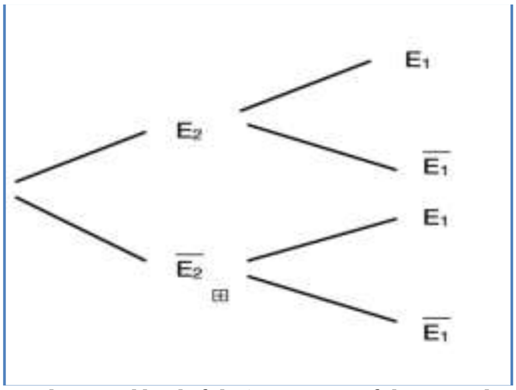


Fig-11: The second level of the inverse part of the Bayes theorem for computing the required conditional probabilities to the convergent part of my proposed philosophy

At the same time, with referencer to the results in Lam Dec [4] & Lam March [5], this author discovers that one may always express the Bayesian probability tree in terms of a matrix. The outcome matrix can finally be approximated by applying the Bayesian Linear regression. Or we may have the following results (assume the square matrix operation while the non-squared one can be converted to the squared by Matlab coding which is out of the scope of the present paper):

$$\begin{pmatrix} P(A_1,\mu_1,\sigma_1^2) \\ P(A_2,\mu_2,\sigma_2^2) \\ . \\ . \\ . \\ P(A_n,\mu_n,\sigma_n^2) \end{pmatrix} = \begin{pmatrix} x_{11} & . & . & . & x_{1n} \\ x_{21} & . & . & . & x_{2n} \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{n1} & . & . & . & x_{nn} \end{pmatrix} \begin{pmatrix} \beta'_1 \\ \beta'_2 \\ . \\ . \\ . \\ \beta'_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{pmatrix}$$

But, the probability matrix A can be expressed as:

$$\begin{pmatrix} P(A_1|B_1) & P(A_1|B_2) & . & . & . & P(A_1|B_n) \\ P(A_2|B_1) & P(A_2|B_2) & . & . & . & P(A_2|B_n) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ P(A_n|B_1) & P(A_n|B_2) & . & . & . & P(A_n|B_n) \end{pmatrix} \begin{pmatrix} P(B_1) \\ P(B_2) \\ . \\ . \\ . \\ P(B_n) \end{pmatrix} = \begin{pmatrix} P(A_1) \\ P(A_2) \\ . \\ . \\ . \\ P(A_n) \end{pmatrix}$$

Therefore, accordingly, we have:

$$\begin{pmatrix} P(A_1|B_1) & P(A_1|B_2) & . & . & . & P(A_1|B_n) \\ P(A_2|B_1) & P(A_2|B_2) & . & . & . & P(A_2|B_n) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ P(A_n|B_1) & P(A_n|B_2) & . & . & . & P(A_n|B_n) \end{pmatrix} = \begin{pmatrix} P(A_1) \\ P(A_2) \\ . \\ . \\ . \\ P(A_n) \end{pmatrix} \begin{pmatrix} P(B_1) \\ P(B_2) \\ . \\ . \\ . \\ P(B_n) \end{pmatrix}^{-1}$$

Substitute back the right hand side of the first equality into the above equation, we have:

$$\begin{pmatrix} P(A_1|B_1) & P(A_1|B_2) & . & . & . & P(A_1|B_n) \\ P(A_2|B_1) & P(A_2|B_2) & . & . & . & P(A_2|B_n) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ P(A_n|B_1) & P(A_n|B_2) & . & . & . & P(A_n|B_n) \end{pmatrix} = \left\{ \begin{pmatrix} x_{11} & . & . & . & x_{1n} \\ x_{21} & . & . & . & x_{2n} \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{n1} & . & . & . & x_{nn} \end{pmatrix} \begin{pmatrix} \beta'_1 \\ \beta'_2 \\ . \\ . \\ . \\ \beta'_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{pmatrix} \right\} \begin{pmatrix} P(B_1) \\ P(B_2) \\ . \\ . \\ . \\ P(B_n) \end{pmatrix}^{-1}$$

But from the result of Lam March, 2020, the conditional probability matrix A (given B) can also be expressed in terms of linear regression through a suitable linear transformation (LT), thus we have:

$$\left\{\begin{pmatrix} x_{11} & . & . & . & x_{1n} \\ x_{21} & . & . & . & x_{2n} \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{n1} & . & . & . & x_{nn} \end{pmatrix}\begin{pmatrix} \beta'_1 \\ \beta'_2 \\ . \\ . \\ . \\ \beta'_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{pmatrix}\right\}\begin{pmatrix} P(B_1) \\ P(B_2) \\ . \\ . \\ . \\ P(B_n) \end{pmatrix}^{-1} = \{\mathbf{X} + (\mathbf{XB}+\mathbf{B}_0) + [(\mathbf{XB}+\mathbf{B}_0)\mathbf{D} + \mathbf{D}_0]\}*(\mathrm{LT})^{-1}$$

The above outcome implies that when the Bayesian linear regression is multiplied by the inverse of the event B's probability matrix, it can be well approximated through an ordinary linear regression equation multiplies with the inverse of the linear transform of conditional probability matrix A (given B). Or one can find out the inverse of the event B's probability matrix if we know the inverse of Bayesian regression.

Hence, one can apply both linear regression and the Bayesian regression (for predictive model) methods for the approximation to a random matrix and more about the inverse of probability matrix B:

$$\begin{pmatrix} P(B_1) \\ P(B_2) \\ . \\ . \\ . \\ P(B_n) \end{pmatrix}^{-1} = \left\{\begin{pmatrix} x_{11} & . & . & . & x_{1n} \\ x_{21} & . & . & . & x_{2n} \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{n1} & . & . & . & x_{nn} \end{pmatrix}\begin{pmatrix} \beta'_1 \\ \beta'_2 \\ . \\ . \\ . \\ \beta'_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{pmatrix}\right\}^{-1} * \{\mathbf{X} + (\mathbf{XB}+\mathbf{B}_0) + [(\mathbf{XB}+\mathbf{B}_0)\mathbf{D} + \mathbf{D}_0]\}*(\mathrm{LT})^{-1}$$

Substitute back into the equation: [**Pr(A|B)]** = **X β' β** $^{-1}$ + **ε**', We have:

$$[\mathbf{Pr(A|B)}] = \mathbf{X\beta}'* \left[\left\{\begin{pmatrix} x_{11} & . & . & . & x_{1n} \\ x_{21} & . & . & . & x_{2n} \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{n1} & . & . & . & x_{nn} \end{pmatrix}\begin{pmatrix} \beta'_1 \\ \beta'_2 \\ . \\ . \\ . \\ \beta'_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{pmatrix}\right\}^{-1} * \{\mathbf{X} + (\mathbf{XB}+\mathbf{B}_0) + [(\mathbf{XB}+\mathbf{B}_0)\mathbf{D} + \mathbf{D}_0]\}*(\mathrm{LT})^{-1}\right] + \varepsilon''$$

To sum up, one can always express the Bayesian Linear regression of the probability matrix of event A in terms of the probability matrix B, an ordinary linear regression and the inverse of the linear transform to conditional probability matrix A (given B). At the same time, the conditional probability matrix A (given B) can be expressed as the substitution back of the product among (the inverse of Bayesian linear regression, an ordinary linear regression and the inverse of the linear transform) into the Bayesian linear regression equation of conditional probability matrix A. One of the applications of the above matrix equalities (or equations) is in the field of quantum computing. One can always use the causal relationships (quantum switch) and a random matrix etc to model different quantum systems in an approximated manner. The final result is shown in the above matrix expression (*) for our conditional probability A (and B) — a random matrix modelled to a quantum system.

In much a similar manner, one may also use the aforementioned matrix expression (*) to model our living organisms' mutation together with its cause and effect. The result can then be applied in tissue engineering or regeneration. In practice, the method to find out the inverse of a matrix is an interesting topic for our university's advanced mathematics study (mathematics major students) or the course like Matrix Theory and its application etc.

MATERIALS AND METHODS

This research aims to answer the following questions:

1. What is the level of tooth decay observed in K1 kindergarteners at the time of the first examination?
2. What is level of tooth decay observed in older [lower kindergarten (K2), K3] kindergarteners after fluoride treatment?
3. Why we need fluoride treatment and what are the implications of it?
4. What the predicted outcomes will see from K2 and K3 children fluoride treatment during the prevention of tooth decay?
5. How should the authorities develop corresponding strategies for regenerative medicine of teeth?

The investigation will comprise the correlation and survey methods to answer these questions. First, one will use quantitative statistical research methods to determine the relationships between the different variables. Second, one will also use a qualitative survey to improve the understanding of the implications of fluoride treatment in kindergarteners.

Both, the correlation studies and survey will find out the implications of fluoride treatment before analysing the causes. Finally, the correlation design will help in predicting the incidence of tooth decay after fluoride treatment.

The study will begin by making records of local Hong Kong kindergartens belonging to K1 and K2 [6]. Subsequently, one will use tailor-made software to pull the necessary data into MS Excel. One will record down the variables like name, exam and birth dates, the number of decayed and infected teeth, treatment required etc. From, the collection of these data, one may apply them to Bayesian inference and Bayesian regression from the correlation relationship via simulation using software such as R. Hence, one can set up the needed predictive model for caries etc. While it is critical that sample size, the value of precise measurements, and the use of unbiased samples during the quantitative research. This study will use both explanatory and predictive designs. We propose a few variables for the former: level of tooth decay observed at the first examination, the status of cavities after fluoride treatment, and the number of missing teeth. Subsequently, the researcher will use a model comprising these variables, and the number of decayed teeth. The researcher will compare one group with decayed teeth after fluoride treatment with another group with decayed teeth without fluoride treatment, i.e., the control group. The total number of participants will be 180,000. Indeed, there is a prediction research design. The investigator can identify those specific variables to help predict the behavioral outcomes from the prediction model setting up before we have mentioned. This study concerns with fluoride treatment, the level of tooth decay, and how it can aid in the prediction of healthy teeth. Tooth decay is a time-related variable. The research will correlate the predictor and outcome variables. We will use multiple regressions to determine the predictor's effects on the variables. Hence, one can find out the desired causal relationships with a structural equation model or partial least square method from the collected data. Thus, one can predict the outcomes of after the fluoride treatment.

A longitudinal study [7] involves repeated observations of the same variables (e.g., population) over a certain period. Thus, we will also conduct qualitative yearly interviews (until K3) with the previous (K1) kindergarten students. The researcher will ask questions from an interview guide, listen to the answers and/or see behavior, and record the responses. The interviewer asks open-ended questions, without options for responses, and records the answers provided in a qualitative interview. We usually use a longitudinal study to identify the reasons for the development of outcome implications (e.g., the implications of fluoride treatment). This information can guide the plan to develop better strategies for regenerative medicine for teeth.

There is being some ethical considerations while handling the mankind type of big data investigation. The concerns may include privacy, informed consent, data ownership and transparency etc. This applicant believes that the researchers can solve those public worries if they can eliminate those misunderstandings among the investigated respondents.

RESULTS AND DISCUSSION

This applicant tries to reference the wanted research results (as suggested taking place in Hong Kong Kindergarten School) through using those data obtained from the United State Centre for Disease Control. The major working steps are:

1. Developing a suitable predictive philosophy such as the Butterfly one (as shown in the previous section) and rationalising it by my Net-Seizing Theory;
2. Using the Structural Equation model method to find out the hypothesis's causal relationships, so as the one indicated by Australia's influenza case in Lam June 2019;
3. Convert the data (collected from U.S.A CDC Department) into a time series[2]. Perform the Granger Causal Connectivity Analysis [8] by the software EViews [3]. Fit those data into the maximum likelihood estimation (or using Hayes' mediated-moderated model) to approximate the causality behavior with multilevel modelling whenever necessary. Then, one can continue the Bayesian estimation;
4. Applying Bayesian Inference to find the Prior and Likelihood distribution of the collected data. Hence, one can find the deserved Posterior distribution from the United States' tooth cavity's data. SPSS software can perform all the mentioned work;
5. After finding the parameters (mean and standard derivation) for the posterior distribution, one may then use the Bayesian Regression for the following step of posterior's estimation. This applicant suggests one should use the software JASP for the calculation of the wanted Bayesian regression model (or

[2] Indeed, the transformed time series data can also be used for prediction and can be used for verification of my proposed forecasting method here. The verification is out of the scope of this paper. (Ref: Babii.A, (2020) Machine learning time series regressions with an application to nowcasting, Social Science Research Network)

3 One may also perform Causal Inference through the Bayesian Additive Regression Tree.

equation), hence one may predict the behavior of a child with tooth cavities;

6. From the predicted posterior distributions (models) together with the reversed Bayesian tree, one may further construct the forecasting prior distribution (model) for our child tooth cavities. This is indeed a kind of recursive formalism;
7. Establish the corresponding (probability) decision tree (or random forest) or even Bayesian Adaptive tree for Bayesian Causal Inference from our tooth statistics; the aim is to determine how well is the tooth situation in the countries like the U.S.A. and thus create some suitable strategies in the regenerative medicine for the case of teeth;
8. Implement the strategies or policies found in the previous step.

The following is the outcomes for Bayesian Inference, where we can obtain the outputted tables from using the software SPSS:

群組統計量

	Indicator_Value	數目(C)	平均值	標準偏差	標準誤平均值
Data_Value	= 1	72	50.110	15.5512	1.8327
	= 2	46	45.896	12.8593	1.8960

獨立樣本平均值的事後分布特徵[a]

		後段		95% 信賴區間	
	眾數(O)	平均值	變異(V)	下限	上限
Data_Value	-4.214	-4.214	7.218	-9.490	1.062

a. 變異的事前：Diffuse。平均值的事前：Diffuse。

群組統計量

	Indicator_Value	數目(C)	平均值	標準偏差	標準誤平均值
Data_Value	= 1	72	50.110	15.5512	1.8327
	= 3	72	20.844	8.8841	1.0470

獨立樣本平均值的事後分布特徵[a]

		後段		95% 信賴區間	
	眾數(O)	平均值	變異(V)	下限	上限
Data_Value	-29.265	-29.265	4.584	-33.470	-25.061

a. 變異的事前：Diffuse。平均值的事前：Diffuse。

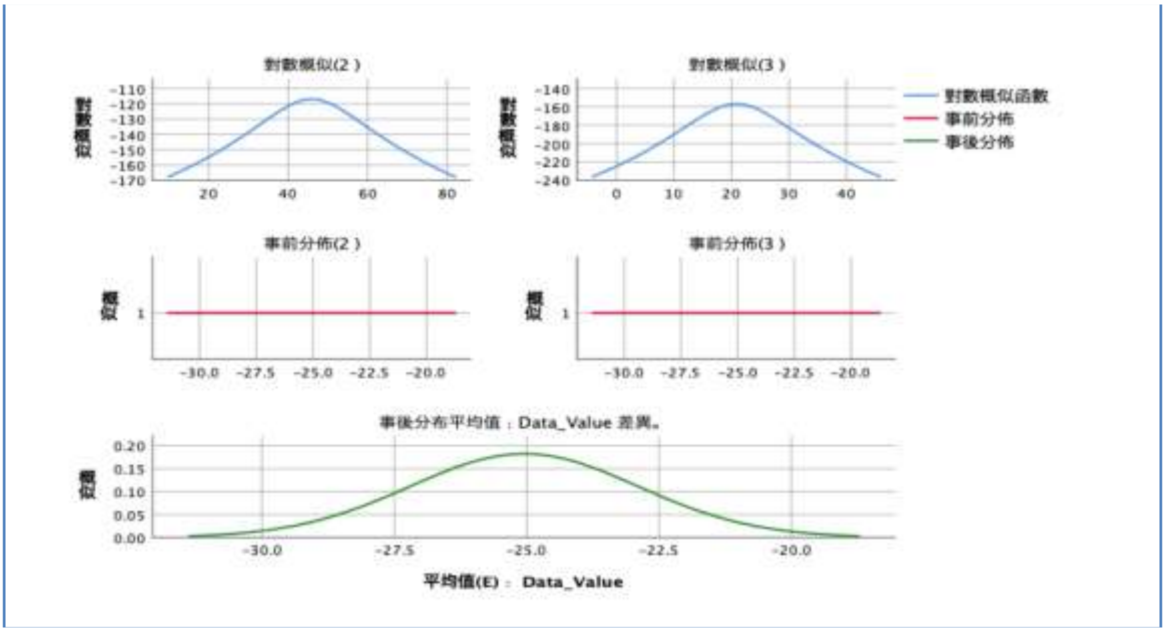
群組統計量

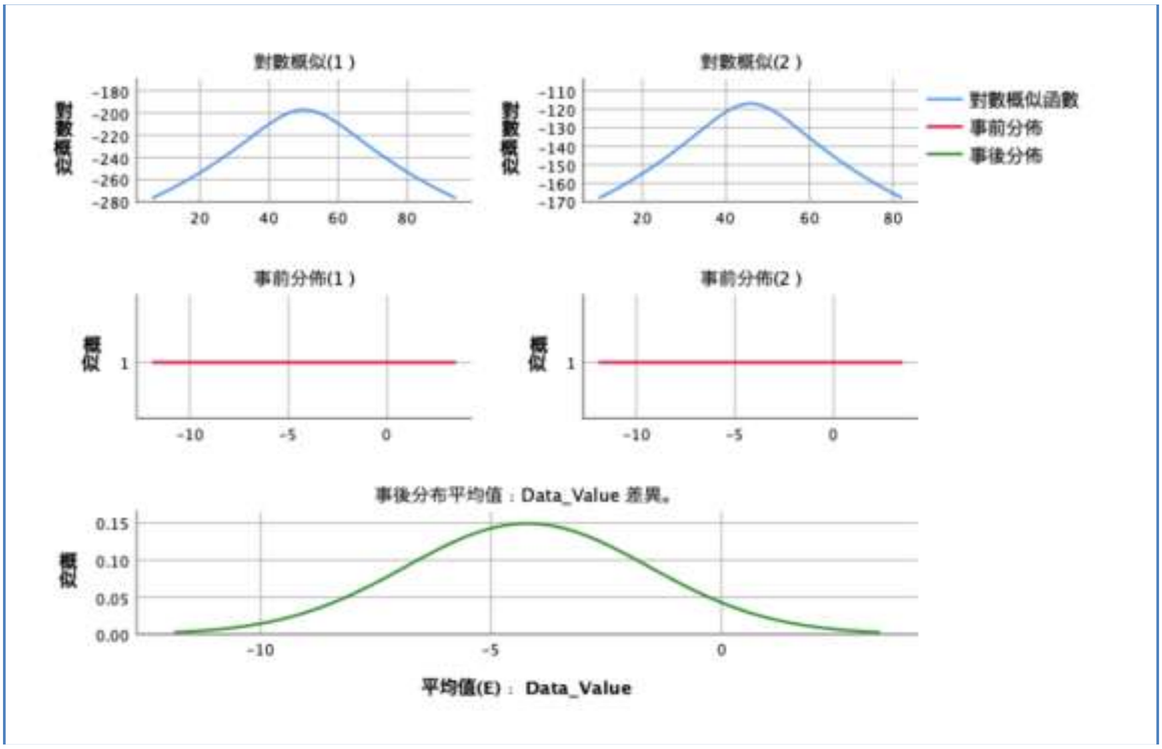
	Indicator_Value	數目(C)	平均值	標準偏差	標準誤平均值
Data_Value	= 2	46	45.896	12.8593	1.8960
	= 3	72	20.844	8.8841	1.0470

獨立樣本平均值的事後分布特徵[a]

		後段		95% 信賴區間	
	眾數(O)	平均值	變異(V)	下限	上限
Data_Value	-25.051	-25.051	4.890	-29.398	-20.704

a. 變異的事前：Diffuse。平均值的事前：Diffuse。


對數概似(2)
對數概似(3)
對數概似函數
事前分佈
事後分佈
事前分佈(2)
事前分佈(3)
事後分佈平均值：Data_Value 差異。
平均值(E)： Data_Value


對數概似(1)
對數概似(2)
對數概似函數
事前分佈
事後分佈
事前分佈(1)
事前分佈(2)
事後分佈平均值：Data_Value 差異。
平均值(E)： Data_Value

Bayesian Linear Regression ▼

Model Comparison

Models	P(M)	P(M|data)	BF_M	BF_{10}	R^2
Null model (incl. Response_Rate, Percent_eligible_for_the_NSLP_SampleStudents)	0.500	0.620	1.632	1.000	0.353
Sample_Size	0.500	0.380	0.613	0.613	0.379

Note. All models include Response_Rate, Percent_eligible_for_the_NSLP_SampleStudents.

Posterior Summary

Posterior Summaries of Coefficients

						95% Credible Interval	
Coefficient	Mean	SD	P(incl)	P(incl|data)	$BF_{inclusion}$	Lower	Upper
Intercept	24.881	2.300	1.000	1.000	1.000	19.980	29.783
Response_Rate	0.135	0.096	1.000	1.000	1.000	−0.071	0.340
Percent_eligible_for_the_NSLP_SampleStudents	0.196	0.215	1.000	1.000	1.000	−0.261	0.654

The regression equation (or model) is: Data_Value = 0.135*(Response_Rate) + 0.196*(Percent_eligible_for_the_NSLP_SampleStudents) + 24.881

Bayesian Linear Regression ▼

Model Comparison

Models	P(M)	P(M|data)	BF_M	BF_{10}	R^2
Null model (incl. Sample_Size, Response_Rate)	0.500	0.647	1.832	1.000	0.369
Percent_eligible_for_the_NSLP_SampleStudents	0.500	0.353	0.546	0.546	0.379

Note. All models include Sample_Size, Response_Rate.

Posterior Summary ▼

Posterior Summaries of Coefficients ▼

						95% Credible Interval	
Coefficient	Mean	SD	P(incl)	P(incl|data)	$BF_{inclusion}$	Lower	Upper
Intercept	24.881	2.271	1.000	1.000	1.000	20.041	29.722
Sample_Size	0.001	0.001	1.000	1.000	1.000	−0.001	0.004
Response_Rate	0.060	0.140	1.000	1.000	1.000	−0.238	0.358

The regression equation (or model) is: Data_Value = 0.001*(Sample_Size) + 0.06*(Response_Rate) + 24.881

Bayesian Linear Regression ▼

Model Comparison

Models	P(M)	P(M|data)	BF_M	BF_{10}	R^2
Null model (incl. Sample_Size, Percent_eligible_for_the_NSLP_SampleStudents)	0.500	0.635	1.737	1.000	0.362
Response_Rate	0.500	0.365	0.576	0.576	0.379

Note. All models include Sample_Size, Percent_eligible_for_the_NSLP_SampleStudents.

Posterior Summary

Posterior Summaries of Coefficients

						95% Credible Interval	
Coefficient	Mean	SD	P(incl)	P(incl|data)	$BF_{inclusion}$	Lower	Upper
Intercept	24.881	2.284	1.000	1.000	1.000	20.013	29.750
Sample_Size	0.001	0.001	1.000	1.000	1.000	−6.889e −4	0.004
Percent_eligible_for_the_NSLP_SampleStudents	0.073	0.262	1.000	1.000	1.000	−0.485	0.631

The regression equation (or model) is: Data_Value = 0.001*(Sample_Size) + 0.073*(Percent_eligible_for_the_NSLP_SampleStudents) + 24.881

This applicant notes that one can use the similar techniques as listed above in modelling the data_value corresponding to three different indicators of the tooth treatment data. Hence, one can predict the tendency of tooth decay that relates to the children daily behavior. One of the tooth indicators — Untreated Tooth Decay Bayesian Linear Regression table (from statistical software JASP) is listed below in the next page.

Bayesian Linear Regression ▼

Model Comparison ▼

Models	P(M)	P(M|data)	BF_M	BF_{10}	R^2
Sample_Size + Response_Rate + Percent_eligible_for_the_NSLP_SampleStudents	0.250	0.544	3.578	1.000	0.202
Response_Rate + Percent_eligible_for_the_NSLP_SampleStudents	0.083	0.430	8.292	2.371	0.196
Response_Rate	0.083	0.020	0.227	0.111	0.157
Sample_Size + Response_Rate	0.083	0.006	0.067	0.033	0.162
Null model	0.250	6.809e −9	2.043e −8	1.252e −8	0.000
Percent_eligible_for_the_NSLP_SampleStudents	0.083	2.643e −9	2.908e −8	1.458e −8	0.020
Sample_Size + Percent_eligible_for_the_NSLP_SampleStudents	0.083	6.303e −10	6.933e −9	3.476e −9	0.021
Sample_Size	0.083	3.674e −10	4.041e −9	2.026e −9	0.001

Posterior Summary

Posterior Summaries of Coefficients

Coefficient	Mean	SD	P(incl)	P(incl|data)	$BF_{inclusion}$	95% Credible Interval Lower	95% Credible Interval Upper
Intercept	23.410	0.555	1.000	1.000	1.000	22.416	24.526
Sample_Size	6.302e −5	8.615e −5	0.500	0.550	1.222	−3.111e −5	2.419e −4
Response_Rate	−0.143	0.021	0.500	1.000	9.569e +7	−0.182	−0.102
Percent_eligible_for_the_NSLP_SampleStudents	−0.079	0.028	0.500	0.974	37.069	−0.132	−0.021

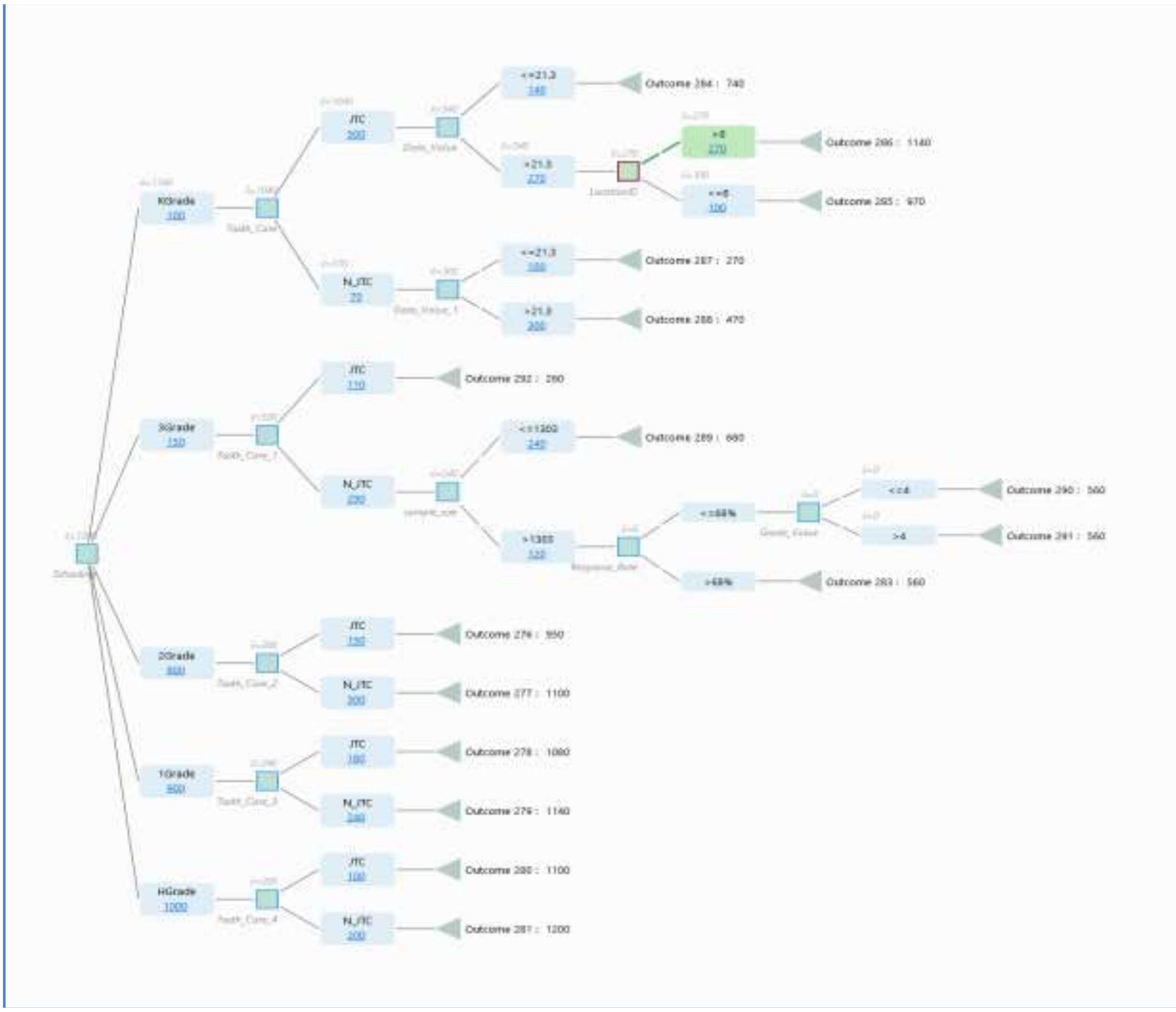
With reference to the above result, one may obtain the Bayesian (predictive) regression equation (model) as:

Data_Value = 6.302e-5*(Sample_Size) - 0.143*(Response_Rate)
0.079*(Percent_eligible_for_the_NSLP_SampleStudents)

Hence, it is obviously that the percentage of untreated tooth decay is indirectly proportional to the response rate. This event suggests that the investigated population may be among the low socio-economic position [9]. My predictive models (or equations) tell us how these low socio-economic people will take less care (behavior) to their tooth decay. This event is because they may have more tooth decay without any medical treatment. Hence, one may perform a similar technique for the prediction of behavior to another two types of tooth indicators.

A Decision Tree with expected values for finding a suitable Strategy in teeth, regenerative medicine — a thought experiment

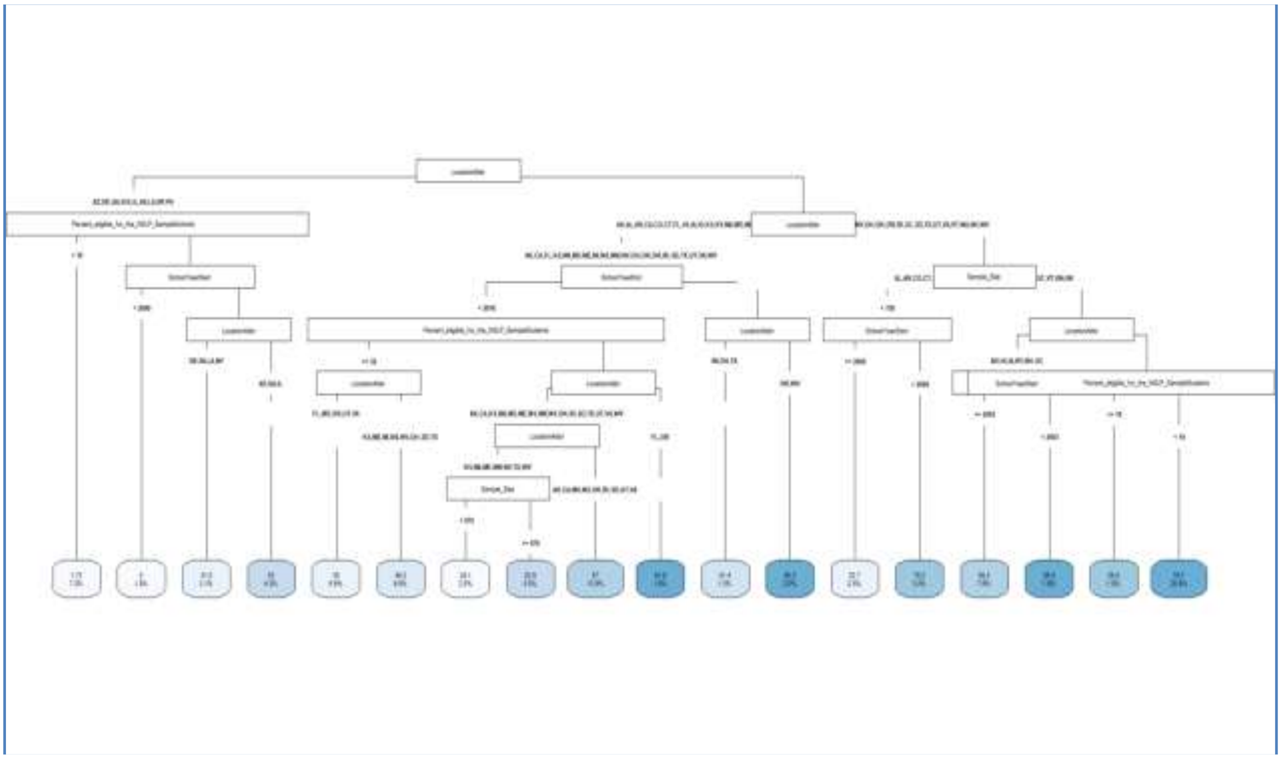
From my Butterfly Effect philosophy, one can further develop the Bayesian Decision Theory for finding the best strategy during decision making. Thus, another thing that we need to discuss is finding the most suitable strategy for our regenerative medicine. Suppose there is a thought experiment including five grades of children with different cost for making decisions such as whether one should join the tooth care scheme (JTC) etc. This event (the thought experiment) is shown in the following diagram (Figure 12) in the next page by employing the software — SpiceLogic Decision Tree Analysis.



The software finds that locationID > 8 is the evaluated policy for stochastic dominance (to determine the expected utility maximiser). One may also get all the feasible paths with their corresponding payoff:

Terminal	Total Payoff
Schooling → KGrade → Tooth_Care → JTC → Data_Value → <=21.3	740
Schooling → KGrade → Tooth_Care → N_JTC → Data_Value_1 → <=21.3	270
Schooling → KGrade → Tooth_Care → N_JTC → Data_Value_1 → >21.3	470
Schooling → KGrade → Tooth_Care → JTC → Data_Value → >21.3 → LocationID → >8	1140
Schooling → 3Grade → Tooth_Care_1 → JTC	260
Schooling → 3Grade → Tooth_Care_1 → N_JTC → sample_size → <=1303	680
Schooling → 3Grade → Tooth_Care_1 → N_JTC → sample_size → >1303 → Response_Rate → >68%	560
Schooling → 3Grade → Tooth_Care_1 → N_JTC → sample_size → >1303 → Response_Rate → <=68% → Grade_Value → <=4	560
Schooling → 3Grade → Tooth_Care_1 → N_JTC → sample_size → >1303 → Response_Rate → <=68% → Grade_Value → >4	560
Schooling → 2Grade → Tooth_Care_2 → JTC	950
Schooling → 2Grade → Tooth_Care_2 → N_JTC	1100
Schooling → 1Grade → Tooth_Care_3 → JTC	1080
Schooling → 1Grade → Tooth_Care_3 → N_JTC	1140
Schooling → HGrade → Tooth_Care_4 → JTC	1100
Schooling → HGrade → Tooth_Care_4 → N_JTC	1200
Schooling → KGrade → Tooth_Care → JTC → Data_Value → >21.3 → LocationID → <=8	970

With reference to both of the results (paths' payoff and the decision tree), the total payoff for the suggested path LocationID > 8 is only 1140, which nearly attains payoffs highest value (1200 for the JTC with Tooth_Care_4). The minimum value of the payoff is 270 and the medium value is 560 corresponding to the path with Data_Value smaller than 21.3 and also there are some other paths that lead to the same payoff. If one selects the proposed path, this event shows that when there is a tooth care scheme, the society can always obtain the best return. Moreover, the thought experiment implies that when the researcher collects enough daily data, one can find the most suitable strategy. Hence, one can get the best path under the condition that there is a need to maximise the payoff. I remark here that there is a public policy selection theory [10] which is named as rational choice theory. Its main context is explaining social phenomena as the outcomes of a personal action that is considered to be rational. The advantage of rational theory is that it can guarantee the decision making (selection) of a suitable policy [11] can result in a maximum social gain or benefit. Obviously, my thought experiment fits well with the rational theory.



Regression Tree for Predicting the trend

One can also use a regression tree for the foreseeing the trend features of collected dental data from the U.S. CDC. In the following section, this author applies the “rpart” and the “rpart.plot” functions of the R program to find our predicted Response_Rate percentage and hence determines the mean absolute error from the difference between actual and predicted values. In the next page, we can show a regression tree in terms of the Response_Rate.

The highest predicted Response_rate is 85.9% with the probability of only 0.02. This event corresponds to the “locationabbr” of OR and WV together with the SchoolEndYear greater than year 2018. Another highest predicted Response_rate is 78% with the probability of nearly 0.26. The event respects to the Percent_eligible_for_the_NSLP_SampleStudents of less than 19% and “locationabbr” other than AR, HI, IA, KY, NH, SC. This event is also the most probably expected one that occurred. It seems that areas with “locationabbr” not included in AR, HI, IA, KY, NH and SC have greater willingness of response. Thus, the complementary areas may be richer or have some kind of fundings in dental care also. However, as the probability is not too low which tells us that the population of this group of people is about one in a fourth. The lowest predicted Response_rate is zero with the probability of only 0.018. This event also corresponds to SchoolYearStart less than year 2000 with the Percent_eligible_for_the_NSLP_Sample Students of less than 19%. The likelihood of the rest predicted Response_Rate is around 0.02 to 0.08, which are quite evenly distributed.

Indeed, the importance of difference variables (the first five) w.r.t the Response_rate is:

```
Console   Terminal

> rpart.plot(m1, type=3, digits=3, fallen.leaves=TRUE)
> m1$variable.importance
                    LocationAbbr                                   SchoolYearStart
                      254358.005                                         62695.057
                   SchoolYearEnd             Percent_eligible_for_the_NSLP_State
                       58749.680                                         54805.074
                     Sample_Size   Percent_eligible_for_the_NSLP_SampleSchools
                       48135.642                                         46054.775
                      LocationID  Percent_eligible_for_the_NSLP_SampleStudents
                       43844.791                                         37203.053
                    LocationDesc                                             Grade
                       28024.528                                         27336.657
                     GeoLocation                           Low_Confidence_Interval
                       26162.158                                         13017.378
                      Data_Value                                           GradeID
                       11586.950                                         11243.915
                     Grade_Value                                LatestDataforGrade
                       11029.378                                          9856.254
                     SortGradeID
                        9315.438
```

The above result tells us that the five importances of variables which include the location (i.e., different states), the start and end of the school years, the percent of eligible for NSLP state and the sample size do determine the response_rate. In other words, where the respondents live, when the research starts and how large is the sample size with the percent of eligible are the key factors (or the social-economic background behind the population) that related to the respondent's population (or who will be investigated). It seems that area with code AR, HI, IA, KY, NH and SC are much richer or having some kind of funding scheme in fighting tooth decay that starts from 2018.

This author also remarks that the corresponding R code for the regression tree is shown in below:

```
> library(readxl)
> imputation_1 <- read_excel("C:/Users/User/Desktop/imputation_1.xlsx")
> View(imputation_1)
> library("rpart")
> library("rpart.plot")
> m1 <- rpart(Response_Rate ~ ., data = imputation_1, method="anova")
> m1
> rpart.plot(m1, type=3, digits=3, fallen.leaves=TRUE)
> m1$variable.importance
```

This author remarks that one may compare different Bayesian Linear Models from their Bayes Factors with respect to their BF_M and BF_{01}. The smaller the value of BF_{01}, the higher the chance for the model hypothesis H_1 than H_0 to be likely occurred. This event is because BF_{01} is just the reverse of BF_{10}. In addition, one may find out the expectation values from the different BF_Ms. Hence, we can select the highest BF_{10} with well-balanced BF_Ms in finding the best Bayesian prior model. (Notes: Both the P(M) and P(M|Data) may need to be considered for selecting the best model also.) Next, from the best prior model selected, we can calculate the corresponding posterior model through Bayesian estimation with the calculated mean as the expectation values. Finally, one may discover the policy that relevant to that posterior model. Actually, the symmetrical theory is applied whenever one needs to public relation problems [4].

The following is the practice of what the above mentioned in theory:

The above data shows that the best model found is:

Response_rate ~ Sample_Size + Percent_eligible_for_the_NSLP_SampleSchools + Grade_Value

4 https://www.researchgate.net/publication/331581188_IMPLICATIONS_OF_SYMMETRY_FOR_A_THEORY_OF_ETHICS_AND_SOCIAL_RESPONSIBILITY_IN_PUBLIC_RELATIONS

Bayesian Linear Regression ▼

Model Comparison

Models	P(M)	P(M|data)	BF_M	BF_{10}	R^2
Indicator_Value + Sample_Size + Percent_eligible_for_the_NSLP_SampleSchools + Grade_Value + Data_Value + LocationID	0.143	0.189	1.394	1.000	0.065
Sample_Size + Percent_eligible_for_the_NSLP_SampleSchools + Grade_Value	0.007	0.188	32.254	19.982	0.058
Sample_Size + Percent_eligible_for_the_NSLP_SampleSchools	0.010	0.115	13.531	9.161	0.047
Indicator_Value + Sample_Size + Percent_eligible_for_the_NSLP_SampleSchools + Grade_Value + Data_Value	0.024	0.097	4.428	3.103	0.064
Sample_Size + Percent_eligible_for_the_NSLP_SampleSchools + Grade_Value + Data_Value	0.010	0.082	9.235	6.489	0.060
Sample_Size + Percent_eligible_for_the_NSLP_SampleSchools + Grade_Value + LocationID	0.010	0.063	6.972	4.999	0.059
Sample_Size + Percent_eligible_for_the_NSLP_SampleSchools + Grade_Value + Data_Value + LocationID	0.024	0.058	2.533	1.852	0.062
Indicator_Value + Sample_Size + Percent_eligible_for_the_NSLP_SampleSchools + Grade_Value	0.010	0.051	5.644	4.096	0.058
Indicator_Value + Sample_Size + Percent_eligible_for_the_NSLP_SampleSchools + Grade_Value + LocationID	0.024	0.036	1.534	1.148	0.059
Sample_Size + Percent_eligible_for_the_NSLP_SampleSchools + LocationID	0.007	0.019	2.674	2.003	0.047

The corresponding BF_{10} is 19.982 which is highly expected to occur than the null hypothesis. Moreover, the BF_M is 32.254 which is the expected value for that particular model. In the coming paragraph, this author will perform the Bayesian estimation for the posterior model.

When one is performing Bayesian Estimation for the posterior model by using statistical software SPSS, the result is listed below:

Bayesian Estimates of Coefficients[a,b,c]

Parameter	Posterior Mode	Posterior Mean	Posterior Variance	95% Credible Interval Lower Bound	95% Credible Interval Upper Bound
(Intercept)	39.535	39.535	15.547	31.804	47.266
Sample_Size	.001	.001	.000	.000	.001
Percent_eligible_for_the_NSLP_SampleSchools	.223	.223	.002	.136	.310
Grade_Value	1.762	1.762	.868	-.065	3.589

a. Dependent Variable: Response_Rate

b. Model: (Intercept), Sample_Size, Percent_eligible_for_the_NSLP_SampleSchools, Grade_Value

c. Assume standard reference priors.

From the above data, it is observed that Grade_Value has the highest mean value. This event implies that when the tooth healthy scheme is implemented in K1 or the earliest, the expected outcome and return to prevent tooth decay in U.S. kindergarten students will be better. Since for those students in the higher grade like K3 will soon have their milk tooth replacement. If the youngest kindergarten students can protect their tooth during the earliest stages, there will be a less expenditure that spent before the permanent tooth. Thus, this author suggests that there is a need to have tooth heathy checking schemes — a public policy for our kindergarten students in Hong Kong. My proposed public policy also agrees well with the aforementioned rational theory in the previous paragraph.

Predictive models and their implications

It should be note that from the above regression tree model, one can further perform the prediction with the R program code "predict" to obtain the wanted forecast values. In an example, if one's initial assumption believes that a particular coefficient (β) is positive but the fitting result gives a negative one, a contradiction appears. One of the case study is the fitting of physical experiment data that is used in the nuclear databases — inconsistent data that causes system error [12]. This event shows the philosophy, logic-ism behind the model [13]. Georg in 2018 notes that one can finally celebrate these experiment data back to become best fitted and analysed through the Bayesian methods [12]. In such case, those prescribed data without precise fitting celebrations may be acted as the deserved working case for our intelligent machines. A practical logic-ism model case study is the Chromatic Induction model for the integrated visual system. Simultaneously, when we continue to feed back the predictive values into the Bayesian tree, (say for 5 times) according to the procedural rules, then the outcome implies the existence of philosophy, formalism can be used to model [14] the nested causal relationships behind. A generalised case study is the predictive loop model that works under procedural laws. This event is used for the prediction of the earthquake when it cooperates together with the database management system. Finally, when one develops the Hidden Markov Model w.r.t. the data (such as in this thesis one) given, then this event provides us with the philosophy, intuition-ism behind [5]. Once we can compare the RMSE values calculated from the above three models mentioned and thus decide the philosophy indicated. This event is because the RMSE values usually tell us the level of model fitting. Obviously, the best fitted model (with the above three cases) will then demonstrate which kind of the philosophy laying backward among the data collected. In this thesis, if the Hidden Markov Model computed has the lowest RMSE value when one compares it with the other two, the Markov model will be the best fitted one, the tooth care behaviour of our investigated people is thus only motivated according to their intuition.

To sum up, one can always express implicitly the three types of philosophy (can be extended to various ones) in terms of three kinds of mathematical model during their prediction respectively. In other words, one may always relate the models to philosophies. This event encourages us to model our body, mind and spirit. It is because in Lam July, 2016, I connect the body, mind and spirit with formalism, intuitionism and logic-ism. If it is true that I can also mathematically model the above three philosophies, the implication is one can also indirectly model our body, mind and spirit. Another side effect is we can also teach our computer what philosophy is [18]. Thus, the final result that associated with the above two implications may be the development of human-like artificial intelligent machine.

In particular, our planets can be viewed as the body since they are governed up by nature rules. While for our universe, the observations give us the clue of the natural constants. This event is the source of intuitionism or our mind. Finally, whenever there is a contradiction to these constants during computations, the universe's spirit — consciousness or the logic-ism shows itself.

Actually, a person's lifestyle has a close relationship with philosophy. One of the case is a healthy philosophy (somehow a person's belief) of dieticians may lead to a healthy life-style [6]. This event will then give a satisfaction in his/her life and is linked with a group of per-son or a social class's consciousness. I note that one may apply the cluster analysis [15] to the qualitative data for the classification of them with similar characteristics or patterns. This implies the possibility of modelling different social classes' consciousness. If we model our human's conscious, the result is one may have an in-depth understanding of the integrated information theory or even the generalised one. The outcome is advancement in the field of quantum physics, computer science and mathematics. One of the interesting application is the feasibility to obtain a communication based model of consciousness [7] or even to go ahead a step — the rationalisation to the philosophical model of time consciousness [16]. It is true that there are various types of consciousness model. It includes the functionalism or the philosophy of mind. The result of these consciousness models such as the architectural one may finally lead to the conclusive integrative theory of consciousness. One of the possible case to model these human's conscious-ness can be referenced with the article Lam, March 2020 and will be left to the next cycles of investigation for those researchers interested.

CONCLUSION

The most significant aspect of this research is to investigate caries of prevalence in local kindergarten students and determine the effectiveness of fluoride treatment in pre-venting tooth decay. This event can be accomplished by verifying the proposed Butterfly Effect philosophy (during the prediction procedure), using the data collected from the re-search. The outcome of the study may better equip teachers to educate young children regarding the importance of using toothpaste with suitable fluoride content, or to undergo fluoride treatment for tooth decay. Education is the best method to prevent caries in kindergarteners since they are most likely to listen to a trusted adult (i.e., their teachers). Moreover, the collected data could be used to help predict behavioural patterns which cause caries (this event can be examined by hypothesis testing: post-test and pre-test), thereby, preventing caries-causing practices. Finally, one may obtain the deserved Bayesian regression from the simulation of Bayesian inference, with estimated mean and variance that show normal distribution. Hence, a prediction model for caries development will be obtained. Indeed, with the use of the data (associated to different variables) that collected from the research, one can establish the corresponding decision trees. If one can further compare these trees and select the best-optimised path, the most feasible choice for a better decision making will be applied. Then, one may develop and find the most suitable strategy for the regenerative medicine in teeth through our prediction model such as the need of subsidy for kindergarten tooth care — a machine generated policy.

5 HTTPS://WWW.FRONTIERSIN.ORG/ARTICLES/10.3389/FPSYG.2013.00573/FULL

6 HTTPS://WWW.NCBI.NLM.NIH.GOV/PMC/ARTICLES/PMC5746676/

7 HTTPS://MODELS-OF-CONSCIOUSNESS.ORG/BOOKLET/MOC19_CONFERENCE_BOOKLET.PDF

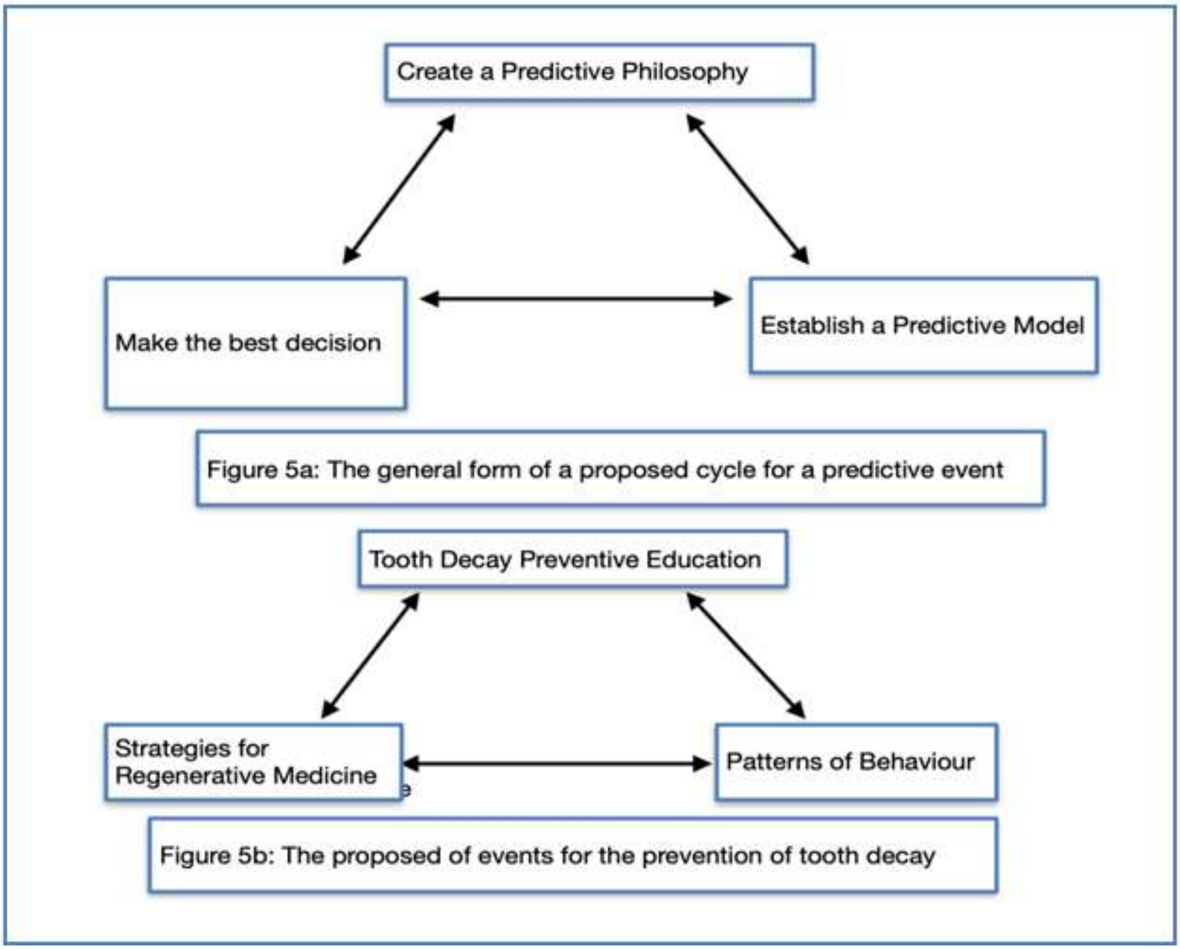


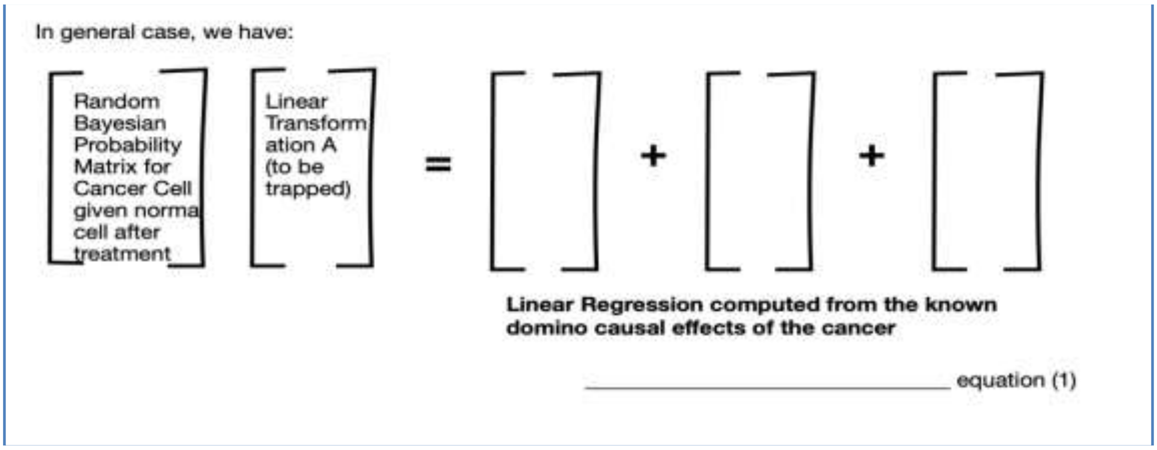
Fig-5a and 5b depicts the flow of development for the research proposal as shown in the following

REFERENCE

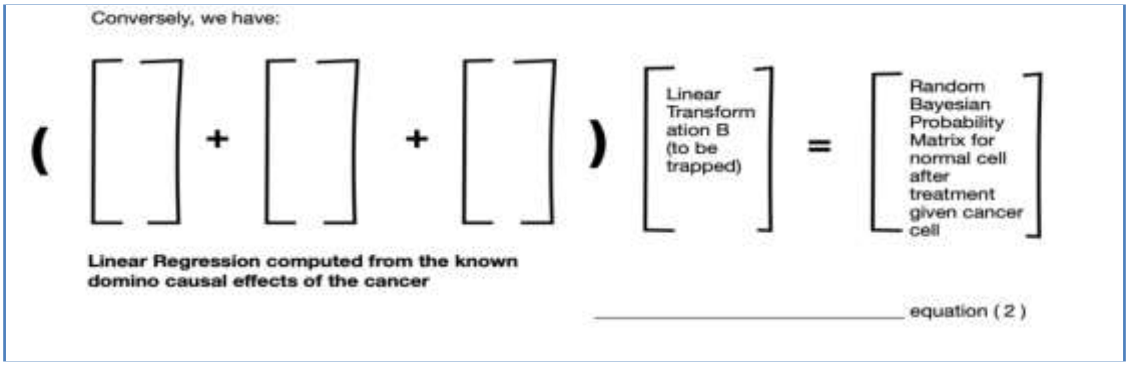
Remarks

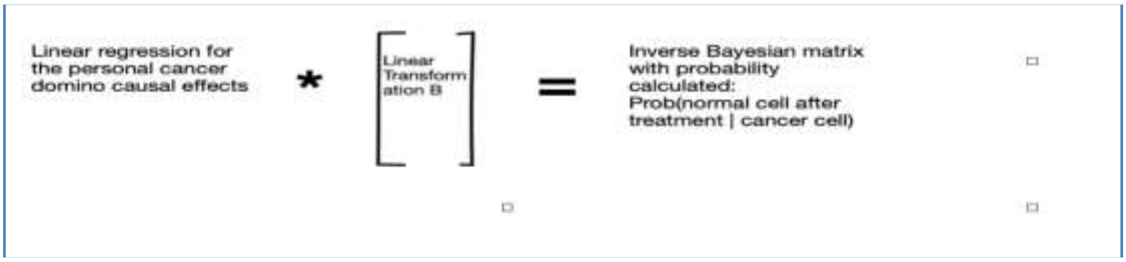
In order to compute the linear transformation (or the linear mapping), I suggest theoretically, we may fit our data in forms of Bayesian matrix on one side of the equality. At the same time, we should obtain the corresponding domino effects expressed in terms of causality through vector algebra on the other side of the equality. The outcome is that we may "trap" the immediate part of the linear transformation. Certainly, there are defects of such method, to overcome it, we may fit millions of Bayesian matrix and the linear regression for the related (areas or subjects) domino into such equation (like the training process in machine learning) for calibration so that we may select the best wanted linear transformation. Then we need to find the wanted (or the specified) area or subject's domino effect in terms of linear regression with the previously computed linear transformation. Therefore, theoretically, we can calculate the new inverted Bayesian matrix and get the wanted or questioned probabilities. The aforementioned process theoretically constitutes the flow chart (or the algorithm) for finding the inverse of the random matrix.

Practically, we can use a similar "trap" method for the converse part of the proposed Butterfly Effect philosophy. With the old inverted Bayesian matrices on one side and the old (in other related areas or subjects) linear regression obtained by the causality of the domino effects (of the equality), we may trap the immediate linear transformation. By using the necessary machine learning (or machine training) techniques and fits millions of the random matrices and the corresponding linear regressions, we may finally get the best corresponding linear transformation for the inverted random matrix. Next, we need to find out the inverse of the linear transformation and multiply it with the present research area or subject's linear regression of the causality for the domino effect. The conclusive outcome is that we can get the newly inverted Bayesian (random) matrix with what we expected the predictive probability to the research subject areas. Besides language translation, the case can also be implemented in the earthquake; we may come with a computed predictive probability for a particular geographical area. One of the samples in cancer treatment likes the following:



In particular, for an individual patient or a group of them with similar background, we may apply this event in the precision medicine for cancer treatment:





________________________ Equation (3)

As shown in the covering letter of Lam Dec, 2019 one may also use the topology projection and 3D printer technique etc to design a tailor-made treatment in precision medicine for an individual cancer patient. From the collection millions of the treatment, one may further genetic engineer a universal T-cell that with the suitable antigen to act on the MR1 of cancer cells and destroy them. The above method together with the equation (3) just constitute a cancer treatment philosophy (universal to individual and vice versa) in the precision medicine.

In addition, one can also use the suggested steps in this essay — Bayesian Inference, Bayesian Regression to predict people's corresponding responses from the treatments and the behaviour that lays behind. Hence, medical professionals can select the best option and develop the most suitable novel therapeutic according to the regression tree's results obtained.

Another remark is that one can also apply the regression model method such as the one shown in Lam March, [17] to approximate the proposed mutation causality (chain). The causality event occurs in infectious diseases like influenza (human flu) and (Wu Han) coronavirus-19 which may be possibly depicted as an example below:
Bats —> Rats —> Monkey —> Human Beings

To sum up, for the former part of the Bayes' tree, one should first compute the predicted values from my Net-Seizing Theory through correlation relationships — linear regression as shown in chapter 15. Then, one should feedback these calculated forecast values into the former Bayes part. The result is one can obtain the posterior distribution by the Bayesian Regression approximation.

This author also remarks that for the probability matrix [Pr(A|B)]:

$[Pr(A|B)]^2 = X' \boldsymbol{\beta}' X (LT)^{-1} + X' \boldsymbol{\beta}' X B (LT)-^{1} + X' \boldsymbol{\beta}' B_0(LT)^{-1} + X' \boldsymbol{\beta}' X B D (LT)^{-1} + X' \boldsymbol{\beta}' B_0 D (LT)^{-1} + X' \boldsymbol{\beta}' D_0 (LT)^{-1} + \boldsymbol{\varepsilon}\|'$

Where $[Pr(A|B)]^2$ is now expressed in form of a linear polynomial. Hence, the random matrix say Pr(A|B) can be found from its square root. $[Pr(A|B)]^2 = X' \beta' (I + B + BD) X (LT)^{-1} + X' \beta' (B_0 D) (LT)^{-1} + [X' \beta' (B_0+D_0)(LT)^{-1} + \varepsilon''']$

In fact, we have found a mediator D for the A^2. Or we have just established a Hayes model for $[Pr(A|B)]^2$ (– or even the square of matrix A). Hence, we can perform further forecasting from the model.
The precise expression of A in terms of Bayesian linear regression is assumed to be:
$A = X' \beta' + \varepsilon$. In addition, the $[Pr(A|B)]^2$ is only a conditional random matrix. But if both of the events A and B are independent, then it will reduce to the case like my tossing coins as described above. This gives us the true random matrix.

Indeed, no matter how we are finding the matrix A through the inverse of Bayesian linear regression or the square root of the linear polynomial expressed, both of the cases are not easy tasks.

Another remark is that it may be a case study for my HKLam Theory in this paper and may not be a generalised validation for the theory. However, when one discovers that most swan are white and concludes it as a theory, there is the possibilities that we can find the swan with other colour. Practically, what can determine the fate of my HKLam theory is whether or not it represents the true reality [8]. Hence, inductively from the case study of the child tooth care (c.f. Grounded Theory), we may upgrade the former part of HKLam Net-Seizing Theory (Bayesian Network version) as follows:

[Matrix from the ANCN] [LT] = [linear regression of the causal domino events]
[Linear regression of the causal domino events] [LT] = [Matrix from the ANCN]

Where ANCN means the artificial neural causal network.

To be precise, the former part can be generalized by the multiple level of network methodology – MLNM (multiple layers of Bayesian, Neural or Causal networks etc.,) i.e.,

[Matrix from the MLNM] [LT] = [Linear regression of the causal domino events]
[Linear regression of the causal domino events] [LT] = [Matrix from the MLNM]

If my HKLam (generalised) theory is correct, we may deductively (c.f. grounded theory) apply it in the case like the earthquake and predict the most feasible places (with the highest probabilities) that will be affected for the future or the trend of the disturbance to track the source.

HKLam theory implies that we can obtain the required model of the investigated field from the linear regression of causal dominos. Hence, the prediction is possible from the model. While for the converse, we can find the original source(s) of the study field or foresee the evolution trend of the topic. When we compare the Lorenz system's equations (we use the differentials first with the respective linear equations) to the domino's linear regression equation. Then we substitute the corresponding gradient descent for the linear regression cost function [9] into the aforementioned domino's linear regression equation. The outcome is the unknown parameters σ, β, ρ can be expressed in terms of the summation of x_i, y_i z_i — the true values and the $h_\theta(x_i)$, $h_\theta(y_i)$, $h_\theta(z_i)$ — the predicted values together with the known linear regression equation parameters etc [10]. By selecting a suitable value — x_i, y_i, z_i, and $h_\theta(x_i)$, $h_\theta(y_i)$, $h_\theta(z_i)$ that we want, we can control the (values of) statically object — σ, β, ρ, then I may manage the relaxation and tension of my net-seizing for the wanted level of chaos [11]. There is also a data population reduction[12] method (or the hydra paradox/effect) to control chaos. According to Eduardo I in 2011, harvesting can decrease such effect[13]. Instead of eliminating all chaos by HKLam theory, the key focus is how one should maintain the equilibrium between chaos and ordering through game theory as I have mentioned earlier in this essay.

To go a further step, the Lorenz system (or attractor) implies the Butterfly Effect. Since my HKLam theory has the ability to manage the chaos that caused by the Butterfly Effect (by controlling the values of σ, β, ρ), I can actually capture (or catch) the butterfly.

8 Emma S.W., (2015) Generalising from Case Studies, Halmstad University

9 https://ml-cheatsheet.readthedocs.io/en/latest/linear_regression.html
10 https://www.geeksforgeeks.org/gradient-descent-in-linear-regression/
11 https://www.sciencedirect.com/topics/mathematics/lorenz-equation
12 https://www.sciencedirect.com/science/article/abs/pii/S0375960109014960
13 Eduardo, L (2011) The hydra effect, bubbles, and chaos in a simple discrete population model with constant effort harvesting, Journal of Mathematical Biology, DOI: 10.1007/s00285-011-0489-2

This is because the Lorenz system is now expressed in terms of the computed values of σ, β, ρ or the summation of x_i, y_i z_i and the $h_\theta(x_i)$, $h_\theta(y_i)$, $h_\theta(z_i)$ etc. Once if the parameters σ, β, ρ are determined (or computed previously), the predicted values $h_\theta(x_i)$, $h_\theta(y_i)$, $h_\theta(z_i)$ can also be expressed back by the domino's linear regression equations and the summation of x_i, y_i z_i etc. Or we may perform the deserved prediction for the corresponding Butterfly model. This is the breakthrough of my HKLam Theory such that it has the characteristics of net-seizing and is a kind of recursive formalism. In other words, we may express σ, β, ρ in terms of the summation of x_i, y_i z_i and the $h_\theta(x_i)$, $h_\theta(y_i)$, $h_\theta(z_i)$ etc. Then we can substitute back the computed σ, β, ρ to obtain the predicted vales. The process can be circulated for thousands of time until we finally approach the inflection point for the Lorenz system.

The aforementioned is just the conceptual part (philosophy) of how Lorenz system interacts with my HKLam Theory. One may verify it through suitable computer programming and stimulation (with sufficient data provided). When this is applied to the common diseases like COVID-19 or influenza, we may have the chance to control the spread of them according to the computed results. However, this paper only concerns the conceptual part of how the control should be implemented. The causality dominos regression has been shown in Lam March, 2020. At the same time the gradient descent was written in Lam July, 2019 while the cost function is described in the present paper. I note that we may hybrid the "algorithmic and dialectic" proof in Lam May, 2016 with the Bayesian optimisation that shown in Lam Dec, 2019 for the machine learning method. Then we may accelerate the learning rate to find the minimum in the cost function and obtain the least error. However, the style of this paper is not like Lam July, 2017, thus the mathematical computation details of the control is out of the scope of this paper. Indeed, one already has a case study in the establishing of suitable policies with the corresponding human behaviour as shown in this article. Therefore, my present study can act as a reference for finding the strategies in how one should manage these diseases. Indeed, in this paper, I only want to discuss how human make their decisions and the respective government policies that are associated. One of the major application of machine learning (HKLam Theory) and policy making is our (future) intelligent weather prediction [14]. Besides that, there is also the possibility for our humans to manage the weather by capturing the butterfly (or the Lorenz attractor) in it according to the HKLam Theory, Lorenz system, cost function and regression, Bayesian optimisation with dielectric and decision making etc [15]. Indeed, weather prediction and management are not the main discussion topics in this paper.

14 https://www.itproportal.com/features/advancing-weather-forecasting-with-machine-learning/
15 https://www.nature.com/articles/s41598-019-57162-8

Finally, one of the major applications of artificial intelligence is to have the cancer test through the blood screening [16]. This event is first experimented by the Harvard University, United States of America. With the help of such technology, doctors can forecast the early status of human cancer and propose a suitable treatment to their patients. Hence, this action will increase rate of those having cancers' survival well before the disease becoming serious.

REFFERENCE

1. Gorka N and David RM. Improving Bayesian Reasoning: What works and Why? Frontiers in Psychology.2016.
2. Thierry M. Engineering Risk Management, Walter de Gruyter GmbH, Berlin/Munich/Boston.2016.
3. Leung KT. Linear Algebra and Geometry, Hong Kong University Press. 1974.
4. Lam, Shun, Charging Our Electrical Devices in Anywhere and at Any Time (January 31, 2019). Available at SSRN: https://ssrn.com/abstract=3326476 or http://dx.doi.org/10.2139/ssrn.3326476
5. Lam, Shun, Evaluation of the Weather-Influenza Pattern with a Regression Model Approximation to Causality (March 29, 2020). Available at SSRN: https://ssrn.com/abstract=3563601 or http://dx.doi.org/10.2139/ssrn.3563601
6. Chu CH. Management of early childhood caries, *Oral Health Magazine* 2001;
7. Creswell JW. *Educational research: planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle, NJ: Pearson Education International; 2012.
8. Seth AK. A MATLAB toolbox for Granger causal connectivity analysis *Journal of Neuroscience Methods*. 2010; 186(2), 262–273.
9. Ekholm O, Gundgaard J, Rasmussen NK, Hansen EH. The effect of health, socio-economic position, and mode of data collection on non-response in health interview surveys. Scandinavian journal of public health. 2010 Nov;38(7):699-706.
10. Ebenezer OO. Public Policy Analysis, Concepts Publication Ltd.2016. https://www.researchgate.net/publication/334749461_PUBLIC_POLICY_ANALYSIS
11. Emeka AO, Nwachukwu CL, Obiora AC, Lfeoma EU. (PDF) public policy analysis and decision making. (n.d.). 2008.Retrieved from https://www.researchgate.net/publication/332523852_PUBLIC_POLICY_ANALYSIS_AND_DECISION_MAKING
12. Georg S. Fitting and Analysis Technique for Inconsistent Nuclear Data.2018. https://arxiv.org/abs/1803.00960

13. Siu MK. The algorithmic and dialectic aspects in proof and proving. InProof and proving in mathematics education. ICMI study 19 conference proceedings 2009 (Vol. 2, pp. 160-165).
14. Casini L, Illari PM, Russo F, Williamson J. Models for prediction, explanation and control: recursive Bayesian networks. THEORIA. Revista de Teoría, Historia y Fundamentos de la Ciencia. 2011 Mar 1;26(1):5-33.
15. Yuen AH, Park J, Chen L, Cheng M. The significance of cultural capital and parental mediation for digital inequity. New Media & Society. 2018 Feb;20(2):599-617.
16. Mölder B. How philosophical models explain time consciousness. Procedia-Social and Behavioral Sciences. 2014 Mar 21; 126:48-57.
17. Lam, Shun, Predictive and Regenerative Medicine for Humans (December 28, 2019). Available at SSRN: https://ssrn.com/abstract=3510415 or http://dx.doi.org/10.2139/ssrn.3510415
18. Lau WW, Yuen AH. Exploring the effects of gender and learning styles on computer programming performance: implications for programming pedagogy. British Journal of Educational Technology. 2009 Jul;40(4):696-712.
19. Graff HJ, Christensen U, Poulsen I, Egerod I. Patient perspectives on navigating the field of traumatic brain injury rehabilitation: a qualitative thematic analysis. Disability and rehabilitation. 2018 Apr 10;40(8):926-34.