

Failure of Medicine Is Attributed to the Flawed Research Model: A Comprehensive Study

Jianqing Wu^{1*}, and Ping Zha²

¹Healthier World (Independent researcher for cause), P. O. Box 689, Beltsville, MD 20704

²Independent Researcher

*Corresponding author: Jianqing Wu, Healthier World (Independent researcher for cause), P.O.Box 689, Beltsville, MD 20704, USA, E-mail: tempaddr2@atozpatent.com

Citation: Wu J, Zha P. Failure of Medicine Is Attributed to the Flawed Research Model: A Comprehensive Study. J Math Stat Comput. 2022; 3(1): 001.

Copyright: © 2022 Jianqing Wu. This is an Open Access article published and distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

Received Date: 23rd March 2022

Accepted Date: 4th April 2022

Published Date: 11th April 2022

Abstract

Medicine has failed to find cures for chronic diseases, but nearly all diseases can spontaneously resolve in reality. The riddle between lack of cures for chronic diseases and disease spontaneous resolution is known for centuries. We suspect that lack of cures and inability to understand disease spontaneous resolution is due to the inherent limitations in the medical research model. We show that chronic diseases are reflection of disturbed metabolic networks, and accuracy and reliability of the medical research model are insufficient for characterizing chronic diseases. We examined implied presumptions in the research model and assumptions used in clinical trials and statistical models, and conducted numerical simulations by using hypothetical model data. We found that clinical trials and statistical analysis introduce too many errors and too bigger inaccuracies that tend to conceal weak and slow-delivering effects of treatments and thus could reject lifestyle factors that would have power to cure chronic diseases. We further found that a large number of prevalent uncontrolled co-causal or interfering factors can greatly enlarge the means and the variances or experimental errors, and the use of high rejection criteria (e.g., small p values) further raises the chances of failing to find treatment effects. We concluded that the research model is biased towards rejecting weak lifestyle factors. We believe that misusing statistical models to characterize human metabolic networks is most probably responsible for failure to identify weak treatment effects for chronic diseases and failure to detect harmful effects of toxic substances on humans. We proposed alternative experimental models involving the use of single-person or mini optimization trials for studying weak and slow-delivering treatments.

Keywords: Population medicine; Randomized controlled trial; Statistical model; Health optimization; Incurable chronic diseases; Failure of medicine

Introduction

Medicine started emerging as modern medicine after the Industrial Revolution in the 18th century. Over the last 150 years, medicine has accomplished some astonishing achievements. Most achievements are in the areas of treating acute diseases such as bodily injures, infections, poisoning, pains, trauma, etc. In each of those cases, drugs are not used to restore impaired or lost balance in the body. Despite the success in treating acute diseases, medicine has failed to find cures for chronic diseases. Main evidence for its failure includes: (1) Nearly half of all adult Americans suffer from at least one chronic disease. This is equivalent to approximately 45% or 133 million of the population; (2) nearly all chronic diseases are officially listed as incurable diseases in medical references. A long list of chronic diseases is still without cure. In addition, many types of cancer are considered as incurable and terminal; (3) in 2009, 7 out of 10 deaths in the U.S. can be attributed to chronic diseases. Heart disease, cancer and stroke account for more than half of all deaths each year. We estimated that the total number of premature annual deaths attributed to

chronic diseases is about 30 million in the world based on total death data [1,2].

The failure of finding cures is best reflected in cancer. A systematic review concluded the complete response of rates of chemotherapy for later stage of cancer have remained static and locked at about 7.4% [3]. The response rate of thyroid cancer treatment was 22.1%-27.1%, with complete response rates being 2.5%-2.8% [4]. A recent study examined the most promising cancer treatment methods, and concluded: "The claimed 'targeted' therapies that may or may not extend remission of cancer for a few months should not be accepted any longer as 'cure' by oncologists, scientist or patient" [5]. The prevalent chronic diseases in the U.S. has become a huge burden on the U.S. economy. In a study done by the Milken Institute, the annual economic impact on the U.S. economy of the most common chronic diseases is calculated to be more than \$1 trillion, which could balloon to \$5.7 trillion by 2050 [6].

We see that medicine advances on two distinctive tracks. It is capable of achieving good achievements in the treatment of acute diseases. However, it fails to find cures for chronic diseases. The clear separate line between the two kinds of diseases seems to indicate that the performance difference is related to the medical research and practicing models. In this article, we explore if the population-based model such as clinical trials has some inherent limitations that prevent medical researchers from finding cures.

Methods

Our purpose is to examine the performance of clinical trials and statistical methods in the context of characterizing chronic diseases.

Basic Model Assumptions

We suspect that human individuals introduce very large variances to any measured health properties so that clinical trials are unfit for studying chronic diseases. To prove it, we will use following model assumptions:

Treatment: $s_1 \sim N(\mu_1, \sigma_1^2)$ that affects a trial outcome

True error: $\varepsilon \sim N(0, \sigma_\varepsilon^2)$

Other causal or interfering factors: s_2, s_3, \dots, s_k .

$s_2 \sim N(\mu_2, \sigma_2^2)$

$s_3 \sim N(\mu_3, \sigma_3^2)$

....

$s_k \sim N(\mu_k, \sigma_k^2)$

s_2, s_3, \dots, s_k represent anything that could influence measured health properties relevant to the trial outcomes. They may be substantial cause

factors, causal factors, indirect causal factors, covariates (independent factors or confounding factors), etc. The only qualification criterion is that they can affect the intended treatment so that they must be considered in practice.

In a clinical trial, treatment s_1 must be much larger than the total combined effects of ε and all s_2, s_3, \dots, s_k so that s_2, s_3, \dots, s_k can be ignored or treated as part of the error ε for convenience.

The model assumption in our study is that s_1 is close to ε and also close to one or all of s_2, s_3, \dots, s_k . For example, in a clinical trial to study a cancer treatment, the trial outcome may be judged by observing patients' average survival times. A large number of factors shown in Table 3 are known to affect patients' survival times. Those factors may be traced to genotypes, lifestyle, diets, physical activities and exercise, toxic compound levels in the body, viral infections, gut microbiota, other health problems, etc. In this study, it is further assumed that they affect patients' survival times randomly. Each of such factors may appear in some patients but not in other patients.

Our question is whether a randomized controlled trial can accurately detect the effects of s_1 and what could be done to increase the chance of actually detecting the treatment effect which is similar to or weaker than other causal and interference factors. To answer this question, we used a randomized controlled trial model and a mini optimization trial model to evaluate their respective performance. The basic design of the two types of trials are shown in following Table 1 and Table 2.

Two Hypothetical Experimental Models

Model A. Randomized Controlled Trial is shown in the below table.

Table 1. Randomized Controlled Clinical Trial With 3 Randomized Interfering Factors

Treatment Arm (TX: s_1)	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
Interference Factors	s_2		s_3		s_4	s_2			s_3		s_2
Control Arm (CA)	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	
Interference Factors	s_3		s_2	s_4			s_3	s_2		s_2	

The human subjects are allocated to the two arms randomly. The table shows only one potential way of allocation solely for illustration purpose. The effects of s_1 on health properties is closer or even smaller than any of those 3 interfering factors s_2, s_3, s_4 . For example, s_2 may be exercise, s_3 is a dietary adjustment, s_4 is stress management, etc. They affect patients survival times like chemotherapy or other treatment (s_1).

Model B. An Optimization Trial is used where all s_1, s_2, s_3, s_4 causal factors and interfering factors are used as one single treatment package for the purpose of raising total treatment effects.

Table 2. Optimization Trial Design With All Four Factors Used As A Treatment (No Randomization)

Treatment Arm (TX: s_1, s_2, s_3, s_4)	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Control Arm (CA)	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10

In this optimization design, all other non-treatment causal and interfering factors (s_5, s_6, \dots, s_k) must be sufficiently small and thus can be bundled into the error term. We call this design as an optimization trial because as many important factors are used as the treatment to deliver the maximum treatment effects. Here, all important causal factors and interfering factors (s_1, s_2, s_3, s_4) that would be identified and used are bundled as one treatment package.

We then evaluate how the optimization trial increases the chance to determine true effects of the treatment package and how to increase the chance of finding cures for chronic diseases.

Our Analysis

Our focus is on how to determine true treatment effect when the treatment is influenced by one or more interfering factors. Our initial focus is the accuracy and reliability required to detect the true effect of the treatment.

1. We examined the machine-repairing model (such as car repairing model) to understand why a population-based method similar to clinical trials cannot be used in diagnosing and repairing machines. Attention is directed to accuracy requirement for repairing complex machines and restoring structural and functional balances in machines. We found that an implied requirement for conducting a population-based trial is that all trial subjects must be “nearly identical units.”

2. We explored accuracy and reliability required to accurately characterize chronic diseases and their changes. One key fact considered is that chronic diseases development speeds. Slow disease development speeds imply small changes in biochemical processes and organ structure in a given unit time. The slow disease development rates and small structural changes further imply that high accuracy and reliability are required to accurately characterize chronic diseases, as compared those for studying acute diseases.

3. We examined personal differences in light of genotypes, phenotypes, and emotional states, or treatment-relevant factors such as race, personal genotype, age, sex, diet, lifestyle, medication use condition, etc. In addition, we further examined massive differences in health properties found in reference ranges of laboratory tests for human beings to determine whether humans can be treated as “nearly identical units”. This determination was made in light of high accuracy and reliability required for accurately characterizing chronic diseases and their development. Many aspects of the massive personal differences work like causal and interfering factors on disease outcomes.

4. We determined whether variances in measured health properties attributed to personal differences are too high to satisfy the requirement of “nearly identical units.”

5. We collected a number of exemplar causal and interfering factors from medical literature to show how they raise the variances of measured health properties. We paid attention to variances that arise from race, personal genotype, age, sex, diet, lifestyle, physical activity, exercise, medication history, etc. because they have been found to be causal factors, risk factors, or associated factors of chronic diseases including cancer.

6. We examined the logic used in statistical analysis methods such as t-test, z-test f-test, Chi test, frequency test, etc. to determine whether they could remedy the flaw that clinical trials are unable to offer the minimum accuracy and reliability that are required to accurately characterize chronic diseases. Assumptions used in statistical analysis were examined in light of our models.

7. We determined that if the sources of variances from other causal and interfering factors are merged into the experimental error term as an apparent error as in case in Table 1, how the raised variances affect hypothesis test outcomes, and produce biased results by failing to detect treatment effects on chronic diseases. We then showed a pattern of biases by conducting

hypothetical tests using hypothetical data as our model data comprising a weak treatment and at least one interfering factor with similar effects on the measured health properties.

8. We examined whether health properties from different patients can be added up and averaged as in a statistical analysis by using a multiple causes and treatments model.

9. We finally conducted a comparative analysis for the two models: a randomized controlled trial and an optimization trial. We showed why randomized control trial is invalid for studying chronic diseases, and showed that optimization trials could offer far better chances for finding treatment effects for chronic diseases.

Results

The clinical trial developing history reveals that most of early clinical trials were used to investigate malnutrition, infections, and wounds (except rheumatism). No effort has been made to understand inherent limitations of clinical trials in the history. We also note that the functional approach used in machine is inherently incompatible with the population approach (C, Sup.). The population approach cannot be used in diagnosing and repairing machines made of different blueprints. A population approach may be used to study properties of only “nearly identical units.” Differences, if any, must not cause any functional imbalance, structural misfits, fuel imbalance, flow imbalance, heat imbalance, etc. The population approach has not been used to address mechanical problems.

Whether a health problem can be studied by clinical trials depends on the purpose of the study. A threshold requirement is that the effect of treatment’s on health property is sufficiently larger than the experimental error. This requirement can be satisfied in cases studying strong treatment effects such as pain-killers, surgery, antibiotic drugs, sedative drugs, etc. In those cases, differences among persons will not significantly alter results.

High Accuracy and Reliability Required for Studying Chronic Diseases

“Chronic diseases are defined broadly as conditions that last 1 year or more and require ongoing medical attention or limit activities of daily living or both.” [2]. We show the level of balance required in a human body is much higher than the degree of matches between parts in a machine. Health problems can arise from small biochemical imbalances, which result in small changes in structure, shape, and capacity of body components (A, Sup.). As shown in those examples, the deviations in biochemical and cellular processes for causing chronic diseases are “infinitesimally small.” Net departures from ideal numbers are often in a tenth percent to a few percent of the ideal personal number. Most net conversion rates must be of right values, and small departures from ideal numbers in either way can be the cause of chronic diseases.

Clinical Trials Do Not Support Accurate Evaluation of Health Properties for Chronic Diseases Due to Massive Personal Differences

The population approach is extended to all areas of medicine, but one problem that has never been studied is personal variations. Two big sources are different genotypes and phenotypes [7,8]. The chance of match between two unrelated persons is like that of a DNA match (1 in 113 billion based on 9 loci; 1 in 400 trillion in 13 loci). In addition, personal differences further arise from different emotional states. The personal differences that are important to health may be expressed, in the alternative, as diet, lifestyle, emotional state, culture, environment, sex, medication history, etc. Personal differences are reflected in reference ranges of laboratory tests for human beings, which are established by empirical methods. The reference ranges, which reflect measured variances in any health properties in a population, depend on personal differences in genotypes, phenotypes, daily fluctuations, and measurement error. The number of combinations could be infinitely large. Each of the health properties of a person may fall a distinctive position of the correspondent populations reference range. No person would have all of his health properties match the population's means. Differences between two persons can be inferred from differences in body size/shape, organ size/shape, structural strengths, skin colors, physical capacities, emotional conditions, etc. Differences are also reflected in diagnostic data, image data, health conditions, disease histories, etc. Reference ranges of more than five hundred health properties are published [9]. The measured value of each health property for any person will fall a distinctive point of the range shown in Table S1 (D, Sup.). If each reference range is divided into N levels which could be resolved by detection resolution, the total number of variations of all health ranges would be the product of all possible numbers of all reference ranges. The recognized chemicals in the reference range table are not complete. All departures of a person's measurements of health properties from the population's means are necessary to correct genetic weakness or to maintain the phenotype, and thus are presumed to be important in maintaining health and prevention of diseases.

Personal differences must be considered in treating chronic diseases. First, when persons are sufficiently different, they cannot be treated as same or similar units in a clinical trial because their differences can interfere with the measured health properties in the trial. Second, the values of health properties cannot be used as parameters for predicting chronic diseases. Such health properties cannot be correlated to conversion rates of metabolites and final size of tissue structures. Conversion rates of metabolites normally depend on multiple parameters. Health properties may fluctuate in daily, weekly, monthly or yearly basis within the lowest and highest ranges. Chronic diseases may arise when health properties in a person depart from optimal values for sufficient duration. Cures to such diseases would require correcting such small departures. Finally, personal differences, which are reflected in the health properties shown in Table S1, affect both disease process and healing process. Personal numbers such as vitamins levels, heavy metals, HDL, LDL, cholesterol, platelet count, red blood cell count, white blood cell count, fatty acids levels, glucose levels, triglycerides, etc. can be altered by changing a large number of lifestyle factors.

From conducting a review of clinical trial development history, we found that none of old studies we could find discussed personal differences, interfering factors, and their effects on a weak treatment effect (B, Sup.). In a traditional clinical trial, the treatment effect is much stronger than the experimental error so that interfering factors will not alter trial outcome (Figure 1). Absolute reference in our figures is an imagined health property that could be measured when the treatment is not applied. Since a chronic disease is caused by lost balances, the absolute references define the disease state when those deviated balances such as excessive omega 6/3 fatty acid ratio, excessive heavy metal levels, physical inactivity, abnormal gut microbiata, lack of dietary fibers, abnormal emotional state, etc. are not corrected. An absolute reference exists in a patient, but could not be applied to a population. It may be used to a small number of "sufficiently similar patients" only if research focus is limited to a small number of interfering factors.

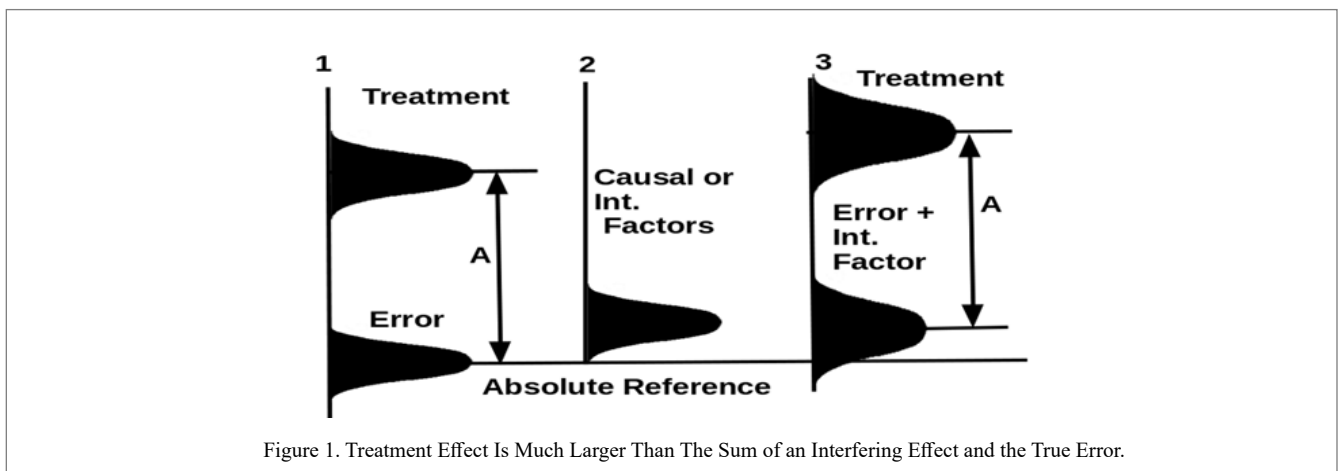


Figure 1. Treatment Effect Is Much Larger Than The Sum of an Interfering Effect and the True Error.

In studying a strong treatment effect (Figure 1), an assumption can be made that all persons can be treated as identical units because the strong treatment effect cannot be distorted by interfering effects in meaningful amounts. Any differences caused by personal differences are so small that they can be properly neglected. In this situation, randomization is sufficiently good. The justification of use of clinical trials is good approximation. After the error

and interfering factors are summed up, resulting in a new distribution under the line 3 (E, Sup.), the treatment effect is still much stronger than combined effects of the error and the interfering factor. Even if many interfering factors exist, their effects could still be neglected.

In studying a chronic disease (Figure 2), the treatment effect is weak relative to two interfering factors shown under the line 2. When the two interfering factors and the error are summed up, they generate an apparent error under the line 3. The mean of this apparent error are the sums of the means of the error and means of the two interfering factors. Without considering the interfering factors, the trial is to find the differences

between the treatment and the error under line 1. If the interfering factors are considered, the trial actually determines the treatment effect under the line 4 relative to the apparent error under line 3. The trial may be unable to find the treatment effect if the data comes out with the treatment's effect at a lower tail region and the error at the upper tail region.

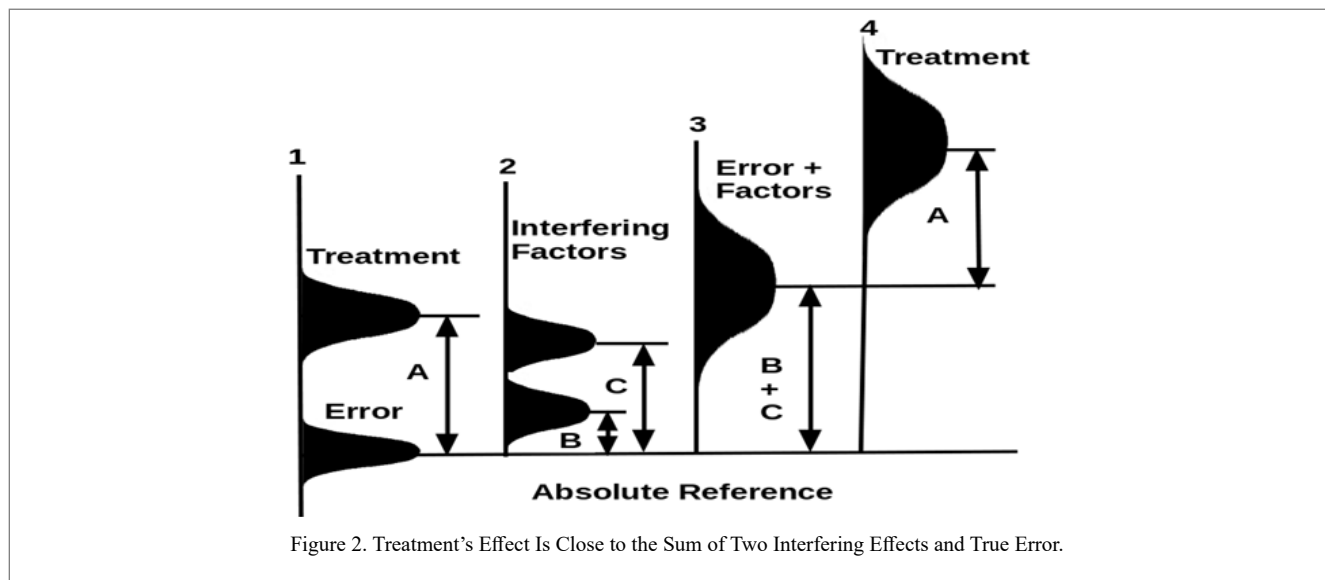


Figure 2. Treatment's Effect Is Close to the Sum of Two Interfering Effects and True Error.

In a worst situation (Figure 3), the effect of one or more interfering factors is larger than the effect of a treatment. In this case, the error under the line 1 and the interfering factor under the line 2 merges to become a large apparent error with large variances under line 3. The treatment and the apparent error have a large overlap region (if the profile under 3 is moved onto line 4 horizontally).

A trial may come out with the treatment effect falling at the lower tail region while the apparent error at the upper tail region, resulting in a finding that treatment is negative relative to the control. This result is clearly against the model assumption that the treatment has a weak effect indicated by letter A.

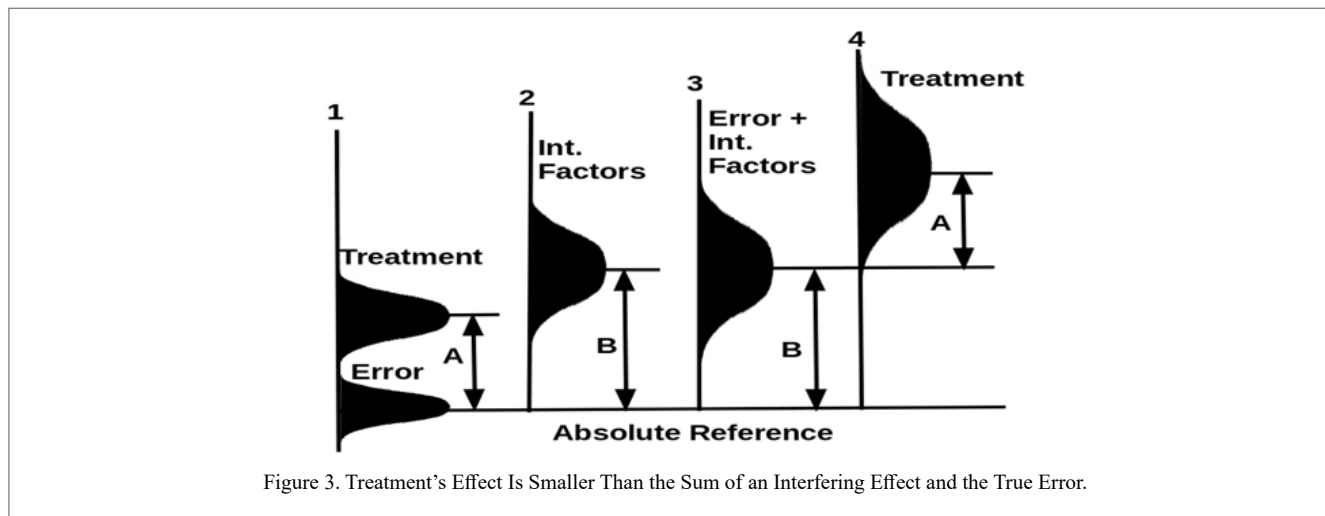


Figure 3. Treatment's Effect Is Smaller Than the Sum of an Interfering Effect and the True Error.

When the weak treatment overlaps the apparent experimental error as shown in Figures 2 and 3, the trial is meaningless. Nothing can correct this problem that arises from breaching the basic presumption that the treatment effect must be much larger than the experimental error.

the chance to determine the treatment effect. Optimization with both the original treatment and the interfering factor will reduce the variances of the apparent error and increase the difference (designated by $A+B$) between the mean of the whole treatment package and the control.

Figure 4 shows how an optimization trial by including the interfering factor (which appears in Figure 3) as part of the treatment will dramatically improve

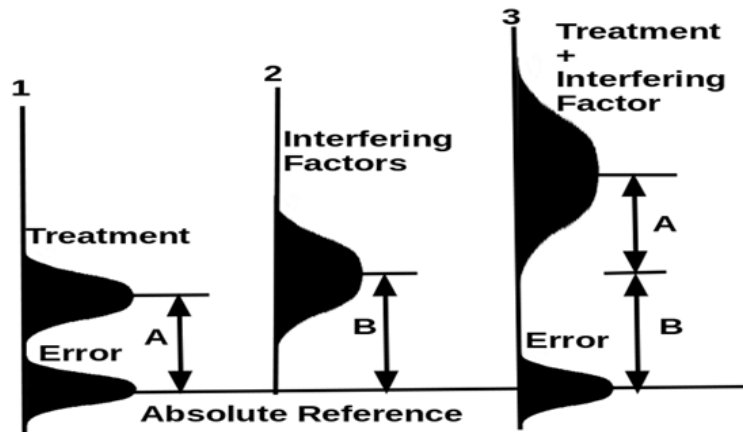


Figure 4. Treatment's Effect Is Increased by an Interfering Effect When the Interfering Factor is Used as Part of the Treatment in an Optimization Trial

In studying a chronic disease, all persons must be regarded as different. Cancer provides the best example in this regard. Each tumor is unique due to the genetic and epigenetic basis and exogenous exposures such as dietary and lifestyle factors [8]. If a treatment protocol developed from the population data can be used to cure the disease of a particular person, one would have to wrongly argue that the health properties of the person are unimportant to diseases, and phenotypes can be freely changed. Health properties are not quantities that can be summed up and averaged, and a treatment protocol based on population data cannot be applied to any specific person as cures for chronic diseases. This might be the reason why medicine could not find cures by using clinical trials.

If a statistical analysis of clinical trial data yield a “significant difference” over a large number of interfering factors, such a treatment must be very strong. It could be unlikely for such a strong treatment to correct many weak causal causes for chronic diseases. This might be a reason that treatment protocols from clinical trials are able to control symptoms quickly, but are unable to restore sophisticated balances in human bodies.

Multiple Co-Casual and Interfering Factors Are Prevalent in Human Beings, But Often Could Not Be Controlled in Clinical Trials

Massive differences among individual persons are anticipated to affect the accuracy and reliability of clinical trials required for studying and characterizing chronic diseases. In a large clinical trial, a measured health property such as survival time or hazard ratio depends on nature of the disease, the effect of the treatment, all uncontrolled interfering factors, and their interactions. Naturally, all those factors are added into the error term. The final conclusion of the trial depends on the treatment effect relative to the bloated error term. If many factors are not controlled, the presumption that the treatment effect is much larger than the experimental error fails and result is incorrect. We will show a list of uncontrolled factors that can be seen in clinical trials.

Table 3. Factors That Are Known to Influence Chronic Diseases; Most of Them Normally Are Not Controlled in Clinical Trials

Case No.	Classes	Impacts and Mechanisms	Implied Effects and Significant Degrees	References
A1	Genetic (mutations).	Cancer initiation, development and metastasis.	As well accepted somatic evolution theory (driving gene)	[Nowell, 1976; Loeb, 1991][10]
A2	Genetic (angiogenesis)	19 Endogenous angiogenic polypeptides (VEGF, APN, etc.).	Tumor mass is limited to 1–2 mm without neovascularization.	[Sagar et al. 2006] [11]
A3	Genetic (apoptosis)	Inflammation, apoptosis and autophagy.	54 genes related to those properties in breast cancer.	[Schuetz et al. 2019] [12]
A4	Genetic (race).	Among Asian people: Stroke is more prominent than CHD.	Different characteristics.	[Ueshima et al. 2008] [13]
B1	Age and aging on cancer incidence.	Cancer incidence rate is related to sixth power of age.	Huge impacts on prevalence and cancer death rates.	[Armitage & Doll, 1954] [14]
B2	Age on inflammation	Aging and hormonal changes.	Age has great impacts on inflammation.	[Prasad et al. 2012] [15]
C	Sex.	Males had higher age-adjusted death rates for 12 of the top 15 causes of death.	Sex's effects depend on diseases.	[Kalben et al. 2000] [16]

D1	Chronic stress.	Cancer initialization, growth and metastasis.	Affect immunity; neuroendocrine/ β -adrenergic signaling.	[Segerstrom and Miller, 2004; Sloan et al. 2010] [17][18]
D2	Chronic stress.	Increased levels of platelet-leukocyte aggregates, and vascular diseases.		[Brydon et al. 2006; Sundquist et al. 2005; Nemeroff et al. 1998] [19] [20][21]
E1	Diet (vitamins, fibers, minerals, etc.).	Nutrition affects tissue ecosystem and cancer proliferation.	A massive number of studies show diet's impacts on various aspects of health and cancer.	[Ogino et al. 2012, 2013] (too many to include).[7,8]
E2	Diet (natural compounds).	Targeting apoptosis pathways in cancer.	A massive number of natural compounds.	[Millimouno et al. 2014] [22]
E3	Diet and gut microbiota	An impaired microbiota dysbiosis linked with cancer.	Probiotics, potential corrective diet, Fecal Microbiota Transplantation	[Vivarelli et al. 2019] [23]
F1	Systemic inflammation	Affect the tissue ecosystem and inflammation.	Through age, body mass index, dietary saturated fat, and EPA and DHA omega-3 fatty.	[Navarro et al. 2016][24]
F2	Diabetes and Cancer.	Potential diabetes and cancer association.		[Giovannucci et al. 2010] [25]
F3	Viruses	DNA viruses and RNA viruses	Cause of about 15% of all cancers in the world.	[Liao, 2006] [26]
G	Prior chemo.	Cancer repopulation.	Mechanism is unknown.	[Salani et al. 2011] [27]
H	Surgery.	Cancer repopulation.	Systemic inflammation; promote M2 Tumor Associated Macrophages.	[Kral et al. 2018; Colleoni et al. 2016; Cheng et al. 2012; Demicheli et al. 2007][28][29][30][31]
I	Radiotherapy.	Tumor cells repopulation.	Mechanism is unknown.	[Salani et al. 2011] [27]
J1	Exercises.	Cancer initiation, growth and metastasis.	28%–44% reduced risk of cancer-specific mortality.	[Cormie et al. 2017] [32]
J2	Exercises.	Affect 35 Chronic diseases.	Good rehabilitative therapy.	[Booth et al. 2012] [33]
K	Temperature, vibration, etc.	Enzyme activity; cell division apparatus.	(Influence cancer by time-averaged effects.)	[Levine & Robins, 1970] [Yeung et al. 2003] [34] [35]
L	Lifestyles	Association of risk factors to myocardial infarction.	Changing lifestyle could prevent at least 90% of myocardial infarction.	[Yusuf et al. 2004] [36]

The above table shows only a few exemplar factors that normally influence chronic diseases including cancer. The exact working mechanisms are unimportant to our analysis. Those factors affect a treatment result for a chronic disease if the treatment is evaluated by measuring a health property such as survival time, hazard ratio, chemical analysis data, structure's size, biochemical process speeds, etc. They affect measured health properties by causal effects or by influencing one or more causal factors. Some factors may work like confounding factors.

Variances of each factor arise also from an error in measuring the factor and the mechanisms at which the factor affects the measured health property. For example, it is impossible to accurately measure intensity, amount, and duration of exercise. Even if exercise were used in a precise accuracy, actually delivered effects on the health property would depend on personal conditions. Surgery is considered as a factor of influencing cancer cell re-population by different mechanisms. Exercise is found to be an important adjunct therapy in the management of cancer based on a large number of studies [32]. Physical

inactivity is one important cause of most of 35 chronic diseases [33]. Chronic stress can dramatically speed up cancer metastasis [17,18]. A prior surgery can dramatically alter the body's ability to resist cancer return growth speed [28-31]. Age affects cancer incidence rate by a sixth power [14]. Age, body mass index, dietary saturated fat, and EPA and DHA omega-3 fatty can affect the body's inflammation potential [24]. Many uncontrolled factors may be magnitudes stronger than treatment's effects when their effects are evaluated in long terms.

In clinical trials, most of those factors are not controlled or cannot be accurately controlled. For example, surgery cannot be well controlled. If patients in a typical trial have been operated previously, the amount of tissue loss and surgical locations are dictated by medical needs. Ages may be classified by age groups but their effects cannot be well controlled due to personal differences. Besides, two persons at the same age may have very different biological ages. Most lifestyle factors cannot be measured accurately and thus are anticipated to have different effects. Since people

have different lifestyles, their prior lifestyles may have residual effects on health properties after their lifestyles are changed per required treatment.

If a clinical trial is designed to study a weak factor, tens to hundreds of other uncontrolled factors with similar levels of effects are “bundled” into the error term. All of those factors affect human subjects in both the treatment and the control; and due to randomization, they do not cause meaningful difference between the treatment’s mean and the control’s mean. Each interfering factor raises both the mean and variances of the apparent experimental error term (See Figures 2, and 3). We will show that statistical analysis not only fails to correct such a problem, but makes the problem worse by failing to recognize weak treatment effects.

Randomization and Statistical Analysis Cannot Correct Biases Caused By Personal Differences, but Increase the Chances of Accepting Null Hypothesis

Randomized control trial does not automatically deliver a precise estimate of the average treatment effect, and it yields an unbiased estimate for the sample selected for the trial [37]. Past studies discussed many problems, but did not discuss the inherent biases that arise when a treatment effect is weak while multiple interfering factors exist. Accordingly, no attempt has been made to understand the merit of using multiple factors optimization method.

One common type of statistical analysis is to compare the mean of a treatment with a control by conducting a hypothesis test. Our simulation shows that the statistical outcome depends on degree of data dispersion. In cancer cases, if the survival times become more widely dispersed, the point (often call t- statistic or F-statistic, etc.) for rejecting the null hypothesis will shift toward to a high value. This means that a weak treatment effect will be rejected as random errors at high chances (see all examples in F-I, Sup.). This can be seen from Figures 2, 3 as well.

In conducting a hypothesis test by using t distribution, a health property is observed before a treatment and after a treatment. The paired difference is used in conducting a hypothesis test. The rejection point depends on how patients respond to the treatment similarly. If they respond to the treatment in the exactly same way, even a small treatment’s effect can be recognized. However, large differences in patients’ responses will cause the rejection point to move toward to a large value for the same p value, and thus fails to recognize the effect of the treatment (F, Sup.). In conducting two populations’ mean test, large differences within each treatment group will cause the rejection point to shift toward to a large value (G, Sup.).

In conducting a variance analysis, uncontrolled interfering factors affect the health property to be measured. The test outcome depends on ratio of the variances of the treatment to the variances of the random error. If interfering factors are not controlled, they will go into the error term and thus reduce the ratio of treatment variances to error variances. The uncontrolled factors cause F statistic to shift to lower value so that the F test will be more likely to accept the null hypothesis (H, Sup.).

Interfering effects of uncontrolled factors cannot be corrected by any other statistical analysis method including χ^2 goodness-of-fit test, common frequency test (J-K, Sup.). Some statistical methods take into account only sampling drawing error, and others may address specific problems, but none have the power to correct this fundamental flaw that must be

addressed by raising measurement accuracy. The problem cannot be cured by any methods such as randomization and stratification (L-M, Sup.). Simpson's Paradox is also powerful proof that different persons cannot be treated as same in a clinical trial (N, Sup.).

Prior studies on the benefits of randomization in clinical trials are focused on how randomization reduces systematic biases [38,39] and prevents selection biases [40]. When human subjects are randomized, all interfering factors that affect trial outcomes can be similarly allocated to the treatment group and the control group. This similarity in their effects allows for statistical inferences on the treatment effects [41]. While those points are correct in the context of studying a strong treatment as shown in Figure 1, the method does not work when their effects are closer to the treatment’s effect. They did not consider how combining multiple factors as a single treatment can dramatically raise the capability to detect treatment effects.

Simpson's Paradox can be fully explained by the interference factors. An effect that occurs when the marginal association between two categorical variables is qualitatively different from the partial association between the same two variables after controlling for one or more other variables. The real cause of Simpson's Paradox is large variances at personal levels, and health properties from different persons cannot be studied in a model. In characterizing a chronic disease, each person must be treated as a unique system. A distinctive regression curve is presumed to exist for each person. When data from different people are pooled in conducting a regression analysis, it is an attempt to find a regression curve among different systems. Such a regression curve cannot be right except by accident. The regression pattern will change when one or more important factors is controlled. Parameters from the regression may be applicable to a population, but the population does not have diseases. Thus, any treatment developed using population data cannot cure diseases for any specific person. The pattern of Simpson's Paradox implies that such a regression data is improperly combined.

The problem discussed above is rooted in the fact that massive personal differences in clinical trials affect a measured health property. No statistical analysis, and nor any other methods under the Sun can correct this problem, which is like a bad laboratory report which is based on data generated by using an erratic household scale. Causal and interfering factors include health factors that patients can correct and factors that patients cannot change. Some of interfering factors are called as covariates; and some examples include sex, age, trial site, disease characteristics, disease prognosis, etc. [42]. A presumed fix is to achieve balance among treatment and control arms with the hope that the conclusions of a clinical study are not sensitive to covariates. However, none of the proposed methods in the CHMP Guideline can actually correct the biases of clinical trials because those measures cannot reduce the variances of the error term. In another study, attempts have been made to evaluate different methods for correcting baseline imbalances [43]. They focused only pre- and post-treatment scores and how different analytical methods affect biases, but did not address how interference factors distort the true effects of weak factors. The problem cannot be addressed by co-variance analysis.

We also show that health properties are not the types of things that can be summed up and averaged (O, Sup.). Good personal health is achieved by maintaining sophisticated balances. Beneficial effects and adverse or negating effects happen in different patients, and they cannot be averaged in

reality. This unique problem arises in the context of characterizing chronic diseases. It is safely assumed that chronic diseases are caused by imbalances, which can be caused by disturbance in two opposite directions. Each biochemical pathway must be maintained at a proper speed relative to other pathways, and changing the speed of this pathway in either way can disturb this balance. A same amount of qualitative change from a right pathway speed in one person cannot be used to compensate for the same amount of change in an opposite way in another person. However, statistical analysis is based on an assumption that health properties is fungible and transferable and thus can be summed up and averaged. This assumption cannot hold in reality. An identical amount of departure from the population's mean has different impacts on different patients. The same amount change may cure, hurt or kill a patient, depending on the specific conditions of the person.

Statistical analysis is based on an oversimplified and unrealistic assumption that health properties can be treated as a fungible property like weight and volume. Statistical analysis adds negating effects to the sum of the treatment and thus lowers the treatment's mean. This also results in wrong results like that the sum of 20% positive effects and 20% negative effects is equal to no effect. In reality, one can avoid negating effects by avoiding applying the treatment to mismatched patients and can actually deliver 20% positive effects. For those obvious reasons, a treatment protocol developed from a clinical trial predictably fails to work on real human beings. Optimization focusing on a single patient is the only way to avoid this fundamental flaw. This problem is less critical when health properties among "sufficiently similar subjects" are summed up or averaged to get rid of fluctuations caused by uncontrollable errors.

Based on a hypothetical model study, where each of k factors can influence the health property by a same amount, using k factors as a treatment is superior to the treatment using a single factor (P, Sup.). If each of the k factors has a same treatment effect and same variances in the health property and is similar to the experimental error, using an optimization trial to optimize the health property by using k factors will raise treatment effect by k times, and raises the T statistic, Z statistic and F statistic by about $k \cdot \sqrt{k}$. The sensitivity and ability of a hypothesis test to detect true treatment effect increases with the number of interfering factors. When the total number of factors is increased from 1 to 2, 5, 10, and 100, all test statistics will increase by 2.8, 11.2, 32, and 1000 times.

Discussion

Multiple Sources of Big Errors in Clinical Trials

In conducting a valid experiment, one fundamental requirement is that accuracy and reliability of detection technologies for detecting a treatment's effect must be sufficiently higher than those for detecting the experimental error. This requirement can hold only in studying strong treatments that can stand out over the apparent experimental error. In medical research, this requirement becomes that detectable treatment's effect must be much larger than the apparent experimental error. The failure of this presumption in clinical trials can be rephrased as one that the effect attributed to a treatment is too close to the apparent experimental error so that data set tends to come out with its test statistic falling within an acceptance region. This results in an outcome of failing to recognize weak treatment effect.

All statistical analysis methods are premised on the model assumptions, and every model assumption including the test hypothesis must be correct [44]. The fluctuations caused by beneficial and adverse/negating effects among different patients are not same as drawing error or true random errors in typical statistical models. The effects of uncontrolled factors may be merged into the experimental error only if the total experimental error is still sufficiently smaller than the treatment's effect. Big data dispersion in statistical analysis may not be ignored [45]. When the experimental error in a clinical trial is close to the treatment's effect, such a trial will generate meaningless results.

Lack of required accuracy and reliability is inherent in clinical trials used to characterize chronic diseases. Chronic diseases, by definition, progress slowly. This means that changes in any measured health property such as hazard ratio, organ function, survival time or other measured chemical data in any given time interval is infinitesimally small. Thus, the accuracy and reliability required to accurately characterize chronic diseases is much higher than those for studying acute diseases.

Compared with mechanical systems such cars, planes, etc, human beings are the most unfit subjects for clinical trials because of a massive number of genetic differences and phenotypes [7,8]. In addition, the personal differences are further increased by different emotional states of human beings. Since the massive personal differences in clinical trials interfere with accurate assessment of any health properties, it is impossible to detect weak and slow-delivering treatment effects. By using clinical trials, medical researchers cannot accurately determine what can cure chronic diseases and what harm personal health on long terms.

Statistical analysis has been widely abused in a long history [45]. Misuse of statistical analysis in medical research is a well known problem which has been discussed in a large number of studies [46-50]. Problems discussed in those cited studies are in addition to the model flaws we have discussed above.

One-Way Biased Conclusions of Clinical Trials and Their Severe Adverse Impacts on the Global Health Landscape

Our simulation results from all different models consistently show that the effects of clinical trials are one-way biased when the trial is used to evaluate a weak treatment. The averaging operation tends to reduce the treatment mean and this effect is not reflected in any assumption in basic statistical models. Statistical mean, μ_s , must be smaller or much smaller than μ_b , the actual beneficial mean when the treatment is correctly used only to be matched patients (For example, vitamin D is used only on those with Vitamin D deficiency but not on those with Vitamin D excess). This effect is described by a degrading factor $g = \mu_s / \mu_b$, which reflects the degree of "indiscriminate application" of the treatment. This value is in the range of 0 to 1. Statistical analysis is unfit for studying chronic diseases. If a measured health property is influenced by multiple interfering factors, a study focusing on a single treatment with other factors randomized will increase the chance to reject the treatment as having no effects on the health property. Hundreds of trials, with each focusing on one single factor, will result in failure to find effect for any of the factors.

Clinical trials distort hypothesis tests by enlarging the error term and statistical analysis reduces the treatment's effects by averaging effect. They both work in favor of rejecting the treatment. If a clinical trial results in rejection of the null hypothesis, the finding will likely stand except that the true treatment effects may be actually larger than determined values. However, if a hypothesis test outcome is acceptance of the null hypothesis, it may be wrong due to the negating effects and interfering effects. Therefore, conclusions in a good number of published studies should be interpreted differently. This one-way bias can be traced to the irreconcilable conflicts among massive personal differences, required high measurement accuracy and reliability, weak and slow effects of treatments for chronic diseases, and the unique roles of imbalances in chronic diseases.

Personal health is influenced by diets, nutrition, exercises, mind regulation, chronic stress, fears, etc. Many of those factors work like double-edged swords: they can benefit some patients, but hurt others if they are misused to destroy some established balances. The effects of nutrition and diets are expected to be highly random and unpredictable due to different personal lifestyles. In such a trial, the apparent error is inflated by dietary factors as uncontrolled interfering factors. Findings from a clinical trial represent only an abstract population, and are inapplicable to real patients as far as chronic diseases are concerned. A large number of factors in diet, lifestyles, exercise, emotional states, etc. can affect cancer outcomes, and thus, each study focusing on one single or a few factors will result in rejecting each factor as a potential treatment.

Being biased towards false acceptances, misused statistical analysis keeps rejecting weak and slow-delivering treatment effects. This explains why a clinical trial could not positively affirm single lifestyle factor's curative benefits even though it is found to be a significant risk factor of the disease in other types of long-term studies. Clinical trials are primarily responsible for promoting mainly surgery, synthetic drugs, radiation as "scientifically valid" treatments and rejecting potentially tens of thousands of non-medical weak and slow treatments, which would be one to two orders magnitude more powerful if they are used collectively in optimization trials. Clinical trials are most probably the main culprits that preclude the mankind from finding cures for chronic diseases. It is reasonable to infer that clinical trials are in part responsible for creating current national health epidemics in the U.S., China and many other nations in the world.

A serious problem is the combined toxic effects of environmental pollutants, contaminants, food additives, pesticide residues, herbicides, industrial chemicals, etc. By focusing on a single toxic agent in each trial, each such study cannot catch a weak and slow-delivering toxic agent. However, multiple toxic agents always work together in human bodies. A null finding could be "caused" by interference of other similar or stronger toxic agents and similar or stronger interfering effects. Most known toxic substances co-exist in human bodies. If a hundred similar factors are studied at the same time, Z statistic, T statistic and F statistic could be 1000 times more than counterpart in the clinical trial focusing on a single factor. When a large number of similarly harmful factors attack the human in the control, each of toxic agents is naturally hidden as "the experimental error." However, several, tens, or even hundreds of toxic agents can slowly damage human bodies. This single toxic agent can be identified only if all those toxic substances are not present in trial subjects. Findings from studying one or few toxic agents a time do not reflect the real damages of multiple toxic agents to the human body.

To find cure for a chronic disease, a required capability is determining which factors can speed up the disease's progression and which can slow down or reverse its progression. Considering massive differences among human subjects and a large number of interfering factors, clinical trials are unfit for establishing treatment protocols. Optimization trials using multiple factors as a treatment provides much better chances for finding cures for chronic diseases. We will show three huge gains below.

First, the biggest gain from using an optimization trial is to avoid negating effects caused by indiscriminate application of the treatment. For a single factor treatment, an optimization trial can raise beneficial effects by $(1/g)$, where $g = \mu_s / \mu_b$. It can be 1 to any reasonable number (See treatments C to G in Table 7S, Sup.). In clinical trials, the same treatment is indiscriminately used on all patients in the treatment group, many lifestyle factors can disturb various balances in two opposite directions. If those factors are randomly used against all patients in the treatment group or subgroup, their true beneficial effects on some patients can be "nullified" by their negating effects on other patients (per the analysis for the model in Table S7). In an optimization trial, controllable factors are used as part of treatment and are used on only the patients who need them. Sufficiently similar patients are selected in such a trial.

Second, we have shown that a large number of interfering factors directly interfere with clinical trials. They have different levels of effects and different variances. They can be used as part of treatment package for chronic diseases. Thus, a wise strategy is including multiple factors that would affect disease outcomes as a treatment package. The apparent error distribution in a patient can be estimated by the mean, $\mu_t = \mu_1 + \mu_2 + \dots + \mu_k$, and variances, $\sigma_t^2 = \sigma_e^2 + \sigma_1^2 + \sigma_2^2 + \dots + \sigma_k^2$ if all interfering factors are not used as part of the treatment. An improvement can be achieved by using the Model B. By bundling all controllable co-causal and interfering factors as a treatment package, the total treatment's effect is raised by k times while the error variances are reduced by about \sqrt{k} according to Medal B analysis.

The total gain in treatment effects existing in an optimization trial over a clinical trial is $(1/g)*k$ while the all test statistic such as T statistic, Z statistic, and F statistic used in hypothesis tests are increased by $(1/g)*k*\sqrt{k}$, where $1/g$ is attributed to avoiding negating effects, k is attributed to the additive effect of multiple treatment factors, and \sqrt{k} is attributed to reduction in the error variances. Their collective impacts could be huge. This conclusively shows why medicine could not find "scientific evidence" for any treatment based on a single lifestyle factor.

This conclusion is backed up surprisingly by all of simulation results for three hypothesis tests in every model we used in Supplement. Thus, we assume the gain is an inherent estimate (but not a precise number due to complexity of the human body). Moreover, the actual gains in mean or test statistics are predicted to be more than $(1/g)*k$ or $(1/g)*k*\sqrt{k}$. If interfering factors are matched to patients, true gain in the treatment effect is more than k times. We assume that some adverse effects which cannot be directly measured are not reflected in the negating effects and thus could not get into the g value. The true treatment effects could be further raised by avoiding adverse side effects. In contrast, optimization trials are good for using lifestyle factors, natural remedies, and mild or safe environmental factors, they do not implicate serious side effects.

Even though the variances of the treatment's mean \bar{X} could approach zero, the \sqrt{k} term most probably cannot be ignored by approximation in studying chronic diseases.

The inevitable conclusion of clinical trial's invalidity is strongly resonant with ancient medical practices such as herbal formulations and practices under the ancient holistic model. This ancient holistic model stresses the need to work on the whole body by using a large number of natural compounds or multiple treatment methods.

Based on the strength of our evidence as a whole, we reject clinical trials as a misused wrong experimental method for studying weak and slow-changing health properties and propose optimization trials as replacement. Optimization trial is suitable for studying weak, slow-delivering and natural remedies, but may not be used to study the side effects of single or a few synthetic drugs.

One solution is using a single human subject in a clinical trial. In this case, a control cannot be found because no two persons are similar in the world. Thus, the person's condition before the treatment is used as a control. This is essentially what was once used in ancient medical systems. The treatment effect is assessed by comparing the health properties before the treatment and after the treatment. One problem is that if the treatment lasts a long time, aging process can interfere with trial results and other previously used treatments may influence the current treatment. Some practical adjustments must be made. Trustworthiness of findings should be established by replicating the same trial for several times. Acceptance of this approach would require examining the rationale of using clinical trials. The notion that a treatment is good for all people in the population is clearly incorrect as far as chronic diseases are concerned.

An alternative solution is controlling all influencing factors in a mini trial so that the variances from interfering factors are minimized. It is difficult to get rid of the massive personal differences which are presumed to interfere with trial outcomes. What could be achieved in practice is using "sufficiently similar subjects" in the mini trial. To investigate a treatment in a trial, all significant co-causal and interfering factors including those listed in Table 3 and other known factors should be controlled. For example, relevant genetics, diet, exercises, toxic agent levels, medication use history, sex, age, race, emotional states, etc. are controlled. When the variances from those factors are controlled, the trial's sensitivity will be dramatically increased. By using sufficiently similar subjects, weak, slow-delivering and natural treatment effects can be detected with increased sensitivity. To see whether the treatment works on patients with similar important health properties, a second or third mini trial is conducted. After a series of mini trials have been done, a researcher can see when the treatment works and under what conditions the treatment works.

Personal genetics and emotional states are difficult to control. Personal genetics can be controlled by selecting human subjects. To control those factors, one should focus on their nexuses to measured health properties. If a treatment works on a particular biochemical process, subjects with known genes that control the process should be selected or voided, but other genes with little effects may be neglected. Emotional states should be stressed if they are predicted to play significant roles in influencing measured health

properties. When human subjects are nearly identical, variances attributable to personal differences will be dramatically reduced. In personalized medicine, randomization, subject selection bias, and statistical analysis have limited utility and should not control experimental designs.

When clinical trials involve a small number of sufficiently similar patients, statistical analysis should not be concerned. When all significant factors are controlled, measured health properties may be treated as ordinary variables and thus statistical analysis can be avoided or used as a mere causal checks. P-values 0.5 ± 0.15 (or any other suitable numbers) may be used because trustworthiness of trial findings are established by replicating mini trials. For a single person trial, statistical analysis cannot be used in most situations unless study purpose is examining things caused by instruments or sampling technologies, and trustworthiness of findings should be established by repeating the same or similar trial. All details on controlled factors should be documented for replicating the trial.

The single-person or mini optimization trial can be used to study combination of factors. Cancer is clearly responsive to lifestyle changes involving a large number of factors. When tens to hundreds factors are controlled, their combination effects are added up in some ways while co-causal and interfering factors are dropped out from the error term. All co-causal and interfering factors are used to promote healing in the treatment arm. Such an experimental design will dramatically increase the detection sensitivity of the trial and raise the treatment's effect.

All inaccuracy problems cannot be corrected by doing more studies under the same model, and the validity of clinical trials cannot be decided by examining the outcomes of studies. Similarly, the validity of our challenges cannot be evaluated by the same standards used in the regular manner. To correct such foundation error, one must examine presumptions, medical model development history, facts that support the use of the medical model, and newly discovered facts that are in conflict with the medical model. The widespread misuse of clinical trials has been driven by the incentive for avoiding selection bias, which results in much worse research culture that turning each patient into a statistical number. Rather than focusing on personal care and emotion, patients are treated like statistical figures. The deeply flawed research model has driven medical research into a wrong track for more than one century, and resulted in a research model barring the use of emotional component as part of cures.

As a result of influences of clinical trials, a notion that diseases are cured by indiscriminate exposure to approved treatments becomes dominate in medicine. Even though completely response rates of chemotherapy are 7.4% for cancer and virtually no cures are found for most chronic diseases, medicine is still unable to examine its foundation. Despite frequent criticisms by non-medical professionals, medicine is unable or unwilling to examine this flawed research model, and continues using the peer review system to maintain accepted wrong standard and suppress discoveries that would expose the deeply flawed foundation. By blocking the reforming path, medicine continues producing greatly biased, incorrect, or irrelevant research findings. Even after more than a century of failure to find cures and that U.S. medicare is predictably facing bankruptcy, the idea of "avoiding selection bias" still control research practices. While such measures kill

every chance of finding workable solutions to chronic diseases, they cannot stop researchers from altering data at trial levels. Besides, a better approach to preventing research fraud and bias is repeating the same trial by one or more times.

Optimization trials are superior to clinical trials for studying weak effects. By recognizing the validity of single-person trials and mini optimization trials, personal medical miracles can be conveniently studied. Research focus is not about experimental designs, evidence quality, statistical analysis, selection bias, rejection criteria, etc, but delivery of predictable cures which can be tailored to all specific patients including “minority patients.” This mission cannot be accomplished by indiscriminate application of treatments in clinical trials.

Limitations of This Study

Our findings are not applicable to clinical trials, the findings of which are not used as the basis for treating diseases. If the purposes of research are to explore costs and resource allocations, their validity are not subject to the same analysis. Also, if clinical trials are used to study disease mechanisms as a way to control health costs, they still provide useful information for policy makers.

It has routinely assumed that measured health property in a trial is mainly attributed to a treatment. However, this presumption is always breached in studying chronic diseases. Thus, our findings are applicable to any clinical trial. When a weak treatment plus at least one interfering factor affect the measured health property, the validity of trial outcomes depends on the relative size of the treatment to those of the interfering factor. Moreover, concerning chronic diseases, health properties are different from person to person. This implies that true cure must be formulated for each specific patient, and treatment established by population trials cannot restore balance for specific patients except by accident.

We note that the effects of interfering factors are not linearly additive, their effects may vary in degrees, their variances are not similar, their distributions may be not normal, and many factors may interact with each other in complex ways. However, they affect the mean and variances of the experimental error in certain ways. The effects from all interfering factors are added up linearly or non-linearly. When the causal and interfering factors are bundled into the error term, they ruin the trial. If they are bundled into a treatment, test statistics increase as a result of addition of all co-causal and interfering factors, and are further enlarged by an empirical multiplying factor that is attributed to the reduced variances of the apparent experimental error.

If a clinical trial’s design breaches any core assumption, its findings are incorrect for that reason. If the breach is sufficient to change trial outcomes, the trial is invalid without regarding the validity of our findings. Thus, whether or not those assumptions used in our models hold will not affect our conclusions. Our findings underscore the importance of adhering to model presumptions in designing clinical trials and conducting statistical analysis.

Conclusion

By examining machine repairing model and accuracy and reliability requirements for studying chronic diseases, we found that the one-treatment-for-a-population approach is flawed as far as it is used in studying chronic diseases. Clinical trials are good only if the treatments are sufficiently strong or when all human subjects can be treated as “nearly identical units” as in classical probability trials or classical clinical trials. None of the two conditions are met when clinical trials are used to characterize chronic diseases. Randomized clinical trails are unable to deliver required accuracy and reliability due to the massive personal differences attributable to genotypes, phenotypes and emotional states of individual persons.

We further found that clinical trials and statistical analysis are fundamentally flawed because they are in conflict with the multiple factors disease causes model. Their flaws are also revealed in the outcomes of all our simulation studies using hypothetical model data such as two population means hypothesis test, paired data hypothesis test, F-test in variance analysis, etc. We found that health properties are not the types of fungible things that can be summed up and averaged because all human beings must be treated as different things. Beneficial effects and adverse effects happen in different persons with different meanings, and cannot be averaged in reality. Statistical analysis degrades performance of the treatment by averaging beneficial and negative effects within each treatment or subgroup. This averaging operation dramatically degrades the treatment effects. In conducting statistical analysis, the poor accuracy problem becomes one that the total experimental error is closer or even larger than the treatment’s effects under the alternative hypothesis. Both the means and the variances of randomized and uncontrolled co-causal and interfering factors are added to those of the error term as an apparent error. When the apparent “error” is far too large relative to the effects of the treatment, data set tends to come out with test statistics falling on the region of acceptance of the null hypothesis, thus resulting in false acceptance of the null hypothesis or false rejection of true treatment effects. Those fatal flaws are expected to be present under most circumstances. No statistical method, no any other methods under the Sun, can ever correct this great bias that arises from breaching the core presumption used in the statistical model. Thus, clinical trials are invalid and have been misused in studying chronic diseases. “Incurable diseases” is manufactured by the wrong research model.

Our model analysis shows that optimization trials can dramatically increase chances to determine treatment effects than randomized clinical trials. Based on a multiple interfering factor model, where k co-causal or interfering factors can influence a measured health property by a same degree, a treatment package using all k factors is much better than using a single treatment. If each of the k factors has a same treatment effect and same variances, an optimization trial to evaluate the health property by using all k factors will raise the total treatment effect by $(1/g)^k$ times than a randomized trial (where g is a degrading factor caused by misapplication of a treatment to patients, with its value from 0 to 1), and raises T statistic, Z statistic or F statistic by about $(1/g)^k \sqrt{k}$. Assuming that a treatment has no negating effects, when the total number of the factors is increased from 1 (without any interfering factor) to 2, 5, 10, and 100, T statistic, Z statistic

and F statistic will increase by approximately 2.8, 11.2, 32, and 1000 times. Moreover, by avoiding negating effects, an optimization trial using k factors as a treatment package can raise treatment effect potentially by one to several orders of magnitude relative to randomized clinical trials. The gain cannot be eliminated by increasing the patient number in the trial. The findings show why studies using clinical trials cannot produce “scientific valid” evidence in support of using single lifestyle factor as cure for a chronic disease. The total effect magnitude of lifestyle factors implies that lifestyle factors are extremely powerful measures for chronic diseases and cancer.

Clinical trials have been correctly used in evaluating strong treatments for centuries, but widely misused in studying chronic diseases in the last century. No prior study has paid attention to the serious conflict between the massive inaccuracies caused by personal differences and the required high accuracy and reliability for characterizing chronic diseases. The misuse of randomized clinical trials was mainly driven by a misplaced incentive to avoid so-called selection bias and quality of evidence. The breached core presumption (or lack of accuracy and reliability), plus improper averaging of positive-and-negative treatment’s effects, plus interference of multiple interfering factors, plus the stringent rejection criteria or low p values inevitably resulted in biased or wrong conclusions in past medical studies. The research model is unable to reveal the true benefits of any weak and slow-delivering treatment which could be vitally important for correcting subtle imbalances in the human body. The misuse of clinical trials are predictably responsible for the failure to find treatment effects for chronic diseases and failure to identify harmful effects of toxic compounds in environment. In sum, clinical trial should be rejected because it offers no chance to find cures under any of our theoretical and practical models mimicking real chronic diseases. Our findings may be similarly applicable to randomized controlled trials used in social sciences, environmental studies, life sciences, etc. as long as those required conditions are met. However, the validity of trials depends on whether a particular trial can offer enough accuracy required for addressing the research question. It is not strange the same study finding can be valid for one purpose but completely wrong for another purpose.

Additional Information:

Supplementary information is provided.

Updates and corrections will be provided in the supplement.

Funding Statement:

The author(s) declared that no grants were involved in supporting this work.

Prior Publication History:

The article was first published in the SSRN e-journal on November 17, 2019, and revised for republication due to its impacts in medicine.

References

1. Tinker A. How to improve patient outcomes for chronic diseases and comorbidities. *Health Catalyst*. 2017.
2. Raghupathi W, Raghupathi V. An empirical study of chronic diseases in the United States: A Visual Analytics Approach to Public Health. *Int. J. Environ. Res.* 2018 Mar;15(3):431.

3. Ashdown ML, Robinson AP, Yatomi-Clarke SL, et al. Chemotherapy for late-stage cancer patients: meta-analysis of complete response rates. *F1000Res.* 2015;4:232.
4. Albero A, Lopéz JE, Torres A, et al. Effectiveness of chemotherapy in advanced differentiated thyroid cancer: a systematic review. *Endocr Relat Cancer.* 2016 Feb;23(2):R71-84.
5. Maeda H, Khatami M. Analyses of repeated failures in cancer therapy for solid tumors: poor tumor-selective drug delivery, low therapeutic efficacy and unsustainable costs, *Clin Transl Med.* . 2018 Dec;7(1):1-20.
6. Milken Institute. Milken Institute Study: Chronic Disease Costs U.S. Economy More Than \$1 Trillion Annually: July 2, 2019.
7. Ogino S, Fuchs CS, Giovannucci E. How many molecular subtypes? Implications of the unique tumor principle in personalized medicine. *Expert Rev Mol Diagn.* 2012 Jul 1;12(6):621-628.
8. Ogino S, Lochhead P, Chan AT, et al. Molecular pathological epidemiology of epigenetics: emerging integrative science to analyze environment, host, and disease. *Mod Pathol.* 2013 Apr;26(4):465-84.
9. Access Medicine. ClinicalLaboratory. ReferenceValues
10. Nowell PC. The clonal evolution of tumor cell populations. *Science.* 1976 Oct 1;194(4260):23-28. Nowell PC.
11. Sagar SM, Yance D, Wong RK. Natural health products that inhibit angiogenesis: a potential source for investigational new agents to treat cancer-part 1. *Curr Oncol.* 2006 Feb;13(1):14-26.
12. Schuetz JM, Grundy A, Lee DG, et al. Genetic variants in genes related to inflammation, apoptosis and autophagy in breast cancer risk. *PLOS ONE* 2019;14(1):e0209010.
13. Ueshima H, Sekikawa A, Miura K, et al. Cardiovascular disease and risk factors in Asia: a selected review. *Circulation.* 2008 Dec 16;118(25):2702-2709.
14. Armitage P, Doll R. The age distribution of cancer and a multi-stage theory of carcinogenesis. *Br J Cancer.* 1954;8(1):1-12.
15. Prasad S, Sung B, Aggarwal BB. Age-Associated Chronic Diseases Require Age-Old Medicine: Role of Chronic Inflammation. *Prev Med.* 2012 May 1;54:S29-37.
16. Kalben BB. Why men die younger: causes of mortality differences by sex. *North American Actuarial Journal.* 2000 Oct 1;4(4):83-111.
17. Segerstrom SC, Miller GE. Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry. *Psychological Bulletin.* 2004;130(4):601-630.
18. Sloan EK, Priceman SJ, Cox BF, et al. The sympathetic nervous system induces a metastatic switch in primary breast cancer. *Cancer Research.* 2010 Sep 15;70(18):7042-7052.
19. Brydon L, Magid K, Steptoe A. Platelets, coronary heart disease, and stress. *Brain Behav Immun.* 2006 Mar 1;20(2):113-9.
20. Sundquist J, Li X, Johansson SE, Sundquist K. Depression as a predictor of hospitalization due to coronary heart disease. *Am J Prev Med.* 2005 Dec 1;29(5):428-33.
21. Nemeroff CB, Musselman DL, Evans DL. Depression and cardiac disease. *Depression and Anxiety.* 1998;8(S1):71-79.
22. Millimouno FM, Dong J, Yang L, et al. Targeting Apoptosis Pathways in Cancer and Perspectives with Natural Compounds from Mother Nature. *Cancer Prev Res.* 2014 Nov 1;7(11):1081-107.
23. Vivarelli S, Salemi R, Candido S, et al. Gut Microbiota and Cancer: From Pathogenesis to Therapy. *Cancers (Basel).* 2019 Jan;11(1):38.
24. Navarro SL, Kantor ED, Song X, et al. Factors associated with multiple biomarkers of systemic inflammation. *Cancer Epidemiology Biomarkers and Prevention.* 2016 Mar 1;25(3):521-31.

25. Giovannucci E, Harlan DM, Archer MC, et al. Diabetes Care. 2010 Jul; 33(7):1674-1685.
26. Liao JB. Cancer issue: Viruses and human cancer. Yale J Biol Med. 2006 Dec;79(3-4):115.
27. Salani R, Backes FJ, Fung MF, et al. Posttreatment surveillance and diagnosis of recurrence in women with gynecologic malignancies: Society of Gynecologic Oncologists recommendations. Am J Obstet Gynecol. 2011 Jun 1;204(6):466-478.
28. Krall JA, Ferenc R, Mercury OA, et al. The systemic response to surgery triggers the outgrowth of distant immune-controlled tumors in mouse models of dormancy. Sci. Transl. 2018 Apr 11;10(436):eaan3464.
29. Colleoni M, Sun Z, Price KN, et al. Annual hazard rates of recurrence for breast cancer during 24 years of follow-up: results from the international breast cancer study group trials I to V. J. Clin. Oncol. 2016 Mar 20;34(9):927-935.
30. Cheng L, Swartz MD, Zhao H, et al. Hazard of recurrence among women after primary breast cancer treatment-A 10-year follow-up using data from SEER-medicare. Cancer Epidemiol. Biomarkers Prev. 2012;21:800-809.
31. Demicheli R, Retsky MW, Hrushesky WJ, Baum M. Tumor dormancy and surgery-driven interruption of dormancy in breast cancer: learning from failures. Nat. Clin. Pract. Oncol. 2007 Dec;4(12):699-710.
32. Cormie P, Zopf EM, Zhang X, Schmitz KH. The impact of exercise on cancer mortality, recurrence, and treatment-related adverse effects. Epidemiologic reviews. 2017 Jan 1;39(1):71-92.
33. Booth FW, Roberts CK, Laye MJ. Lack of exercise is a major cause of chronic diseases. Compr Physiol. 2012;2(2):1143-1211.
34. Levine EM, Robbins EB. Differential temperature sensitivity of normal and cancer cells in culture. Journal of Cellular Physiology. 1970 Dec;76(3):373-379.
35. Yeung PK, Wong JT. Inhibition of cell proliferation by mechanical agitation involves transient cell cycle arrest at G1 phase in dinoflagellates. Protoplasma. 2003 Mar;220(3):173-8.
36. Yusuf S, Hawken S, Ounpuu S, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. The lancet. 2004 Sep 11;364(9438):937-952.
37. Deaton A, Cartwright N. Understanding and misunderstanding randomized controlled trials. Soc Sci Med. 2018 August;210:2-21.
38. Kalish LA, Begg GB. Treatment allocation methods in clinical trials a review. Stat Med. 1985 Apr;4(2):129-44.
39. Fleiss JL, Levin B, Park MC. Statistical Methods for Rates and Proportion. 3rd Ed. Hoboken NJ, John Wiley & Sons; 2003.
40. Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. The Lancet. 2002 Feb 16;359(9306):614-618.
41. Altman DG. Randomisation. BMJ. 1991;302:1481-2.
42. CHMP (Committee for Medicinal Products for Human Use). Guideline on adjustment for baseline covariates in clinical trials. 26 February 2015.
43. Egbewale BE, Lewis M, Simcorresponding J. Bias, precision and statistical power of analysis of covariance in the analysis of randomized trials with baseline imbalance: a simulation study. BMC Med Res Methodol. 2014 Dec;14(1):1-2.
44. Greenland S, Senn SJ, Rothman KJ, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol. 2016 Apr;31(4):337-350.
45. Campbell SK. Flaws and Fallacies in Statistical Thinking. Prentice-Hall Inc, 1974.
46. Strasak AM, Zaman Q, Pfeiffer KP, Göbel G, Ulmer H. Statistical errors in medical research-a review of common pitfalls, SWIS SMED WKLY 2007 Jan 27;137(0304).
47. Gore SM, Jones IG, Rytter EC. Misuse of statistical methods: critical assessment of articles in BMJ from January to March 1976. Br Med J. 1977 Jan 8;1(6053):85-7.
48. Kim JS, Kim DK, Hong SJ. Assessment of errors and misused statistics in dental research. Int Dent J. 2011 Jun;61(3):163-7.
49. White SJ. Statistical errors in papers in the British Journal of Psychiatry. Br J Psychiatry 1979 Oct;135(4):336-42.
50. Hall JC, Hill D, Watts JM. Misuse of statistical methods in the Australasian surgical literature. Aust N Z J Surg. 1982 Oct;52(5):541-3.

Submit your manuscript at
<http://enlivenarchive.org/submit-manuscript.php>
New initiative of Enliven Archive

Apart from providing HTML, PDF versions; we also provide **video version** and deposit the videos in about 15 freely accessible social network sites that promote videos which in turn will aid in rapid circulation of articles published with us.