

Predicted Cancer Risk Probability using Linear and Nonlinear Regression Models with Medical Condition Inputs from Chronic Diseases and Lifestyle Details Collected Data of a Type 2 Diabetes Patient Based on GH-Method: Math- Physical Medicine (No. 550)

Gerald C Hsu

EclaireMD Foundation, USA

*Corresponding author

Gerald C Hsu, EclaireMD Foundation, USA

Submitted: 01 Dec 2021; Accepted: 08 Dec 2021; Published: 20 Dec 2021

Note: Readers who want to get a quick overview can read the abstract, results and graphs sections.

Citation: Gerald C Hsu (2021) Predicted Cancer Risk Probability using Linear and Nonlinear Regression Models with Medical Condition Inputs from Chronic Diseases and Lifestyle Details Collected Data of a Type 2 Diabetes Patient Based on GH-Method: Math- Physical Medicine (No. 550). J App Mat Sci & Engg Res, 5(4), 01-11.

Abstract

The author has spent approximately 40,000 hours over the past 12 years self-studying and researching internal medicine branches, with a focus on endocrinology and diabetes. Since 2018, he has expanded his interest, learning and research work into other medical branches related to lifestyle, metabolism and immunity. Currently, his type 2 diabetes (T2D) is well under control, where the HbA1C level decreased from 10% in 2010 down to 5.8% in 2021 without medication intervention. Naturally, he is concerned about other life-threatening diseases of the elderly population, specifically cancers and dementia.

Over the past decade, he has written and published more than 540 medical papers in various medical journals. In total, he applied about 30 different research methodologies based on his developed GH-method: math-physical medicine, including physics theories, engineering modeling, mathematical equations, computer science tools of big data analytics and artificial intelligence (AI), as well as some traditional statistical approaches to explore and interpret various biomarkers and their biophysical phenomena. However, the majority of published medical research papers he has read to date are primarily based on statistics (~90% of his total reading volume of ~2,000 papers). In this particular article, he decides to follow the majority of other medical scientists' footsteps, to use the traditional statistical regression model with linear and various nonlinear formulas involving multiple independent variables to investigate his overall risk probability of developing cancer versus 4 categories of metabolic disorder induced chronic diseases (obesity, diabetes, hypertension, and hyperlipidemia) and 6 categories of lifestyle details (food, water, exercise, sleep, stress, and daily life routines).

In this paper, he will not repeat the detailed introduction of the regression analysis models in the Methods section because it is available in many statistics text books. It should be noted that in regression analysis, the correlation coefficient R should be > 0.5 or 50% to indicate a strong inter-connectivity and the p -value should be < 0.05 or 5% to be considered as statistically significant.

The author recently studied a consensus report published jointly by the American Diabetes Association (ADA) and the American Cancer Society (ACS) in 2010 regarding relationships between cancers and diabetes. Based on the information from this report plus ~3 million collected data of his own overall metabolism situation, including medical conditions and lifestyle details, he decided to conduct a research study regarding the estimation of his overall and relative risk probability of having cancer over two time periods: the longer 12-year period from 2010 to 2021 and the shorter 9-year period from 2013 to 2021. The reason for using two time-periods is due to his insufficient data gathering and guesstimated data from 2010 and 2012. He would like to investigate the prediction difference.

In summary, from the time-domain analysis, the two cancer risk waveforms, metabolism index (MI) based cancer risk curve and regression predicted cancer risk curve, are similar to each other in terms of shape similarity. They have an extremely high of 92% - 98% correlation between the curves and with a 100% prediction accuracy of averaged cancer risks.

For the longer period of 12 years, the space-domain cancer analysis results show that the variances of data distribution are 90% for the linear medical case and 95% for the nonlinear medical case, while 84% for the linear lifestyle case and 98% for the nonlinear lifestyle case. **This has proven that a selected nonlinear model, polynomial order 2, provides a higher and better variance value, especially for the lifestyle case, for this 12-years cancer risk study.**

For the shorter period of 9 year, the space-domain cancer analysis results reveal that the variances of data distribution are 90% for the linear medical case and 95% for the nonlinear medical case, while 82% for the linear lifestyle case and 97% for the nonlinear lifestyle case. **This has also proven that a selected nonlinear model, polynomial order 2, provides a higher and better variance value, especially for the lifestyle case, for this 9-years cancer risk study.**

Furthermore, in terms of the overall numerical value ranges, the shorter time period's variances are almost the same as the longer time period's variances.

There are 2 specific conclusions. First, the nonlinear regression model gives a better and higher variance value than the linear regression model (note: this finding is not a universal conclusion for all cases).

Second, by using the nonlinear results, as a tool to obtain an improved predicted cancer risk value, there are no significant difference between medical conditions and lifestyle details.

However, this conclusion is quite different from his previous research findings of the CVD/Stroke risk study using the same regression models. The CVD risk study has found higher variances from the lifestyle details (99% for linear case and 97-99% for nonlinear case) than the medical conditions (77% for linear case and 88% for nonlinear case). These variance comparisons between cancer risk versus CVD/Stroke risk have shown that 13% medical difference for the linear case and 7% medical difference for the nonlinear case. In this paper, Figure 3 will demonstrate a variance numerical comparison between the cancer risk versus CVD/Stroke risk for both linear and nonlinear cases of medical conditions and lifestyle details.

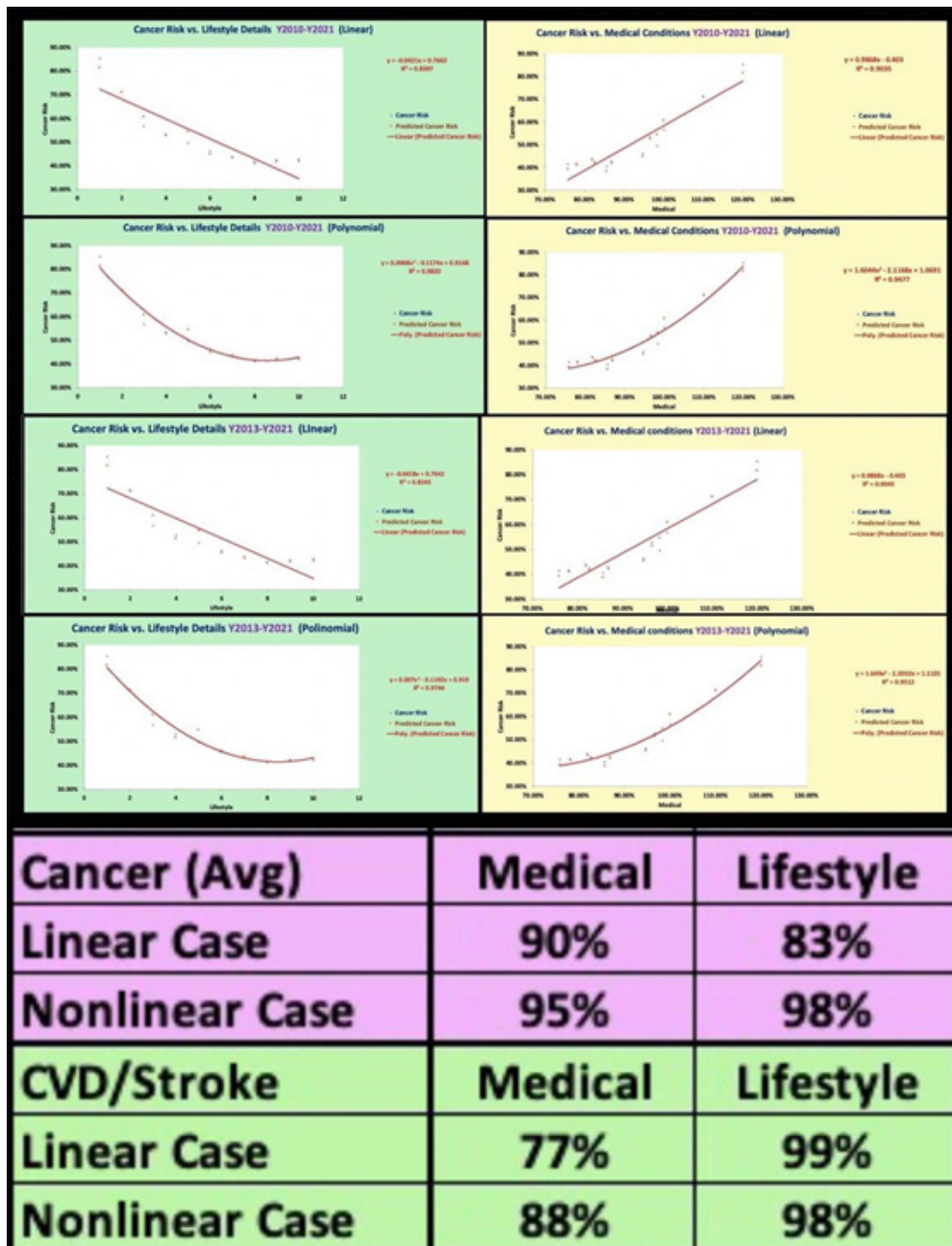
We should focus on the nonlinear case results since it is a better-fitting model for this cancer risk study. The following data table outlines the conclusion of the nonlinear variance comparison in the format of (Cancers% vs. CVD%):

Medical Condition: (95% vs. 88%)

Lifestyle Details: (98% vs. 99%)

It is obvious that "the CVD risk is influenced more by the lifestyle details than the chronic disease medical conditions, while cancers have almost equal weight of influences from both medical conditions and lifestyle details".

Therefore, "lifestyle details are the common root-causes for metabolic disorder induced chronic diseases along with their complications and all different kinds of cancers".



Introduction

The author has spent approximately 40,000 hours over the past 12 years self-studying and researching internal medicine branches, with a focus on endocrinology and diabetes. Since 2018, he has expanded his interest, learning and research work into other medical branches related to lifestyle, metabolism and immunity. Currently, his type 2 diabetes (T2D) is well under control, where the HbA1C level decreased from 10% in 2010 down to 5.8% in 2021 without medication intervention. Naturally, he is concerned about other life-threatening diseases of the elderly population, specifically cancers and dementia.

Over the past decade, he has written and published more than 540 medical papers in various medical journals. In total, he applied about 30 different research methodologies based on his developed GH-method: math-physical medicine, including physics theories, engineering modeling, mathematical equations, com-

puter science tools of big data analytics and artificial intelligence (AI), as well as some traditional statistical approaches to explore and interpret various biomarkers and their biophysical phenomena. However, the majority of published medical research papers he has read to date are primarily based on statistics (~90% of his total reading volume of ~2,000 papers). In this particular article, he decides to follow the majority of other medical scientists' footsteps, to use the traditional statistical regression model with linear and various nonlinear formulas involving multiple independent variables to investigate his overall risk probability of developing cancer versus 4 categories of metabolic disorder induced chronic diseases (obesity, diabetes, hypertension, and hyperlipidemia) and 6 categories of lifestyle details (food, water, exercise, sleep, stress, and daily life routines).

In this paper, he will not repeat the detailed introduction of the regression analysis models in the Methods section because it is

available in many statistics textbooks. It should be noted that in regression analysis, the correlation coefficient R should be > 0.5 or 50% to indicate a strong inter-connectivity and the p -value should be < 0.05 or 5% to be considered as statistically significant.

The author recently studied a consensus report published jointly by the American Diabetes Association (ADA) and the American Cancer Society (ACS) in 2010 regarding relationships between cancers and diabetes. Based on the information from this report plus ~3 million collected data of his own overall metabolism situation, including medical conditions and lifestyle details, he decided to conduct a research study regarding the estimation of his overall and relative risk probability of having cancer over two time periods: the longer 12-year period from 2010 to 2021 and the shorter 9-year period from 2013 to 2021. The reason for using two time-periods is due to his insufficient data gathering and guesstimated data from 2010 and 2012. He would like to investigate the prediction difference.

Methods

Metabolism Index (MI) Model

From 1/1/2012 to 11/15/2021, he has collected around 3 million data of his own biomedical conditions and personal lifestyle details. This big dataset is stored on the cloud and managed by Amazon data service. He can access and process them using his developed software on his iPhone.

In 2014, the author applied mathematical topology concept, engineering finite- element method technique, and nonlinear algebra operations to develop a complex mathematical model of metabolism index (MI).

This MI model contains ten specific categories, including four output categories of medical conditions (body weight, glucose, blood pressure, and lipids), and six input categories of lifestyle details (food quantity and quality, water intake, physical exercise, sleep, stress, and daily life routines). These 10 categories are comprised of approximately 500 detailed elements. He has also defined two new resulting parameters: the metabolism index or MI, as the combined score of the above 10 metabolism categories and 500 elements using his developed algorithm, along with the general health status unit (GHSU), as the 90-days moving average value of MI.

A physical analogy of this mathematical metabolism model is similar to “using multiple nails that are encircled by many rubber bands”. For example, at first, we hammer 10 nails into a piece of flat wood with an initial shape of a circle, then take 3,628,800 (=10!) rubber bands to encircle the nails, including all 10 nails. These ~3.6 million rubber bands (i.e., big number of relationships) indicate the possible relationships existing among these 10 nails (i.e., 10 original metabolism data). Some rubber bands encircle 2 nails or 3 nails and so on, until the last rubber band encircles all of these 10 nails together (no rubber band to encircle a single nail is allowed). Now, if we move any one of the nails outward (i.e., moving away from the center of the nail circle), then this moving action would create some internal tension inside the encircled rubber band. Moving one nail **outward**

means one of these ten metabolism categories is becoming **unhealthy**” which would cause some stress to our body. Of course, we can also move some or all of the 10 nails outward at the same time, but with different moving scales. If we can measure the summation of the internal tension created in the affected rubber bands, then this summarized tension force is equivalent to the metabolism value of human health. The higher tension means the higher metabolism value which creates an unhealthy situation. The author uses the above-described physical scenario of moving nails (categories) and their encircled rubber bands (inter-relationships) to explain his developed mathematical metabolism model of human health.

The Consensus Report of Cancer and Diabetes

The following is a rather long excerpt (~2,419 words) from the Reviews/Commentaries/ADA Statements, “**Diabetes and Cancer, A consensus report**” by Edward Giovannucci, MD, and many other authors, published by the American Diabetes Association and the American Cancer Society jointly. The original paper has more than 8,000 words without counting its 123 references. The author considers this paper as a trove of knowledge; therefore, he has kept ~30% of its original words in this excerpt for his future easy-access.

“Diabetes and cancer are common diseases with tremendous impact on health worldwide. Epidemiologic evidence suggests that people with diabetes are at significantly higher risk for many forms of cancer. Type 2 diabetes and cancer share many risk factors, but potential biologic links between the two diseases are incompletely understood. Moreover, evidence from observational studies suggests that some medications used to treat hyperglycemia are associated with either increased or reduced risk of cancer. Against this backdrop, the American Diabetes Association and the American Cancer Society convened a consensus development conference in December 2009. Following a series of scientific presentations by experts in the field, the writing group independently developed this consensus report to address the following questions:

1. *Is there a meaningful association between diabetes and cancer incidence or prognosis?*
2. *What risk factors are common to both diabetes and cancer?*
3. *What are possible biologic links between diabetes and cancer risk?*
4. *Do diabetes treatments influence risk of cancer or cancer prognosis?*

1. Is there a meaningful association between diabetes and cancer incidence or prognosis?

Both diabetes and cancer are prevalent diseases whose incidence is increasing globally. Worldwide, the prevalence of cancer has been difficult to establish because many areas do not have cancer registries, but in 2008 there were an estimated 12.4 million new cancer cases diagnosed. The most commonly diagnosed cancers are lung/bronchus, breast, and colorectal, whereas the most common causes of cancer deaths are lung, stomach, and liver cancer (1). In the U.S., the most commonly diagnosed cancers are prostate, lung/bronchus, and colon/rectum in men and breast, lung/bronchus, and colon/rectum in women. Of the world

population between the ages of 20 and 79 years, an estimated 285 million people, or 6.6%, have diabetes (2). In 2007, diabetes prevalence in the U.S. was 10.7% of persons aged 20 years and older (23.6 million individuals), with an estimated 1.6 million new cases per year. Type 2 diabetes is the most common form, accounting for 95% of prevalent cases (3). Worldwide, cancer is the 2nd and diabetes is the 12th leading cause of death (4). In the U.S., cancer is the 2nd and diabetes is the 7th leading cause of death; Cancer and diabetes are diagnosed within the same individual more frequently than would be expected by chance, even after adjusting for age. Both diseases are complex with multiple subtypes. Diabetes is typically divided into two major subtypes, type 1 and type 2 diabetes, along with less common types, while cancer is typically classified by its anatomic origin (of which there are over 50, e.g., lymphoma, leukemia, lung, and breast cancer) and within which there may be multiple subtypes (e.g., leukemia). Further, the pathophysiologies underlying both cancer and diabetes are (with rare exceptions) incompletely understood.

For more than 50 years, clinicians have reported the occurrence of patients with concurrent diabetes and cancer. However, as early as 1959, Joslin et al. (5) stated, "Studies of the association of diabetes and cancer have been conducted over a period of years, but evidence of a positive association remains inconclusive." Subsequently, an association between the two diseases was identified in the 1960s in population-based studies. More recently, the results of several studies have been combined for meta-analytic study (6), indicating that some cancers develop more commonly in patients with diabetes (predominantly type 2), while prostate cancer occurs less often in men with diabetes. The relative risks imparted by diabetes are greatest (about two-fold or higher) for cancers of the liver, pancreas, and endometrium, and lesser (about 1.2–1.5 fold) for cancers of the colon and rectum, breast, and bladder. Other cancers (e.g., lung) do not appear to be associated with an increased risk in diabetes, and the evidence for others (e.g., kidney, non-Hodgkin lymphoma) is inconclusive.

Diabetes-related factors including steatosis, nonalcoholic fatty liver disease, and cirrhosis may also enhance susceptibility to liver cancer. With regard to pancreatic cancer, interpretation of the causal nature of the association is complicated by the fact that abnormal glucose metabolism may be a consequence of pancreatic cancer (so-called reverse causality). However, a positive association between diabetes and pancreatic cancer risk has been found when restricted to diabetes that precedes the diagnosis of pancreatic cancer by at least 5 years.

Only for prostate cancer is diabetes associated with a lower risk. This association has been observed both before and after the advent of screening with prostate-specific antigen (PSA). Some metabolic factors associated with diabetes, such as reduced testosterone levels, may be involved. While obesity has not been associated, and in some studies is even inversely associated, with prostate cancer incidence, obese men with prostate cancer have higher cancer mortality rates than those of normal weight (7). In addition to metabolic factors such as hyperinsulinemia, obesity may be associated with clinical factors (such as delayed diagnosis,

poorer treatment) that may underlie the worsened prostate cancer prognosis.

Results of some, but not all, epidemiological studies suggest that diabetes may significantly increase mortality in patients with cancer (8).

Unanswered Questions

Diabetes has been consistently associated with increased risk of several of the more common cancers, but for many, especially the less common cancers, data are limited or absent (6) and more research is needed. Uncertainty is even greater for the issue of diabetes and cancer prognosis or cancer-specific mortality. It remains unclear whether the association between diabetes and cancer is direct (e.g., due to hyperglycemia), whether diabetes is a marker of underlying biologic factors that alter cancer risk (e.g., insulin resistance and hyperinsulinemia), or whether the cancer-diabetes association is indirect and due to common risk factors such as obesity.

In view of the variable associations between diabetes and cancer risk at specific sites, the authors discourage studies exploring links between diabetes and **risk of all cancers combined**. For example, since lung cancer does not appear to be meaningfully linked with diabetes, including this common cancer in studies will dilute observed associations, should they exist.

2. What risk factors are common to both cancer and diabetes?

Potential risk factors (modifiable and nonmodifiable) common to both cancer and diabetes include aging, sex, obesity, physical activity, diet, alcohol, and smoking.

Nonmodifiable Risk Factor

Age

Although the incidence of some cancers peaks in childhood or in young adults, the incidence of most cancers increases with age. In economically developed countries, 78% of all newly diagnosed cancer occurs among individuals aged 55 years and older (11). Diabetes also becomes increasingly common with age: Prevalence is 2.6% in U.S. adults 20–39 years of age, 10.8% in those 40–59 years of age, and increases to 23.8% in those 60 years of age or older (3). In parallel with the obesity epidemic, type 2 diabetes is becoming more frequent among adolescents and young adults (12,13), potentially adding years of additional risk from diabetes to the population.

Sex

While certain cancers are sex-specific (e.g., cervix, uterine, testicular, prostate), or nearly so (breast), overall cancer occurs more frequently in men. Men have slightly higher age-adjusted risk of diabetes than women (3).

Race/Ethnicity

In the U.S., African Americans are more likely to develop and die from cancer than other race or ethnic groups. Following African Americans are non-Hispanic whites, with Hispanics, Native Americans, and Asian Americans/Pacific Islanders having lower

cancer incidence and mortality (14). While incompletely understood, genetic, socioeconomic, lifestyle, and other environmental factors are thought to contribute to these disparities.

Modifiable Risk Factors

Overweight, Obesity, and Weight Change

Overweight (BMI ≥ 25 and < 30 kg/m²) or obese (BMI ≥ 30 kg/m²) individuals have a higher risk for many types of cancer compared with individuals whose BMI is considered within the normal range (18.5 to < 25 kg/m²) (16,17). The cancers most consistently associated with overweight and obesity are breast (in postmenopausal women), colon/rectum, endometrium, pancreas, adenocarcinoma of the esophagus, kidney, gallbladder, and liver. Obesity may also increase risk of mortality from some cancers, such as prostate (7). A growing body of evidence suggests that weight gain is associated with an increased risk of some cancers, breast cancer in particular (17). Increases in body weight during adulthood largely reflect increases in adipose tissue rather than lean mass, so total body fat may be a better measure of the risk for cancer than BMI.

Studies over decades have consistently shown a strong association between obesity and both insulin resistance and type 2 diabetes incidence (18), with risk of diabetes and earlier age at onset directly linked to obesity severity (19). For type 2 diabetes (20) as well as certain cancers (e.g., colon) (21), some studies suggest that waist circumference, waist-to-hip ratio, or direct measures of visceral adiposity are associated with risk independently of BMI.

The case for a causal relationship between obesity and disease is strengthened by evidence that weight loss lowers disease risk. In the randomized, prospective, multicenter Diabetes Prevention Program trial, an intensive lifestyle intervention of diet (targeting 5–7% weight loss) and physical activity was associated with a 58% reduction in diabetes incidence in high-risk individuals (22), and weight loss accounted for most of the effect (23). In addition, weight loss may also limit the risk of developing gestational diabetes (24).

The association between weight loss and subsequent cancer risk is less clear. Weight loss may be a sign of undiagnosed cancer.

Diet

A majority of studies suggest that diets low in red and processed meats and higher in vegetables, fruits, and whole grains are associated with a lower risk of many types of cancer (17,28,29). Diets that are low in red and processed meat but high in monounsaturated fatty acids, fruits, vegetables, whole grain cereals, and dietary fiber may protect against type 2 diabetes, possibly through improving insulin sensitivity (30,31). Low-carbohydrate diets (which often include greater consumption of red meats and fat) have also been associated with weight loss and improvements in insulin sensitivity and glycemic control. However, randomized controlled trial evidence of dietary interventions and diabetes prevention only exists for low-fat, low-calorie, plus/minus high-fiber diets (22,32). Several studies suggest that diets high in foods with a high glycemic index or load are associated with an increased risk of type 2 diabetes (28,33). However, evidence of their associations with cancer risk is mixed (28,34,35).

Regardless, to the extent that energy-dense and sugary foods contribute to overweight and obesity, the American Cancer Society, the World Cancer Research Fund, and the American Institute for Cancer Research recommend limiting consumption of these foods (17,29).

Physical Activity

Evidence from observational epidemiologic studies consistently shows that higher levels of physical activity are associated with a lower risk of colon, postmenopausal breast, and endometrial cancer (17,36,37). Physical activity may also help prevent other cancers, including lung and aggressive prostate cancer, but a clear link has not been established. Some evidence also suggests that physical activity postdiagnosis may improve cancer survival for some cancers, including breast (38) and colorectal (39).

A protective role for increased physical activity in diabetes metabolism and outcomes has been demonstrated. Data from observational and randomized trials suggest that 30 min of moderate-intensity exercise, such as walking, at least 5 days per week substantially reduces (25–36%) the risk of developing type 2 diabetes (40).

Tobacco Smoking

It is estimated that worldwide, tobacco smoking accounts for 71% of all trachea, bronchus, and lung cancer deaths (41). Other cancers strongly associated with smoking are larynx, upper digestive, bladder, kidney, pancreas, leukemia, liver, stomach, and uterine cervix.

Alcohol

Alcoholic beverage consumption, even in moderate amounts, increases the risk of many types of cancer including those of the oral cavity, pharynx, larynx, esophagus, liver, colon/rectum, and female breast (45).

Unanswered Questions

A critical question is whether the associations between diabetes and risk of certain cancers is largely due to shared risk factors (obesity, poor diet, physical inactivity, and aging), or whether diabetes itself, and the specific metabolic derangements typical of diabetes (e.g., hyperglycemia, insulin resistance, hyperinsulinemia), increase the risk for some types of cancer. While it is clear that lower levels of adiposity, healthy diets, and regular physical activity are associated with reduced risk for type 2 diabetes and for several common types of cancer, these factors are generally interrelated, making the contribution of each factor difficult to assess.

3. What are possible biologic links between diabetes and cancer risk?

Carcinogenesis is a complex process. Normal cells must undergo multiple genetic hits before the full neoplastic phenotype of growth, invasion, and metastasis occurs. This process of malignant transformation can be divided into multiple steps: initiation (irreversible first step toward cancer), promotion (stimulation of the growth of initiated cells), and progression (development of a more aggressive phenotype of promoted cells).

Hyperglycemia and Cancer

In considering the complexity of interactions between diabetes, diabetes treatments, and cancer, it is important to not overlook glucose as a potentially relevant mediator. The recent resurgence of interest in the Warburg hypothesis and cancer energetics (66) emphasizes the dependence of many cancers on glycolysis for energy, creating a high requirement for glucose (or even glucose addiction”) Insulin receptor activation may be a more important variable than hyperglycemia in determining tumor growth.

Major Unanswered Questions

As previously outlined, there is a growing body of epidemiologic evidence supporting a link between diabetes and the incidence and/or prognosis of some cancers. It is recognized the association may not be causal; diabetes and cancer may be associated simply because they share common predisposing risk factors such as obesity.

Individuals with type 1 diabetes represent 5% of the diabetes population worldwide. The autoimmune destruction of the pancreatic β -cells results in the loss of insulin production and the need for immediate and lifelong insulin therapy. In contrast, type 2 diabetes is much more common and accounts for 95% of the diabetes population. Type 2 diabetes is generally associated with overweight and obesity (in an estimated 80% of cases) and commonly advances from a pre-diabetic state characterized by insulin resistance (hyperinsulinemia) to frank diabetes with sustained insulin resistance accompanied by a progressive reduction in insulin secretion.

Insulin and Insulin Analogs

Insulin is required for all patients with type 1 diabetes. It is also necessary for many patients with type 2 diabetes to treat hyperglycemia, in part due to the progressive loss of β -cell function over time. Between 40–80% of individuals with type 2 diabetes will ultimately be considered for insulin therapy in an effort to achieve glycemic targets (77).”

The Author S Learned Knowledge from this Consensus Report

After reading this consensus report ten times, the author derived his own conclusive knowledge.

Cancers and diabetes have some statistical links, but their biological relationships are still inconclusive. Diabetes has been consistently associated with increased risk for several of the more common cancers, but they are not for all different types of cancer. Although with rare exceptions, pathophysiologies underlying both cancer and diabetes are still incompletely understood, the identification of some “clear and detailed” connections between cancers and various metabolic disorders are also incomplete. However, at least, we have already identified some common risk factors between diabetes and cancers, such as chronic medical conditions (e.g., obesity) and particularly in the areas of lifestyle details (e.g., inactivity) and life-long bad habits (e.g., smoking).

It seems that insulin resistance situation (hyperinsulinemia) and insufficient insulin secretion have some influences on certain types of cancer development. Diabetic hyperglycemia is also the

direct result of insulin resistance and insufficient insulin secretion.

Diets that are low in red and processed meat but high in mono-unsaturated fatty acids, fruits, vegetables, whole grain cereals, and dietary fiber may protect against type 2 diabetes, possibly through improving insulin sensitivity. Furthermore, low-carbohydrate diets have also been associated with weight loss and improvements in insulin sensitivity and glycemic control. These findings have provided the significance of food quality on cancer development.

There is no doubt about the obvious link existing between cancers and lifestyle details or long-term bad habits in life, especially diet (particularly food quality, including types of food, processed food, food additives, and food poisons), physical exercise, sleep, and stress. For example, the shared risk factors of certain cancers and diabetes are obesity (including food quantity), poor diet (food quality), physical inactivity, and aging. Exercise and sleep have been proven to be important for both health and healing process. The author has also self-studied psychology and psychiatry for 9 years; therefore, he understands the importance of stress on the overall physical health. Certain life-long unhealthy habits such as tobacco smoking, excessive alcohol drinking, and illicit drug use can eventually cause certain types of cancer.

Body weight is a strong influence factor on both diabetes and cancer development due to being overweight or having obesity. Since food portion, food quantity, directly contributes to medical situations when a person is overweight or has obesity, food quantity should be considered together with body weight.

Other metabolic biomarkers, such as blood lipids and blood pressure (BP), are important as well for developing various metabolic disorders & diabetic complications, such as cardiovascular disease (CVD), congenital heart disease (CHD), stroke, chronic kidney disease (CKD), and diabetic retinopathy (DR). However, the direct connection between BP/lipid and cancers are still inconclusive in general, except for a few special cases. For example, it has been proven by a Japanese doctor that high triglycerides indeed has a strong association with prostate cancer for 60+ year old males. As a result, even with insufficient evidence for general concerns of cancers, there is still linkage between chronic diseases, obesity, diabetes, hypertension, hyperlipidemia, and certain types of cancers.

The authors of the ADA/ACS consensus report discourage studies exploring **links between diabetes” and risk of all cancers combined**. However, the author of this particular paper still conducted his own research in exploring the general relationships between diabetes and cancers, **the links between general metabolic conditions (including both chronic diseases and lifestyle details) and the risk of all cancers combined**”. He has already learned that metabolism and immunity are the two fundamental cornerstones of our health, where they are two sides of the same coin.

In the ADA/ACS consensus report of diabetes and cancers, the original paper’s authors have repetitively used certain phrases like, “*lacking of epidemiological evidence, having incomplete biological links, or facing unclear pathophysiologies underlying of the association between diabetes and cancers directly*”. This has caused the author of this paper to rethink the types of descriptions and meanings deeper by using his lifelong learned knowledge of applied mathematics and physics as well as a profes-

sional engineer. Even though various cancers and diabetes have their own specific causes, they do share common root-causes. After all, cancer is also a type of chronic disease; therefore, the majority of root-causes for chronic diseases and cancers overlap with each other. In order to identify the direct relationship between diabetes and cancers using *symptoms* alone is a more difficult and wasteful task. It may be easier to start investigating the overlapping *causes or common root causes*, for example, lifestyle, life-long unhealthy habits, environmental factors, such as toxin, pollution, and radiation. The overall metabolism is a good starting point. This situation can be illustrated using the author's engineering and physics background. The tensile stress (stretching force) and strain (longitudinal deformation) are dependent on the Young's modules, while the shear stress (shear force) and strain (shear deformation) are dependent on the shear modules. However, both the Young's modules (similar to the cancers relationship between their causes and symptoms) and shear modules (similar to the diabetes relationship between its causes and symptoms) are directly related to the actual engineering material of the different study subject which is the "body" of human health study. This situation is similar to the different pathophysiologic process of our internal organ diseases. The engineering material (or human organs) contains both of Young's modules and shear modules which is similar to our human body being under the influences of common root-causes, such as lifetime unhealthy habits, lifestyle details, environmental factors, and overall metabolism and immunity. Therefore, we need to start with the understanding of the *internal material* first (i.e., the underlying root-causes), instead of searching for the relationship of various *outside symptoms* between two chronic disease families (cancers and diabetes).

However, *medicine is one of the most complex subject which includes many aspects that other academic subjects do not have. The main reason of these complexities are resulted from human cells are "organic" while engineering materials are "inorganic". Foreexample, the ambiguity between root cause and symptom can be switched under certain circumstances. This role-switching between cause (input) and symptom (output) are not frequently observed in the engineering world. Foreexample, metabolic disorder induced chronic diseases are the symptoms of poor lifestyles (root-cause). However, some of these metabolic disorder induced chronic diseases could be turned into causes of certain cancers, e.g., obesity (functioning as both symptom and root cause).* Another example, CVD and CKD are complications resulting from diabetes (symptom turns into cause) and poor lifestyle (root-causes). Cancers, at least for certain cases, are also related to metabolic disorder induced chronic diseases, including obesity and diabetes, and poor lifestyle. Furthermore, cancer patients are at a higher risk of dying from heart disease and stroke. These findings have demonstrated this "role-switching" phenomenon from the comparison between CVD/Stroke risk versus Cancer risk.

Somemore Details of Cancers

From the annual US death case study, three major causes are chronic diseases with various complications (50%), cancers (29%), infectious diseases (11%), along with non-diseases relat-

ed death (10%). The annual death percentages (prior to 2020's COVID-19 pandemic) mentioned above can be observed in Figure 1.

2017 Death	Category	percentage
Heart	647,457	31%
Cancer	599,108	29%
Accidents	169,936	8%
Respiratory	160,201	8%
Stroke	146,383	7%
Alzheimer's	121,404	6%
Diabetes	83,564	4%
Pneumonia	55,672	3%
Kidney	50,633	2%
Suicide	47,173	2%
Total	2,081,531	100%
Chronic Related		
	1,049,441	50%
Cancer		
	599,108	29%
Infectious		
	215,873	11%
Accidents & Suicide		
	217,109	10%

Figure 1: 2017 death case % of cancers which has 599,108 death cases and 29% of total death cases

Cancer is an exceedingly difficult and complicated disease that can affect any organ within the body, where abnormal cells divide and mutate rapidly, destroying healthy normal cells in the process. The possible cause of cancer can result from a combination of many different reasons. The author has dedicated the past decade on researching endocrinology and metabolism. He considers endocrinology and cancer as being similar from the viewpoint of "digging into a black box of the inner universe in the human body". However, based on his rudimentary understanding of cancer, he also feels that the diseases caused by cancer are probably at least 10 times more complicated than endocrinology. Although he is not an oncology expert, only a patient and research scientist on chronic diseases and metabolism, he has a strong curiosity and motivation in wanting to know more about his own risk probability of developing into cancers (there are more than 50 different types of cancers). This reason inspires his cancer research work by using the learned knowledge on metabolism and his ~30 different math-physical research methodologies to conduct his own assessment on the relationship between the overall cancer risk versus the overall metabolism situation.

One information to support his findings is that there are 23 cancer factors causing a total of 45.2% of entire cancer cases in China (around 2.3 million cases per year). Most of these 23 influential factors happened to be a part of the components which have already been identified in his developed mathematical model of metabolism.

Figure 2 demonstrates six different cancer influential factors from a summarized cancer table developed by the author which connect certain cancer causing or influencing factors and organ systems affected by cancer. This is an ongoing project. He still has a lot to learn about cancer diseases, for example, which organs in one particular organ system are most likely to be affected by an influential factor. Therefore, this article only serves as the beginning of his long journey in cancer research using his developed GH-Method: math-physical medicine.

Cancer Factors	Items	Cancer in Organ System	Cancer in Organ
Genetic (5%)	Age		
	Gender		
	Race		
	Family History		
Personal Bad Habits (25%)	Smoking Cigaretts	respiratory, digestive, endocrine, reproductive, urinary	
	Drinking Alcohol	respiratory, digestive, reproductive	
	Use Illicit Drugs	respiratory, digestive, lymphatic	
	Obesity	digestive, reproductive, endocrine, urinary, cardiovascular, lymphatic	
	Prior Personal Medical Conditions		
	Over-weight& large waistline factors		
	Bad food & diet selection (less plant, more animal)		
	Too much Processed food, fermented food, pickled food		
	Infections & Chronic Inflammations		
	Insufficient daily exercise		
	Take too many medication drugs		
	High stress life & poor sleep conditions		
Environmental Factors (15%)	Toxic chemicals		
	Air pollution (e.g. PM 2.5)		
	Water pollution		
	Expose to Food pollution		
	Poison	digestive	stomach
	Hormone therapy		
	Nuclear Radiation (X-ray, CT, etc.)	lymphatic	
	UV Radiation	integumentary, lymphatic	skin
	Infections (Fungi, Parasites, Bacteria)		
	Cancer-causing chemicals		skin
Viral Infection Factors (10%)	Helicobacter Pylori	digestive	stomach
	Hepatitis B Virus	digestive	liver
	Hepatitis C Virus	digestive	liver
	HIV Virus	digestive, lymphatic	anus, immune system
	Human Papilloma Virus	digestive, reproductive, lymphatic	mouth, throat, tonsils, anus, cervix, vulva, vaginal, penis
	Epstein-Barr Virus	respiratory, digestive	nose, throat
	Paragonimus Sinensis		
	Human Herpes Virus Type 8		
	Kaposi's Sarcoma	digestive	bile duct
	Hodgkin's Lymphoma	lymphatic	lymph nodes
Metabolism Related Factors (45%)			
Chronic Diseases Conditions (15%)	Obesity	digestive, reproductive, endocrine, urinary, cardiovascular, lymphatic	colon, breast, pancreas, gallbladder, Liver, esophagus, kidney, thyroid, ovaries, blood cells, prostate
	Diabetes (Insulin resistance)		
	Hypertension		
	Hyperlipidemia		
	Hyperthyroidism		thyroid
	Infections (foot ulcer)		
	CVD & Stroke		
	CKD		
	Immunosuppression (after organ transplant)		
	Chronic Inflammation (viral or bacterial)		
Lifestyle Details (30%)	Food & Diet	respiratory, digestive, endocrine	
	Exercise & Physical Activities		
	Water Drinking		
	Sleep		
	Stress		
	Regular Daily Life Routines		

Figure 2: Six influential factor groups of cancers

He started his investigation from identifying major causes and the possible organ systems affected by cancer. He describes those steps as follows:

First, like many other branches of medical research, he started with the sub-area of genetics, including his age, race, gender, and family genetic background. He has assigned 5% of weight to this sub-area of genetic factors.

Second, he delved into the sub-area of personal unhealthy habits including smoking or chewing tobacco, drinking alcohol, and/or taking illicit drugs that would lead into various types of cancer affecting different organ systems. In addition, he also investigated other components, such as having an inadequate diet, inactive lifestyle, high stress life, poor sleep quality, and personal medical history along with types, amounts, and duration of medication intake that would also lead into different types of cancers. He assigned 20% of weight to this sub-area of personal unhealthy habit factors.

Third, the sub-area of environmental factors includes toxic chemicals, air pollution (e.g., PM 2.5), water pollution, food pollution, poison, hormone therapy, nuclear radiation (e.g., X-ray, CT), UV radiation, infection from parasites and bacteria, or

other cancer-causing chemicals, and more. He assigned 15% of weight to this sub-area of environmental factors. As an example, relatively speaking, China can have a higher percentage of cancer cases in this sub-area due to its highly polluted environment, including land, water, and air.

Fourth, the sub-area of viral infection factors includes Helicobacter Pylori, Hepatitis B Virus, Hepatitis C Virus, HIV Virus, Human Papilloma Virus, Epstein-Barr Virus, Paragonimus Sinensis, Human Herpes Virus Type 8, Kaposi's Sarcoma, Hodgkin's Lymphoma, and others. He assigned 10% of weight to this sub-area of viral infection factors.

Fifth, the sub-area of metabolic disorder induced chronic diseases and their various complications include obesity, diabetes, hypertension, hyperlipidemia, CVD, stroke, CKD, bladder infection, hyperthyroidism, bladder infection, foot ulcer, diabetic retinopathy, and more. He assigned 15% of weight to this sub-area of chronic diseases factors.

Sixth, the most important sub-area of lifestyle details which count for 35% of the total weight, should be the foundation of the root-causes mentioned in the above situations except for the genetic factor. This sub-area include six categories, food and diet, exercise, water intake, sleep, stress, and daily routine life pattern. These categories combined with the fifth sub-area of chronic diseases have approximately 500 detailed elements (from finite "element" method of engineering).

The above-described 6 influential factor groups and assigned weighting scales for cancers are quite different from his risk assessment model of other chronic disease induced complications, such as CVD/Stroke, CKD, and DR, which are mainly based on 2 major influential factor groups, medical conditions and lifestyle details. It should be noted that CVD/Stroke, CVD, and DR have their certain specific individual influential biomarkers included with their medical conditions.

The author spent 12 years to develop and continuously enhance a sophisticated and customized software program to collect all types of input data regarding health. He then processed them dynamically in order to provide a daily guideline for the purpose of improving his overall metabolism. Once his metabolism is in good condition, then his immune system will be strong enough to defend against most of those infectious diseases, such as COVID-19 pandemic.

The above paragraphs have described the backbone of his mathematical model to calculate his risk probability of having different cancers which are closely related to his metabolism model.

Results

Figure 1 shows 2017 death case %. It should be pointed out that the cancer group has 599,108 death cases and 29% of the total 2,081,531 death cases in 2017.

Figure 2 is a cancer information table which shows 6 influential factor groups of different cancers.

Figure 3 is the final conclusion data table of this cancer risk study which illustrates the comparison of variances between medical conditions and lifestyle details with both linear case and nonlinear case.

Vriances		
Cancer (12-yrs)	Medical	Lifestyle
Linear Case	90%	84%
Nonlinear Case	95%	98%
Cancer (9-yrs)	Medical	Lifestyle
Linear Case	90%	82%
Nonlinear Case	95%	97%
Cancer (Avg)	Medical	Lifestyle
Linear Case	90%	83%
Nonlinear Case	95%	98%
CVD/Stroke	Medical	Lifestyle
Linear Case	77%	99%
Nonlinear Case	88%	98%
Diff (Cancer-CVD)	Medical	Lifestyle
Nonlinear Case	7%	-1%

Figure 3: Comparison of variances of both linear and nonlinear cases for medical conditions and lifestyle details

Figure 4 reflects the input and output data table of this cancer risk study.

Year	Y2010*	Y2011*	Y2012*	Y2013	Y2014	Y2015	Y2016	Y2017	Y2018	Y2019	Y2020	Y2021
Cancer Risk (%)	85.46%	71.07%	56.69%	52.77%	49.56%	46.19%	43.28%	41.04%	42.39%	42.00%	41.40%	40.44%
Medical Conditions	120.00%	110.00%	100.00%	96.59%	94.71%	94.71%	91.88%	78.09%	82.64%	86.85%	75.81%	85.62%
Lifestyle Details	88.00%	78.00%	68.00%	50.89%	54.07%	45.10%	43.41%	41.74%	42.54%	39.21%	38.67%	

Year	Cancer Risk	Medical	Lifestyle	Year	Calculated Risk	Predicted Risk
Y2010*	85.46%	120.00%	80.00%	Y2010*	85%	82%
Y2011*	71.07%	110.00%	70.00%	Y2011*	71%	71%
Y2012*	56.69%	100.00%	60.00%	Y2012*	57%	61%
Y2013	52.77%	96.59%	52.77%	Y2013	53%	51%
Y2014	49.56%	94.71%	54.07%	Y2014	50%	55%
Y2015	46.19%	94.71%	45.10%	Y2015	46%	45%
Y2016	43.28%	81.88%	43.41%	Y2016	43%	44%
Y2017	41.04%	78.09%	41.32%	Y2017	41%	41%
Y2018	42.39%	82.64%	41.74%	Y2018	42%	42%
Y2019	42.00%	86.85%	42.54%	Y2019	42%	43%
Y2020	41.40%	75.81%	39.21%	Y2020	41%	39%
Y2021	40.44%	85.62%	38.67%	Y2021	40%	39%
Average	51%	93%	51%	Average	51%	51%
R	94%	98%		R	98%	
Variance	88%	97%		Variance	97%	

Year	Cancer Risk	Medical	Lifestyle	Year	Calculated Risk	Predicted Risk
Y2013	52.77%	96.59%	50.89%	Y2013	53%	51%
Y2014	49.56%	94.71%	54.07%	Y2014	50%	55%
Y2015	46.19%	94.71%	45.10%	Y2015	46%	45%
Y2016	43.28%	81.88%	43.41%	Y2016	43%	44%
Y2017	41.04%	78.09%	41.32%	Y2017	41%	41%
Y2018	42.39%	82.64%	41.74%	Y2018	42%	42%
Y2019	42.00%	86.85%	42.54%	Y2019	42%	43%
Y2020	41.40%	75.81%	39.21%	Y2020	41%	39%
Y2021	40.44%	85.62%	38.67%	Y2021	40%	39%
Average	44%	87%	44%	Average	44%	44%
R	93%	93%		R	92%	
Variance	72%	86%		Variance	86%	

	Timr R	Time R^2	Linear medical	Nonlinear medical	Linear lifestyle	Nonlinear lifestyle
Y10-Y21	98.0%	97.0%	90.4%	94.8%	84.0%	98.2%
Y13-Y21	92.0%	86.0%	90.5%	95.1%	82.5%	97.4%

Figure 4: Input and output data table of this cancer risk study

Figure 5 reveals time-domain analysis comparison between MI-based cancerrisk curve and regression predicted Cancer risks curve of both the longer time period of 12 years and the shorter time period of 9 years. Both cancerrisk curves have extremely high correlations (98% for longer period and 92% for shorter period).

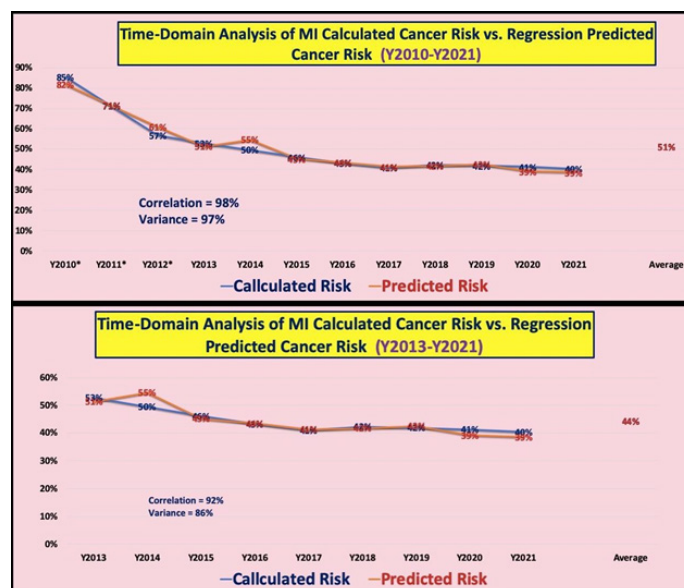


Figure 5: Time-domain analysis of cancer risk study

Figure 6 demonstrates the regression analysis results of the longer period of 12 years (2010-2021). The space-domain cancer analysis results indicate that the variances of the data distribution as 90% for the linear medical case and 95% for the nonlinear medical case, while 84% for the linear lifestyle case and 98% for the nonlinear lifestyle case. This has proven that a selected nonlinear model, polynomial order 2, provides a high-erand better variance value, especially for the lifestyle case.

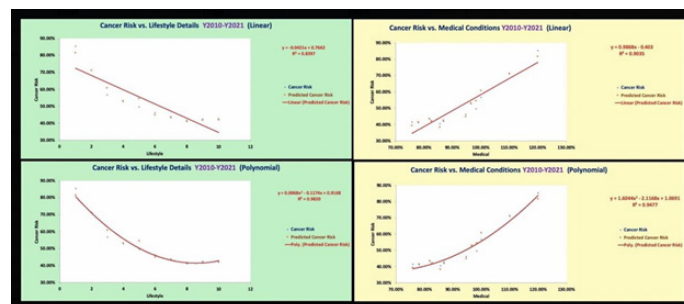


Figure 6: Space-domain regression analysis results of cancer risk vs. medical conditions and lifestyle details for the longer 12-year period from 2010 to 2021

Figure 7 shows the regression analysis results of the shorter period of 9 years (2013-2021). The space-domain cancer analysis results imply that the variances of the data distribution are 90% for the linear medical case and 95% for the nonlinear medical case, while 82% for the linear lifestyle case and 97% for the nonlinear lifestyle case. This also proves that a selected nonlinear model, polynomial order 2, offers a higher and better variance value, especially for the lifestyle case.

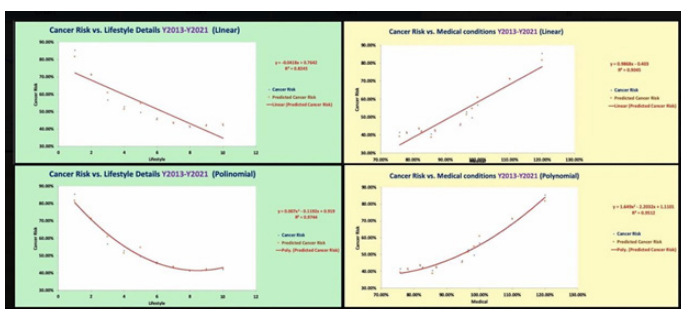


Figure 7: Space-domain regression analysis results of cancer risk vs. medical conditions and lifestyle details for the longer 12-year period from 2013 to 2021

Furthermore, in terms of the overall numerical value ranges, the shorter time period's variances (Figure 5) are almost the same as the longer time period's variances (Figure 6).

Conclusion

In summary, from the time-domain analysis, the two cancer risk waveforms, metabolism index (MI) based cancer risk curve and regression predicted cancer risk curve, are similar to each other in terms of shape similarity. They have an extremely high of 92% - 98% correlation between the curves and with a 100% prediction accuracy of averaged cancer risks.

For the longer period of 12 years, the space-domain cancer analysis results show that the variances of data distribution are 90% for the linear medical case and 95% for the nonlinear medical case, while 84% for the linear lifestyle case and 98% for the nonlinear lifestyle case. **This has proven that a selected nonlinear-model, polynomial order 2, provides a higher and better variance value, especially for the lifestyle case, for this 12-years cancer risk study.**

For the shorter period of 9 year, the space-domain cancer analysis results reveal that the variances of data distribution are 90% for the linear medical case and 95% for the nonlinear medical case, while 82% for the linear lifestyle case and 97% for the nonlinear lifestyle case. **This has also proven that a selected nonlinear model, polynomial order 2, provides a higher and better variance value, especially for the lifestyle case, for this 9-years cancer risk study.**

Furthermore, in terms of the overall numerical value ranges, the shorter time period's variances are almost the same as the longer time period's variances.

There are 2 specific conclusions. First, the nonlinear regression

model gives a better and higher variance value than the linear regression model (*note: this finding is not a universal conclusion for all cases*).

Second, by using the nonlinear results, as a tool to obtain an improved predicted cancer risk value, there are no significant difference between medical conditions and lifestyle details.

However, this conclusion is quite different from his previous research findings of the CVD/Stroke risk study using the same regression models. The CVD risk study has found higher variances from the lifestyle details (99% for linear case and 97- 99% for nonlinear case) than the medical conditions (77% for linear case and 88% for nonlinear case). *These variance comparisons between cancer risk versus CVD/Stroke risk have shown that 13% medical difference for the linear case and 7% medical difference for the nonlinear case.* In this paper, Figure 3 will demonstrate a variance numerical comparison between the cancer risk versus CVD/Stroke risk for both linear and nonlinear cases of medical conditions and lifestyle details.

We should focus on the nonlinear case results since it is a better-fitting model for this cancer risk study. The following data table outlines the conclusion of the nonlinear variance comparison in the format of (Cancers% vs. CVD%):

Medical Condition: (95% vs. 88%)

Lifestyle Details: (98% vs. 99%)

It is obvious that "the CVD risk is influenced more by the lifestyle details than the chronic disease medical conditions, while cancers have almost equal weight of influences from both medical conditions and lifestyle details".

Therefore, "lifestyle details are the common root-causes for metabolic disorder induced chronic diseases along with their complications and all different kinds of cancers".

References

- Edward Giovannucci, MD and others, "Diabetes and Cancer, A consensus report", this article is jointly published by the American Diabetes Association and the American Cancer Society. © 2010 by the American Diabetes Association. Author Affiliations

Corresponding author: Edward Giovannucci, egiovann@hsph.harvard.edu. Diabetes Care 2010 Jul; 33(7): 1674-1685.

<https://doi.org/10.2337/dc10-0666>

Diabetes Care 2010 Jul; 33(7): 1674-1685.

Copyright: ©2021 Gerald C Hsu. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.