Journal of Science and Healthcare Exploration (JSHE)
Monthly, Peer-Reviewed, Refereed, Indexed Journal
Volume - 5, Issue - 12, Dec – 2023
ISSN(O): 2581-8473
[Impact Factor : 5.273]
Publication Date: 31/12/2023

DOIs:10.2019/JSHE/202312003 --:-- Research Paper / Article / Review

Classification of data in Clinical research

[1]Javed Ali Khan, [2]Pathan Jamal Khan, [3]Dr.Ghazala Javed

[1.] Investigator, Biostatistics, National Research Institute of unani medicine for skin disorders, Hyderabad , Telangana.
Email: crihydstats@gmail.com,

[2.] Ex HOD , Orrata University, South Africa

[3.] Research officer CCRUM NEW DELHI.

Abstract: *Many investigations in the biological sciences are quantitative, with observations consisting of numerical facts called data. As biological entities are counted or measured, it becomes apparent that some objective methods are necessary to aid the investigator in presenting and analysing research data. In this paper let us discuss about types of clinical data used to draw statistical inferences. To draw statistical inferences, we need to understand the data first. The article discusses classification of data in clinical research.*

Key Words: *Types of data in clinical research.*

1. INTRODUCTION:

A characteristic that may differ from one biological entry to another termed as variable. Different types of variables may be encountered by biologists, and it is desirable to be able to distinguish among them. The data which is directly collected from the primary source of clinical experiment is known as primary data. Data collected from published records is known as secondary data. Both types of data having equal importance in drawing statistical inferences about the efficacy of the drug.

2. METHODS & DISCUSSION:

Primary data:

Data which is collected for the first time directly from the source are known as primary data.

Methods of collecting Primary data:

1. Direct personal interviews
2. Indirect oral investigations

3. Information received through local agencies.
4. Mailed questionnaire method
5. Scheduled sent through enumerators.

Secondary data:

Secondary data are those that have been already collected and analysed for some agency. This may be published or unpublished.

- official publications of International bodies like UNO or its subsidiaries, foreign Governments etc.,
- Official publications of Central and State Governments.

- Semi-official publications of varies local bodies such Municipal corporations and districts boards. - Private publications such as publications of Trade and Professional bodies like Institute of Charted Accountants, Financial and Economic Journals, Annual Reports of Joint Stock companies, Publications brought by Research Institutes etc.,

Further classification as follows

Ordinal Data Definition:

Ordinal data is a statistical type of quantitative data in which variables exist in naturally occurring ordered categories. The distance between two categories is not established using ordinal data.

In statistics, a group of ordinal numbers indicates ordinal data and a group of ordinal data are represented using and ordinal scale. The main difference between nominal and ordinal data is that ordinal has an order of categories while nominal doesn't.

Normal data Vs Ordinal data

Likert scale is a popular ordinal data example. For a question such as: "Please express the importance pricing has for you to purchase a product.", a Likert Scale will have the following options which are coded to 1,2,3,4 and 5 (numbers). 1 is lesser than 2, which is lesser than 3, which is lesser than 4, which in turn is lesser than 5.

Very Important	Important	Neutral	Unimportant	Very Unimportant
1	2	3	4	5

Ordinal data is thus a collection of ordinal variables, i.e., if you have variables in a particular order – "low, medium, high", they can be represented as ordinal data. There are two important factors to consider for ordinal data –

There are multiple terms that represent "order" such as "High, Higher, Highest" or "Satisfied, Dissatisfied, Extremely Dissatisfied".

- The difference between variables is not uniform.

Ordinal Data Characteristics:

For a question such as the following, here are five ordinal data characteristics:

- Which of the following categories best describes your last purchasing experiences with a product/service?
  - Very Pleasant
  - Somewhat Pleasant
  - Neutral
  - Somewhat Unpleasant
  - Very Unpleasant

➢ **Establish a relative rank**: In the above-mentioned example, Somewhat pleasant is definitely worse than very pleasant or very unpleasant is worse than somewhat unpleasant. There clearly is a rank within the options – which is a sign of ordinal data.

➢ **Value of interval is unknown:** The variation between very pleasant and somewhat pleasant need not be the same as the difference between somewhat unpleasant and very unpleasant. This interval can't be concluded using the ordinal scale.

- **Measure non-numeric traits:** In the given example, all the answer options are non-numeric and similarly ordinal data can be used to capture feelings such as satisfaction, happiness, frequency etc.

- **Add-on to nominal data:** Nominal data is "labeled" data. Ordinal data is labeled data in a specific order. In the above mention sample, there is a notable order in the options which makes it a classic case of ordinal data.

- **Ordinal data has a median:** Median is the value in the middle but not the middle value of a scale and can be calculated with data which has an innate order.

Ordinal Data Analysis:

- **Easy methods of Ordinal Data analysis:**

Ordinal data is presented in a tabular format which makes analysis easier for the researcher. Mosaic plots are also used to establish the relationship between nominal and ordinal data.

For instance, if an organization intends to analyse the number of employees in each hierarchy to make a systematic hiring process for the upcoming year – they can put this data in an ordered tabular format. HR executives will find this data extremely easy to refer to and analyse for any future updates.

Methods of Data collection

Essentially there are several ways one can collect data but the most familiar ways are: Observation and controlled experiments.

Observational study—measures the characteristics of a population by studying individuals in a sample or population.

Examples *of Observational data collection*

- Census, economic and educational data
- Local and government data, surveys
- Sports data
- Stock, commodity, bond, option and currency market data

Examples *of Experimental data collection*

- Comparative experiment
- The before and after study
- Simulation (mathematical Model. Ex y = Ax+B

The data collection method is related to the nature of the problem to be solved and the ethical and practical constraints of collecting data in some particular environment. Based on this thought, there are several ways to collect data. Mostly, the focus (interest) of the statistical study dictates the selection of the best way to collect data. The following is a brief summary of four methods of data collection, which are components of the two main ways of collecting data.

Data collection is done using the Case Report Format(CRF) that may exist in the form of a paper or an electronic version. The traditional method is to employ paper CRFs to collect the data responses, which are translated to the database by means of data entry done in-house. These paper CRFs are filled up by the investigator according to the completion guidelines. In the e-CRF-based Computer Data Management (CDM), the investigator or a designee will be logging into the CDM system and entering the data directly at the site. In e-CRF method, chances of errors are less, and the resolution of discrepancies happens faster. Since pharmaceutical companies try to reduce the time taken for drug development processes by enhancing the speed of processes involved, many pharmaceutical companies are opting for e-CRF options (also called remote data entry). From the industry perspective, the biggest hurdle would be the planning and implementation of data management systems in a changing operational environment where the rapid pace of technology development outdates the existing infrastructure. In spite of these, CDM is evolving to become a standard-based clinical research entity, by striking a balance between the expectations from and constraints in the existing systems, driven by technological developments and business demands.

3. CONCLUSION:

CDM (Computer data management) has evolved in response to the ever-increasing demand from pharmaceutical companies to fast-track the drug development process and from the regulatory authorities to put the quality systems in place to ensure generation of high-quality data for accurate drug evaluation. To meet the expectations, there is a gradual shift from the paper-based to the electronic systems of data management. Developments on the technological front have positively impacted the CDM process and systems, thereby leading to encouraging results on speed and quality of data being generated. At the same time, CDM professionals should ensure the standards for improving data quality. CDM, being a speciality in itself, should be evaluated by means of the systems and processes being implemented and the standards being followed. The biggest challenge from the regulatory perspective would be the standardization of data management process across organizations, and development of regulations to define the procedures to be followed and the data standards.

REFERENCES:

1. Gerritsen MG, Sartorius OE, vd Veen FM, Meester GT. Data management in multi-center clinical trials and the role of a nation-wide computer network. A 5 year evaluation. Proc Annu Symp Comput Appl Med Care. 1993:659–62.
2. Lu Z, Su J. Clinical data management: Current status, challenges, and future directions from industry perspectives. Open Access J Clin Trials. 2010;2:93–105.
3. CFR - Code of Federal Regulations Title 21 [Internet] Maryland: Food and Drug Administration. [Updated 2010 Apr 4; Cited 2011 Mar 1]. .
4. Study Data Tabulation Model [Internet] Texas: Clinical Data Interchange Standards Consortium. c2011. [Updated 2007 Jul; Cited 2011 Mar 1].
5. CDASH [Internet] Texas: Clinical Data Interchange Standards Consortium. c2011. [Updated 2011 Jan; Cited 2011 Mar 1].
6. Fegan GW, Lang TA. Could an open-source clinical trial data-management system be what we have all been looking for? PLoS Med. 2008.