

Reflections on the Effects of Data Pooling

Catharina Dahlbo, Capish Nordic AB, Malmö, Sweden
(catharina.dahlbo@capish.com)

INTRODUCTION

Pooling data from disparate studies and domains could 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

Access to data

Pooling legacy data is on its own valuable but the value can increase if it is complemented with other clinical data made available through data transparency initiatives or other publicly available data. De-identification is a necessity for the protection of patient privacy; information may in no way be traced to the individual patient.

Preparation of data

Efficient and effective data pooling can be achieved when data from multiple sources are combined into one unified view and data is mapped to an ontology. This includes utilization of a standardized terminology as well as adaptation to common measurement units and value lists in the data curation process.

Data repository

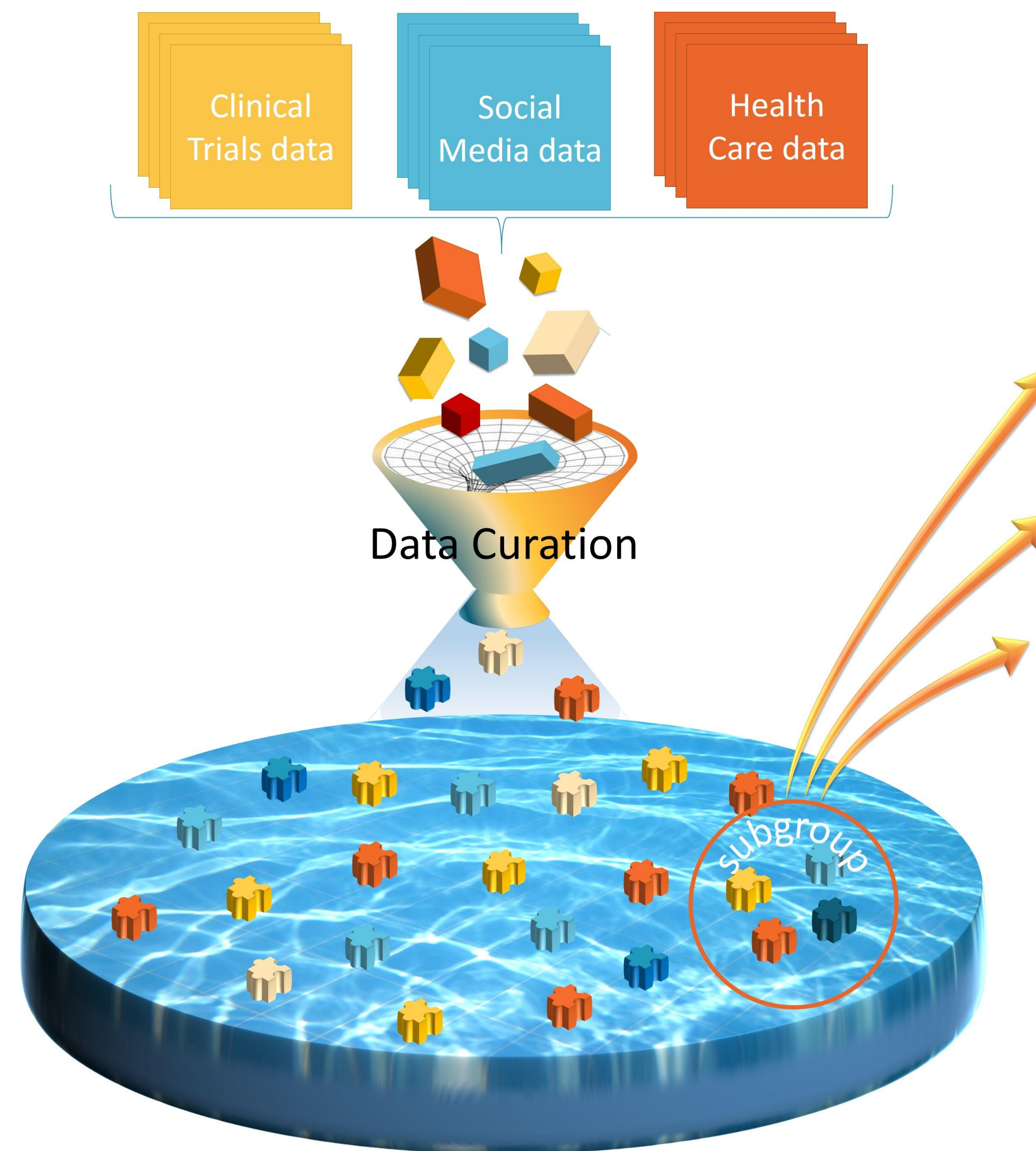
In order to take full advantage of curated data it needs to be stored in a central repository that can be accessed independent of location and type of device. Data should be accessible easily and securely for users with proper access rights.

Exploring data

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.

Winning score

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.



CONCLUSION

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

EFFECTS OF DATA POOLING

Improved collaboration and communication

As experts from different domains have the possibility to share a common repository for their exploratory analyses, a greater understanding of data is achieved. Internal and external communication can be based on a common platform.

Improved design of trials

Upcoming trials can benefit from experience from previous trials, e.g. better defined patient population, adjusted follow-up time or more adequate endpoints or response criteria.

Identifying safety signals

A repository with pooled data opens up for early detection of safety signals. The details of each patient in the clinical development program can be made available in the repository. Outliers can thus be identified and explored further.

Personalized medicine

With pooled data, it is easier to identify subgroups of patients with similar characteristics that clearly benefit from the drug. Thus in the long run individual patients can be offered better individual 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.

BENEFITS FOR

... agencies

Faster and smoother review of NDAs as companies have identified and dealt with findings prior to submission.

... companies

Improved communication and collaboration resulting in better knowledge of the drug and earlier market authorization.

... researchers

Reuse of data can be the engine for new insights and new trials without repeating research unnecessarily.

... patients

In an extension, patients get better individual treatments faster. Also with access to de-identified data from others, patients can learn about the development of the disease and treatments.