PHUSE 2016

Paper PP21

Reflections on the Effects of Data Pooling

Catharina Dahlbo, Capish Nordic AB, Malmö, Sweden

ABSTRACT

There is an incredible amount of data generated every day. This, in combination with a changed attitude for making data transparent, increases the amount of data available for pharma companies. Companies can make use of this data to better understand drug effects. By providing domain experts with user-friendly access to legacy data, in combination with data from other sources, such as real-world data generated in e.g. social media or healthcare or academic clinical trials, important insight can be gained regarding e.g. safety signals. Here we explore prerequisites for data pooling and how the pharma industry can benefit from having access to pooled data.

INTRODUCTION

Today, personal and mobile devices generate a vast amount of information about patients' health, giving physicians new opportunities to customize treatments based on an individual patient's situation. In general, patients are becoming more and more active and engaged in their own health, and will in turn demand higher quality of care and information. Pharma will be affected as patients' consciousness rise and their loyalty to brands and companies decrease (1).

Janus Clinical Trials Repository (CTR), has been developed at FDA to support the regulatory agency in fulfilling their responsibility of protecting public health and patient safety. One use of the Janus CTR is to aggregate/pool data across studies to detect trends as well as address clinical hypotheses.

Most pharma companies only have access to their own legacy data but what would be different if they would have access to other clinical and complementary data, and what benefits would there be for the industry, regulatory agencies and also for the individual patient? And how could companies make use of pooled data in a more transparent and collaborative environment?

Data pooling is not a new concept, most pharmaceutical companies do it, for example in the form of meta-analysis. However, meta-analysis based on aggregated data is not feasible for exploring safety signals found in outliers. Pooling data from disparate studies and domains can make e.g. exploration of links between safety and efficacy data very powerful.

Here we discuss data pooling of individual, raw and detailed data and how data that is treated and stored under specific conditions can be even more efficient.

PREREQUISITES FOR POWERFUL DATA POOLING

DATA TRANSPARENCY AND PATIENT PRIVACY

In order to be able to pool data we need to have access to it. Data transparency has different meanings depending on context. Is it about making government data available for the public or is it in the context of making company specific data transparent for a company's employees or sharing of data between companies?

An internal company-based repository containing data from a full clinical program can in itself be of great value. However, amended with other clinical data made available through data transparency initiatives or other publicly available data or real world evidence data, the value and opportunities increase even further.

Closely tied to data transparency is the issue of patient privacy. Data protection legislation implies strict requirements regarding how data needs to be de-identified to ensure that information in no way can be traced to the individual patient. This goes further than just removing a social security number. Localization of servers and data transfer across borders is also an issue that needs to be considered in this context.

DATA CURATION

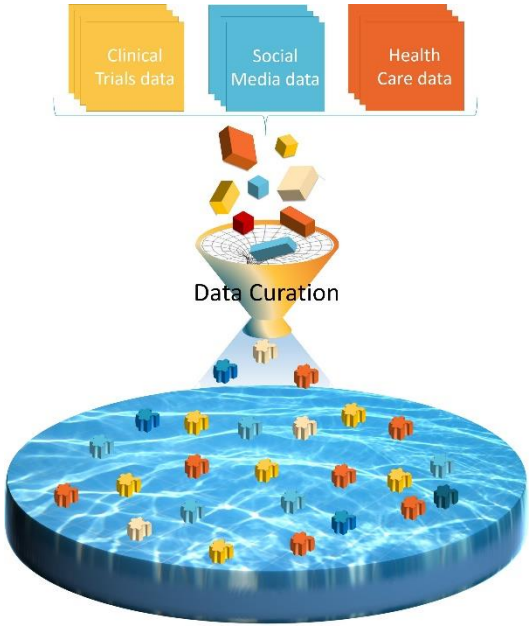
In order to effectively combine data from multiple sources into one unified view, it is necessary to prepare data before pooling. A prerequisite for an efficient and effective data pooling is a standardized terminology as well as common measurement units and value lists. One way to achieve this is to map data to an ontology. This ensures that parameters with the same meaning are mapped to a common definition throughout the repository. For example, a data curation process ensures that Blood Pressure and BP are given a common name and have a common measurement unit.

TECHNOLOGY

In order to take full advantage of curated data it needs to be stored in a central repository that can be accessed independent of location and device. Data should be accessible easily and securely for everyone with proper access rights.

Data exploration tools, making the curated data available and searchable, increase the possibilities for interesting discoveries. Functionality such as graphic displays, filtering and highlighting makes it possible to compare and explore data efficiently.

Clinical trial data as well as non-clinical data can be accessible in the same tool. The winning score of this procedure is not the separate steps alone but rather the total concept. Preparing the data from disparate sources and domains for analysis given the prerequisites above gives the user flexibility when exploring the data.



EFFECTS OF DATA POOLING

In the pharma industry, many different functions or roles need access to data but not all employees in a company need or should have access to the same data/information. Even with self-explanatory data described in a context, set by an ontology needs domain experts for proper understanding. However, users always need data that is easy to access, easy to search in, easy to share, easy to survey and easy to explore. To get there, the toolbox has to give you the power to answer all the questions you are asking.

IMPROVED COLLABORATION AND COMMUNICATION

If more people have the possibility to do exploratory analyses in an open, controlled, and reliable environment with high quality data, it will give greater understanding of data and new knowledge. A common data source is also important for communication within an organization e.g. in preparation for a new drug application but also when responding to questions or statements from regulatory bodies or of the research community. With standardized data in an objective terminology, available in a repository, a change of focus does not pose a challenge. Data is available for any purpose.

IMPROVED TRIAL DESIGN

With access to pooled data from multiple clinical trials, patient populations can better be defined, follow-up time adjusted, more adequate endpoints selected or response criteria revised and thus upcoming clinical trials can be better designed. Improved design can lead to a shorter time to market and cost savings, which is good not only for the pharma company but also for the patients who will get access to new treatments faster. However, it can also mean that fewer patients will need to be included in the development program.

IDENTIFYING SAFETY SIGNALS

A repository with pooled data is an efficient starting point for detecting safety signals early. Here the details of each patient in the clinical development program is available and outliers can be identified and explored further. Data from patients with similar characteristics can be examined as a subgroup or in detail. With all collected data in the repository, it is possible, not only to explore safety data but also link it to efficacy data.

PERSONALIZED MEDICINE

With pooled data, it will be easier to find subgroups of patients with similar characteristics that clearly benefit from the drug and thus that will gain the individual patients to find better individual treatments. In an extension, patients can have access to de-identified data from other patients, with the same disease and learn about the development of the disease and treatments.

RESEARCH FACILITATION

After pooling data from various open and/or public data sources, researchers will have a better base of information to study and understand e.g. classes of drugs. Discovering trends, similarities and differences between different classes of drugs or between drugs within in a class can provide important information for the medical community. “Old” data can be the engine for new insights and new trials without repeating research unnecessarily.

CONCLUSION

Transparent and powerful data pooling gives the industry better conditions to know their own data, for example in a submission. The companies can on their own identify issues that the agencies otherwise would find because of their access to bigger data repositories where data can be pooled. Data pooling may also give more opportunities for collaboration. Larger sample sizes improve the ability to identify statistical differences between subgroups and identifying rare adverse events. Companies that realize that the reality probably will look much different in some years and dare to change the company’s culture and systems will probably have an advantage for many years (2).

Data accessibility, data curation and effective technology in combination are prerequisites for effective data pooling which can be very important for both patients, researchers and the industry. This has the potential to help companies get drugs to market faster.

REFERENCES

1. “How pharma can win in a digital world”, December 2015, McKinsey&Company
2. “Digital transformation: The three steps to success”, April 2016, McKinsey&Company

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Catharina Dahlbo
Capish Nordic AB
Carlsgatan 3
211 20 MALMÖ, Sweden

Work Phone: +46 (0)40 10 88 80
Email: catharina.dahlbo@capish.com
Web: www.capish.com