## Big Data

Big data is the capability to collect and analyze huge volume of data gathered from multiple sources. Big data is commonly described using terms like volume, velocity, variety, veracity, and value (Ref-u214420). Volume indicates the data volume, velocity is the data generation speed, variety is the multiple resources or diverse data used, veracity is the authenticity of the gathered data, and value is the new insights gained from big data (Ref-u214420). Big data in healthcare is the huge volume of health and care-related data, like data related to disease monitoring in communities or nursing research data. Big data can help gather a huge volume of better-quality data faster, and cheaper. It has the potential to help gain valuable insights from the gathered data, contain the spreading of diseases, minimize or eliminate resource wastage and provide cost-effective and efficient patient care.

### Current Issues of having Big Data in Healthcare

Big data is poised to revolutionize the field of healthcare and inspire innovation and new thinking, but there can be many pitfalls of using it. Problems caused by incomplete data and unreliable standards and processes used to gather, store, analyze and present data can give rise to credibility issues (Lee et al., 2020). There can be other development limitations as well. The five significant facts about Big Data issues in healthcare are discussed below:

## No indication of Causality

Big data can help understand relationships by revealing the "what " part of the data but not their causality as it cannot explain the "why" part. Big data can help uncover large trends and patterns but fail to identify pathbreaking innovations (Ref-f618755). Big data can also help provide suitable solutions, but it cannot meet new expectations.

## Failure to show the Whole Data and potential for Selection Bias

Big data can only approximate the whole, but fail to represent the whole data. When there is a huge volume of data, cluster analysis can be tough due to unwanted noise or interference issues (Ref-f913926). The possibility of gathering a large quantity of unwanted data makes data cleaning and padding of missing data necessary.