Validity and Reliability, Causality, Causal Relationship Possibilities, External Validity

See online here

This article discusses the measures of association, causality relationship, and different possibilities in a causal relationship. The Bradford Hill criteria comprising nine principles are also discussed. The possibilities of a causal relationship include four principles. This article also explains reliability, validity, and related characteristics. The major threats to reliability include poor sampling, instability, and divergence.

Measures of Association

This term refers to a wide variety of coefficients that are required to measure the statistical strength of different variables. There are many statistical distinctions associated with the understanding of the relationship between statistical measures. **Statistical measures are different from statistical significance**. Measures of association assume a categorical or continuous level of data. Perfect predictive monotonicity and perfect ordered monotonicity are impossible to achieve at the same time in a measure of association.

Categorical data include data at nominal or ordinal levels, whereas the causal direction followed by the measure of association is based on symmetry or asymmetry. Measures of association can be categorized into four types, including, concordant pairs, discordant

pairs, tied pairs on one variable, and tied pairs on the other variable.

Causality

Causality refers to **any reason that leads to a specific disease in order for timely diagnosis and to take preventative or curative measures**. A cause can be sufficient, necessary, or both, in order to create an effect leading to a specific disease. There can be more than one causal mechanism leading to a single disease.

The causality relationship refers to the association between cause and its effect on an individual and includes predisposing, enabling, precipitating, and reinforcing factors.

Bradford has defined criteria to define a cause and effect relationship.

Bradford Hill Criteria

In 1965, Austin Bradford Hill introduced certain criteria to **provide evidence for the causal relationship between a presumed cause and an observed effect**. These criteria are currently used in public health research processes. It is helpful in the epidemiological research process by addressing different areas involved. The nine principles underlying the Bradford Hill criteria are as follows:

Temporality

It refers to the **time factor between cause and effect**. Primarily, for a disease to occur, the cause should precede an effect, and thus, exposure, disease, treatment, and resolution should occur in that order. The effect that occurs due to a cause by the subject is addressed in the context of time under this principle. A delay in the cause and effect relationship is addressed under this category. It states that if there is a delay in an effect related to the cause, the effect should have occurred after the delay.

Strength and association

This criterion refers to **the effect size created due to the epidemic cause**. The size of an association impacts the intensity of the effect. The cause and effect relationship is usually seen in the context of a statistical correlation between repeated events. The complete correlation between these variables is denoted by 1. In the case of a weak association, the cause and effect relationship will show higher variations and vice versa. This principle has been used by physicians to point out factors that increase the incidence of a disease condition and help guide patients to follow a healthy lifestyle.

Biological gradient (dose-response)

In case a patient is given a dose of a drug, there exists a **relationship between drug dose and the patient's reaction to the administered dose**. It does not indicate a simple linear relationship due to minimum and maximum thresholds. An association can have a causal relationship if there exists a biological gradient between exposure and disease. The higher the exposure, the greater the effect of the cause. In some circumstances, the mere presence of a biological gradient can trigger a large effect.

Consistency

In order to determine the reproducibility of a research process, the principle of consistency is mandatory to achieve accurate repetition and maintain ruggedness. In order to prove the usefulness of a treatment modality, the **consistency principle contributes to its productivity** in a wide range of circumstances; thus, the more firmly a principle is established in numerous studies with different methodologies, the higher the chances of an effect being verified.

Plausibility

The cause and effect relationship should be sensible and logical in the context of all related theories, concepts, and results. In case the causal relationship between the cause and effect of a subject indicates the occurrence of factors outside the science of research, it may create a hindrance in the accurate analysis of the causal relationship.
It **investigates the plausible mechanism employed by the causal relationship between cause and effect**. The principle also advocates for allowing what is not yet known to be construed as possible if more research is conducted; thus, new information should not be dismissed or discarded without due verification.

Specificity

In case there is no other plausible explanation, it **explains the specificity of a population**. For example, if there is a specific population of patients suffering from asthma in a town in California, there will be a specific association between the disease factor and its effect.

The more specific this association, the higher the probability of the existence of a causal relationship. It is not always possible in medical research that the symptoms of a disease are a result of a wide range of causative conditions. The fact that diseases have multiple etiologies and remedial therapies weakens this criterion, which is rectified with technological developments that allow isolation so as to measure specificity.

Evidence

Experimental evidence offers strong proof for the cause and effect relationship between a disease and the causative factors. Several significant variables are held stable in order to prevent them from interfering with the experimental results.

Analogy

This principle **considers the effects of similar factors in order to create a logical relationship between the suspected cause and its effect**. The other related factors should create a logical sense with the research subject; otherwise, they should not be considered in the investigation process.

Coherence

The **likelihood of the effect of a cause and its effect increases when there is coherence between the epidemiological and laboratory findings of a study**. Despite the coherence criterion, Hill noted that if laboratory evidence is unavailable or insufficient, it cannot completely nullify the epidemiological effect on associations.

Possibilities in a Causal Relationship

The four possibilities of a causal relationship to create an association are as follows:

1. Necessary and sufficient
2. Necessary, but not sufficient
3. Sufficient, but not necessary
4. Neither sufficient nor necessary

One of the four possibilities stated above should be fulfilled.

Necessary and sufficient

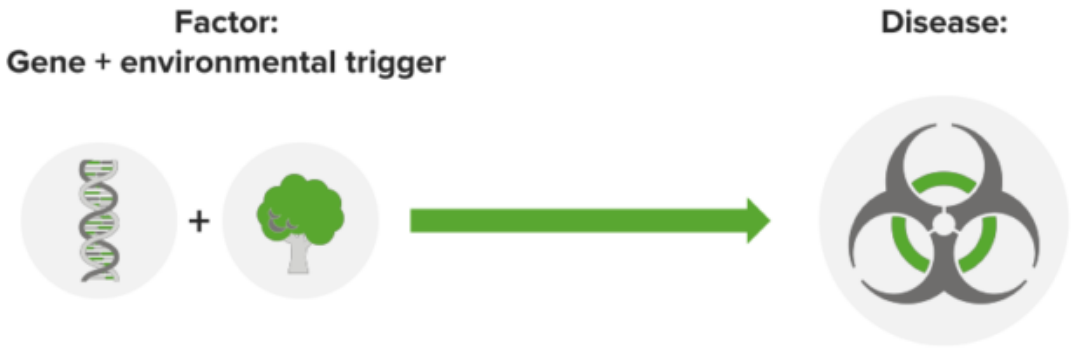
A particular condition is necessary for the occurrence of a dependent condition, because, without the former, there is no possibility of the occurrence of the latter. This condition should also be sufficient to produce a cause and effect. In such a situation, both necessary and sufficiency requirements should be fulfilled.

Example: The Coronavirus causes SARS; therefore, the necessary condition for SARS is the Coronavirus.

Necessary, but not sufficient

This is the condition where the **existence of a situation is adequate to cause a problem**. In this situation, it is not necessary to measure the sufficiency of the condition for the occurrence of the related effect.

Example: In case a gene is activated by an environmental trigger, such as pollution or other harmful factors, it can cause disease.
In this case, just the trigger is insufficient, only the existence of the particular gene can cause the problem.

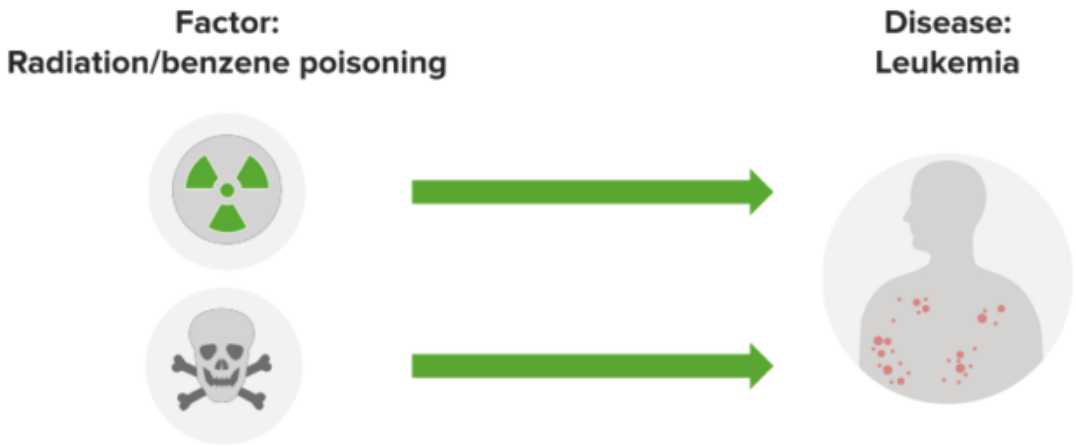


“Necessary, but not Sufficient” by Lecturio

Sufficient, but not necessary

In this situation, the sufficiency of a factor is necessary to create an effect.

Example: Both radiation and benzene poisoning can lead to leukemia. In this situation, both leukemia and benzene alone are sufficient to cause leukemia, but none of them are necessary for the calamity.

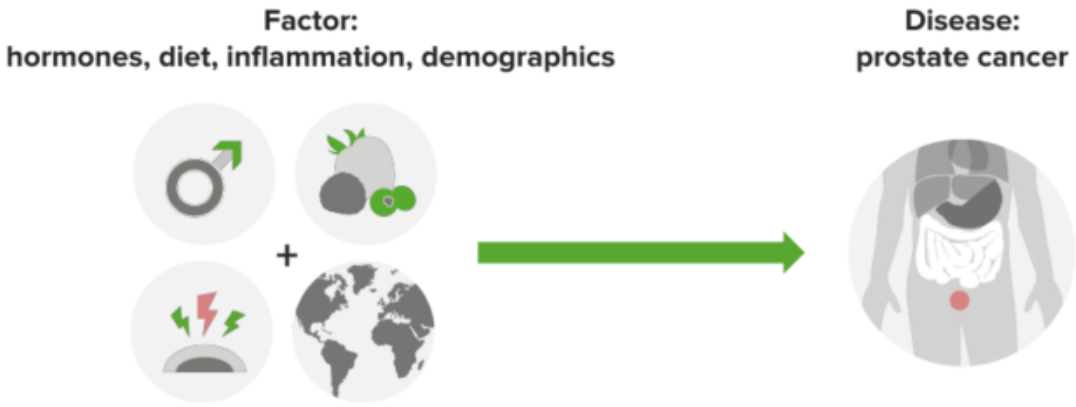


"Sufficient, but not Necessary" by Lecturio

Neither sufficient nor necessary

In this case, none of the factors are mandatory for the occurrence of a condition.

Example: Being tall is neither necessary nor sufficient for a person to become educated in life.
In the case of epidemics, an effect that is caused by a damaging factor requires no specific sufficient or necessary conditions for the occurrence of the disease.



"Neither Sufficient nor Necessary" by Lecturio

Reliability and Validity

Reliability

It refers to the **degree to which a method or tool is used to generate stable and consistent results**. Reliability demonstrates how accurately and consistently a particular test can measure a given characteristic. It is indicated using a reliability coefficient. There are several types of reliability, some of which are as follows:

Test-retest reliability: A test is repeated to measure the reliability of results. Repeatability is important to attain reliability.

Parallel forms reliability: Different versions of assessment tools are used to generate desirable results. A high-reliability coefficient indicates that the outcome will yield similar results regardless of the test that is chosen, while a low-reliability coefficient indicates that the tests are not similar and thus cannot be interchanged.

Inter-rater reliability: Different raters are approached to determine accurate research results. Raters need to be well trained for these reliabilities to be stable.

Internal consistency reliability: This aspect measures the degree to which different test samples generate the same result. This variant of reliability can be affected by the length of a test and results effectively reveal the homogeneous or heterogeneous nature of the test items.

Validity

It refers to the **ability of a test measure to estimate a result that is desired to be measured**. Key points include the characteristics of the test measure and how well it can measure outcomes, such that predictions can be made based on the test scores. Reliability alone is not sufficient to evaluate the required results. **For assured reliability, the test measure should be valid**. Validity tests can be based on a criterion, content, construct, or characteristic and is indicated by a validity coefficient.

Example: Suppose a weighing scale is off by 5 lbs. During daily weight measurements, the weighing scale indicates the weight in excess of 5 lbs. The scale measures the weight reliably and consistently; however, it does not give a valid result.

Internal validity

Internal validity **refers to a 'zero generalizability' concern**. It shows that the researcher has evidence that the measures taken in an investigation or research study have caused what was observed in the study.

The major requirements of internal validity include **temporality, strength, and plausibility**. It revolves around the question of the application of scientific research methods in experimental design.

Major drawbacks include **confounding and selection bias**.

External validity

It is the **degree to which the results of a study can be generalized at a large extent or for the general population**. Research and experimental findings are measured such that they are sufficient for a large population to conclude a specific or required result. The requirements of external validity include minimized observer effects and parsimonious exclusion criteria. It generally allows identifying causal relationships that may be applied to other possibilities.

There are several drawbacks to be countered by external validity, including overly-specific study characteristics, the **Hawthorne effect (humans tend to alter their behavior when being studied)**, and the **Rosenthal/experimenter expectancy effect** (**researchers may develop bias regarding the expected outcome of their experiments)**. External validity is of two types, i.e. population and ecological validity.

Population validity

It refers to the extent to which the results of an experiment can be generalized to the whole population. In case the sample population is representative of the reference population, it is known as population validity.

Ecological validity
In this case, the environment of the study resembles real-world conditions. It is the extent to which the conclusions of a research study match the generalized findings in the context of the whole population.

Validity and reliability are the philosophical cornerstones of what is accepted as scientific proof.

Threats to Reliability

The major threats to reliability include the following:

Environmental changes

The time between measurements may be subject to environmental changes, which could consequently affect measurements.

Observer/researcher error

It can occur during reading and recording of measurements and can alter the reliability coefficient significantly.

Poor sampling

If the selected sample is not representative of the whole population, it can lead to inappropriate and unreliable results in an experiment.

Example: The mean age of a non-random sample represents an inappropriate population.

Instability

It refers to the inconsistent characteristics of measurement that are used to evaluate the required result.

Example: If blood pressure is measured multiple times a day, the blood pressure, here, is the unstable variable that changes every time it is measured. It can, therefore, lead to unreliable results.

Divergence

Mood divergence of raters and evaluators can lead to unreliable results.

References

B.M.Moskal & Leydens, J., 2000. Scoring rubric development: Validity and reliability. Practical Assessment, Research & Evaluation, 7(10), pp. 25-27. Available at: http://pareonline.net/getvn.asp?v=7&n=10

Hill, A. B., 2017. The Bradford Hill Criteria. Available at: http://changingminds.org/disciplines/argument/types_reasoning/bradford_hill.htm

Norman Swartz, 1997. Necessary Conditions and Sufficient Conditions.

Available at: https://www.sfu.ca/~swartz/conditions1.htm

Phelan, C. & Wren, J., 2017. Exploring Reliability in Academic Assessment.
Available at: https://chfasoa.uni.edu/reliabilityandvalidity.htm

Statistics Solutions, 2017. Measures of Association.
Available at:
http://www.statisticssolutions.com/directory-of-statistical-analyses-correlation-measures-of-association/

Notes