Research & Reviews: Journal of Statistics and Mathematical science

Hard science, Soft data, and the difficulties of selecting Clinical Variables in Study

Karrie Williams *

Editorial office, Statistics and Mathematical Science, India.

OPINION

Received date: 15/08/2021
Accepted date: 22/08/2021
Published date:28/08/2021

***For Correspondence**

Karrie Williams, Editorial office, Statistics and Mathematical Science, India

E-mail: physicalmath@enginsights.org

OPINION

The investigating clinician has not been idle during these mathematical operations. He is usually in charge of developing the basic ideas to be tested in the research, assembling any additional clinicians who collaborated, recruiting the patients who volunteered to be observed, providing or arranging for suitable care for those patients, and recording the information that describes what happened to them. At important points in the evaluation process, the clinician is responsible for determining whether a "significant" finding is substantively rather than just stochastically significant. Although this split of research territory appears to be well suited to the diverse abilities, backgrounds, and interests of the clinical and statistical partners, a critical scientific zone of the area has yet to be explored. This zone is concerned with the strategy for selecting the clinical variables that will be observed throughout the research, as well as the methods for converting clinical observations into analyzable data. Most of the poorly conceived research initiatives that have caused so much unhappiness, conflict, and debate in clinical epidemiologic inquiry today are the result of poor management of these techniques and approaches.

The first problem to explore is the separation between clinical variables and all other variables used to communicate data about medical occurrences in persons. Demographic variables are fundamentally personal characteristics such as age, gender, ethnicity, employment, marital status, or religion. The information received from technologic techniques, such as roentgenograms, histologic or cytologic exams, and laboratory findings for chemical, microbiologic, and electrographic data, is referred to as paraclinical variables. Therapeutic variables provide information on the dose, duration, and other aspects of therapy. In normal medical practise, a physician recognises all of a patient's clinical characteristics and may even make notes on many or all of them in the patient's medical record. However, for the majority of research initiatives, the patient's original medical record is not the official document that is evaluated. Instead, a portion of the patient's data is placed into a unique format known as a case report form. The data obtained on this form will be the fundamental information recorded in the research based on the data collected on this form. Furthermore, the information included in the case report form does not comprise the fundamental data that will be evaluated. For most current studies, the gathered case report data is translated into entries that are punched on Hollerith (IBM) cards or fastened to magnetic tape for processing with a digital computer using different categorization and coding methods.

There are various approaches for determining if a variable is significant from a statistical standpoint. Based on the standard deviation and coefficient of variation for a single variable's data, one strategy is to assign little value to data with minimum fluctuation. For example, if everyone in a study were 62, 63, or 64 inches tall, the coefficient of variation for height would be relatively minimal, and height would most likely not be a major discriminator among the patients. Based on statistical analyses of correlation coefficients for pairs of variables, a second method assumes that if two variables are highly connected, one of them is probably useless and may be deleted.

For example, suppose we discovered with a correlation value of 0.99 between the variables colour of shoes and colour of

shoelaces, we may infer that the colour of shoelaces (or colour of shoes) might be excluded from the next study.

Even if all of the preceding proofs and arguments are accepted, a follower of the clinico-statistical faith has a significant fallback position. This fourth line of defence is the claim that the necessary clinical data cannot be obtained because physicians will either be unwilling or, if cooperative, will be inconsistent, imprecise, or unstandardized in how they make observations and record their results. The argument is based on a bias that resembles a self-fulfilling prophesy in many ways. Even if all of the preceding proofs and arguments are accepted, a follower of the clinico-statistical faith has a significant fallback position. This fourth line of defense is the claim that the necessary clinical data cannot be obtained because physicians will either be unwilling or, if cooperative, will be inconsistent, imprecise, or unstandardized in how they make observations and record their results. The argument is based on a bias that resembles a self-fulfilling prophesy in many ways.

The issues of enhancing clinical data quality are wide enough to constitute a separate dissertation, which will be left for a future part. The challenges will necessitate attention to methods for improving observer variability; strategies for developing effective indexes, scales, and criteria for classifying clinical phenomena; tactics for "measuring" entities such as chronometry and severity; and an understanding of the importance of evaluating transitions as transitions, rather than as two separate states of existence. All of these activities are part of clinimetrics, a sadly undeveloped topic that is presently in critical need of imaginative clinico-statistical attention.