5.5 Value

Importantly, expanding on the IBM 4Vs of Big Data[6], CDM needs to maximize the relative value of any one data point in this ocean of data. Given the rapid influx of information from such a large volume of highly diverse sources in today's trials, achieving this requires management and oversight.

In the context of CDS, value goes beyond integrity and quality. To leverage the full potential of the data we have, we must look beyond its original purpose. During a clinical trial, we collect data to validate the hypothesis of the protocol and, ultimately, obtain market authorizations. Once databases have been locked, most pharmaceutical companies will only re-use them for regulatory purposes (e.g., annual safety updates, integrated efficacy and safety summaries, market authorization in other countries, etc.). However, to unleash the full value of clinical trial data, sponsors must pro-actively anticipate what will be needed in the future and seek patient authorization up-front, through unambiguous informed consent forms.

Some companies are beginning to re-use clinical data in new ways and this has influenced others to seriously consider it. Examples include creating synthetic arms, identifying trends from audit trails, feeding back data to patients during study conduct to boost retention, creating machine learning training datasets, and extracting real world evidence from real world data.

As these examples show, emerging technologies need to be leveraged to extract the full value of data for all stake holders – patients, sites, sponsors, regulators, caregivers and payers – and stop creating data silos.