Exploring heterogeneity of treatment response: assumptions, logic, algorithm, computations

Lev S Sverdlov, M.D., Ph.D. [i]

Introduction

The ideas of individualized, personalized treatment designate one of the major trends in modern medicine. This motivates a growing interest in the study of heterogeneity of the population exposed to treatment. In this paper, we consider two different views on the heterogeneity of the population: the heterogeneity of the *effect of treatment*, and the heterogeneity of *response of the subject to the treatment* among the members in this population.

Traditionally, the heterogeneity of *treatment effect* is understood as "the nonrandom, explainable variability in the direction and magnitude of treatment effects for individuals within a population,"[1] where the treatment effect is defined as reduction of risk for negative outcome. An amount of studies in this area is growing, as well as a level of their sophistication.

This paper focuses on the *treatment response*, which is another, substantially less studied aspect of the treatment process. With regard to treatment, the term "response" is used out of established clinical and terminological tradition and often with no satisfactory definitions. Sometimes, the "treatment effect" and "response to treatment" are used as synonyms. In many clinical publications, the patients with a negative outcome are labeled as "non-responders," and those with a positive outcome - "responders," which has some misleading aspects.

People react to treatment differently: while being exposed to the same treatment, various individuals have different outcomes, and while being exposed to different treatments, various individuals can have either the same outcome, or different outcomes. It moves a focus of attention toward heterogeneity of phenomena and factors of the *individual ways of reacting* to a specified treatment, and the heterogeneity of *treatment response* can be defined as qualitative and quantitative diversity of response to treatment, and factors determining this response among individuals within a population.

In this paper, we use the same set of definitions and identical notation for describing the *effect of treatment*, and *response of the subject to the treatment.* It makes the approaches comparable by their major characteristics, and compatible in analysis of the heterogeneity of a population. Also, we consider these two approaches using the same model a randomized clinical trial with a mortality as a primary outcome as a model of the trial population. It makes the approaches comparable by their major characteristics, and compatible in analysis of the heterogeneity of a population.

The difference in starting positions and focusing on different aspects of the treatment process dictates the differences in the methodology. *Heterogeneity of treatment effect* is thoroughly studied within a paradigm of a statistical approach, using classical and modern statistical and computational methods. Our paper is focused primarily on the *heterogeneity of treatment response,* which is not studied systematically yet. The *heterogeneity of treatment response* also can be subject to statistical analysis, especially given the capacity of modern statistical and computational approaches, but originally the approach towards analysis of the *heterogeneity of treatment response*

[i] E-mail: lev@redmondanalytics.com

derives from a tradition of a clinical method, which essentially is a qualitative, exploratory approach, and which relies on logic as its primary instrument of analysis.

1. Definitions and Notation

1.1 Population

Per Jerzy Neyman, at a logical level, a population is defined as "categories of entities satisfying certain definitions but varying in their individual properties." (ref. 9, p.1). In our analysis, we consider a population of a placebo controlled randomized clinical trial with a mortality as a primary outcome as a model.

The model of a clinical trial we use consists of four major components: treatment population, treatment, outcomes, and covariates (conditions). These components are expressed in a form of binary variables following a terminological and methodological tradition for clinical trials on the effects of treatment on morbidity and mortality. Also, it is natural for the compared modalities of treatment to be expressed as binary or categorical variables. A positive outcome vs. negative one, i.e., having a heart attack vs. not having it, or recovery vs. death, are the most clear-cut clinical presentations of this dichotomy. Analysis of continuous outcomes may require more logical steps and complex computations, but the principles of the approach will remain the same.

At the analytical level, the study population is described as a set of individuals: $\mathrm{P} = (I_i)$, where an individual case I_i is described by a set of variables (conditions) $I_i = (T_i, Y_i; A_i, B_i, C_i \ldots; \alpha_i, \beta_i \ldots)$; or alternatively $T(I_i) = T_i,\ A(I_i) = A_i$, etc., and presented in the format of the data matrix (Fig 3A).

1.2 Data matrix

The data matrix is a description of a study population in the format suitable for statistical analysis and as such it is a framework for analysis of experimental and observational data. The data matrix consists of n rows and p columns[2] as shown on Fig 3A. The observations (cases, individuals) presented as the rows are assumed to be separate units (I_j). The variables describing the individual including treatment (Tx), outcome (Y) and conditions $(A, B, C, \ldots)$, physical, clinical, demographic, environmental, etc., characteristics of each case, i.e., conditions $(A, B, C, \ldots)$ are listed consecutively such that the values of each variable in several cases create a separate column.

The observations (cases, individuals) presented as the rows are assumed to be separate units (I_j). The variables describing the individual including as treatment (T), outcome (Y) and conditions $(A, B, C, \ldots)$, physical, clinical, demographic, environmental, etc., characteristics of each case, i.e., conditions $(A, B, C, \ldots)$ are listed consecutively such that the values of each variable in several cases create a separate column.

1.3 Treatment (Tx/tx)

In this article, the meaning of the term "treatment" (Tx) is as it is understood in medicine and public health, i.e., an action or a complex of actions to cure a disorder or improve a state of a patient. Depending on the context, it can be an independent variable (e.g., in the context of the assessment of the treatment effect), or a dependent variable (e.g., in the context of exploring an indication for the treatment). It is presented as

- $Tx = 1$, which is active treatment (also labelled Tx_+);
- $Tx = 0$ is a lack of active treatment (also labelled Tx_-);

1.4 Outcome (Y/y)

In probability theory, an outcome is a possible result of an experiment (in our case it is treatment). In this meaning the term is used for designating the outcome in the individual exposed (or not exposed) to treatment during a defined period of time

- $Y = 1$ designates survival (also labelled Y_+);
- $Y = 0$ designates death (also labeled Y_-).

At the population level, the outcome designates a proportion of the sum of individuals with the outcome of interest in the defined population.

1.5 Conditions

The term "conditions" is used broadly as it is used in the probability theory. In the context of this article, it covers the variables referring to characteristics and properties of an individual, internal or external environments, events, and relationships. The conditions also will be presented as binary variables.

- A, B, C ...: include external and internal environments, and a time factor; ($A = 1{,}0; B = 1{,}0; C = 1{,}0;$ ... ; also labeled $A = A_+, A_-; B = B_+, B_-; C = C_+, C_-;$ ...)
- α, β ...: unknown or unobserved variables.

1.6 Determinism

The term "determinism," "deterministic" is used in a narrow sense, as it is used in statistics, i.e., "opposite to random," "related," with the probability 1 or 0.

2. Objectives

Using the model described in the previous section, we will consider the heterogeneity of a population of a randomized clinical trial. Using the same definitions and notation above, the relationships between treatment, outcome and conditions will be considered from two positions: from a position of statistical analysis, and from the position of exploratory analysis deriving from a tradition of clinical method. We will describe a heuristic for logical analysis leading to a possibility of statistical examination of the developed hypotheses. Also, we will explore formal relationships between the heterogeneity of treatment response and treatment effect. Thus, we hope to set a bridge between these two hardly compatible approaches in the hope to better understand the treatment process and to make a step towards the individualization of treatment.

3. Heterogeneity of treatment effect

3.1 Treatment effect

A randomized clinical trial assigns people randomly to treatments. The groups are (on average) equivalent, and the difference in the outcome can be attributed to the treatment since that was the only difference between the groups. An estimate of the difference is called the Average Treatment Effect.[3] The effect of treatment is measured using indices of absolute and relative risk reduction (*ARR* and *RRR* respectively)

- absolute risk reduction: $ARR = Pr(Y_-^c) - Pr(Y_-^a)$;
- relative risk reduction: $RRR = \frac{Pr(Y_-^c) - Pr(Y_-^a)}{Pr(Y_-^c)}$;

where risk is understood as a probability of a negative outcome estimated as a proportion of this outcome in the studied population, a stands for an active arm of the trial, and c - for control.

Thus, the treatment effect, as it assessed in a clinical trial, is a statistical concept, and the *heterogeneity of treatment effect* is commonly considered in the frame of a statistical paradigm as "the nonrandom, explainable variability in the direction and magnitude of treatment effects for individuals within a population."

3.2 Assumptions

For our discussion, it is necessary to emphasize some of the premises implicit in traditional statistical analysis of the results of a randomized clinical trial as follows.

1. *Subjects are anonymous and interchangeable*. Personal information may move from one group to another for the reasons not related to analysis. Exclusion of any single subject from the study, moving him or her from one group to another, and/or trading single subjects between groups does not affect substantially the result of the study.
2. *Two or more events co-occur by chance unless the contrary is proven.* Statistical models inherently involve randomness. The hypothesis of random co-occurrence of the events (null hypothesis) is accepted or rejected because of its correspondence (or a lack of correspondence) to the observed relationships.
3. *Numerous subjects are required for making valid inferences.* A special chapter in any manual for clinical research is devoted to estimating a sample size to obtain a desired level of power before data collection.

3.3 Progress in studying heterogeneity of treatment effect

In virtually any population of patients, whether it is a "real world" population of treated patients, or a thoroughly selected clinical trial population, there is a considerable variation in the risk of the outcome of interest. The "average" benefit observed in the summary result of a clinical trial may even be non-representative of the treatment effect for a typical patient in the trial to the extent that some subgroups of patients can benefit when summary results of trials are negative.[4]

A typical approach for examining HTE is subgroup analysis,[5] which studies risk of a negative outcome conditional on some selected variables. Importantly, "selection of subgroups should be based on mechanism and plausibility (including clinical judgment), taking into account prior knowledge of treatment effect modifiers." [1]

During a last decade, one can observe a great progress in studying the heterogeneity of treatment effect. Apparently, the progress is motivated by the ideas of the individualization of treatment and supported with the modern statistical and computational methods.[6] Having started from subgroup analysis, the methodology of the studies on the heterogeneity of treatment effect is becoming more sophisticated and complex. For instance, using deep learning, the identification of the variability of the treatment effect was performed across clinically relevant subgroups associated with clusters of variables.[7] Lamont, et al.,[8] identified predicted individual treatment effect in a randomized clinical trial using Monte Carlo simulation, multiple imputation, non-parametric random decision trees.

3.4 Limitations in analysis of heterogeneity of treatment effect

Investigation of heterogeneity of treatment effect has some limitations related to the nature of the approach.

The decisions on the individualized treatment are to be made based on the assessment of the effectiveness of the planned treatment. The indices of risk, which are the integral component of the indices of the treatment effect are computed as a proportion of negative outcome in the population of the clinical trial, or the set of observational data. Therefore, the assessment of the heterogeneity of treatment effect requires data on the entire trial population for analysis of the variability of risk for a negative outcome; also, for identification of the segments of the population with outstanding values of the risk the entire set of data of the completed trial are required. In the "real world," however, the decisions on the individualized treatment often are to be made based on information from populations with unknown distributions of variables, from groups with incomplete data, small samples, and in non-randomly selected

individuals, etc., i.e., systematized data on risk are not necessarily available, and when they are available, we face even more fundamental limitations.

In the article discussing major theoretical aspects of heterogeneity of population in the context of statistically understood causality, Yu Xie states that in the context of "population thinking," "the ubiquitous presence of individual-level variability makes it impossible to study individual-level causal effects." [9] It corresponds with the notion of many good statisticians that the assessment of the effect of treatment in an individual case is impossible.

Per Richard von Mises, "When we speak of the "probability of death," the exact meaning of this expression can be defined in the following way only. We must not think of an individual, but of a certain class, e.g., "all insured men forty-one years old living in a given country and not engaged in certain dangerous occupations." A probability of death is attached to the class of men or to another class that can be defined in a similar way. We can say nothing about the probability of death of an individual even if we know his condition of life and health in detail. The phrase "probability of death," when it refers to single person, has no meaning for us at all."[10]

Richard von Mises introduced the concept of "collective" to emphasize that probability does not deal with individual cases. (An) "… example of a collective is a whole class of insured men and women whose ages at death have been registered by an insurance office."…"The definition of probability which we shall give is only concerned with 'the probability of a certain attribute of this collective'."

In subgroup analysis, risk of the negative outcome is assessed conditional on a selected condition, i.e., we reduce a size of the "collective," but the relationships between the single person and the "collective" remains the same.

In more sophisticated analyses [1,6,7,8] using multivariable models and accounting for a combination of several selected conditions, a size of the "collective" can be reduced even more.

Paradoxically, the size of the "collective" can be reduced to just a single person,[8] but the individual treatment effect in this case is predicted based on analysis of the distribution of the predictive variables within the entire analyzed population, and the meaning of risk for an individual patient becomes fundamentally unclear. Also, predicting the effect in a different population with a different distribution of the predictive variables is problematic similar a typical problem of applying the result of a clinical trial to the "real world" population.

4. Heterogeneity of treatment response

Determining, whether the observed outcome related or not related to the treatment is an everyday analytical task in medicine, drug safety and other areas. In the frame of a binary model, a response of a patient to treatment can be either recovery, or death. The patient could recover either because of treatment, or regardless of treatment because some patient could recover without any treatment.[11] In the case of a negative outcome, the death could be caused by the target disorder, but also it could be due to other causes: accident, suicide, stroke, infection, etc. Under the *treatment response* we mean a reaction of a member of the population to the treatment, to which he or she has been exposed, expressed as an outcome attributable to the treatment.

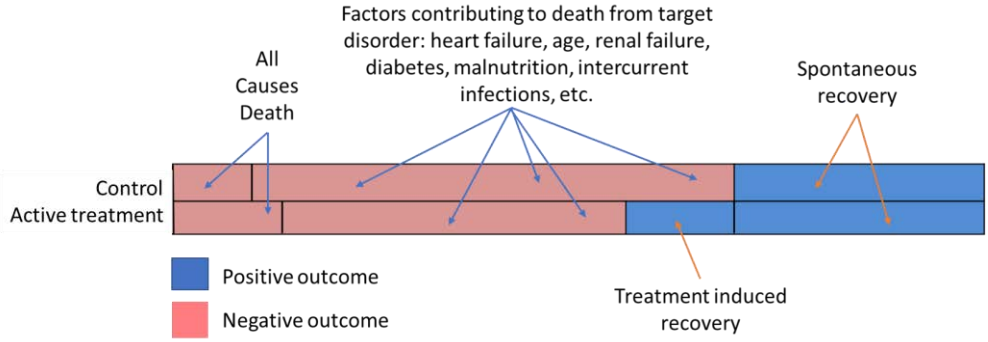
4.1 Clinical perspective on heterogeneity

The relationships between treatment, outcome and conditions in each individual case are routinely evaluated using the clinical method, which exposes the heterogeneity of the treatment response. Analysis of the methodology of clinical method is far beyond our

objective. We refer to the approach utilizing a clinical method for analysis of single cases and small groups which is routinely used in clinical practice and in monitoring safety of clinical trials, analysis of data in post marketing surveillance, etc.[12] If used for analysis of the heterogeneity of the trial population, this approach, reveals the following.

Both the populations of active and control (placebo) arms of the trial are heterogeneous (Fig. 1). While having been exposed to respective treatment, each of them is divided to unequal groups with negative and positive outcome. Regarding the outcome, each of these groups is phenomenologically homogeneous, but each of them consists of diverse subgroups differing by substantial factors.

Fig.1. Histogram of randomized clinical trial



4.1.1 Negative outcome

In virtually any trial with a mortality being a primary outcome, among subjects with the negative outcome there is a subgroup of All Causes Deaths (ACD), i.e., those who have died not from a target disorder, but from different causes, such as accidents, suicides, other disorders, etc. It means that both in active arm and control, the groups with negative outcome are heterogeneous since they include the All Causes Death subgroups as well as the subgroups of those died from the target disorder. In turn, these subgroups are heterogeneous too: one of them is heterogeneous by definition ("All Causes Death"), and in the subgroup of those having died from the target disorder, the factors contributing to the deaths (e.g., age, heart failure, diabetes, stroke, infection, etc.) are numerous and diverse.

Note, that in clinical trials, the cause of death as well as the contributing factors are usually identified with all possible accuracy, using clinical, pathology, biochemistry, etc. state-of-the-art methods, such that these data are considered final and used as a foundation for all other inferences.

Within the relevant subgroups, the identified causes and contributing factors can be distributed at random, or by clusters, or they can have a hierarchical structure, i.e., various patients and subgroups of patients with the negative outcome can be associated with different "generating conditions" or subsets of the "generating conditions." Still, all these diverse patients and groups of patients had one common characteristic: by various reasons, under influence of diverse factors, they were not sensitive to the treatment they were exposed to.

The ACD subgroups are not necessarily equal in the active arm and in control. This usually poses substantial challenge to the assessment of the effect of treatment. This problematic situation does not have a satisfactory analytical solution, and usually in

analysis of the treatment effect it is accepted under a strong assumption that the impact of this factor is not significant.

4.1.2 Positive outcome

The group of patients with a positive outcome also is not homogeneous. Obviously, the patients with positive outcome from the control have recovered spontaneously since they were not exposed to the active treatment. Our trial is randomized. Therefore, in the active arm, we should have an approximately equal, but not directly identifiable group of patients, whose recovery cannot be attributed to the effect of active treatment – apparently, they would recover spontaneously. Only the rest of the positive outcomes could be directly attributed to the effect of active treatment, but neither of these groups can be directly identified. There are reasons to believe that the capacity to recover spontaneously is related to numerous factors – immunological, nutritional, genetical polymorphism, environmental factors, etc. - these subgroups are in turn heterogeneous.[11]

5. Conceptual apparatus of heterogeneity of treatment response

Thus, following an exposure to the same treatment, various individuals can produce different outcomes; in turn, following two different treatments, various individuals can produce the same outcome or different outcomes, i.e., the individual outcome depends on the treatment, but rather it depends on the way the individual responds to the treatment he or she is exposed to. From a clinical standpoint, each patient reacts to the treatment in his or her individual way. Factors forming the response to the treatment can be unique for an individual or prevalent in the population; they can cluster, and create hierarchy, i.e., the trial population, as far as it is comprised of human beings suffering from real disorders, is intrinsically heterogeneous regarding their response to treatment, such that for various individuals *the ways to death can be diverse, as well as the ways to recovery*.

Analysis of heterogeneity of treatment response derives from a set of assumptions which are not acceptable for statistics but are natural and necessary for qualitative explorations and inferences, and particularly, for clinical thinking.

5.1 Assumptions

I. *Each subject is a unique individual.* Changing, removing, and/or adding individuals might completely change the subject and result of analysis.

II. *Two or more co-occurring events are related unless the contrary is proven.* This assumption is the major motivating factor for analysis, and it can be accepted if the possibility of the co-occurrence of the events by chance is small, and it should be rejected when this possibility is large.

III. *Valid inferences potentially can be made from single cases and small number of cases.*

It must be strongly emphasized, that these are not positive assertions, but rather the assumptions in a genuine meaning of this term, designating starting positions of the approach, which is explorative by its nature. The inference deriving from analysis based on these assumptions is relevant for this instance (case, cases) only. However, the conclusion on this instance can be an element of further analysis, for instance, comparing with other similar and different elements, producing inductive and deductive inferences, and prone to comparing the inferences with other hypotheses, and thus correcting and further developing the obtained knowledge.

5.2 Aggregation

In two or more cases (sets of variables describing individuals) placed in juxtaposition, some subsets of variables may be identical (“similar”), while the others may be not. In

statistics, a correlation (in a broad sense) can be one of the measures of the similarity/dissimilarity of individuals (more precisely, descriptions of). For our binary model, a relevant measure for two binary variables $\boldsymbol{V}_{(i)}$and $\boldsymbol{V}_{(i+1)}$ can be a *Phi* coefficient

$$r_{\varphi} = \frac{n_{11}n_{00} - n_{10}n_{01}}{\sqrt{n_{1.}n_{0.}n_{.0}n_{.1}}};$$

where, n_{11} stands for a number of "similar" pairs $V_{(i)+}V_{(i+1)+}$, n_{00} for a number of "similar" pairs $V_{(i)-}V_{(i+1)-}$, n_{10} for a number of "dissimilar" pairs $V_{(i)+}V_{(i+1)-}$ and n_{01} for a number of "dissimilar" pairs $V_{(i)-}V_{(i+1)-}$.

For instance, on Fig 2a, two individuals, I_1 and I_2, are described with a set of variables *V = A, B, C, …, I, J*. The similarity of these two individuals in statistical terms can be measured as $r_{\varphi} = \sim 0.41$. Note that for computing r_{φ} only the amount of "similar" and "dissimilar" pairs matters. It does not matter, which of the pairs specifically are identical and which are not. Thus, this measure converts information on the similarity/dissimilarity of individual pairs into the characterization of the similarity/dissimilarity of the compared individuals, which is informative for the entire population, but does inform any more about similarity/dissimilarity of each variable in the pair.

As far as we are going to build our analysis on the assumptions I,II, and III, we are - contrary to the above - interested in identifying a type of the association, which is directly observable as the "similar" segments of the descriptions in some members of the population, assuming that these variables, in these individuals can be related to each other.

Therefore, we introduce a concept of aggregation. Unlike the correlation, the aggregation refers to a subset of variables about which it is thought or known from observation *that this subset of variables* in the *individuals of this subset* is identical.

For instance, the aggregation, $Ag = \begin{vmatrix} I_j & |V_{i-}, V_{(i+1)+} \\ I_{j+1} & |V_{i-}, V_{(i+1)+} \\ & \dots \\ I_{j+k} & |V_{i-}, V_{(i+k)+} \end{vmatrix}$ is the *known* segment of the population with identical values of the variables $\boldsymbol{V}_{(i)}, \boldsymbol{V}_{(i+1)}, \dots,$ and $\boldsymbol{V}_{(i+k)}$ in the individuals $I_j, I_{j+1}, \dots, I_{j+k}$. Hereafter, we use the term "aggregation" if the assembly of elements gathered non-randomly. For the assemblies gathered by chance we use the term a "case-variable association."

The elements of the aggregation assumed related to each other. As noted above, this is the assumption, not assertion. This is only a starting position for analysis, a hypothesis, which must be scrutinized and compared to other hypotheses in the process of analysis.

5.3 Factors of treatment response: categories

The infinitely numerous factors forming the individual treatment response can be grouped, clustered for the sake of analysis in many ways, but ultimately, they can be generalized into two not mutually exclusive categories. One of them is the capacity (or a lack of capacity) of a subject to respond positively to the treatment, i.e., be sensitive or not sensitive to the treatment. Another one is the capacity (propensity, predisposition, readiness) to recover spontaneously, i.e., without, or regardless of treatment. These

categories were thoroughly considered in our work[11] from historical, clinical, epidemiological, and experimental positions.[ii]

We call a patient *"Sensitive"* (St_+) to a specified treatment if there are reasons to believe that this treatment has imposed the outcome of interest, or it would have imposed it if applied.

We call a patient *"Spontaneous"* (Sp_+) if a positive outcome has developed without of treatment, or there are reasons to believe that it would have developed if the treatment was not applied.

Under each of the categories, *"Spontaneous," and "Sensitive,"* we mean a relevant property (attribute, quality, characteristic, ability or trait) of a patient determined by a single factor inherent to a relevant category of patients (e.g., polymorphism), or by a confluence of multiple internal or external factors, either prevalent or rare.

Essentially, under the assumptions I, II, III, the identification and prediction of the treatment response in an individual becomes possible if at least some of the conditions and/or markers covered by the categories "Sensitive" and "Spontaneous" are known, with two reservations:

- at this point we do not account for possible interactions between the factors;
- unknown factors (conditions) capable of modifying the response can exist.

5.4 Treatment-Outcome Complex

Classification of patients by treatment and outcome creates a 2×2 table (Table 1) which is typically used as a framework for the assessment of the treatment effect.

Table 1. Combinations of Treatment and Outcomes

	Tx_+	Tx_-
Y_+	$Tx_+\ Y_+$	$Tx_-\ Y_+$
Y_-	$Tx_+\ Y_-$	$Tx_-\ Y_-$

In our analysis, we consider treatment and outcome not separately, but rather as a unit designated with a category of the treatment-outcome complex[13] ($Tx_+Y_+;\ Tx_+Y_-;\ Tx_-,Y_-;\ Tx_-Y_+$), and use it as a ground for classification of patients. Importantly, the category of the treatment-outcome complex is logically related to the categories of "sensitive" and "spontaneous," which makes a framework for possible inferences regarding treatment response.

5.5 Inferences on treatment response

Table 2 schematizes the relationships between the treatment and outcome using the categories of factors forming a response of the individual to the treatment (St_+, St_-, Sp_+ and Sp_-).[11]

The left side of the table indicates a presence of the categories of "Spontaneous" and "Sensitive" in all possible variants of the treatment-outcome complex. The right side of the table demonstrates the outcome prospectively expected in individuals having all possible combinations of the categories "Spontaneous" and "Sensitive."

Summarizing this section, the major logically related concepts of analysis of the treatment response has been presented. Having assumed that *Two or more co-occurring events are related unless the contrary is proven,* we have introduced the concept of aggregation and further assume that the elements of the aggregation are related to each other. As emphasized above, this is the assumption, not assertion. This is only a

[ii] Spontaneous recovery, i.e., recovery (or remission, or intermission) without treatment, or regardless of treatment is well documented in numerous severe disorders including, but not limited to smallpox, plague, cholera, anthrax, typhus, Ebola virus disease (EVD), myocardial infarction, cancer, asthma, pernicious anemia, disseminated sclerosis, rheumatoid arthritis, schizophrenia, depression, etc. The list can be expanded indefinitely[11]

hypothesis, which is subject to scrutiny and comparison to other hypotheses, for instance, to the hypothesis of random co-occurrence of these elements.

Table 2. Relationships between the Categories of Treatment-Outcome Complex, "Sensitive" and "Spontaneous"

	Y-	
	St-	St+
Sp-	OK	No
Sp+	OK	No

	Y+	
	St-	St+
Sp-	No	OK
Sp+	No	OK

	St-	
	Y-	Y+
T-	OK	No
T+	OK	No

	St+	
	Y-	Y+
T-	No	OK
T+	No	OK

	Y-	
	St-	St+
Sp-	OK	No
Sp+	No	No

	Y+	
	St-	St+
Sp-	No	OK
Sp+	OK	OK

	St-	
	Y-	Y+
T-	OK	No
T+	No	OK

	St+	
	Y-	Y+
T-	No	OK
T+	No	OK

Analysis of the treatment effect requires establishing relationships between three components: outcome, treatment, and conditions, which requires a statistical approach. Having considered the treatment and outcome not separately, but rather as a unit, as a treatment – outcome complex, we obtained a possibility of exploring relationships between this complex and various conditions and clusters of conditions, under which the treatment was applied, and the outcome was developed. The concept of the treatment - outcome complex is logically related to the categories of the capacity of spontaneous recovery (Sp) and sensitivity to treatment (St) which provides a possibility for exploring the heterogeneity of the treatment response, via the inferences based on analysis of these logical relationships.

6. Heterogenous structure of population and identifying aggregations

6.1 Small number of observations

Fig 2 is to illustrate the approach towards the identification of the aggregations. If the order of the variables does not matter, the Fig. 2a can be re-written as Fig. 2b., where the subset of variables *B, C, E, F, H* is identical in both individuals I_1 and I_2, which corresponds to our definition of the aggregation.

Fig. 2. Aggregations and case-variable associations

Fig. 2a

Variable	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
I_1	0	0	1	0	0	0	0	1	0	1
I_2	1	0	1	1	0	0	1	1	1	0

Fig. 2b

Variable	*B*	*C*	*E*	*F*	*H*	*A*	*D*	*G*	*I*	*J*
I_1	0	1	0	0	1	0	0	0	0	1
I_2	0	1	0	0	1	1	1	1	1	0

Fig. 2c

Variable	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
I_1	0	0	1	0	0	0	0	1	0	1
I_2	1	0	1	1	0	0	1	1	1	0
I_3	1	1	1	0	1	1	0	1	0	1

Fig. 2d

Variable	*A*	*C*	*H*	*B*	*E*	*F*	*D*	*G*	*I*	*J*
I_1	0	1	1	0	0	0	0	0	0	1
I_2	1	1	1	0	0	0	1	1	1	0
I_3	1	1	1	1	1	1	0	0	0	1

If number of individuals increased (Fig. 2c), the number of possibly observable aggregations is increasing. If for instance, three individuals were observed (Fig. 2c),

there would be additionally observed (Fig. 2d) the subsets of variables repeated in two $(I_{1,2}V_{C,H,B,E,F};\ I_{1,3}V_{D,G,I,J})$, $(I_{2,3}V_{A,C,H})$, or more individuals $I_{1,2,3}V_{C,H}$.

Typically, a data matrix, which represents our model of the population, consists of the associations or aggregations of various size and form depending on a distribution of the variables within the population. They can be identified programmatically, or manually, using qualitative or quantitative criteria, based on formal or content-wise criteria.

With relatively small binary the matrixes, it can be done manually, based on the visual characteristics of the matrix and using a procedure of sorting. Large matrices require the use of a computational technology for applying a machine leaning process. Particularly, the deep learning procedure of nonnegative matrix factorization[14] appears adequate for this purpose.

Our objective here is to explain the logic of the analytical approach rather than discussing mathematical and programmatic aspects, as well as the effectiveness of various computational procedures for the selection of the aggregation. By this reason, we prefer visual analysis and manual operations to be described in our simulation.

6.1 Visualization of aggregations in multiple observations

The case-variable associations and aggregations objectively exist in the data matrix, but it is not possible to visualize all of them simultaneously. The order of variables, as well as the order of cases in the data matrix is flexible. The aggregations and associations not always can be directly observed in a single layout of the data matrix. In a relatively small data matrix, which is the case in many clinical trials, a visualization of the aggregations can be achieved via multiple targeted sorting of the variables and cases. Each sorting, however, while making visible one aggregation, can destroy a visibility of previous one.

Fig 3. Heterogeneity of treatment response in the trial population

A

Case	Tx/ts	Y/y	A	B	C	D	E	F	G	H	J	I	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Z	..
1	1	0	0	0	1	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	0	1	0	1	0	0	0	..
2	0	0	1	1	1	1	0	0	1	1	0	1	1	0	0	1	1	0	1	1	0	1	1	1	0	1	0	..
3	1	1	1	1	1	0	1	1	1	1	0	0	1	1	1	1	0	0	1	1	0	1	1	1	1	0	1	..
4	0	1	1	1	1	0	0	1	0	1	0	0	1	0	1	0	0	0	0	1	1	1	1	1	0	1	1	..
5	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	..
6	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	..
7	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	1	0	1	0	0	1	..
8	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	0	0	1	1	1	1	0	1	1	0	1	..
9	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	1	0	1	..
10	1	0	0	0	1	0	0	0	0	1	1	0	0	1	0	0	1	1	0	0	0	0	0	1	0	0	1	..
11	1	1	0	0	0	0	1	0	0	0	1	0	1	1	1	1	1	0	0	1	1	0	0	0	0	1	0	..
12	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	0	0	0	0	..
13	0	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	0	1	1	1	0	1	1	1	1	..
14	0	0	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	..
15	0	0	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	..
16	0	1	1	1	1	0	0	1	1	1	1	0	1	0	1	0	0	1	0	1	0	1	0	1	1	1	0	..
17	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	..
18	0	0	0	0	1	0	0	0	0	1	1	0	0	1	1	0	0	0	0	0	0	0	1	0	0	1	0	..
19	0	1	1	1	0	0	0	1	0	0	1	0	1	1	0	0	0	1	1	1	1	1	0	0	0	1	1	..
20	0	1	1	1	1	0	0	1	0	0	1	0	1	1	0	0	1	1	1	1	0	1	0	1	1	0	1	..
21	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	0	0	1	1	1	1	1	0	..
22	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	..
23	1	1	1	1	0	0	0	1	1	0	1	1	1	1	1	0	0	1	1	0	1	1	0	0	0	1	1	..
24	1	1	0	1	0	0	0	0	1	0	0	1	1	1	1	0	0	0	1	0	1	0	1	0	0	0	0	..
25	1	1	0	1	0	1	0	0	1	0	0	0	1	1	1	1	1	0	1	1	1	0	0	0	0	0	0	..
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..

B

Case	Tx/ts	Y/y	A	B	S	C	H	V	O	J	D	E	F	G	I	K	L	M	N	P	Q	R	T	U	W	X	Z	..
2	0	0	1	1	0	1	1	1	1	0	1	0	0	1	1	1	0	0	1	0	1	1	1	1	0	1	0	..
7	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	..
14	0	0	1	1	0	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	0	1	1	1	1	..
15	0	0	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1	0	0	1	0	0	1	1	1	..
18	0	0	0	0	0	1	1	1	1	0	0	0	0	1	0	1	1	0	0	1	0	0	0	0	0	0	1	..
1	1	0	0	0	0	1	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	..
5	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	..
6	1	0	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	..
10	1	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	..
19	0	1	1	1	1	0	0	0	0	1	0	0	1	0	0	1	1	0	0	1	1	1	1	0	0	1	1	..
9	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	1	0	1	..
17	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	..
4	0	1	1	1	1	1	1	1	0	0	0	0	1	0	0	1	0	1	0	1	0	0	0	1	1	1	1	..
13	0	1	1	1	1	1	1	0	1	1	1	0	1	1	1	0	1	1	0	1	0	1	1	0	1	1	1	..
16	0	1	1	1	0	1	1	0	1	1	0	0	1	1	0	1	0	1	0	1	0	1	0	1	1	0	1	..
20	0	1	1	1	0	1	0	1	1	1	0	0	1	0	0	1	1	0	0	1	1	0	0	1	1	1	1	..
21	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	..
22	1	1	1	1	0	1	0	0	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	0	1	1	1	..
3	1	1	1	1	0	1	1	0	0	0	1	1	1	0	1	1	1	1	0	1	1	1	1	0	1	1	1	..
8	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	1	1	1	0	0	1	1	1	0	1	1	..
12	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	..
24	1	1	0	1	1	0	0	0	0	0	0	0	1	1	1	1	1	0	0	1	0	1	0	0	1	0	0	..
11	1	1	0	0	1	0	0	0	1	1	0	1	0	0	0	1	1	1	1	0	0	1	0	0	0	1	0	..
23	1	1	1	1	1	0	0	0	0	1	0	0	1	1	1	1	1	1	0	1	0	1	1	0	0	1	1	..
25	1	1	0	1	1	0	0	0	1	0	1	0	0	1	0	1	0	1	1	1	0	1	1	0	0	0	0	..
..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..

Ag_1 Ag_2 Ag_3 Ag_4 Ag_5 Ag_6 Ag_7

Our exploration strategy is determined by our interest in heterogeneity of treatment response. To illustrate the idea, within the binary data matrix (Fig. 3A), we have selected the aggregations with the largest number of elements, which include the variables designating the treatment-outcome complex $(Tx\_Y\_)$, $(Tx_+Y\_)$, (Tx_Y_+) and (Tx_+Y_+). On Fig. 3B, the aggregations are color coded. Note that some of the aggregations may partially overlap. The list of the aggregations is not exhaustive. Smaller aggregations also can be informative. They can be relatively easy explored and might be considered for relevant analytical purposes and depending on research interest.

The treatment-outcome complexes (Tx_-Y_-), (Tx_+Y_-), (Tx_-Y_+) and (Tx_+Y_+) are indicative of the type of treatment response. The subset of individuals comprising the aggregation are assumed to have a type of treatment response designated by the treatment-outcome complex. Each group delimited by a single variant of the treatment-outcome complex is linked to one or several aggregations. Essentially, it represents a heterogeneous structure of the treatment response in this group, or more precisely, one level of this structure.

According to the logic presented on Fig. 2, the factors belonging to the categories "sensitive" and "spontaneous" are present, or can possibly be present, or should not be expected within the groups designated by the variants of the treatment-outcome complex (Table 3).

Table 3. Aggregations related to various types of treatment response

	Present	Possibly present	Impossible	Aggregations
Tx_-Y_-	Sp_-	St_+, St_-	Sp_+	Ag_1'
Tx_+Y_-	St_-Sp_-	Sp_-	Sp_+, St_+	Ag_1', Ag_2'
Tx_-Y_+	Sp_+	St_+, St_-	St_+, St_-	Ag_3', Ag_4' Ag_5'
Tx_+Y_+	$Sp_+ \cup St_+(Sp_+ \cap St_+)$	$St_- \cup Sp_-$	$St_- \cap St_-$	Ag_5', Ag_6', Ag_7'

The terms "present," "possibly present" and "impossible" mean that the individual with a relevant treatment - outcome complex respectively does have, possibly has, or does not have the factors belonging to a relevant category. Since the aggregations are determined on a subset of "known" variables, there is a possibility that factors of treatment response can belong to the subset of the "unknown" variables and not necessarily (all of them or part of them) should be included into the associated aggregations. It does set some limitations to the exploration. On the other hand, it is a stimulus to increasing a subset of "known" variables, which is an integrative part of exploration strategy.

Nevertheless, under the assumption *Two or more co-occurring events are related unless the contrary is proven,* each of the explored aggregations, can be interpreted as an element of the heterogeneous structure of treatment response and as a hypothesis of the relatedness of the relevant type of the treatment response with a cluster of variables in the subset of individuals defining the aggregation (with a necessary reservation, that relatedness not necessarily means causally).

7. Mapping conditions

Mapping the factors belonging to the categories "sensitive" and "spontaneous" to the variables describing the population is a step toward revealing a content and causal aspect of the heterogeneity of treatment response. The structure presented on Fig.3B and Table 3 can be described in a greater detail as follows.

$$Ag_1 = I_{2,7,14,15,18}V_{(Tx_-Y_-),C_+,F_-,H_+,K_+,O_+,S_+,V_+};$$
$$Ag_2 = I_{1,5,6,10}V_{(Tx_+Y_-),H_+,I_+,J_+,O_+,V_+};$$
$$Ag_3 = I_{9,7,19}V_{(Tx_-Y_+)C_-,D_-,E_-,G_-,H_-,I_-,M_-,O_-,V_-,U_-};$$
$$Ag_4 = I_{4,13,16,20}V_{(Tx_-Y_+)A_+,B_+,C_+,S_-,W_+,M_+,T_+,W_+,Z_+};$$
$$Ag_5 = I_{3,8,21,22}V_{(Tx_+Y_+),A_+,C_+,E_+F_+,G_+,L_+,M_+,R_+,Q_+,W_+,Z_+};$$
$$Ag_6 = I_{3,8,12,24}V_{(Tx_+Y_+),J_-,O_-,X_-};$$
$$Ag_7' = I_{11,12,23,24,25}V_{(Tx_+Y_+)C_-,H_-,O_-,V_-,W_-};$$

where I_j is a case, and $V_i -$ variable.

A procedure of mapping the factors belonging to the categories "sensitive" and "spontaneous" is based on the logic of pairwise comparison of individuals or comparing

groups of individuals. The comparisons can be made within and across the subpopulations delimited by the treatment-outcome complexes.

7.1 Pairwise comparisons

Pairwise comparisons of individuals allow for mapping the properties "Spontaneous" and "Sensitive," i.e., ascribing a status of potential carriers of these properties to some subsets of variables. There are 10 unique types of the pairs, each with unique respective logic of inferences regarding the mapping.

Let us consider the following set of individuals:

$$I_1 = (1,0, A_1 \dots, \alpha_1 \dots)$$
$$I_2 = (0,0, A_2 \dots, \alpha_2 \dots)$$
$$I_3 = (1,0, A_3 \dots, \alpha_3 \dots)$$
$$I_4 = (0,1, A_4 \dots, \alpha_4 \dots)$$
$$I_5 = (1,0, A_5 \dots, \alpha_5 \dots)$$
$$I_6 = (0,0, A_6 \dots, \alpha_6 \dots)$$
$$I_7 = (1,0, A_7 \dots, \alpha_7 \dots)$$
$$I_8 = (0,1, A_8 \dots, \alpha_8 \dots)$$
$$\dots$$

Table 4 shows all 32 possible types of comparisons (C_i) by the treatment-outcome complex. Ten unique types of comparisons are highlighted, and the remaining are types are redundant.

Table 4. Pairwise Comparisons

		I1 T+ Y+	I2 T- Y-	I3 T+ Y-	I4 T- Y+
I1	T+ Y+	N/A	C1	C2	C3
I2	T- Y-	C1	N/A	C4	C5
I3	T+ Y-	C2	C4	N/A	C6
I4	T- Y+	C3	C5	C6	N/A
I5	T+ Y+	C7	C1	C2	C3
I6	T- Y-	C1	C8	C4	C5
I7	T+ Y-	C2	C4	C9	C6
I8	T- Y+	C3	C5	C6	C10

Unique types of comparisons highlighted in Table 4 are shown separately in Table 5, demonstrating similarity-dissimilarity of the compared treatment-outcome complexes.

Table 5. Unique Types of Pairwise Comparisons

C1	T- Y-	vs.	T+ Y+
C2	T+ Y-	vs.	T+ Y+
C3	T- Y+	vs.	T+ Y+
C4	T+ Y-	vs.	T- Y-
C5	T- Y+	vs.	T- Y-
C6	T- Y+	vs.	T+ Y-
C7	T+ Y+	vs.	T+ Y+
C8	T- Y-	vs.	T- Y-
C9	T+ Y-	vs.	T+ Y-
C10	T- Y+	vs.	T- Y+

7.2 Logic of mapping via pairwise comparison

In a small group of observations, we can compare them pairwise, explore similarity and dissimilarities in their descriptions, and link them to the treatment-outcome complexes, which in turn enable mapping the properties associated with the categories "Spontaneous" and "Sensitive" among the sets of the variables describing the individuals.

Within the binary model it can be done via eliminating the variables that cannot belong to the relevant category. For instance, in the pair of individuals I_j and I_{j+1}, the treatment-outcome complex $Tx\_Y\_$ indicates, that the individual I_j does not have the property designated by the category "Spontaneous." It means that in individual I_{j+1}, the

subset of variables describing "similar" to those in the individual I_j, also does not contain the variables belonging to the category "Spontaneous," but among the "dissimilar" variables these factors can be, but not necessarily, present.

Each one of the 10 possible unique pairwise comparisons has a specific pathway for mapping the properties Sp and St. The algorithm covering the mapping in all these variants should utilize the logic, examples of which (C1 and C2) are shown below.

As a reminder, all logical operations are made under the assumption that two or more co-occurring events are related *unless the contrary is proven.* Each individual (case) is described with a set of binary variables including treatment, outcome and a set of known (I_i) and unknown (α_i) conditions.

Example 1 (C2)

$$\begin{Bmatrix} (Tx_+Y_+), I_1 \\ (Tx_+Y_-), I_2 \end{Bmatrix};$$

where $I_i = (A_i, B_i, C_i \ldots; \alpha_i)$.

In these individuals, the treatment was identical, but the outcomes were different. Comparing these individuals does not allow any reasonable guess regarding the effect of treatment. The only possible explanation of difference in the outcomes could be the difference in the response of the individuals to the treatment.

Each member of the pair has a subset of "known" variables identical ("similar") to the counterpart
$I_1' = I_3'$; such that $I_1' + I_3' = 0{,}2$; and a subset of complementary ("dissimilar") conditions

$$I_1'' \neg I_3'';$$
such that $I_1'' + I_3'' = 1.$
$$I_1 = (Tx_+Y_+), I_1', I_1'', \alpha_1;$$
$$I_3 = (Tx_+Y_-), I_3', I_3'' \alpha_3.$$

Since the treatment was identical in both individuals, the difference in the outcome $(Y - y)$ can be attributed only to the difference in the properties of the individuals and conditions, under which the treatments have been conducted $(I_1 - I_3)$. Since $(Tx_+Y_-) \not\subset Sp_+$, and $(Tx_+Y_-) \not\subset St_+$

Then

$$I_1 = I_1', I_1'', \alpha_1;$$
$$I_3 = I_3', I_3'', \alpha_3;$$
$$I_1', I_3', I_3'', \alpha_2 \not\subset Sp_+;$$
$$I_1', I_3', I_3'', \alpha_3 \not\subset St_+.$$
$$I_1'', \alpha_1 \sqsubset \ni [Sp \cup St \cup (Sp \cup St)];$$ [15]

or

$$I_1'' \diamond \ni [Sp_+ \cup St_+ \cup (Sp_+ \cap St_+)];$$

i.e., at least one or more variables exerting the property Sp or St or both can possibly be located to the segment I_1''. Neither of variables describing I_3, as well as I_1' represent these properties.

Example 2 (C1)

$$\begin{Bmatrix} (Tx_+Y_+), I_1 \\ (Tx\_Y\_), I_2, \end{Bmatrix};$$

where $I_i = (A_i, B_i, C_i \ldots; \alpha_i)$.

Under the accepted assumptions, the difference in the outcomes can be explained either by the difference in treatment, or the different conditions. The subsets of "similar" variables can be eliminated from the mapping.

Similar to C2,

$I_1 = I_1'^a, I_1''^a, \alpha_1;$
$I_2 = I_2'^a, I_2''^a, \alpha_2;$
$I_1'^a = I_2'^a;$
$I_1''^a \neg I_2''^a;$
$I_2, \alpha_2 \not\subset Sp_+;$
$I_1'^a, I_2'^a, I_2''^a, \alpha_2 \not\subset Sp_+;$
$I_1''^a, \alpha_1 \square \ni [Sp_+ \cup St_+ \cup (Sp_+ \cap St_+)];$
$I_1''^a \Diamond \ni [Sp_+ \cup St_+ \cup (Sp_+ \cap St_+)];$
$I_1'^a, I_2'^a, I_2''^a, \alpha_2 \Diamond \ni St_+.$

i.e., the set I_2 as well as the subset $I_1'^a$ do not include a variable exerting the capacity for spontaneous recovery Sp_+; the variable presenting the property of sensitivity St_+ to the treatment Tx_+ can reside both in the set I_2, α_2 and I_2, α_2.

7.3 Mapping in multiple observations

Let us consider the set of individuals $(\boldsymbol{Tx}, \boldsymbol{Y}), I_J, \alpha_j$, where $(\boldsymbol{Tx}, \boldsymbol{Y})$ is the variable of the treatment-outcome complex, I_J – a set of known variables, and α_j - a set of unobserved or unknown variables describing individual I_J:

$$(\boldsymbol{Tx}, \boldsymbol{Y}), I_J, \alpha_j = \left\{ \begin{array}{c} (Tx_+Y_-), I_k, \alpha_k \\ (Tx_+Y_-), I_{k+1}, \alpha_{k+1} \\ \dots \\ (Tx_-Y_-), I_{k+2}, \alpha_{k+2} \\ (Tx_-Y_-), I_{k+3}, \alpha_{k+3} \\ \dots \\ (Tx_-Y_+), I_l, \alpha_l \\ \dots \\ (Tx_+Y_+), I_m, \alpha_m \\ \dots \end{array} \right\}.$$

The variables potentially presenting the categories “Sensitive” and “Spontaneous” can be found via elimination of the variables, which are known that they are not related to these categories. For example, a description of individuals I_k and I_{k+1} who have the treatment-outcome complex Tx_+Y_- includes neither property St_+, nor Sp_+, variables. The subset of known variables potentially related to these properties, can be found via eliminating the variables of the subset $I_k + I_{k+1}$ from the set of variables describing the population:

$$S_1 = (\boldsymbol{Tx}, \boldsymbol{Y})(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \dots, \boldsymbol{Z})\alpha_j - (I_k + I_{k+1}).$$

Apparently, S_1 includes the variables potentially presenting St_+, Sp_+, or both. Considering individuals with the treatment outcome complex Tx_-Y_-, e.g., I_{k+2} and I_{k+3}, a description of which does not include the property Sp_+, makes possible mapping out the property (Sp_+) to the subset

$$S_2 = (\boldsymbol{Tx}, \boldsymbol{Y})\big(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \dots, \boldsymbol{Z}\big)\alpha_j - (I_{k+2} + I_{k+3}\big);$$

Considering both the individuals with both Tx_+Y_- and Tx_-Y_- makes possible narrowing down the set of variables possibly related to Sp_+:

$$S_3 = (\boldsymbol{Tx}, \boldsymbol{Y})\big(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \dots, \boldsymbol{Z}\big)\alpha_j - (I_k + I_{k+1} + I_{k+2} + I_{k+3}\big).$$

The variables from the sets S_1, S_2 and S_3, can be unevenly distributed among the individuals creating either associations and/or aggregations in some subsets of the individuals, or being single entities in other cases, and having some of individuals or subsets of individuals spared.

8. Interpretation of aggregations

Until this point, we considered the sets of variables constituting aggregations primarily from a quantitative angle. Such a set can be a cluster of symbols not having an interpretable meaning, but also it can have an identifiable connotation and to be a

readable "word," or to be a combination of meaningful and "meaningless" subsets. Revealing the meaning and content of the aggregations may lead to reducing the dimensionality of the description of the population. Besides, it might facilitate understanding the structure of heterogeneity and mechanisms of treatment response.

The process of the interpretation of the "words" in many instances is similar, but not identical, to the process we described for the interpretation of principle components. [16] There is a similarity of the operations in the frame of heuristics, [iii] as well as psychological and epistemological aspects of the interpretation. The major difference is that within the aggregation, the "word" is identical in all individuals, i.e., a correlation between the "letters" across the individuals is 1.0. Therefore, unlike the principle components, the content of aggregations does not present contrasting positive and negative "images." The meaning of the "word" should be surmised, revealed, directly from the combination of the events, things, or properties denoted by the "letters" comprising the "word." A further investigation of this insufficiently studied area is critical for the progress in the "human-machine" interaction in the process of analysis of complex relationships and mechanisms of the treatment response and treatment effect.

9. Constructing hypotheses via aggregations

9.1 Deterministic hypothesis

Above, we have described a heuristic to identify the variables potentially indicative of the sensitivity to the treatment and capacity for spontaneous recovery - the properties which determine the treatment response. This identification is based on analysis of aggregations, members of which assumed related to each other. The aggregation including a specific treatment-outcome complex and a subset of variables (conditions, properties) is a diagrammatic or symbolic expression for a data driven hypothesis about the relatedness of the outcome of the specified treatment with the subset of "generating conditions" in the subset of individuals comprising the aggregation.

Our explorations stem from logic of Francis Bacon,[17] who believed that a cause underlying a phenomenon should be deduced by elimination of the factors not matching the occurrence of the phenomenon, and inductive reasoning should account for the agreement, difference, and concomitant variation of the factors. For example, "…if an army is successful when commanded by Essex, and not successful when not commanded by Essex: and when it is more or less successful according to the degree of involvement of Essex as its commander, then it is scientifically reasonable to say that being commanded by Essex is causally related to the army's success."[18] Importantly, Bacon emphasized that this inference is not a final conclusion. Rather this is only a hypothesis, which must be scrutinized and compared to other hypotheses.

Using this logic, in our analysis of aggregations the hypothesis is deduced from a series of eliminations of the conditions not coinciding with the outcome of interest, and the inclusion of the remaining conditions based on the assumption that *two or more co-occurring events are related unless the contrary is proven*, i.e., essentially, the generated hypothesis is deterministic. The existence of the subset of unknown or unobserved variables α_i implies that under any circumstances some share of uncertainty remains despite the apparent determinism.

[iii] The major elements of the strategy of the interpretation of principal components include: 1) presenting variables in a comparative way, e.g., "younger" vs. "older" instead of "age;" 2) using mathematical and logical inversion, the development of two contrasting (positive and negative), but still logically equivalent representations of each retained component; 3) sorting absolute values of loadings in a descending order; 4) contrasting positive and negative characteristics [16]

9.2 Quantitative characteristics of aggregations and alternative hypothesis

Thus, the aggregation represents a deterministic hypothesis of the relatedness between its elements, and this hypothesis can be and should be contrasted with an alternative hypothesis of random association of the elements. Analysis of the heterogeneity of treatment response in this dualistic manner raises numerous mathematical, statistical and computational problems, which are beyond the objectives of this paper. We limit ourselves to a description of a combination of indices including the size of the aggregation, the probability of random gathering of the elements of the aggregation, and "density" of the aggregation, which in our opinion could be helpful and informative for the analysis, and also, which opens a field for classic and modern statistical and computational methods.

The alternative to the deterministic hypothesis above is the hypothesis of random gathering of the elements comprising the association/aggregation:

$$Pr(Ag_1') = [Pr(Tx_+Y_+) * Pr(C_-) * Pr(H_+) * Pr(O_+) * Pr(S_+) * Pr(V_+)]^{|I_{3,8,12}|}.$$

It can lead to computing a size (Sz) of the aggregation expected under a condition of random gathering of the elements comprising the association/aggregation:

$$Sz_{Ag_1(exp)} = n_{Ag_1'} Pr(Ag_1');$$

where $n_{Ag_1'}$ – a size of the matrix.

An observed size of the aggregation is a product of the cardinalities of the set of variables and set of individuals comprising the association/aggregation:

$$Sz_{Ag_1(obs)} = \left|I_{2,7,14,15,18}\right| * \left|\boldsymbol{V}_{(Tx\_Y\_),C_+,F\_,H_+,K_+,O_+,S_+,V_+}\right|.$$

A comparison of the expected and observed size of the aggregation can be a basis for examining the hypothesis of random gathering of the elements comprising the association/aggregation.

To quantitatively assess on average the relatedness of the elements within the aggregation in the frame of a random hypothesis, the index of "density" (Dn) of the aggregation can be suggested, which is

$$Dn = \frac{Pr(Ag_1')}{S_{Ag_1(obs)}}.$$

Also, it can be used to assess separately any part of the aggregation or any pair of the elements.

Apparently, deriving from these indices, it is possible to obtain a comprehensive statistical description of the heterogeneous population, which can be further analyzed using classic and modern statistical and computational methods.

10. Relationships between treatment effect and treatment response

An individual outcome is a result of an individual response of a subject to the treatment. The treatment effect is the integration of individual responses to the treatment. While being generalized and summarized, i.e., abstracted from the individuality, the individual and/or small group outcomes developed in response to the treatment lay the ground for the concept of the treatment effect as it is defined in a clinical trial, and operationalized as the indices of absolute and relative risk reduction. Therefore, the relationships between treatment effect and treatment response are not symmetrical. The indices of treatment effect can be *directly* computed from the data on the treatment response, while the opposite operation cannot be performed *directly*.

The heterogeneity of treatment response and heterogeneity of treatment effect, therefore, are two different ways of looking at the diversity of the ways of interaction of

the individuals comprising a population with the treatments, both necessary to contribute to solving the problem of individualization of treatment.

10.1 From treatment response to treatment effect

The heterogeneous structure of a population exposed to treatment Tx_- and Tx_+ includes the following subsets.

$$S_{Tx_-Y_-} = \{Ag_{Tx_-Y_-} | (Tx_-Y_-) \in Ag_{Tx_-Y_-}\};$$
$$S_{Tx_-Y_+} = \{Ag_{Tx_-Y_+} | (Tx_-Y_+) \in Ag_{Tx_-Y_+}\};$$
$$S_{Tx_+Y_-} = \{Ag_{Tx_+Y_-} | (Tx_+Y_-) \in Ag_{Tx_+Y_-}\};$$
$$S_{Tx_+Y_+} = \{Ag_{Tx_+Y_+} | (Tx_+Y_+) \in Ag_{Tx_+Y_+}\};$$

where $S_{Tx_*Y_*}$ is a set of aggregations including the variable Tx_*Y_* and describing the population with the following heterogeneous structure.

$$S_{Tx_-Y_-} = Ag_i, \dots, Ag_{i+a};$$
$$S_{Tx_-Y_+} = Ag_k, \dots, Ag_{k+b};$$
$$S_{Tx_+Y_-} = Ag_l, \dots, Ag_{l+c};$$
$$S_{Tx_+Y_+} = Ag_m, \dots, Ag_{m+d}.$$

The size of the sets $S_{Tx_*Y_*}$ is

$$|S_{Tx_-Y_-}| = |Ag_i + \cdots + Ag_{i+a}|;$$
$$|S_{Tx_-Y_+}| = |Ag_k + \cdots + Ag_{k+b|};$$
$$|S_{Tx_+Y_-|} = |Ag_l + \cdots + Ag_{l+c}|;$$
$$|S_{Tx_+Y_+}| = |Ag_m + \cdots + Ag_{m+d}|.$$

The index of absolute risk reduction (ARR), which is one of the major indicators of the treatment effect used in clinical trials, is

$$ARR = Pr(Y_-^c) - Pr(Y_-^a) = \frac{|S_{Tx_-Y_-}|}{|S_{Tx_-Y_-}| + |S_{Tx_-Y_+}|} - \frac{|S_{Tx_+Y_-}|}{|S_{Tx_+Y_-}| + |S_{Tx_+Y_+}|};$$

i.e., the data on the treatment response can be directly "translated" into the language of the treatment effect. It is true for the entire population if the description of the heterogenous structure of the population is complete. Also, it is true for a segment of the population delimited by a certain condition (e.g., subgroup analysis).

10.2 From treatment effect to treatment response

Contrary, the "back translation" from the language of treatment effect to treatment response cannot be performed *directly* from the indices of the treatment effect and its variability: parts cannot be *directly* identified from the sum. They can be identified empirically, or such identification requires a chain of analytical steps including using complex classic and modern methods. In the cited article, using a "training set - validation set" design and complex modern methods, Lamont, et al. (2018), identified the predicted individual outcomes in a randomized clinical trial. The individual prediction of the treatment effect is one (but not only) of important steps in the individualization of treatment, and this study, undoubtedly, is a step in a technological progress in this area.

It should be noted, however, that since 1960s, numerous studies were conducted using a "training set - validation set" design and classic statistical methods to identify the individual outcome in the context of computational diagnostics, automatic classification, image recognition, etc. It became clear that although the individual effect (diagnosis, classification, image, etc.) can be satisfactory re-produced in the validation

set, the extrapolation of the results to a "real world" population ("generalization" of results of a clinical trial) is problematic because a distribution of the variables designating treatment, outcomes and conditions in these populations may have substantial differences. The response to the treatment in the non-experimental population is not necessarily correspondent to the treatment effect defined in the experimental population.

10.2 Treatment effect, treatment response, individualized treatment

The problem of predicting the treatment effect in the non-experimental population is substantially mitigated with propensity scoring, [19] or propensity score matching (PSM). Note, however, that the propensity scores could be computed, and the predictions could be made based on analysis of the distributions in the *entire* experimental (trial, training set) and *entire* non-experimental populations, and the effect can be predicted for the "collective" rather than individuals. The identification of the predicted individual effect will take the efforts and have the limitations described above.

In the "real world," the decision on the individualized treatment, however, are to be made when the data on the entire population are not necessarily available. Rather these decisions are being made in the populations with unknown distributions of variables, in the groups with incomplete data, small samples, and in non-randomly selected individuals, under conditions that the rules established on the trial population or a training set simply cannot be effective, or even applicable.

Intuitively, predictions on the *treatment response* for individual cases and small groups can be made on a basis of "similarity" of the relevant cases and small groups in the experimental and non-experimental populations, and not necessarily with a reference to the *entirety* of these populations.

At a first glance, it may appear paradoxical. However, it should be argued that the trial population and the non-experimental population, in which the predictions are to be made, are the samples of a general population of the patients. The aggregation identified in the trial is a group of elements *non-randomly* related to each other. It is reasonable to suggest then, that these non-random relationships may exist in the general population and, therefore, they may exist in the non-experimental population, in spite the natural differences in the distributions between the samples. The interpretation of the "similar" clusters of the conditions can provide additional content-wise support for the inference.

If this suggestion is correct, the predictions can be made for the cases in non-experimental populations, in which there are combinations of conditions are identical to those detected in the aggregations of the experimental (trial, training set) population.

The methods of Bayes statistics with a reference to the estimates of the incidence of relevant factors in the general population could be potentially effective in quantitative characterization of these predictions, and modern methods of stochastic modeling can be helpful in mitigating possible power problems.

The predictions of this kind, however, cannot be made for the cases, where such identical clusters of conditions were not observed. They can be possible only after additional variables become "known" and potentially related to the outcome. With this substantial limitation, analysis of the treatment response can be the instrument of choice for the assessment of factors and prediction of the results of treatment in single cases and small groups, which is a routine objective in clinical practice. Also, this approach can apply for safety monitoring of clinical trials[12] and for analysis of the factors

influencing the heterogeneity of treatment response in a clinical trial population as far as the trial is complete. Generally, it can be applicable in populations with incomplete and non-systematically collected data, which is the case in a numerous "real world" clinical and social situations.

Discussion

The treatment effect and treatment response are the interrelated, but not identical aspects of the treatment process. While the treatment effect is thoroughly studied, phenomena and factors of the response of subjects to treatment are not systematically studied yet.

As noted above, people react to treatment differently: while being exposed to the same treatment, various individuals have different outcomes, and while being exposed to two different treatments, various individuals can have either the same outcome, or different outcomes. Under the term of *treatment response,* we mean the way the individuals react to treatment qualitatively and quantitatively. At a population level, any "real world" population exposed to treatment or clinical trial population is inherently heterogeneous with regard to the treatment response, meaning qualitative and quantitative diversity of the ways of reacting to treatment, and factors determining this response among individuals within a population.

Statistical analysis of the results of a randomized clinical trial, which is a "gold standard" for the assessment of the treatment effect, derives from a set of assumptions including, but not limited to the following.

1. *Subjects are anonymous and interchangeable*.
2. *Two or more events co-occur by chance unless the contrary is proven.*
3. *Numerous subjects are required for making valid inferences.*

Our analysis of the heterogeneity of treatment response derives from a set of assumptions which are not acceptable for statistics but are natural and necessary for qualitative explorations and inferences, and particularly, for clinical thinking.

I. *Each subject is a unique individual.*
II. *Two or more co-occurring events are related unless the contrary is proven.*
III. *Valid inferences potentially can be made from single cases and small number of cases.*

Following these assumptions, we have introduced a concept of aggregation as a central concept of our approach. The aggregation presents the association of the elements of the aggregation, but unlike a correlation, it refers not to the entire population, but rather to a directly observed cluster of variables having co-occurred in a subset of individuals.

To avoid any misunderstanding, it is necessary to emphasize one more time that while the assumption *Each subject is a unique individual* should not raise any objective out of the boundaries of formal statistical operations, the assumptions *Two or more co-occurring events are related unless the contrary is proven* and *Valid inferences potentially can be made from single cases and small number of cases* are not assertions, they are the assumptions in a genuine meaning of this term. They designate a starting position of analysis, only a hypothesis, which must be scrutinized and compared to other hypotheses in the process of analysis.

For an unprejudiced investigator, the assumption of relatedness in the concept of aggregation is as legitimate as the assumption of random co-occurrence in traditional statistical analysis. This is a starting position and the framework, in which the phenomena are considered. The hypothesis of relatedness of the elements of the aggregations in analysis of the treatment response subject to contrasting with the

hypothesis of their random co-occurrence. In turn, the hypothesis of random co-occurrence in the assessment of treatment effect subject to contrasting with the hypothesis of the relatedness between subjects, events and relationships in statistical hypotheses testing.

Since the elements assumed related, the aggregation including the variables of treatment, outcome, and conditions can be interpreted as a data driven deterministic (in a sense of "deterministic probability"[20]) hypothesis of the relatedness of the treatment, the outcome, and the conditions *in the given subset of individuals*. The hypothesis of relatedness is subject to contrasting with a hypothesis of random co-occurrence of the elements of the aggregation, which can be examined by the rules of statistics.

Another important concept is the treatment-outcome complex, which considers treatment and outcome not separately but rather as a unit. In our analysis, each patient is classified by four possible types of a treatment-outcome complex ($Tx_+Y_+, Tx_+Y_-, Tx_-Y_+, Tx_-Y_-$). There are two important ramifications.

Within each of the classes, more than one aggregation can be observed. It indicates hypothetically that among individuals exposed to the same treatment the same outcome might be related to different subsets of conditions or causative factors. It is true for both negative and positive outcomes. It corresponds with historical records, epidemiological observations, experimental (trials) data and clinical experience: the subpopulation with a negative outcome includes the subsets, in which the death caused by diverse agents (All Cause Deaths and the death caused by the target disorder). A subpopulation with a positive outcome consists of patients who recovered owing to the treatment, and those who recovered spontaneously. Each of these groups also is heterogeneous.[11] This type of heterogeneity is not a primary focus of studies on the treatment effect. Just the opposite, there are substantial difficulties with accounting for these factors in the methodology and interpretation of randomized clinical trials. Contrary, a distinctive capacity of the treatment response approach is a possibility of exploring *directly* this aspect of heterogeneity, which makes it suitable for the use in clinical settings, in exploring heterogeneity of populations including clinical trials and observational data, for monitoring safety of clinical trials, etc.

The concept of a treatment-outcome complex is logically related to the categories "sensitivity to treatment" (St), and "capacity for spontaneous recovery" (Sp). Together, they create a logical structure that makes possible

- to infer, deriving from the data on the treatment-outcome complex, the presence or absence of the factors covered by the categories St and Sp in an individual patient;
- via comparing individuals, to map the variables potentially indicating the factors included in the categories St and Sp.
- from the data on the presence or absence of the categories St and Sp, or on the factors belonging to these categories, to predict the response to the treatment in an individual patient, or in a group of patients.

Thus, the heterogeneity of treatment response is represented by a set of aggregations. Deriving from the assumptions I, II, and III, the aggregation represents a deterministic hypothesis of the relatedness between treatment, outcome and conditions in the subset of individuals comprising the aggregation.

An alternative to the null deterministic hypothesis of the relatedness of the outcome, treatment, and conditions should be a hypothesis of random gathering of the elements of the aggregation.

The major characteristics of the approaches are compiled in Table 6.

Table 6. The major conceptual characteristics of heterogeneity of treatment effect and treatment response

Concept	Treatment effect	Treatment response
Experimental and observational study objectives	Hypothesis testing	Exploration Hypotheses generation
Assumptions	• *Subjects are anonymous and interchangeable*. • *Two or more events co-occur by chance unless the contrary is proven.* • *Numerous subjects are required for making valid inferences.*	• *Each subject is a unique individual.* • *Two or more co-occurring events are related unless the contrary is proven.* • *Valid inferences potentially can be made from single cases and small number of cases.*
Categories of analysis	Risk	Sensitivity to treatment Capacity for spontaneous recovery
Hypothesis	Theory driven	Data driven
Concept of relatedness	Correlation, i.e., (in a broad sense) any statistical association	Aggregation as a subset of individuals (cases) with an identical set of variables
Measure of relatedness	Absolute and relative risk reduction	Co-occurrence of indices of treatment, outcome and conditions
Categories of heterogeneity	Variability of risk	Individuals and groups with diverse types of treatment response; the same response under different treatment; the same response different conditions
Logic	Neyman-Rubin Causal Model	Agreement, difference, and concomitant variation (Francis Bacon)
Vector of analysis of heterogeneity	From risk for the entire population, to identifying groups, subgroups, and individuals at risk of negative outcome	From identifying individuals with positive and negative outcome, to identifying subgroups and groups with distinct clusters of conditions, to identifying risk for the entire population
Null hypothesis	Random association between treatment, outcome and conditions	Relatedness of treatment, outcome and conditions
Alternative hypothesis	Relatedness of treatment, outcome and conditions	Random association between treatment, outcome and conditions
Uncertainty	Measured in accordance with concepts and rules of statistics.	Accounted for as a fact of limited knowledge, assuming an existence of unknown variables. Assessed statistically in the frame of the alternative hypothesis of random co-occurring the events.
Outcome prediction	Prediction at the level of the entire population and groups. Identifying individual predicted outcome is possible via multivariable analyses	Can be applied to individuals, subgroups and groups of the non-experimental population, in which the clusters of the conditions identical to those in the experimental population are identified.
Extrapolation of results from experimental to	Problematic because of the difference of the	Intuitively, applicable to individuals and groups with combinations of

non-experimental population	distribution of treatment, outcome and conditions in experimental and not-experimental population	conditions identical to those in the aggregations. The outcome in the entities with different conditions cannot be predicted

Thus, analysis of heterogeneity of the population exposed to treatment can be and should be performed from two positions, the treatment effect and treatment response, exploring different aspects of the relationships between the individual, treatment, and conditions. Applying these approaches in parallel, we obtain two not necessarily equal images of these relationships of treatment, conditions and outcome. One of the images is based on the theory driven hypothesis in the context of analysis of treatment effect. Another one is guided by the data driven hypothesis developed in the context of analysis of treatment response, which, upon the interpretation, can give a rise to a new theory-based hypothesis.

The estimates of the treatment effect and treatment response can completely correspond at the population level, as well as at the level of individual level. However, at the intermediate, i.e., small group and group level, the ways and the targets of the analyses are different. Understanding and reconciling the differences defines a content and becomes both the major problem and the major objective of further analysis of heterogeneity of the population exposed to treatment, which should facilitate exploring the mechanisms and to improve prediction of treatment response and treatment effect.

Literature

[1] Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. Velentgas P, Dreyer NA, Nourjah P, et al., editors. Rockville (MD): Agency for Healthcare Research and Quality (US); 2013 Jan.

[2] Johnson, Richard A; Wichern, Dean W (2001). Applied Multivariate Statistical Analysis. Pearson. p. 111-112.

[3] Sekhon, J.S. The Neyman-Rubin Model of Causal Inference and Estimation via Matching Method. The Oxford Handbook of Political Methodology (2007). https://pdfs.semanticscholar.org/cc61/64063760d265513be4462a65d5a607e20992.pdf (Accessed 08.24.2018)

[4] Kent DM, Rothwell PM, Ioannidis JP, Altman DG, Hayward RA. Assessing and reporting heterogeneity in treatment effects in clinical trials: a proposal. Trials. 2010 Aug 12; 11:85. doi: 10.1186/1745-6215-11-85.

[5] Subgroup analysis. Since Direct. https://www.sciencedirect.com/topics/medicine-and-dentistry/subgroup-analysis (Accessed 3/30/2018)

[6] Sies, A., Demyttenaere, K., Van Mechelen, I. (in press). Studying treatment-effect heterogeneity in precision medicine through induced subgroups. Journal of Biopharmaceutical Statistics. https://doi.org/10.1080/10543406.2019.1579220 (Accessed 9/26/2019)

[7] Ran Chen, Hanzhong Liu. Heterogeneous Treatment Effect Estimation through Deep Learning. arXiv:1810.11010v1 [stat.ME] 25 Oct 2018. https://arxiv.org/pdf/1810.11010.pdf (Accessed 3/2/2019)

[8]Lamont A, Lyons MD, Jaki T, Stuart E, Feaster DJ, Tharmaratnam K, Oberski D, Ishwaran H, Wilson DK, Van Horn ML. Identification of predicted individual treatment effects in randomized clinical trials. Stat Methods Med Res. 2018 Jan; 27(1):142-157.

[9] Yu Xie. Population heterogeneity and causal inference. PNAS April 16, 2013 110 (16) 6262-6268

[10] Mises, Richard von. Probability, Statistics and Truth. 2nd Revised ed. New York: Dover, 1981 [1957] p. 11

[11] Sverdlov, L.S. 2015. Increased Chance for Recovery vs. Risk Reduction. In: Proceedings of the American Statistical Association, Section on Health Policy Statistics. Alexandria, VA, pp.3316-3330

[12]Sverdlov, L.S. 2018.Assessment of treatment effects in single cases and small groups: assumptions, logic, algorithm, computations. JSM Proceedings, Biopharmaceutical Section. Alexandria, VA: American Statistical Association. 2491-2506.

[13] Sverdlov, L.S. 2011. Treatment-Outcome Complex and Analysis of Observational Data. In: Proceedings of the American Statistical Association, Section on Statistics in Epidemiology. Alexandria, VA, pp. 5415–5430.

[14] A Practical Introduction to NMF (nonnegative matrix factorization) https://mlexplained.com/2017/12/28/a-practical-introduction-to-nmf-nonnegative-matrix-factorization/ (Accessed 8.29.2019)

[15] Modal Logic. Stanford Encyclopedia of Philosophy. Principal Editor: Edward N. Zalta. *First published Tue Feb 29, 2000; substantive revision Sat Sep 8, 2018.* https://plato.stanford.edu/entries/logic-modal/ Assessed 05/19/2019

[16] Sverdlov, L.S, Interpretation of Principal Components: Methodological Aspect. 2005 Proceedings of the American Statistical Association, Section on Nonparametric Statistics, Alexandria, VA

[17] Francis Bacon. New Organon, English translation, based on the 1863 translation of James Spedding, Robert Leslie Ellis, and Douglas Denon Heath. http://www.constitution.org/bacon/nov_org.htm (Accessed 9/22/2018).

[18] Hesse, M. B. (1964), "Francis Bacon's Philosophy of Science", in A Critical History of Western Philosophy, ed. D. J. O'Connor, New York, pp. 141—52

[19] Paul R. Rosenbaum; Donald B. Rubin. Score in Observational Studies for Causal Effects. Biometrika, Vol. 70, No. 1. (Apr., 1983), pp. 41-55

[20] Strevens, M., Probability out of Determinism. In: Probability in Physics. Ed. by Claus Beisbart and Stephan Hartmann. Oxford Scholarship Online, 2011. https://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199577439.001.0001/acprof-9780199577439-chapter-13 (Accessed 9/20/2018)