



**Research Article**

## Two Coexisting Physiological Changes Which Trigger Depression. Discovery With a Novel Method of Finding Non-Infectious Disease Causes Which Combines Multiple Research

**Alan Olan**

### Abstract

Depressive disorder (also known as depression) is a common mental disorder. It involves a depressed mood or loss of pleasure or interest in activities for long periods of time. It causes severe symptoms that affect how a person feels, thinks, and handles daily activities, such as sleeping, eating, or working. Depression is highly prevalent and often a chronic or recurring problem that interferes with work and family. The depression erodes the motivation, energy and enjoyment needed to nurture and sustain social and marital relationships, parenting, etc.

Risk factors and causal mechanisms involved in depression have implicated a wide range of genetic, neurological, hormonal, and endocrinological factors that may play a role in underlying vulnerability or in the processes by which stressors trigger depression in some people. A dominant model of the neurobiology of depression that has emerged in recent years emphasizes the *underlying dysregulation of the body's response to stress*, involving the neuroendocrine system and brain responses. Depression is most commonly found among those facing chronically stressful conditions, such as social disadvantage and distressed relationships or lack of supportive and intimate relationships.

The literature on biological, environmental, and personal risk factors for depression also shows that not all individuals who have been exposed to risk factors for depression develop the disorder. Overall the research suggest that the causes of depression likely include genetic, biological, psychological, and environmental factors but the exact causes of depression are still unknown.

In this work an author is introducing a novel method of finding non-infectious diseases and then applies it to analyze the causes of Depression. This method has been already published in other article [2] and also applied to other non- infectious diseases, for example to Breast Cancer disease and has shown results which are matching to existing medical facts. The method is using a special algorithm based in math and allows to find disease causes for a specific non- infectious disease using results of multiple researches regarding risk factors of the disease. The method is based on a model presented in the article "*A Connection between Factors Causing Diseases and Diseases Frequencies: Its Application in Finding Disease Causes*" (Alan Olan, Journal of Clinical Trials, Vol.13, Issue 4.) which is confirmed by empirical data. This method allows to combine dozens of such researches together with dozens of researches on biochemistry and physiology to extract implicit information on the causes of non-infectious disease out this entire research.

**Affiliation:**

Independent Researcher

**\*Corresponding author:**

Alan Olan, Independent Researcher, USA.

**Citation:** Alan Olan. Two Coexisting Physiological Changes Which Trigger Depression. Discovery With a Novel Method of Finding Non-Infectious Disease Causes Which Combines Multiple Research. Journal of Psychiatry and Psychiatric Disorders. 10 (2026): 75-94.

**Received:** April 07, 2026

**Accepted:** April 13, 2026

**Published:** May 15, 2026

The use of method requires to find number of causes for a specific non-infectious disease using a special formula and data on disease frequency (usually the incidence rate) in specific population. These disease causes are two or more physiological changes beyond approximately 1-sigma interval which if they are co-existing long enough must trigger the non-infectious disease (triggering is not optional). After that the method requires to find disease causation factors out of multiple risk factors found for a disease. Only some of risk factors are real disease causing factors. This is achieved using a set of special disease causation criteria discussed in the article by Olan A. [1] above and provided here for a reference as well.

The method often allows to find few dozen of disease causation factors and more for a non-infectious disease out of existing medical research. These disease causation factors all point to the same set of limited number of physiological parameters changes beyond 1-sigma. The number of these physiological changes usually vary depending on the non- infectious disease from a minimum of 2 to a maximum of 6 or very rare more. In this case the few dozens of disease causation factors make changes (we can say “point”) to the same set of 3 physiological parameters(for example).The method then allows to find these physiological parameter changes (which are the real cause of the non-infectious disease) using a properties based in math.

After this the method allows to determine which physiological parameter of this group is impacted by each of the dozens of disease causation factors previously found. Then method allows to group these factors according to the physiological parameter they impact. The disease causation factors taken out of each group of these factors and combined together will cause a change beyond 1-sigma to all required for disease triggering physiological parameters.

**These combinations of disease causation factors applied long enough will cause a non-infectious disease.** The occurrence of the disease causation factors is random but once they act together the non-infectious disease **triggering is a must** unless the factors are removed fast enough.

Final step of the method is validation of its results using other research or the already discussed disease causation criteria in order to eliminate any errors in steps of the method which we could potentially make.

Once the simultaneously taking place physiological changes causing a non-infectious disease has been found the method allows to build a hypothesis of the disease pathology by using them and “connecting the dots” using the existing medical research on physiology, immunology, etc. The example of this process shown in the work as

well. The hypothesis of Depression pathology is proposed as one example of this.

The article introduces to the basics of the method, provides required formulas for calculations and then move to a detailed analysis of Depression disease. As the method is novel the appendix has an analogy to explain the idea of the method at “high level”. The author’s introduction to the method will allow other medical researchers to use their own and existing research to determine the causes of non-infectious diseases as per presented model, using a simple algorithm.

**Results:** Using this method and applying multiple *existing selected studies* at the same time an author analyses Depression and as a result **the work gives the causes of Depression** disease as a set of two physiological parameters changes beyond 1-sigma interval (slightly less, actually) and also as a set of disease causing external factors which combinations in an individual *must* cause Depression as per presented model. The research has found that a **reduction of GABAergic inhibition beyond ~1 sigma interval and lowering BDNF level below ~1 sigma interval** will cause the Depression if both conditions coexist long enough. If both of these physiological changes coexist long enough the disease triggering is not optional it is a must. These 2 physiological parameters can be changed by different combinations of factors some of which are found in this research and as the combination of factors can vary greatly this creates an **effect of randomness of Depression causes** which we usually observe while the underlying physiological reasons for the disease stay the same.

Some of the disease causation factors found related to reduction (only under specific conditions) of GABAergic inhibition such as Epilepsy, being a women in late transition to menopause, etc. while factors from the other group reduce (under specific conditions) the BDNF levels such as Insomnia, Social phobias and others. It means that under certain conditions specified in the article the presence of Epilepsy and Insomnia would cause a Depression after some time, but if there is Epilepsy and there is Social phobias (under specific conditions) only the individual will get sick with Depression too as all these combinations affect the real underlying cause - this unique combination of 2 physiological parameters changes which causes the Depression. The facts that triggering Depression requires only a coexistence 2 physiological changes explains the high prevalence of Depression. The model [1] the current research is based on states that **the more incidence rate of non-infectious disease the less number of coexisting physiological changes beyond 1-sigma** (actually, slightly less) **is required to trigger the disease.** This fact connects epidemiologic data and physiological underlying causes of non-infectious disease.

**Keywords:** Disease causes; Depression disease; Non-infectious disease; Multiple research; Frequency; Disease causing factor; Prevention; Root cause; Disease rate; Physiological parameter; Method to find a cause.

## Introduction

In article [1] “*A Connection between Factors Causing Diseases and Diseases Frequencies: Its Application in Finding Disease Causes*” (Alan Olan, Journal of Clinical Trials, Vol.13, Issue 4.) we introduced a model of non-infectious disease. According to this model which is matching to empirical evidence, the non-infectious disease is caused by changes to multiple physiological parameters when their values go beyond 1-sigma interval, slightly less actually. This means a non-infectious disease must occur if 2 or more particular physiological parameters changes beyond 1-sigma exist for while at the same time. Also, based on the model the criteria was introduced to determine if a risk factor is causing a non-infectious disease or not. In order to be a cause of non-infectious disease the risk factor **K** calculated as (RR - 1) or (OR-1) to be **3.55+/-50%** (or 355%+/-50%) if the factor is causing a change in **1** physiological parameter, and should have value of **19.67+/-50%** (1967%+/-50%) if it is causing a change to **2** physiological parameters beyond 1-sigma. We call these conditions a **Disease Causation criteria**. The Disease Causation criteria determines whether a factor is one of multiple which are causing disease. The disease causing factor cannot cause a disease as a standalone as per model in article [1] because it is usually changing only 1 physiological parameter out of multiple required. Non-infectious disease, as per the model presented in article [1], is caused by multiple physiological parameters changed beyond 1-sigma.

The article [1] introduced a formula to calculate number of non-infectious disease causes if the disease rate (incidence rate usually) is known for the disease in a specific population. The formula is shown below ( **n** is number of disease causes (as number of physiological parameter changed beyond 1-sigma):

$$n = \frac{\log P(D)}{0.22}, \text{ where } P(D) \text{ is the frequency of the disease} \quad (1.0)$$

In this formula P(D) is usually represented by **incidence rate** of non-infectious disease (annual disease rate). Using formula 1.0 and Disease Causation criteria it was shown how to find physiological parameters changes which are causing some non-infectious diseases. In this article we introduce a method which is based on this model, which allow researchers to analyze an existing research about a specific non-infectious disease and determine *a full set* of disease causing physiological parameters for a non-infectious disease in much more complicated cases. Also, a method would allow to map a found disease causing factor to a physiological parameter which is changed by the factor beyond 1-sigma interval. For

example, if we know that Diabetes is a disease causing factor for Hypertension there is a need to know which specific physiological parameter impacted by Diabetes is really a cause of Hypertension. There are multiple physiological factors impacted by Diabetes and the method will allow to find a single one out of so many (allows finding “a needle in a haystack”) which is really causing Hypertension.

The method we introduce is based in math and those who mathematically inclined can find its foundation in the appendix of this article below. Here we will introduce a basic idea of the method and steps on how to practically use it.

## Explanation of the method

A method is introducing an *algorithm* based in math (but *not requiring* to use it much) which allows to process data from results of existing medical researches in few steps and *produce a new information* which consist of disease causes represented by a set of **physiological parameters changes beyond 1-sigma** (slightly less, actually) and also, a few **separate groups of disease causing external factors**. These few groups of factors are such that if you take a disease causing factor from each group as a standalone it will *not* cause a disease but combined together with factors from all other groups *must* cause a non-infectious disease.

Let’s say using the formula for a number of Disease causes (1.0) we found the number of causes as physiological parameter changes beyond 1-sigma. We also found from an existing research a few factors which are really causing a specific disease using a Disease Causation criteria provided. Now we need to determine which specific physiological parameters are really causing a disease. We only know their number and don’t know yet what they are. We know the factors which are causing a change to these physiological parameters and we can use them to find these unknown parameters.

In order to do this we need to list many physiological parameters related to the factor and we can find them from existing scientific research (for example, we can look for all biomarkers of Diabetes and list as many as we can find). Once we list the physiological parameters in the header of some table we need to list the factors found to be causing the disease along a most left column (vertically) of the same table. Based on the researches’ data we already know which physiological parameters are related to those factors causing disease and we mark the intersection between the factor and the related parameter with letter “R” (or other letter). Please note, we use a term “related” as it is unknown if the physiological parameter change is a cause of the disease or not yet and all we know is that its change is present if the factor is present. Now we start analyzing which factors are having same parameter in common and we will call it **an intersection if unrelated disease causing factors A and B both having the same physiological parameter changed**.

We need to find all these intersections. Those intersections will be a superset of the physiological parameters which are causing a disease. The details of why this happens are provided in a detailed math explanation in the appendix of the article. Here we can say that a match in just one physiological parameter changed (*important*: we looking for a **match by name**, not by value) between 2 different factors is not a random coincidence but a pattern. This pattern is determined by a fact that factors are causing a change of the physiological parameter beyond 1-sigma interval (or sometimes by other reasons explained later).

Here is a *simplified* explanation how the method works. Suppose there is a non-infectious disease with 2 unknown disease causes (quantity of 2 we have determined using formula 1.0). Suppose we determined already 2-3 disease causing factors (we can do it using Disease Causation criteria). Each disease causing factor is changing physiological parameters randomly beyond *their original(or normal) range* which means a quantity of physiological parameters changes is random, the **parameters' names**, values are different as well, etc. The physiological parameters changes for a disease causing factors are random as they cannot be determined unless we do an experiment and determine what they have changed under influence of each of these disease-causing factors (we assume we do an experiment the first time). Note, the physiological parameter changes are *random for a disease-causing factor* but they may be almost *the same for different people* impacted by this factor.

Suppose via the experiments we found that a disease-causing factor 1 has changed 20 different physiological parameters beyond their original (or normal) range and a disease-causing factor 2 has changed 50 different physiological parameters beyond their original (or normal) range. We just explained that *only 2 unknown* physiological parameters are really causing this disease. In order to determine **their names** we use a fact that due to random "selection" of physiological parameters out of so many (few thousands and more exist in human body) by *unrelated disease causing* factors, a match by **their names** and values between these 2 groups of physiological parameters practically is not possible by accident (this is shown mathematically in this appendix). If we find a **match of physiological parameters by name** it is a **pattern** under these specified conditions, it is not random.

For example, if one type of *disease-causing* factors (suppose a chocolate consumption) is causing a change in hemoglobin levels and a totally different *disease-causing* factor (suppose a long exercising) is causing a change in hemoglobin levels as well then it is a pattern *under these specified conditions*. It becomes a pattern which we are interested in *because* it happens between groups belonging to 2 different *disease-causing* factors. We know these 2 *disease causing factors* are impacting same physiological parameter,

meaning a parameter with the **same name**. For this reason, in order to find a pattern (in our case a disease causation physiological parameter) we need to find a place where 2 physiological parameters changes in both groups are the **same by name** (*important*: not by value) for these 2 different disease causing factors. This place where 2 disease causing factors have **the same name** of physiological parameter changed beyond original (or normal range) we call "an intersection". If we found such a match between physiological parameters **by name** (*important*: not by value) we found a pattern - a physiological parameter which is causing a disease. See a Figure 1. Also, in Appendix we provide a simple analogy which could help to clarify or visualize the explained concept further.

When we find all these intersections we will find a list (or a matrix) of the parameters which are superset of physiological parameters which changes go beyond 1-sigma and are causing a disease. As there might be other patterns causing a match in parameters the number of parameters can exceed the number of parameters we determined via Number of Disease Causes formula (1.0). We will need to eliminate other parameters based on the methods provided in this article to get a final set of Physiological Parameters which changes go beyond 1-sigma (actually slightly less). These changes when they are present at the same time *must* trigger a disease (this requirement comes from the math model). The illustration of this process shown on Figure 1.

Once we know the final set of these physiological parameters and as we know up front which factor is related to which physiological parameter we can find which specific physiological parameter is impacted by the specific factor in such a way that a parameter goes beyond 1-sigma interval. In other words we get an understanding that a factor is impacting a specific physiological parameter. Now, if an individual is affected by a set of factors which are causing a change in ALL these physiological parameters required then the individual will get sick after sometime unless the harmful factors are removed fast enough. This knowledge can help to prevent a disease in the individual or help potentially cure it or reduce severity of the disease.

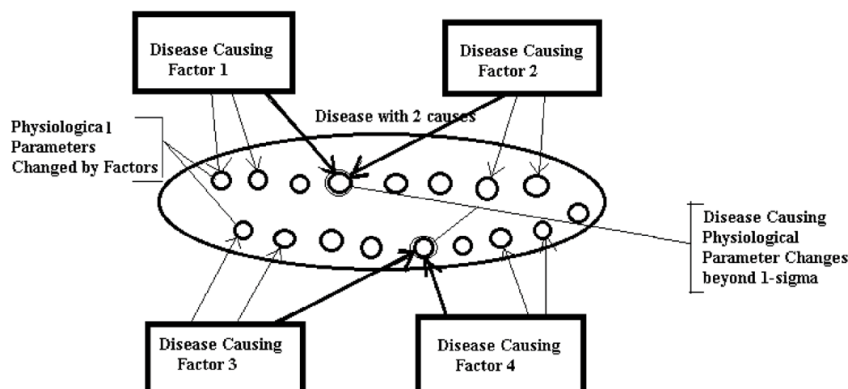
### Using a method in practice

In a nutshell, in order to use this method to find a set of physiological parameters which are causing a disease next steps should be taken:

1. Find the number of causes (as physiological parameters changes) for the non-infectious disease using a formula for a disease causes (1.0) (*see also our article*: "A Connection between Factors Causing Diseases and Diseases Frequencies: Its Application in Finding Disease Causes", Alan Olan, Journal of Clinical Trials, Vol.13, Issue 4.)

2. Find as many as possible, the factors which are causing a disease using a Disease Causation criteria (provided above) to **filter existing experimental data** or new research describing the risk factors for the disease using RR, OR, etc. values. (see our article: “A Connection between Factors Causing Diseases and Diseases Frequencies: Its Application in Finding Disease Causes”, Alan Olan, Journal of Clinical Trials, Vol.13, Issue 4.)
3. For each factor found, you need to list as many as possible parameters related to it in some table. For example, if a factor is a consumption of some food then include which nutrients, chemicals it contains including harmful ones as per existing researches. If a factor is a disease then list which physiological parameters are known to be impacted one way or another as a symptom, a cause, etc. You need to place one factor vertically and appropriate impacted physiological parameters horizontally (in the table’s header) in this table. The column where a factor is related to a physiological parameter can be marked with cross “X” or letter “R”.
4. If the factor is by itself a physiological parameter it can be taken as one of the physiological parameters you are searching for during this step in the method. It can be corrected with next step if needed.
5. Now, you need to find where the columns which contain physiological parameters are causing a crossing between factors. This means you need to **look for a cross (“X” sign) in the same column for 2 or more factors in the table**. We are basically finding which same (by name, not value) physiological parameter is changed by 2 disease causing factors.
6. Build a list or a matrix **Pm** consisting of the parameters which are causing intersections. For example, like Pm: {r1, r19, r23, r45}. In practice Pm can look like **Pm = {“Blood Pressure”, “Estrogen Level Change”, “Oxygen level”, “ROS level”}**
7. Eliminate redundant physiological parameters from matrix **Pm** if required (as per rules provided in this article). If the number of calculated parameters per formula (1.0) is less than number of parameters found via analysis of intersections, then the **redundant parameters need to be eliminated**. The reason of their existence is often a dependence on the disease causative physiological parameter. The number of parameters found via analysis of intersection in a table **should match to the number calculated via a formula** (1.0) or be less if not all parameters can be found due to a lack of experimental data, etc. Also, if some physiological parameters found as result of analysis of intersection in a table are causing some disease causing factors to change more parameters than predicted by disease causing criteria then some physiological parameters may need to be eliminated as well. The disease-causing factor can usually change only 1 or 2 physiological parameters beyond 1-sigma interval (slightly less than this interval, actually) according to multiple experimental data. Extremely **rare** the factors can create 3 or 4 disease causative physiological changes.
8. **Verify the results of using algorithm** by checking if there are appropriate experimental data consistent with the physiological parameters found. Also, you can verify if a physiological parameter found satisfies a Disease Causation criteria above **if** there is any research on the risk of this parameter and correct if any errors in analysis or steps were found.

We can see that a **method allows to use a simple algorithm** on a set of existing experimental data in order to determine disease causes (as a set of physiological parameter changes beyond 1-sigma interval) and **then verify the results using other research** (which was not used doing steps of the algorithm) either by using Disease Causation criteria or by simply checking how this research is consistent with results of your findings.



**Figure 1:** (Disease with 2 causes, as per formula (1.0): Disease causing factors are causing multiple physiological changes. Disease causing physiological changes are common for Factors 1 and 2, for Factors 3 and 4. We find them by determining an intersection in a physiological parameter)

### Example of method’s application for a disease with 2 causes

Let’s look at hypothetical disease with 2 causes and apply to it the method described above. We created a Table 1. We can see in **Table 1** that all physiological parameters related to Factors are listed on the top and all found from the experiments by filtering using the Disease Causation criteria Factors which are causing this hypothetical disease are listed on the left. Now, we can see that parameter **r20** is an intersection for Factor 3 and Factor 4, param **r12** is an intersection for Factor 2 and Factor 3 and Factor 4. A param **r17** is an intersection for Factor 1, Factor 2. We can build our matrix **Pm: {r12, r20, r17}**. In practice, this matrix may look like a list **Pm: {‘Blood Pressure Decrease’, ‘Oxygen Level Decrease’, ‘Estrogen Increase’}**. This matrix contains physiological parameters which changes beyond 1-sigma are triggering a disease if **present at the same time** long enough.

Now, we see the matrix **Pm: {r12, r20, r17}** has 3 elements but our disease only require 2. We need to eliminate the incorrect redundant parameter. We can notice that if we take a valid combination of 2 param as {r12, r17} then a param r12 is showing that *Factor 2 is impacting r12 and r17 parameters but Factor 2 only can cause 1 parameter impact* (as determined by Disease Causation criteria) so the combination is not correct. If we eliminate r12 and take a combination of {r20, r17} then both params are satisfying our requirement (assuming there is no info their changes are within 1-sigma). So from this table we can see that final physiological parameter’s set which is causing this disease is **{r20, r17}**. In this case Factors 3 and 4 are causing a change in parameters **r20**, Factor 1 and 2 are causing a change in physiological parameter **r17**. We has determined 2 physiological causes for a disease with 2 causes.

**Table 1:** (Areas in gray are where an intersections happen in parameters r12, r20 and r17 accordingly. Area surrounded by bold frame is a final set of parameters for matrix Pm)

Factors causing Disease	Physiological Parameters - Entire Set for ALL FACTORS as R1, R2, ... Rn													
	r12	r20	r2	r8	r17	r75	r34	r56	r39	r12	r55	r23	r78	r11
Factor 1			x		x				x	x		x		
Factor 2	x				x			x			x			x
Factor 3	x	x					x							
Factor 4	x	x				x							x	

### Criteria for choosing intersections of physiological parameters

As we show mathematically in the appendix of this article there is a very small chance of random intersection of physiological parameters belonging to 2 different factors (which can be significantly reduced by introducing more factors with same intersection) and also, there are cases when there are more parameters found then determined by disease causation criteria due physiological parameters interdependency, etc. In order to select those parameters which are really causing a disease we need to use these criteria below.

1. We need to check if factors which are impacting physiological parameters found to be intersections are belonging to **same type of disease** or they are **very similar in nature** (like BMI impact and high weight impact) and we can ignore them as the intersection is most likely due to similarity of factors.
2. Make sure **a factor is causing a change the physiological parameter** which intersect (not opposite) As the factors should cause a change in physiological parameters.
3. **Make sure both factors are causing change in the**

**physiological parameter which intersect** or the intersection is not valid.

4. If a **factor causing insignificant change (less 1-sigma interval) in physiological parameter then the parameters need to be ignored as a point of intersection** with this factor. The factors as we discussed need to make change in physiological parameter so its value will be beyond *1-sigma* interval (slightly less, actually).
5. Make sure the **physiological parameters changes in the point of intersection happen in the same direction** ( either increase or decrease). If the changes move in different directions then the parameters are considered as 2 different parameters and not as one.
6. **The intersection of parameters for a similar factor is not considered correct** intersection (as it is the same factor, for example Increased Weigh and BMI over 30 are similar factors).
7. If the experiment found that a factor which causing a disease is impacting 1 physiological parameter but the factor consists of combinations of 2 separate factors (for example observation was done for presence of diabetes or obesity together so they work as a factor) then **the**

**combined factor can be allowed to have 2 intersections despite having impact on 1 physiological parameter** as one of the parts of the factor could be present at the time (for example either a diabetes causing 1 change in physiological parameter or obesity causing the change in another physiological parameter).

8. **Element cannot cross itself**, for example if factor is IgA and physiological parameter IgA then they don't create a crossing. Usually, if a factor found as physiological parameter then we already know this is a physiological parameter causing a disease.
9. In many cases it is possible to **eliminate the redundant intersection by determining** via a known research **if the physiological parameter increases or decreases with the presence of specific factor**. For example, does lipoprotein level decreases or increases with diabetes as causation factor? **If same parameter where we see a intersection, increases in presence of one factor and decreases in presence of another** then the intersection is invalid.
10. Once a set of physiological parameters was found **they can be also validated with already provided Disease Causation criteria to determine if the factor is a cause of disease** ( a parameter should increase a risk of disease 3.5 +/-50% times) using an existing research. For example, if we find that an increased lipoprotein is causing a stroke and if some existing research confirms that a risk of its increase is **3.5 +/-50%** then it confirms a lipoprotein as a parameter was found correctly.
11. In some cases biochemical analysis may help to eliminate an element from matrix Pm.
12. **From practical standpoint**, due to current reliance on experimental confirmation there might be not always a need to remove redundant parameters as per above if their number is determined to be small (for example 1 or 2) as it can be part of observed set of data for the medical experiment and the redundant physiological parameters may be removed via the experimental confirmation of physiological causes which were found using this method.

### Algorithms to eliminate redundancy

Below we provide an approximate algorithm which can be used to efficiently eliminate redundant physiological parameters when the specific disease analysis requires it:

1. Find out if all parameters which intersect in your table satisfy *a hard condition* that a factor only can cause as many changes to physiological parameters as determined via a disease causation criteria (which in most case is 1 and in some 2 or very rare 3). If some physiological parameters are causing not permitted number of changes by a disease causing factor then some parameters need to be eliminated (ignored as intersections).

2. Check the factors which having intersections. If the factors are similar or related then their intersections can be due to similarity and ignored.
3. Select a set of physiological parameters which from a common sense will make most of factors intersect in them. Validate if your selection is following rule number 1) above. If not use the elimination criteria provided in this article to help resolve conflicts.
4. If 3 factors which are not related intersect in the same physiological parameter then treat a physiological parameter as a high priority intersection as it is extremely low probability that this intersection is not a pattern.
5. If a physiological parameter is selected as a real intersection surround it with a bold rectangle to differentiate from original intersections.
6. Remember that if 2 factors intersect it is a pattern which is determined by next facts 1) that factors are causing a same changes beyond 1-sigma to a physiological parameter **or** 2) the intersection occur due to dependency between related factors or related physiological parameters but the changes are not necessarily beyond 1- sigma. In other words in this case the intersection happens due to this relationships not necessarily due to 1- sigma change (but it can be both at the same time: due to relationship and also due to 1-sigma change).

## Methodology

### Analysis of Depression Causes

Now we accomplish the analysis of Depression as per method presented above in order to find its real causes and we also demonstrate how to systematically build a hypothesis of non-infectious disease pathology using this method and provide a hypothesis for Depression using those found physiological triggers of this disease. Using this example we demonstrate in more details how to use a method on relatively *mathematically* simple analysis of Depression which does not require elimination of redundant physiological parameters.

According to existing data Major Depression is affecting more than 8% of American adults each year (<https://mhanational.org/conditions/depression#:~:text=Basic%20Facts%20About%20Depression,are%20affected%20by%20major%20depression>). Using this data as 0.08 and calculating number of causes using formula (1.0) we get:

$$N = \log_{0.22}(0.08) = 1.66 \approx 2$$

As per calculations above the depression is caused by changes to **2** physiological parameters beyond 1-sigma (actually, slightly less than 1-sigma). Despite the well observed fact that the rate of disease is usually different across different states or areas the number of disease causes which will be determined using this formula for other areas

will still be as 2 (except sometimes very rare critical, well explainable and useful for disease prevention cases)! This was shown extensively for multiple diseases in the article [1] “A Connection between Factors Causing Diseases and Diseases Frequencies: Its Application in Finding Disease Causes” (Alan Olan, Journal of Clinical Trials, Vol.13, Issue 4.) and Depression specific data presented in **Table 1.0**. The formula we used above is **connecting epidemiologic data on disease rate with the physiological nature of disease causation**.

**Table 1.0:** Depression disease rate 2008, 2009, 2010 average (Had at Least One Major Depressive Episode in the Past Year among Persons Aged 18 or Older) and its connection to depression causes across different USA states (disease rate sources <https://www.samhsa.gov/data/report/2009-2010-nsduh-state-estimates-substance-use-and-mental-disorders>)

State / Country	Depression rate in percent (%)	Number of depression causes (N=log <sub>0.22</sub> P(D))	Comments
California	5.97	2	Each calculation in column 3 represents a rounded value, the result is a number of physiological parameters' changes which are causing depression if present together.
Connecticut	6.24	2	
Idaho	7.81	2	
Michigan	6.90	2	
Pennsylvania	6.41	2	
Vermont	6.61	2	
Texas	6.03	2	
Depression's 12 months prevalence in countries, 2010 (Data source <a href="https://pmc.ncbi.nlm.nih.gov/articles/PMC4100461/">https://pmc.ncbi.nlm.nih.gov/articles/PMC4100461/</a> )			
Belgium	5.0	2	
France	5.9	2	
Israel	6.1	2	
Italy	3.0	2	
Spain	4.0	2	

Now according to the method we are filtering risk factors of Depression and find the ones which are real Depression causes, list them here and explain how they are determined based on Disease Causation criteria provided above and developed in article “A Connection between Factors Causing Diseases and Diseases Frequencies: Its Application in Finding Disease Causes” (Alan Olan, Journal of Clinical Trials, Vol.13, Issue 4.) :

1. Study shows: “Working 11+ hours a day was related to a 2.43-fold odds of MDE compared to working 7 to 8 hours a day in an analysis adjusted for socio-demographic characteristics”,

It was also robust to additional adjustment for work characteristics (job strain and social support at work), **the**

**odds ratio of MDE being 2.52-fold for 11+ working hours in the final model.**” (“Overtime Work as a Predictor of Major Depressive Episode: A 5-Year Follow-Up of the Whitehall II Study”, Marianna Virtanen, Stephen A. Stansfeld, et al, Plos One, published: January 25, 2012, <https://doi.org/10.1371/journal.pone.0030719>). We get **OR = 2.52 which is the range of 4.5+/-50% (2.27- 6.81)** and it means the factor is one of 2 causes of Depression and impact 1 physiological parameter (as per Disease Causation criteria). We list this factor in Table 1.1 (Factors column).

Another study about Major Depression Disorder (MDD) shows: “individuals with insomnia but without, any psychiatric disorders were also more likely to develop a new-onset MDD in the subsequent year (**OR=5.4,95% CI=2.6-11.3**) compared with individuals with neither insomnia or psychiatric disorders.”, “In a longitudinal study of 979 young adults,<sup>24</sup> **insomnia increased the relative risk for depression fourfold** (95% 0=2.2-7.0) over a 3-year period” (“Sleep disturbances and depression: risk relationships for subsequent depression and therapeutic implications”, Peter L. Franzen, PhD\*, Dialogues Clin Neurosci. 2008 Dec; 10(4): 473–481, doi: 10.31887/DCNS.2008.10.4/plfranzen). As we can see **OR = 4 is the range of 4.5+/-50% (2.27-6.81)** and **OR =5.4** as well, so it means insomnia is a causation factor for a depression and is one of 2 causes of depression and is impacting 1 physiological parameter. We list this factor in Table 1.1 (Factors column)

Study shows: “In addition, initial presentation with **social phobia was associated with a 5.7-fold increased risk of developing major depressive disorder.**” (“The Critical Relationship Between Anxiety and Depression”, Ned H. Kalin, M.D. , The American Journal of Psychiatry, Published Online:1 May 2020 <https://doi.org/10.1176/appi.ajp.2020.20030305>). If **OR =5.7** it is within range of **4.5+/-50% (2.27-6.81)** of Disease Causation criteria for 1 physiological parameter. It means social phobia is a causation factor for major depressive disorder and impacting 1 physiological parameter. We list this factor in Table 1.1 (Factors column)

Study shows: “Subgroup studies indicated that **epilepsy was associated with an increased risk of depression in Asian, African and Caucasian populations** (Asian: **OR/RR = 2.42**; 95% CI: 1.48-3.95; African: **OR/RR = 2.48**; 95% CI: 1.88-3.28; Caucasian: **OR/RR = 1.86**; 95% CI: 1.60-2.15). Subgroup studies showed that **epilepsy was associated with an increased risk of depression among adolescents and adults** (adolescents: **OR/RR = 2.54**; 95% CI: 1.86-3.46; adults: **OR/RR = 2.22**; 95% CI: 1.79-2.75)” (“Association between epilepsy and risk of depression: A meta-analysis”, Chu Chu, Psychiatry Research, Volume 312, June 2022, 114531). We can see that **OR = 2.42, OR/RR = 2.48, OR/RR = 2.54** are within a Disease Causation criteria **4.5+/-50% (2.27-6.81)**. It means that Epilepsy is one of the causes of

depression and impacts 1 physiological parameter. We list this factor in Table 1.1 (Factors column)

Study shows that: **“women in late transition to menopause were nearly 3 times more likely to report these depressive symptoms compared with pre-menopausal women (OR, 2.89; 95% confidence interval [CI], 1.29-6.45; P = .01).”** (*“Hormones and Menopausal Status as Predictors of Depression in Women in Transition to Menopause”*, Ellen W. Freeman, PhD; Mary D. Sammel, ScD; Li Liu, MD, MS and et al, Arch Gen Psychiatry. 2004;61(1):62-70. doi:10.1001/archpsyc.61.1.62). We can see that **OR = 2.89** and it is in within a range of **4.5+/-50% (2.27-6.81)** as per Disease Causing criteria. It means that being a women in late transition to menopause is a disease causing factor which impacts 1 physiological parameter out of 2 which are causing depression. We list this factor in Table 1.1 (Factors column).

Research on Seasonal affective disorder (seasonal depression).” An analysis of sequence variations in three genes that form a functional unit of the circadian clock (**PER2, ARNTL, and NPAS2**) found a **single nucleotide polymorphism** in each gene with a significant association with seasonal affective disorder. The authors also reported an additive effect and identified a risk genotype combination and a protective genotype combination. **They reported that carriers of a risk genotype have a four-fold risk of having seasonal affective disorder** than other genotype combinations and a ten-fold risk compared to those with the protective genotype” (*“Seasonal affective disorder (seasonal depression) – symptoms, risk factors, causes and genetics”*, Medicover Genetics, Medicover Genetics Editorial Team, November 16, 2022). If we take four-fold risk as **OR= 4** then it is within a range for Disease Causation criteria **4.5+/-50% (2.27-6.81)**. It means a particular PER2, ARNTL, and NPAS2 **single nucleotide polymorphism** is a causation factor for a depression and impacting 1 physiological parameter. We list this factor in Table 1.1 ( Factors column).

We know now some of the factors (which we could find from research, there may be others) which are causing Depression but we need to find out: **1)** which physiological parameters are really changed by this factors to cause a disease. We know there are 2 physiological parameters like this via a calculation using formula (1) above but they names are unknown yet. **2)** we need to find whether the causation factors we found impact 1st or 2nd of these physiological parameters. We want to find out for each factor which physiological parameter (1st or 2nd) are impacted by this specific factor. See Figure 2 for illustration.

In order to determine this we have listed the factors above in a Table 1.1 vertically (*as an alternate way*) in factor’s column (1) and will find as many as possible the physiological parameters which are related to these factors and list them horizontally (in the table’s header). We mark the relationship

between a factor and a physiological parameter in the table with letter “R”. It means there is a change to a physiological parameter when the factor is present. As we can see one factor may be related to changes to multiple physiological parameters. We need to find out the intersections of the parameters across the factors as explained in the method above. These intersections, after elimination of invalid parameters (if they exist) will give us a set of 2 physiological parameters changes beyond 1-sigma which **must** cause a Depression when present at the same time. We stress **these physiological parameters changes beyond 1-sigma must cause the Depression if co-exist, it is not optional** as per the model presented in article [1].

Here we provide a list of physiological parameters related to factors of depression as citations from respective studies (note: our bold font):

1. “Two studies showed **increased oxygen consumption across 24 h in insomnia patients** compared with good-sleeping controls.” (“Is Metabolic Rate Increased in Insomnia Disorder? A Systematic Review”, Julia L. Chapman, Maria Comas, Camilla M. Hoyos, et al, Front Endocrinol (Lausanne). 2018; 9: 374, published online 2018 Jul 16. doi: 10.3389/fendo.2018.00374). Based on this we put this parameter in Table 1.1 as “High Oxygen consumption” in column (2) and add “R” letter for Insomnia there.
2. BDNF and Sleep Disorder: “Studies conducted up to this point on the subject of changes in **peripheral BDNF level in insomnia consistently show BDNF reduction in individuals afflicted with this sleep disorder**, regardless of sex or accompanying psychiatric conditions;” (*“Investigating the Role of BDNF in Insomnia: Current Insights”*, Dittmer M, Gabryelska A, Turkiewicz S, Sochal M, Nature and Science of Sleep, 7 December 2023 Volume 2023:15 Pages 1045-1060). Based on this we put this parameter in Table 1.1 as “Low BDNF” in column (5) and place letter “R” for Insomnia
3. Social phobia in animal studies: “According to an experimental study with laboratory animals, **stress by social subjugation is able to foster hyperalgesia with a decrease in the BDNF levels**; this condition is more evident in “susceptible” subjects rather than in “resilient” ones” (*“BDNF Protein and Anxiety Disorders”*, Tatiana Marins Farias, Rebeca Ataíde Cerqueira, Danton Ferraz Sousa, et al, Neurological and Mental Disorders, Published: 28 May 2020 ). Based on this we put letter “R” in Table 1.1 under column (5) for Social Phobia factor as related to it.
4. Long hour and burn out: “...study reported that **the serum BDNF (sBDNF) levels were significantly lower in burnout subjects than those in healthy people.**

Moreover, sBDNF levels were negatively correlated with EE and CY, but positively with PE (Sertoz et al., 2008). Our recent study also showed an inverse association between job burnout and sBDNF level in doctors and nurses (He et al., 2017). (“Interaction between job stress, serum BDNF level and the BDNF rs2049046 polymorphism in job burnout”, Shu-Chang

He, Shuang Wu, Chao Wang, Dong-Mei Wang et al, Journal of Affective Disorders, Volume 266, 1 April 2020, Pages 671-677). Based on this we put letter “R” in Table 1.1 under column (5) for a factor Working long hours (1 cs).

“Mechanistically, single-particle tracking revealed that E2 [Estradiol] treatment selectively reduced the dwell time and thereby decreased the confinement of GABAA Rs at inhibitory synapses.” and also “In summary, our studies provide a molecular mechanism by which estrogen acts to reduce the efficacy of GABAergic inhibition by decreasing the stability of inhibitory synapses.” (“Estradiol modulates the efficacy of synaptic inhibition by decreasing the dwell time of GABAA receptors at inhibitory synapses”, Jayanta Mukherjee, Ross A. Cardarelli, Yasmine Cantaut-Belarif, et al, PNAS, Vol. 114 | No. 44, <https://doi.org/10.1073/pnas.1705075114>). Based on this we put this physiological parameter in Table 1.1, column (4) and put a letter “R” for a factor “Women in late transition to menopause (high estrogen)” as women have significant level of estrogen comparing to men and its levels change periodically.

“the present study reveal that estrogen decreases the GABA<sub>B</sub> receptor-mediated autoinhibition of GABAergic POA neurons. In addition, estrogen induces an apparent decrease in the subsequent ability of these neurons to synthesize GABA.” (“Estrogen Biphasically Modifies Hypothalamic GABAergic Function Concomitantly with Negative and Positive Control of Luteinizing Hormone Release”, Edward J. Wagner, Oline K. Rønnekleiv, Martha A. Bosch, and Martin J. Kelly, J Neurosci. 2001 Mar 15; 21(6): 2085–2093, doi: 10.1523/JNEUROSCI.21-06-02085.2001). Based on this we put “Estrogen Modulation of GABA” in Table 1.1, column (3) and mark the factor of “Women in late transition to menopause (high estrogen)” with letter “R” under this column.

5. Role GABA in Epilepsy: “Experimental and clinical study evidence indicates that GABA has an important role in the mechanism and treatment of epilepsy: (a) Abnormalities of GABAergic function have been observed in genetic and acquired animal models of epilepsy; (b) Reductions of GABA-mediated inhibition, activity of glutamate decarboxylase, binding to GABAA and benzodiazepine sites, GABA in cerebrospinal fluid and brain tissue, and GABA detected during microdialysis studies have been reported in studies of human

epileptic brain tissue; ” (“GABAergic mechanisms in epilepsy”, D M Treiman, Epilepsia 2001;42 Suppl 3:8-12. doi: 10.1046/j.1528-1157.2001.042suppl.3008.x.). Based on this we put a letter “R” under “Reduction of GABAergic inhibition” in Table 1.1, column (4) and place a letter “R” in Epilepsy under this column.

6. Role GABA in Epilepsy: “Reductions of GABA mediated inhibition and decreased activity of GAD has been reported in studies of human epileptic brain tissue” (“Decreased GABA receptor in the cerebral cortex of epileptic rats: effect of Bacopa monnieri and Bacoside-A”, Jobin Mathew, Savitha Balakrishnan, Sherin Antony, Pretty Mary Abraham & CS Paulose, Journal of Biomedical Science, 19, Article number: 25 (2012)). We already placed a letter “R” under “Reduction of GABAergic inhibition” column for Epilepsy. This is 2nd confirmation of relationship between epilepsy and GABA mediated inhibition.

Let’s analyze and find these intersections of the physiological parameters for the factors impacting Depression in Table 1.1. We checking in which column parameters intersect. We can see that only parameters in columns 4 and 5 are causing intersection for all set of factors listed. We mark the intersection in gray color to differentiate them. Remember, we are looking for physiological parameters match by name (not by value).

The areas which are gray out in Table 1.1 show the intersections of physiological parameters but they are not yet a final list. Only once we eliminate redundant parameters (if they exist) then we can surround the final parameters with bold rectangles.

By looking at the intersections of physiological parameters in Table 1.1 we can write down a matrix  $P_m = \{ \text{“Reduction of GABAergic inhibition”, “Low BDNF”} \}$  which has only 2 physiological parameters so we don’t need to eliminate any. We also observe that this matrix satisfies a hard condition that factors should only intersect in 1 physiological parameter. This matrix represent a set of 2 physiological parameters changes beyond sigma which if present simultaneously will trigger depression within some time. We can list them like below:

1. Reduction of GABAergic inhibition beyond ~1 sigma interval
2. Low BDNF ( beyond ~1 sigma)

The columns in Table 1.1 surrounded by bold rectangles represent the physiological parameters found as result of our analysis and which are causing a disease.

There is a factor of Circadian clock genes PER2, ARNTL, and NPAS2 SNP which should be causing impact to 1st or 2nd physiological parameters of the Depression. Which

one is it? Is this **“Reduction of GABAergic inhibition” or “Low BDNF”**? As this factor is causing only 1 impact to physiological parameter then it should be impacting only one of them to be beyond 1-sigma. The existing research below suggests that these Circadian clock genes’ single nucleotide polymorphism (SNP) should cause **“Reduction of GABAergic inhibition”**:

1. One research informs: “These results: (1) implicate Npas2 in the response to stress and the development of anxiety; and (2) provide **functional evidence for the regulation of GABAergic neurotransmission by NPAS2 in the ventral striatum** (*“NPAS2 regulation of anxiety-like behavior and GABAA receptors”*, Ryan Wellington Logan, Frontiers in Molecular Neuroscience, November 2017)
2. Another research states: “Overall, our results provide evidence for a specific **role of glial Per2 in mood-related behavior, accompanied by dysregulation of components of the glutamatergic, GABAergic and dopaminergic signaling**” (*“Deletion of the clock gene Period2 (Per2) in glial cells alters mood-related behavior in mice”*, Scientific Reports, Tomaz Martini, Jürgen A. Ripperger, Jimmy Stalin, et al)
3. Another research states: **“The release of neurotransmitters, such as dopamine, glutamate, and  $\gamma$ -amino-butyric acid (GABA) have been shown to be modulated by circadian rhythms** (Castaneda et al., 2004). **Per2 is associated with the generation of the circadian rhythms** (Arjona and Sarkar, 2006; Sujino et al., 2007) (*“Neurobiological Functions of the Period Circadian Clock 2 Gene, Per2”*, Mikyung Kim, June Bryan de la Peña, et al, Biomol Ther (Seoul). 2018 Jul; 26(4): 358–367.)

On this example of determining the Circadian genes role in Depression you can see that even if we don’t know exact impact of the causation factor on a specific physiological parameter **using results of analysis** we still could determine just 2 possible options for this impact and we are able to suggest which option will likely be correct one based on existing research. Based on this we can place letter “R” as suggested for the “Circadian clock genes **PER2**, **ARNTL**, and **NPAS2** SNP polymorphism” factor in Table 1.1 in column 4 (**“Reduction of GABAergic inhibition”**) to signify the relationship between the genes’ SNP and reduction of GABAergic inhibition,

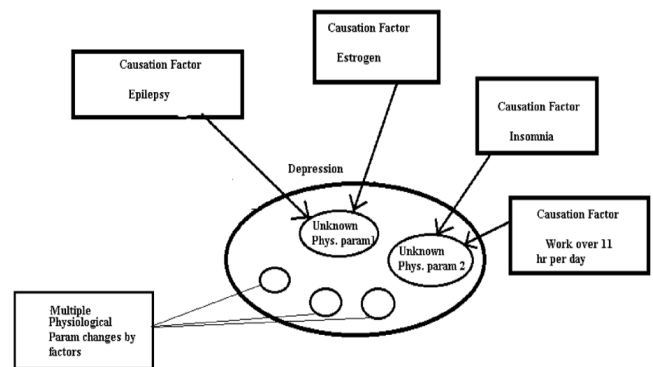
Now let’s take a look at empirical evidence related to the found physiologic causes of Depression by looking at existing researches which are provided below:

- 1) One article informs that: “Studies have demonstrated **low concentrations of  $\gamma$ -aminobutyric acid (GABA) in the plasma and CSF of individuals with major**

**depression**, and low GABA concentrations have also been found in the occipital cortex of depressed subjects” (*“Increased Occipital Cortex GABA Concentrations in Depressed Patients After Therapy With Selective Serotonin Reuptake Inhibitors”* (by Gerard Sanacora, M.D., Ph.D., Graeme F. Mason, et al, The American Journal of Psychiatry, 1 Apr 2002)

and it is consistent with a reduction of GABAergic inhibition found in analysis according to our method. Another article provides a similar view that **“Evidence suggesting that central nervous system  $\gamma$ -**

**aminobutyric acid (GABA) concentrations are reduced in patients with major depressive disorder (MDD)** has been present since at least 1980, and this idea has recently gained support from more recent magnetic resonance spectroscopy data. These observations have led to **the assumption that MDD’s underlying etiology is tied to an overall reduction in GABA-mediated inhibitory neurotransmission.”**(*“Altered  $\gamma$ -aminobutyric acid neurotransmission in major*



**Figure 2:** Illustrates how different causation factors are impacting same 2 disease causing physiological changes in Depression. Notice how different factors “intersect” in the same physiological parameter

*depressive disorder: a critical review of the supporting evidence and the influence of serotonergic antidepressants”* (by Pehrson A, Sanchez C, Drug Design, Development and Therapy, published 19 January 2015 Volume 2015:9). This also consistent with analysis a per our method which has found that a reduction of GABAergic inhibition is a disease cause for Depression by *using not directly related data - from researches on epilepsy and insomnia!*

- 2) An article states: **“The role of BDNF in depression has gained broad attention because many pre-clinical and clinical studies provide direct evidence** suggesting that modulation in **expression of BDNF** could be involved in behavioral phenomenon associated with depression.”(*“Brain-derived neurotrophic factor: role in depression and suicide”* (by Yogesh Dwivedi, Neuropsychiatr Dis Treat. 2009; 5: 433–449.) which is also consistent with analysis’s result made according to our method that a deficiency of

neurprotective peptides (BDNF) is one of the causes of depression. Another article informs: “The **studies mentioned in this review article greatly support the role of BDNF in the pathogenesis of depression** and treatment of this disorder with antidepressants.” (“*Unfolding the Role of BDNF as a Biomarker for Treatment of Depression*” ( by Tarapati Rana, Tapan Behl, et al, Journal of Molecular Neuroscience November 2020, volume 71, pages2008–2021 (2021)) and also consistent with BDNF as disease cause of depression found algorithmically via our method with a use of multiple available researches on factors **not directly related to BDNF** (the research we used was related to insomnia, social phobia and long working hours)!

We can see that existing research generally in agreement with result of our method, the research points to Reductions of GABA mediated inhibition and BDNF levels. Important to notice here that conclusions per our method were done based on researches on risk factors which **were not related to GABA or BDNF**. Using our method we came to these conclusions by using a research for **the factors which were causing disease** and they were not *explicitly referring* to Reductions of GABA mediated inhibition or BDNA level. Our results came from analysis of factors causing disease by using a pretty simple algorithm but these results gave a totally *new set of knowledge* of disease causation for Depression disease and which is consistent with an empirical evidence.

Now we can move data from **Table 1.1 to Table 1.2** which is a table of results and can be used for results validation by checking number of cases (physiological parameters

changes) found for a specific factor with quantity determined by a Disease Causation criteria for the factor.

Now we can move data from the Table 1.2 to Table 1.3 which will show the 2 physiological parameters causing a Depression and also show which factors are impacting them. By looking at this table we can determine which external factors will cause Depression in an individual if combined together with another one

Now let’s recap what we done. We calculated number of causes for Depression based on the incidence rate and determined a number of causes as 2. Then we found the factors which are causing depression using a Disease Causation criteria for a Risk factor (or OR) related to 1 physiological parameter change beyond 1-sigma. We listed the found factors in the Table 1.1 along with related physiological parameters. By looking at the table we found intersections where physiological parameters are same *by name* for different factors. Based on these intersections we built a matrix **Pm** with physiological parameters. We analyzed the **Pm** matrix and found there is no redundancy in parameters as we found 2 and this is expected to be 2. So **we found the physiological parameters causing disease purely from knowing disease causing factors by the simple algorithm** and we can see that **they match to the experimental data!** It shows the algorithm works.

Based on the analysis of external factors **a depression is caused by simultaneous co-existence of 2 physiological parameters changes beyond ~1 sigma long enough:**

**Table 1.1** R - means that the physiological parameter is related to the factor cs - a cause (means a factor is causing a physiological parameter change beyond 1-sigma , 1cs - 1 parameter changed)

Factors causing Depression (1)	Physiological Parameters for Factors ( ALL SET) causing Depression			
	High Oxygen Consumption (2)	Estrogen Modulation of GABA (3)	Reduction of GABAergic inhibition (4)	Low BDNF (5)
Insomnia (1cs)	R			R
Social Phobia (1 cs)				R
Women in late transition to menopause / High estrogen (1 cs)		R	R	
Epilepsy (1 cs)			R	
Circadian clock genes PER2, ARNTL, and NPAS2 SNP Polymorphism 1 cs)			R (suggested)	
Working long hours (1cs)				R

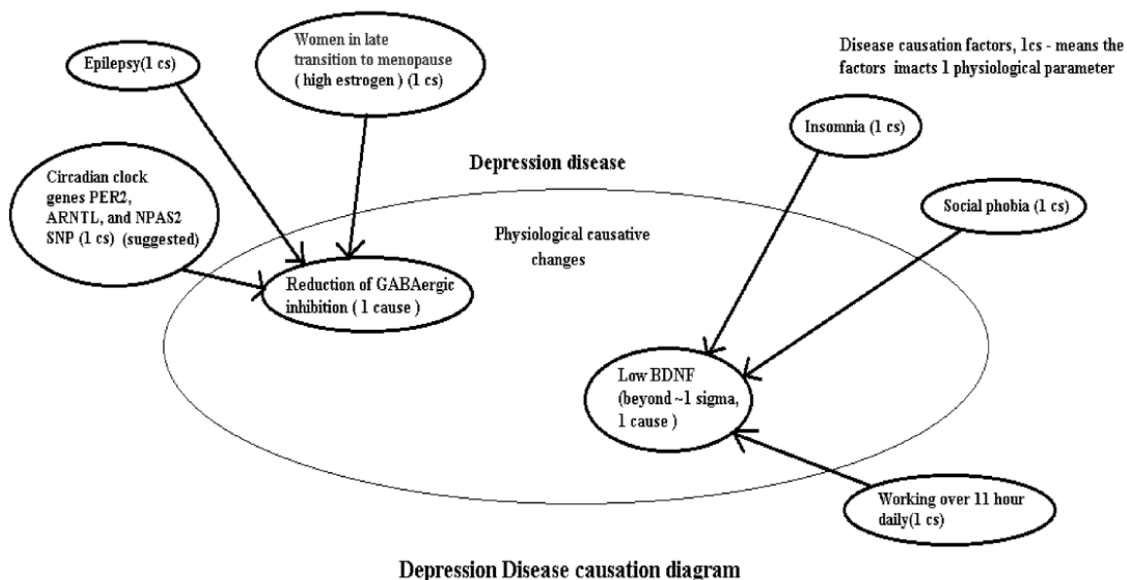
**Table 1.2**

External Factors (1)	Physiological parameters which changes are causing Depression	
	Reduction of GABAergic Inhibition (2)	Low BDNF (3)
Insomnia (1 cs)		x
Circadian clock genes PER2, ARNTL, and NPAS2 SNP polymorphism (1 cs)	X (suggested)	
Epilepsy (1 cs)	x	
Working over 11 hour daily (1 cs)		x
Social phobia (1 cs)		x
Women in late transition to menopause / High estrogen (1 cs)	x	

**Table 1.3:** External factors which can cause a depression shown in Table 1.3. Any factor combination out of 2 different columns should cause depression.

Reduction of GABAergic inhibition (1)	Low BDNF (2)
Epilepsy	Insomnia
Women in late transition to menopause / High estrogen (1 cs)	Social phobia
Circadian clock genes PER2, ARNTL, and NPAS2 SNP polymorphism (1 cs)	Working over 11 hour daily

**Note:** the list of factors in table 1.3 is not complete and based on research known to an author. There can be other disease causing factors but they still should impact these 2 physiological parameters as these parameters are really causing a disease. (There might be few exceptions to this rule though as some factors may have caused a change to a frequency of the disease over the time.)



**Figure 3:** Depression Disease causation diagram shows connection between disease causation factors which here impacting one physiological change so it goes approximately beyond 1-sigma (slightly, less actually) and the impacted physiological parameter. Note: the impacts in some cases is logical (presence of the factor signifies the presence of the physiological change). 2 physiological parameters need to be present at the same time long enough to trigger the Depression.

1. Reduction of GABAergic inhibition beyond ~1 sigma interval
2. Low BDNF (beyond ~1 sigma)

It means a person will have a depression disease if his/her GABAergic inhibition decrease beyond ~1 sigma level AND there will be a low level of BDNF (beyond ~1 sigma).

For example, if we take Epilepsy from the column 1 and combine it with Insomnia from column 2 then we can determine that once an individual has this combination he/she should develop a Depression at some point of time unless one of these factors eliminated. The **Depression Disease**

**causation** diagram in **Figure 3** illustrates this process in more details.

Now the question is how long it takes to develop Depression if both physiological causes are present? This can be estimated at this point experimentally. For example, the research by Peter L. Franzen, PhD\* regarding insomnia which we discussed above points that Depression can develop within a year of presence of those causes.

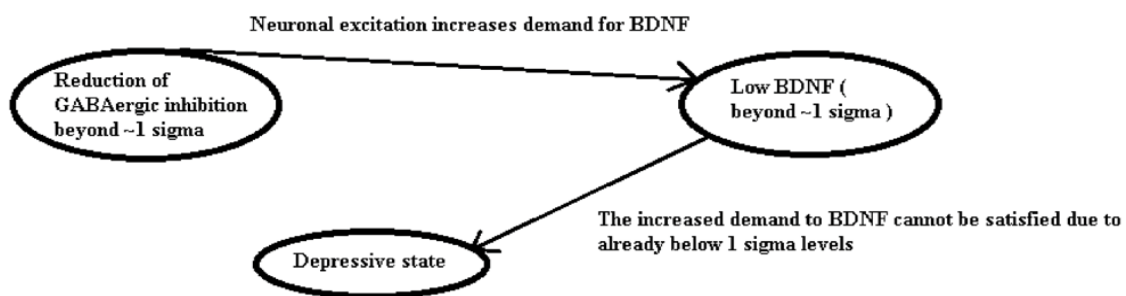
**Building a hypothesis of disease pathology knowing physiological parameters causing the disease**

As we could see we have found a structure of mechanism

of Depression disease - we have determined that it consist of “Reduction of GABAergic inhibition beyond ~1 sigma interval” and “Low BDNF (beyond ~1 sigma)”.But **knowing a structure of mechanism of the disease (the physiological parameters causing the disease) allows us to build a hypothesis of pathology** of the disease by “connecting the dots” of this structure using existing physiological studies and other research. For example, we can build next hypothesis of Depression pathology.

**Hypothesis of Depression pathology** (using 2 found physiological parameters causing a disease): GABAergic inhibition is getting reduced beyond -1 sigma by actions of disease causing factors (epilepsy, estrogen changes, etc) and this should increase excitation of neurons which in sequence should increase a demand for BDNF [20]. At the same time due to a low level of BDNF (beyond 1-sigma) already being present because of other disease causation factors (Insomnia, Working over 11 hours, etc) the increased demand to BDNF

**External factors of Depression are causing changes to 2 physiological parameters so they drop below 1-sigma value**



**Hypothesis of Depression pathology**

**Figure 4: The Diagram of Depression pathology (hypothesis)** illustrates how the drop of GABAergic inhibition below 1-sigma due variety of external factors discussed above increases neural excitation and the required due to this demand the higher BDNF level cannot be satisfied as other external factors already brought BDNF level below 1-sigma level. The brain goes to depressive state.

cannot be satisfied and actually BDNF level is reduced even more and this disrupts the normal work of the brain and causes a Depression.

It is well known that BDNF is playing a crucial role in long-term potentiation (LTP) and synaptic plasticity, etc. Also, low levels of BDNF are reported in major depression. The diagram illustrating the hypothesis and relationship between disease causing physiological changes during disease development can be seen in **Figure 4**.

We have demonstrated here that knowledge of physiological causes of non-infectious disease significantly simplifies creation of hypothesis of disease pathology as it only requires to connect this causes together using usually known physiological and other facts from existing research. We could also notice the method looks at the physiological causes as independent during triggering of disease - the physiological parameters can be changed by external factors independently beyond 1-sigma but obviously we can see the physiological parameters have interactions between themselves during the disease development.

The hypothesis then can be proved or rejected with future research.

**Principle of Indifference**

As we discuss a method to find disease causes via experiments we are proposing to use a simple principle which can be observed in physiological processes. **When regulating its own homeostasis separate physiological systems of the body are indifferent to the side effects of this regulation.** We can call this a *principle of indifference*.

We can observe this principle in multiple cases. For example, if a brain needs to increase a blood pressure to improve a supply of nutrients and oxygen it is *indifferent* to the fact that a heart may not be able to sustain this high blood pressure. Another example could be this. If an intestine is trying to regulate its homeostasis of bacteria it can pass the signals to the brain via different biochemical pathways to reduce appetite so in this case an intestine is *indifferent* to that its actions can eliminate some source of energy and needed nutrients to the brain and other systems and worsen their functioning.

This principle allows to explain that supporting a homeostasis in one physiological system of the body can and often does harm functioning of other systems and this way it can cause a non-infectious diseases. If we find a physiological

system which is regulating its homeostasis with extreme steps and help it to fix the problems we can address a non-infectious disease of another physiological system which can be caused by this indifference.

## Results and Discussion

We introduced a method how to find a disease cause using a multiple researches which give estimates of risks (RR/OR) for different factors in regards to the *same* specific non-infectious disease. Method starts with finding a **number of disease causes using a formula** (1.0) when we know a rate of this non-infectious disease in a population (often we use an incidence rate for this to see annual probability of the disease). According to the method, once the risk factors were found using **disease causation criteria provided above** we need to find out physiological parameters which are changed when these factors are present.

Once we list all disease causing factors vertically and physiological parameters *related* to them horizontally (or vice versa) we mark the relationship between related factors and physiological parameters changed with a letter “R”. Then we find so called intersections between 2 factors - the places where 2 different factors are having same physiological parameters marked as related with letter “R”. The parameter are same by name (not by value).

We mark the column in gray or other color to indicate an intersection. We need to find all intersections like this. The physiological parameters where intersection are happening we list as matrix **Pm**. This list or matrix **Pm** is a superset of the physiological parameters changes which are causing a disease if their changes go beyond 1-sigma interval (actually, slightly less than this interval) and are present at the same time.

We eliminate the redundant parameters from this matrix **Pm** by using *few rules of elimination* and following a recommended algorithm if the count of parameters found exceeds the count determined by formula (1.0) calculating the number of disease causes or if a strict rule that a factor should cause only 1 change (or as determined by Disease Causation criteria) is broken.

One of these elimination rules is to check if the factors where intersection happens are related to each other (for example the same type of disease, or similar experiment like BMI and weight impact on disease, etc) and ignore the intersection for this factors as they are likely due to similarity. Another rule is to check if the physiological parameters for the same factors are related and this causing a change in 2 parameters for the same factor as one valid parameter is changing a dependent one. The dependent parameter need be eliminated from matrix **Pm**.

Remember that most factors in practice are impacting only 1 physiological parameter change which is causing a

disease. If we see a factor with 1 calculated impact is found to be causing 2 or more changes in physiological parameters in our table then the redundant physiological parameters need to be eliminated using the elimination rules.

Once the redundant parameters were eliminated we can check the final list of physiological parameters or matrix **Pm** using our *criteria for disease causing factors* (provided above for your reference) if there is an appropriate research regarding risks of these physiological parameters for the disease and confirm that we did not make any errors in steps of the described method. Also, we can use a related research regarding these found physiological parameters to see significance of their impact on the disease and this way to confirm the significance of the found physiological parameters found via the described method.

As example of method’s application we analyzed Depression disease. We found that according to the method **Depression is caused by 2 physiological parameters changes** beyond 1-sigma interval if they present at the same time. The disease must happen if the disease causing condition is satisfied for some time. We notice that Depression should *not* happen if only 1 of these parameters is beyond 1-sigma interval. **The physiological parameters which must cause a Depression in an individual are Reduction of GABAergic inhibition beyond~1 sigma interval and decrease of BDNF beyond ~1 sigma.** By controlling these 2 parameters we should be able to prevent Depression or possibly cure if the disease is in early stage. How long it takes for Depression to develop can be estimated at this point experimentally, the research on insomnia by Peter L. Franzen, PhD\*, et al indicates it happens within 1 year.

In this research we also showed how a knowledge of disease causes (as physiological parameters changed beyond 1 sigma) allows to build a hypothesis of disease pathology systematically as we know a disease structure - its few physiological causes and just need to connect them using existing vast knowledge on physiology, immunology, etc. We derived the hypothesis of Depression which we state here: GABAergic inhibition is getting reduced beyond -1 sigma by actions of disease causing factors (epilepsy, estrogen changes, etc) and this should increase excitation of neurons which in sequence should increase a demand for BDNF. At the same time due to a low level of BDNF (beyond 1-sigma) already being present because of other disease causation factors (Insomnia, Working over 11 hours, etc) the increased demand to BDNF cannot be satisfied and actually BDNF level gets reduced even more and this disrupts the normal work of the brain and causes a Depression.

It is well known that BDNF is required for long-term potentiation (LTP) and synaptic plasticity, etc.

**The triggering of the disease if causative physiological changes are present long enough is not optional it is a**

**must.** The analogy of this process could be an electrical fuse in a house. We know that if we plug in an electrical iron, air conditioner, etc. and exceed the threshold for the fuse the fuse will burn out. The fuse must burn out, it is not optional. The only random facts here are the electrical appliances which will cause it, these electrical appliances are an analogy of factors causing diseases here. We can have different external factors causing a disease but as long as they impact all required physiological parameters a “fuse” of the body will burn out and the disease will start.

We need to notice here that multiple medical researches **often find risk factors for a specific disease which have values close to RR/OR = 4.6 or RR/OR = 20.6.** These are not just a statistically significant risks but these numbers indicate that 1) the **factors are really disease causing factors** 2) the factors are **causing a change in 1 physiological parameter beyond 1-sigma** (for RR/OR = 4.6 +/-50%) or **2 physiological parameters beyond 1- sigma** (for RR/OR = 20.6 +/-50%). In some rare cases the risk factors can have **RR/OR=93 +/-50%** and this means the disease causing factors are impacting **3 physiological parameters beyond 1-sigma.** The disease causation criteria provided in article “*A Connection between Factors Causing Diseases and Diseases Frequencies: Its Application in Finding Disease Causes*” (Alan Olan, Journal of Clinical Trials, Vol.13, Issue 4) allows to determine more precisely if the risk factors are causing a disease or not and allows to expand validity for these risks beyond the 2 examples of RR/OR above. The referred article is using a risk factor **K** but it is easily converted for RR/OR as **RR = K + 1** (same for OR) and the RR/OR can be used as the theoretical values for RR/OR and a range of +/-50%. A short reference of these criteria was provided in this article as well.

In this article we have shown that **disease causation factors can be separated into few groups** according to a physiological parameter they change beyond 1-sigma. **Any combination of factors** taken as one from each of these groups **must cause a non-infectious disease.** Factors belonging to one group are affecting the reduction (only under specific conditions) of GABAergic inhibition such as Epilepsy, being a women in late transition to menopause, etc. while factors from another group reduce (under specific conditions) the BDNF levels such as Insomnia, Social phobias and others. It means that under certain conditions specified in the article the presence of Epilepsy and Insomnia would cause a Depression after some time, but if there is Epilepsy and there is Social phobias (under specific conditions) only the individual will get sick with Depression too as all these combinations affect the real underlying causes of Depression - 2 physiological changes which trigger it. **This explains why the same risk factor can affect many people but only some get sick with Depression as the 2nd required factor is missing** or the risk factor actually is not impacting the underlying physiological

cause of the disease enough, etc. Epilepsy as we have found **connected to a reduction of GABAergic inhibition** is considered a significant risk factor for Depression but it does not cause Depression in the existing research all the time because the research does not consider the Insomnia in the same research, for example or other factor which reduces BDNF levels in the individuals. This creates an impression of randomness of Depression causes, in one case of Epilepsy patients develop Depression and in other cases they do not.

Now, we need to notice that when we say the factors are causing a physiological change which trigger the non-infectious disease and in our case a Depression this causation is in some cases is **logical not a physical causation.** Logical causation means that the presence of this factor means the presence of physiological change required (important, not a correlation but as a fact) but it is not always the same as physical causation where the factor directly or indirectly causing biochemical change. If we state that if there is **A** condition means there is **B** condition it does not always mean that **A** causes **B.** What important for us is that **A** and **B** conditions **coexist.** For example, Epilepsy is very complex condition which we usually define by symptoms but underneath it has multiple physiological changes and while finding that reduction of GABAergic inhibition is part of physiological changes for Epilepsy the research on Depression cannot say that Epilepsy physically causes GABAergic inhibition. Method’s findings say: **if there is Epilepsy there is GABAergic inhibition.** To determine if GABAergic inhibition is one of the causes for Epilepsy the same method would need to be applied for Epilepsy.

We also introduced a **principle of indifference** which we find as a useful tool in understanding of disease causes as well. The principle of indifference states: **when regulating its own homeostasis separate physiological systems of the body are indifferent to the side effects of this regulation.**

The principle allows to understand that one physiological system can potentially change a physiological parameter beyond 1-sigma which is impacting another physiological system and can be part of the required multiple causes of non-infectious disease.

As we discussed a non-infectious disease always has a minimum 2 physiological causes which triggers it and then it develops and modifies other physiological parameters creating an enormous complexity which the modern research is observing but the triggering point of the disease is just 2 to 6 physiological changes... The analogy of this could be a button on the TV set - once you press it all the complexity of modern TV technology takes place ! If we don’t need the TV to be ON we usually do not press the button, we are not looking to find how complex the internal working of it unnecessary... The introduced method is basically allowing to find this “button” for the non-infectious disease.

## Conclusions

In this article we introduced a mathematically based method of finding non-infectious disease causes. This method is based on a published model [1] which states that non-infectious disease is always triggered by minimum 2 and often up to 6 physiological changes approximately beyond 1-sigma. The number is the same for the same disease. The more the incidence rate of non-infectious disease the less number of physiological changes required to trigger it. Many combinations of external factors can cause the same non-infectious disease but all them causing it by affecting a unique set of 2 or more physiological changes. This creates effect of randomness of non-infectious disease causation but underneath it is the same set of 2 or more physiological changes depending on the disease. This also explains why the risk factor affecting some individuals is not causing disease in them but is causing it in others affected individuals - not all physiological causes present to cause the disease if it is not triggered or the risk is not a real cause of the disease.

The research has shown that depression disease is triggered by a coexistence of **reduction of GABAergic inhibition beyond ~1 sigma** and **a decrease of BDNF beyond ~1 sigma** if they present long enough. Keeping them in correct range is expected to prevent Depression and possibly cure it in early stages. Multiple combinations of factors were found which can create these 2 unique changes and this produce an effect of random causation for Depression.

Overall a method presented here allows to increase efficiency of medical research in finding non-infectious disease causes, allows to develop drugs faster by combining known drugs to address multiple physiological changes and also stresses that non-infectious disease are having a multi-factorial causes. These causes are multiple physiological parameters changes beyond approximately 1-sigma coexisting at the same time and happening due to actions of external disease causing factors. This method helps to move from a search of correlations to a deterministic approach of finding non-infectious disease causes.

## Funding Statement

This work was carried out by the author as independent research.

## Ethical Compliance

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

## Conflict of Interest declaration

The authors declare that they have NO affiliations with or

involvement in any organization or entity with any financial interest in the subject matter or materials discussed in this manuscript. Author Contributions: Alan Olan is the sole contributor to this study.

## References

1. Alan Olan. A Connection between Factors Causing Diseases and Diseases Frequencies: Its Application in Finding Disease Causes. *Journal of Clinical Trials* 13.
2. Olan A. A Unique Set of Physiological Changes which Causes Breast Cancer: A Discovery with a Novel Method of Finding Non-Infectious Disease Causes which Combines Multiple Research. *BJCR* DOI: 10.31488/bjcr.205
3. Franco P G, Silvestroff L, Soto E F, et al. Thyroid hormones promote differentiation of oligodendrocyte progenitor cells and improve remyelination after cuprizone-induced demyelination. *Exp Neurol* 212 (2008): 458-467.
4. Marianna Virtanen, Stephen A Stansfeld, et al. Overtime Work as a Predictor of Major Depressive Episode: A 5-Year Follow-Up of the Whitehall II Study.
5. *Plos One* (2012).
6. Peter L Franzen. Sleep disturbances and depression: risk relationships for subsequent depression and therapeutic implications. *Dialogues Clin Neurosci* 10 (2008): 473-481.
7. Ned H Kalin. The Critical Relationship Between Anxiety and Depression. *The American Journal of Psychiatry* (2020).
8. Vivian N Onaemo, Timothy O Fawehinm, Carl D'Arcy. Comorbid Cannabis Use Disorder with Major Depression and Generalized Anxiety Disorder: A Systematic Review with Meta-analysis of Nationally Representative Epidemiological Surveys. *Journal of Affective Disorders* 281 (2021): 467-475.
9. Chu Chu. Association between epilepsy and risk of depression: A meta-analysis. *Psychiatry Research* 312 (2022): 114531.
10. Seasonal affective disorder (seasonal depression) – symptoms, risk factors, causes and genetics. *Medicover Genetics*, *Medicover Genetics Editorial Team* (2022).
11. Patama Gomutbutra, Nalinee Yingchankul, Siriporn Chattipakorn. Low Psychological Flexibility and high Brain-Derived Neurotrophic Factor (BDNF) associate with Fatigue symptoms. *medRxiv preprint* (2022).
12. Jayanta Mukherjee, Ross A Cardarelli, Yasmine Cantaut-Belarif, et al. Estradiol modulates the efficacy of synaptic inhibition by decreasing the dwell time of GABAA receptors at inhibitory synapses. *PNAS* 114.

13. Edward J. Wagner, Oline K. Rønnekleiv, Martha A. Bosch, and Martin J. Kelly, J. Estrogen Biphaseically Modifies Hypothalamic GABAergic Function Concomitantly with Negative and Positive Control of Luteinizing Hormone Release. *Neurosci* 21 (2001): 2085-2093.
14. D M Treiman. GABAergic mechanisms in epilepsy. *Epilepsia* 42 (2001): 8-12.
15. Gerard Sanacora, Graeme F Mason, et al. Increased Occipital Cortex GABA Concentrations in Depressed Patients After Therapy with Selective Serotonin Reuptake Inhibitors. *The American Journal of Psychiatry* (2002).
16. Pehrson A, Sanchez C. Altered  $\gamma$ -aminobutyric acid neurotransmission in major depressive disorder: a critical review of the supporting evidence and the influence of serotonergic antidepressants. *Drug Design, Development and Therapy* 9 (2015).
17. Yogesh Dwivedi. Brain-derived neurotrophic factor: role in depression and suicide. *Neuropsychiatr Dis Treat* 5 (2009): 433-449.
18. Tarapati Rana, Tapan Behl, et al. Unfolding the Role of BDNF as a Biomarker for Treatment of Depression. *Journal of Molecular Neuroscience* 71 (2020): 2008-2021.
19. Vermehren-Schmaedick A, Khanjian RA, Balkowiec A. Cellular mechanisms of activity-dependent BDNF expression in primary sensory neurons. *Neuroscience* 310 (2015): 665-673.
20. Vermehren-Schmaedick A, Khanjian RA, Balkowiec A. Cellular mechanisms of activity-dependent BDNF expression in primary sensory neurons. *Neuroscience* 310 (2015): 665-673.



This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC-BY\) license 4.0](https://creativecommons.org/licenses/by/4.0/)

## Analogy to clarify a method

A **simplified analogy** which explains why we looking for a match between physiological parameters could be this. Imagine a large town. We observe it from the top. A person arrives to downtown for some *personal business* on a regular day. How likely this person meet a *friend* or colleague in a totally random place of the downtown?..Very unlikely. We observe this person is going to downtown for few days to different places and he never meets a friend or a colleague. Now, we observe from the top that some rare folks meet someone in downtown often and sometime in the same location. We know **there is a pattern explaining these meetings**, they are not random in most cases. There might be someone they have agreed to meet with before (a colleague they travel together with, a friend, etc). If we find these folks meeting we know we found very likely some pattern.

In this analogy, we can treat a downtown as a human body, a person arriving to downtown as a physiological parameter change caused by some disease causing factor. People which meet each other in the downtown are an analogy of physiological parameters which “meet up” as they cause a disease and not just a random meeting. If we find those folks who meet up we know there is some cause there. In the method presented, these “meetings” between physiological parameters are represented by “intersections”.

## Mathematical foundation of the method

Here we will go into a mathematical foundation of the method in details. Let's look at the simple case of a disease where it was prior determined via experiments that multiple factors are causing changes in **only 2** physiological parameters of human body (parameters further).

Let the factors be **F1, F2, ... Fn** and the parameters be **C1, C2**. Now, let's look at the case where *F1, F2, ..., Fn factors separately causing only 1 change* in physiological parameters either C1 or C2 beyond 1-sigma (this actually often takes place in practice). That means that only C1 or C2 changes by some factor Fj (j ∈ {1,2,..,n}). Let's **P1, ..., Pn**, where n > 2 be *the sets of all physiological parameters which are related to factors F1, F2, .. Fn* accordingly. For example, P1: {r12, r15, r29, 43}, P2: {r15, r28, r34, r89, r34, r12, r98}, etc.

Let's look at standalone factor **Fj**. As we know, a factor **Fj** (where j is some integer from 1 to n) impacts the specific physiological parameters either **C1 or C2** then we know that *this params C1 or C2 should be part of its set of Pj as it contains ALL the related to Fj parameters (a complete set)*. Let's take a factor F1 such that its set P1 contains **C1**, and choose some F2 such that its set P2 contains **C2** (it is possible as we know factors impact either C1 or C2), then if we choose any other factor as F3 then its set of P3 must contain either C1 or C2 (as F3 also impacts these physiological parameters - either C1 or C2 and P3 is a

complete set). If P3 contains C1 then it intersect with P1. If P3 contains a C2 it intersects with P2. So **P3 must intersect with either P1 or P2 (either in C1 or C2)**. In similar way we can apply this to P4, P5, ... Pn. So this brings us to conclusion that **a set of physiological parameters Pn, where n > 2 must intersect with either P1 or P2 either in C1 or C2**. This means parameters Pn intersect with each other either in C1 or C2. We can see representation of this set's behavior on the Figure 2. **All sets Pj, Pk, ... matching to Fj, Fk, ... on the Figure 1 are crossing and only in C1 or C2 but not both.**

We don't know the values of C1 and C2 but if we can find where parameters Pn intersects with each other we can determine a **subset of physiological parameters Pm**: {Ry, Rx, Rz, ... , Rt} which contains values of C1 and C2. This subset of Pm will be much smaller than set of all possible params included in P1, P2, .. Pn (as it is a subset and similarities in params of P1, .. Pn are not very probable and that is addressed below) but may contain more than 2 parameters and **only 2 parameters of this subset Pm** can be real physiological parameters causing a disease as they are C1 and C2.

In order to eliminate the incorrect parameters from subset **Pm** we need to notice:

1) that the params C1 or C2 should **be such so all P1, P2, ..., Pn intersect in them** and if some parameters of **Pm**: {Ry, Rx, Rz, .. Rt} don't fit this rule *their need to be eliminated*. Practically it means this. We take random (or using a common sense) a combination of some 2 parameters **Rk** and **Rm** from a set of **Pm** and

*check if the P1, ... P2 all intersect in them if not then the Rk and Rm combination is not a valid set of C1, C2 and we may need to check another set of 2 parameters Rk and Rg*

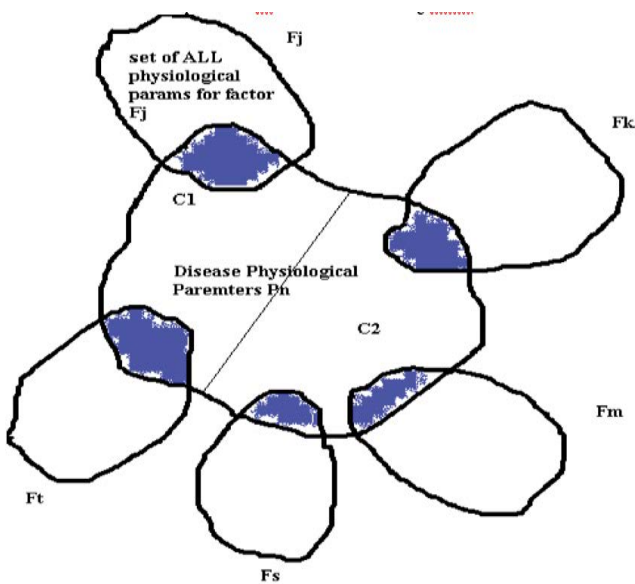
2) if some parameter of set **Pm**: {Ry, Rx, Rz, .. Rt} **is known as not changed beyond 1-sigma it should be eliminated** as disease is caused by change in param beyond 1-sigma (as per our model).

3) if some parameter of **Pm**: {Ry, Rx, Rz, .. Rt} **is causing some set Pn intersect 2 times with some other set Pk then it should be eliminated** as factors F1, F2, .. Fn can *only* impact 1 physiological parameter in this case and cannot impact / intersect 2 or more due to this.

The method above was described for a case of factors F1, ..., Fn impacting **only 2** parameters but it can be extended to 3 and more parameters.

## How likely are random matches between physiological parameters?

As we discussed above the set of physiological parameters **Pm**: {Ry, Rx, Rz, .. , Rt} where we observe intersections may contain more parameters than needed (more than 2 in our case and due to other reasons).



**Figure 2:** (Blue areas are area where Pj, Pk, etc for factors Fj, Fk, etc. intersect with a set of physiological parameters C1, C2 which are part of set Pn and which are causing disease)

We need to be concerned with a question such as if we find one intersection of sets P1 and P2 in a physiological parameter belonging to 2 different external factors how likely it can be a random intersection? To answer this question let's formulate the problem mathematically.

Let's have a set A of integers from  $k = 1$  to very large N. Let's randomly select  $n$  numbers in set of  $P1 = \{Ak, Ag, \dots, At\}$  and then randomly select  $n$  numbers into set  $P2 = \{Af, As, \dots, Al\}$  from our original set A ( $k = 1$  to N) such that each element repeats only once in set P1 and only once in P2 (it is a unique element to sets P1, P2). For example, if we chose a number 3 as part of the set P1 then it only exist one time in the set P1. What is a probability that we find element Ai in set P1 and P2?

To answer this question let's do next steps. Let's limit set A by some top element enumerated by  $t$  (so set is not infinite).

1. We can take  $n$  elements from  $t$  elements of set A with number ways  $tCn$
2. Number of ways to take  $n$  elements with an element Ai equals the number of ways to select  $n-1$  elements (we exclude Ai) from  $t-1$  (set of A elements) and is  $(t-1)Cn-1$
3. Then probability to take  $n$  elements which include element Ai in set P1 (or P2) is  $P(Ai \in Psel) = (t-1)Cn-1 / tCn$ , where Psel is P1 or P2 sets

4. The probability that element Ai will be in P1 and P2 is  $P(Ai \in P1 \text{ and } Ai \in P2) = P(Ai \in P1) * P(Ai \in P2)$  as events independent.
5. So probability  $P(Ai \in P1 \text{ and } Ai \in P2) = P(Ai \in P1) * P(Ai \in P2) = ((t-1)Cn-1 / tCn)^2$
6. Or finally, the probability that element Ai will be in P1 and P2 is  $P(Ai \in P1 \text{ and } Ai \in P2) = ((t-1)Cn-1 / tCn)^2$

Using a formula above let's calculate a probability of match in element Ai if we take randomly elements from a sequence of numbers from 1 to 1000 ( $t = 1000$ , assuming so many physiological parameters exist) and take only  $n = 10$  element into sets P1 and P2 accordingly.  $P(Ai \in P1) = {}_{999}C_9 / {}_{1000}C_{10} = (2.63 * 10^{21}) / (2.63 * 10^{23}) = 1 / 10^2 = 0.01$  the same is fare for  $P(Ai \in P2) = 0.01$  and so the probability of getting element Ai in sets P1 and P2 is  $((t-1)Cn-1 / tCn)^2 = 0.01^2 = 0.0001$

This is a probability of random match. **The probability of non-random match is  $1 - P(Ai \in P1 \text{ and } Ai \in P2) = 1 - 0.0001 = 0.9999 \approx 1$**  so very close to 1. It means *if we see a match between set P1 and set P2 in some element Ai it extremely likely it is not random.* This is an important conclusion. *The only matches we find practically are not random but are caused by some reason* and in our case it is due to same physiological parameter impacted by 2 different factors. We need to notice that number of parameters actually much more as most agree that there are around 20,000 different proteins in our body and each is a potential physiological parameter. So the probability of the match selecting 10 of 20,000 will be much smaller!

In practice there are about over  $t=150$  physiological parameters known to medicine and for a single factor we usually find about  $n=30$  related parameters. Doing calculations for this case we get  $P(Ai \in P1) = {}_{149}C_{29} / {}_{150}C_{30} = (6.43 * 10^{30}) / (3.2 * 10^{31}) = 0.2$  and so the probability of getting element Ai in sets P1 and P2 is  $((t-1)Cn-1 / tCn)^2 = (0.2)^2 = 0.04$

We see the probability of random match is higher in practice (4%) and so in practice we can see more random matches. The probability of at least one random match will also increases as we find intersections for dozens of different causation factors.

This random matches still can be eliminated with methods described in this article by applying other restrictive conditions, including using our criteria for disease causes.