**Disadvantages of using data science in health statistics**

1. Data protection.

A major concern with big data is data security, and who has the authority to de-identify and use the information. HIPAA has guidelines that regulate sharing personal health information across covered entities. However, the guidelines will need additional refinement to take into account changes in the nature of healthcare data, the volume of information and the types of available information. Even with de-identified data, there’s the potential to identify patients, especially when data is associated with a narrowly defined population.

1. Data Ownership.

Another concern is data ownership. When patient information was only on paper, patients had to provide written authorization for paper-based information to be shared with other entities. However, with the advent of electronic health records, patient-generated information and even non-structured data available through social media, determining data ownership has become less clear, the report said.

1. Data Assess and Availability.

Barriers to accessing and sharing data, across health professionals and health systems, remain a significant challenge. Providers have a vested interest in blocking access to data and o sharing it. Healthcare organizations believe that data sharing might cause them to miss chances to make their own discoveries with the information, and, thus,

An opportunity to leverage their health information, whether for financial gain or personal recognition, report authors note. They also fear that shared data may compromise their market share, thus reducing overall revenue.

1. Data Quality.

Researchers have continued to question the reliability of claims data as well as the usability of electronic health records data for predictive analytics. That’s because several factors can compromise data quality, including human error, variability in data capture and storage, and inconsistency or evolution of terminology and practice standards, the report says. In addition, Big Data analyses often use secondary data that’s being used to answer questions for which the data was not initially intended, the authors contend.

1. Data Validity.

Varied, imperfect approaches for de-identifying and aggregating health information gets in the way of effective analysis in healthcare, the report said. These factor combine data across systems, while increasing the likelihood that error will be introduced, thus impacting the validity of the final data set. The work required to minimize inaccuracies and integrate disparate sets of information wastes time and resources.

1. Data Analytics Tools.

The report notes a 2013 study that found, without standardized algorithms and

analyses to support the systematic analysis of Big Data, researchers are likely to

generate spurious results or chase patterns that are not real.

1. Data Analysis Strategies.

A major limitation to the potential of Big Data in healthcare is the lack of strategic approaches for data analysis. Using large data set to identify random finding is both flawed and lazy, the report suggests, adding that only a multidisciplinary team can unlock lessons from the data by using both inductive and deductive reasoning, leveraging a combination of analytical and clinical skills.