PhUSE 2016

Paper DV01

Interactive visualisation for Precision and Translational Medicine

Dr. Christof Gaenzler, PerkinElmer Informatics, Rodgau, Germany

ABSTRACT

Precision medicine requires the use of clinical- as well as real-world data to improve patient outcomes. But information comes from multimedia content in various data sources and locations. This makes it hard for teams to identify issues quickly and easily. Answering important questions from clinical, operational and safety data proves to be time-consuming and expensive as data integration and aggregation are necessary before analysis. PerkinElmer combines TIBCO Spotfire® with the new PerkinElmer Signals™ suite to allow collaboration within and across the corporate firewalls to integrate data from disparate platforms.

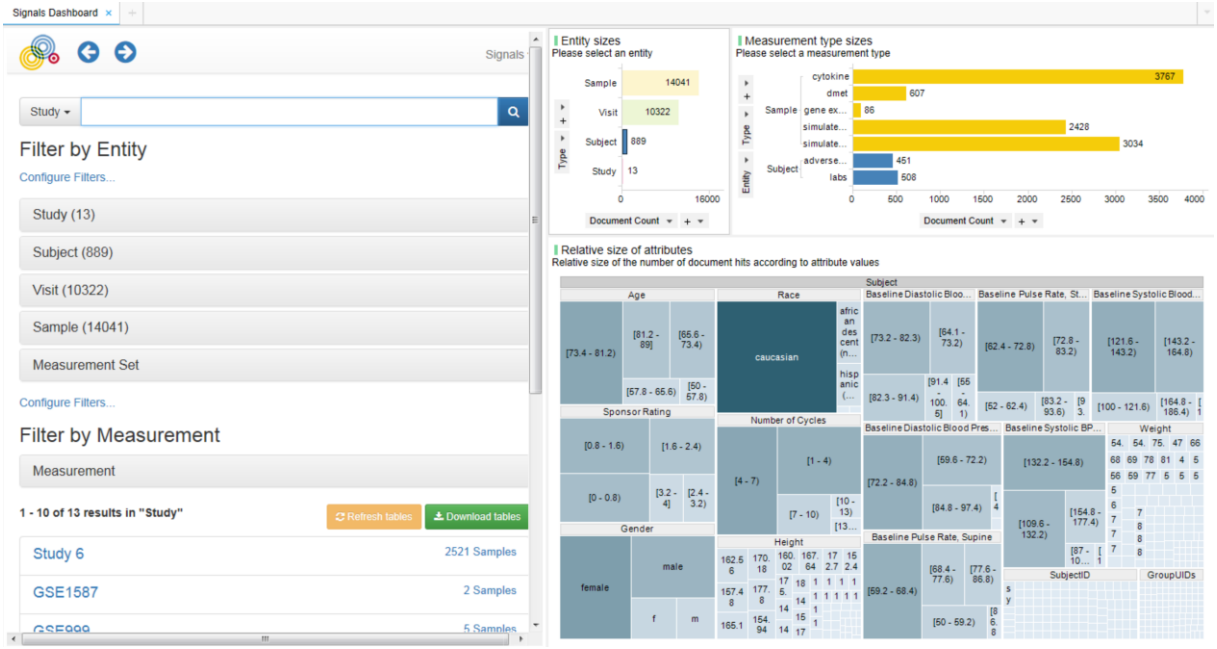
INTRODUCTION

The average industry cost per new prescription drug approval - including failures and capital costs – is estimated to be just below three billion US dollars. [1,2] This number includes discovery research and preclinical development costs as well as clinical and approval costs and the cost of capital. [2] Researching Bio-pharmaceutical companies need to be well aware of the cost of every step along the way and stop or adapt research or development projects as soon as possible if they discover the first warning signs of problems which could lead to a failure in a later stage. Unnecessary cost is the main driver to look out for these warnings, but also other resources e.g. scientists, labs and sites are scarce and need to be treated like perishable goods. This means that control efforts and the assurance of safety and efficacy of the new drug or treatment are of course slowing the whole process down, but through the optimal use of digital tools, this slowdown is minimized. In this article the focus lies on these digital tools which are a combination of data, software and processes in the form of digital workflows and interactions of working groups. These tools need to be formulated as best practices alongside a set of communication standards spanning all work streams of a project. All research conducting life science companies have established a way to constantly optimize their work streams. One piece of this optimization effort is to renew, replace and add digital tools depending on their availability. Another one is the linking of these tools across work streams preventing new and breaking-up already existing silos.

INTERACTIVE VISUALIZATIONS IN PRECISION AND TRANSLATIONAL MEDICINE DATA ANALYSIS

The principles of visual information seeking is based on the key concept of visually browsing and interacting with data to rapidly filter and reduce the result set to identify hidden relationships inside the data. [3] The concept of browsing data by using graphical displays of data and the ability to interact with the data through the progressive drill-down, drill-across and filtering is of extreme importance when the underlying data comes from disparate sources which originally were not designed to be interlinked. This is the case in translation science (https://en.wikipedia.org/wiki/Translational_science). The collection and curation of data from research and development as well as from publicly available sources is the foundation for PerkinElmer's Signals platform. Dimension free visual querying across many data sources is the first step of the workflow of a translational scientist. It starts with selecting clinical subjects and samples. Then processed data and secondary results are added to perform an integrated visual analysis. Fig.1 shows Signals Dashboard a visual query tool based on TIBCO Spotfire in Signals for Translational. It is used to browse the available data and quickly define a dataset to start the analysis with. The underlying data is mapped and organized to be queried as one unified data object. Data will be provisioned for analysis based on the result set of this query step.

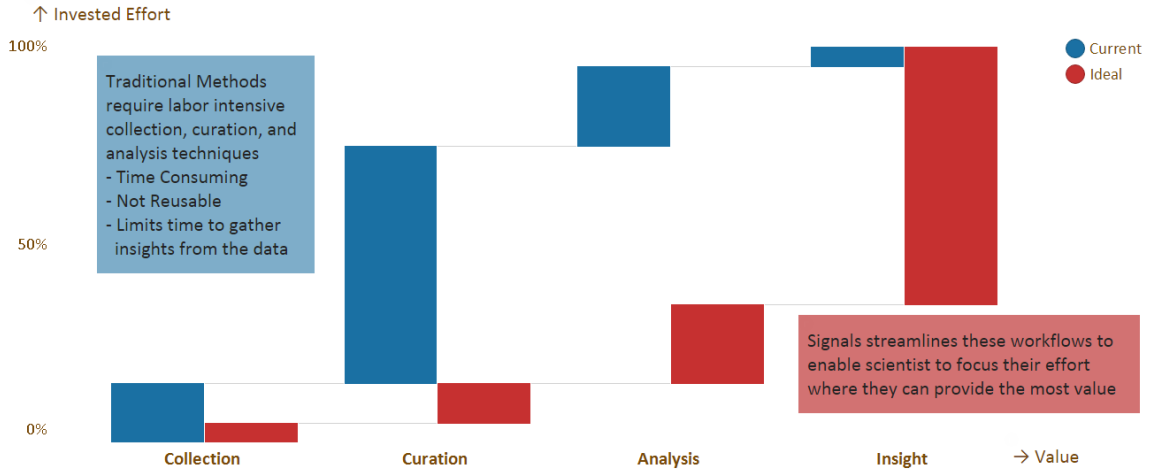
Fig.1: Visual Query Interface of PerkinElmer Signals for Translational



ORGANIZATION OF DATA TO ANSWER TRANSLATIONAL QUESTIONS

Collecting, Curating and Analyzing data was initially done by creating monolithic data warehouses using ETL processes and static tabular reports. This is not practical with translational data, because of its size, update frequency and complexity. Even the creation of dimensions and measures is not appropriate anymore since individual data objects are linked to end user specific actions, not to pre-defined queries. Still today most of the invested effort goes into collecting and curating data and not into analysis and insight coming from the data. [4] The most value for a scientist is created through the insight, but this is hindered by difficult and time-consuming data curation steps which are of low value for the scientist. This is why Signals was designed as a combination of digital tools to support data collection, curation and analysis of relevant data and by that speeding-up the time to the most valuable activity, the finding of insights. [Fig.2]

Fig.2: Shift Effort towards the Higher Value Activities in Translational Medicine Workflows



IMPORTANCE OF INTEGRATION OF REAL WORLD DATA

To achieve this described flexibility of the Signals platform, data sources need to be linked through their context rather than the exact matching of identifiers, e.g. UniProtID, which is of course still the preferred method, but not possible all the time. PerkinElmer Signals is using text-mining methods and workflows to create this context. Data objects in Signals are going through a curation process. The process starts with entity extraction based on simple

hierarchies and the identification of synonyms. Signals then unifies the different data objects by mapping them based on their content. In this step not only company internal siloed data is combined, but also publicly available real world data is integrated. This is important, because it immediately adds more context to the analysis and ultimately it adds more value to the generated insights. These insights can be used within all aspects and phases of therapeutic development.

CONCLUSION

The process leading to an approval to market a new therapy and the process of monitoring its long-term behavior in humans are at the core of every life science organization. [5] But there are huge costs and high risks associated with it. Precision medicine is one of the major trends to mitigate the risk of therapy development. Translational scientists are developing strategies to precisely target new therapies to a very well defined group of patients. The combination of research data with real world data is the foundation for establishing solid models for precision medicine. Through the use of interactive visual analytics data from various sources is brought into context for all project teams. This concept and the tight coupling of the underlying technologies make insights into data and processes more efficient and more secure. Communication between the different work streams of a project is facilitated through the use of interactive visual analytics since humans are still the main interface between work streams.

REFERENCES

1. Transforming Clinical Development, B McGraw (http://www.perkinelmer.com/PDFs/downloads/PerkinElmer-White-Paper-Transforming-Clinical-Development.pdf)
2. Cost of Developing a New Drug, J A DiMasi (http://csdd.tufts.edu/files/uploads/Tufts_CSDD_briefing_on_RD_cost_study_-_Nov_18,_2014..pdf)
3. Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays, C Ahlberg, B Shneiderman (T.R. 93-71, http://drum.lib.umd.edu/bitstream/handle/1903/5410/TR_93-71.pdf?sequence=1)
4. Accelerate BI Initiatives with Self-Service Data Discovery and Integration, Forrester Research 2015, (https://www.attivio.com/learn/resources/white-paper/forrester-data-democratization)
5. PerkinElmer Informatics Solutions for Therapeutic Development (http://www.perkinelmer.com/clinical-development-analytics/index.html)

ACKNOWLEDGEMENTS

I thank Dr Mark Demesmaeker and Dr James Powers for their support during the preparation phase.

RECOMMENDED READING

PerkinElmer Informatics Solutions for Therapeutic Development
http://www.perkinelmer.com/clinical-development-analytics/

CONTACT INFORMATION

Dr. Christof Gaenzler
Manager Scientific Analytics
PerkinElmer GmbH
Ferdinand-Porsche-Ring 17
63110 Rodgau

christof.gaenzler@perkinelmer.com