



International Journal of
Medicinal Chemistry & Analysis

www.ijmca.com

e ISSN 2249 - 7587

Print ISSN 2249 - 7595

MANAGING R&D DATA IN A VIRTUAL WORLD – A REVIEW

*Muthukumar S¹, Shanthi Reddy¹, Selvakumar D², David Banji¹

¹Department of Pharmaceutical Analysis & Quality Assurance, Nalanda College of Pharmacy,
Cherlapally, Nalgonda, Andhra Pradesh, India - 508001.

²Department of Pharmaceutical Chemistry, School Of Pharmacy, Taylor's University, Subang Jaya - 47500, Malaysia.

ABSTRACT

Pharmaceutical and biotech companies are moving from centralized organizations to a virtual network of CROs, academic partners, internal labs, and government agencies. A “virtual” pharmaceutical company is able to access data from its multiple academic partners, CROs, and CMOs from a single hosted data management platform, with each supplier having a secure area for its project data but no view of the others. This approach is almost identical to the current model but with external partners, not internal groups. In the past, collaborations were conducted with document- and Excel®-based exchanges for findings. Document based collaboration misses the context of an experiment that is so crucial to scientific understanding and translation of that data into knowledge. As a result, it is vital that scientific partnerships, both today and in the future, be data centric, context rich, and provenance aware, enabling all relevant data and information about the data to be captured. Technology considerations have been developed such as Hosting and Cloud, Software Infrastructure Management, Security and Audit, Domain Flexibility, Data Capture and Signing, Data Quality, Data Analysis, Dashboards. By capturing this information by default, organizations reduce the risk of losing the value of the collaborative data they have invested so much in generating.

Keywords: Flexibility, Data Capture and Signing, Data Quality, Data Analysis, Dashboards.

INTRODUCTION

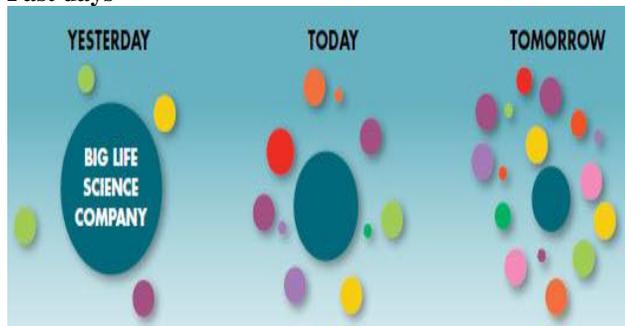
Access to real-world patient data, supporting precision or stratified drug discovery, and the general trend to externalize services all require sophisticated data management that enables the right mix of access and security. This article will look at the different research and development (R&D) processes in life science organizations where data is central to collaboration, and how it needs to be consistently captured, integrated, managed, tracked, and analyzed. Technical considerations for supporting this changing environment will also be explored and, as pharmaceutical companies are already in this increasingly complex network of data and partners, this will be done in a pragmatic way [1].

The glory days of pharmaceutical double-digit growth and megamergers resulted in huge organizations spanning the globe with billion-dollar budgets dedicated to

R&D. The majority of the work was carried out in-house to theoretically protect critical IP around lead compounds, driving innovation from an internal perspective and maintaining oversight and control via portfolio management. The concept of a pharmaceutical company's data going outside its firewall was taboo to these security- and IP-conscious organizations. Departments were relatively siloed and were often following a best-of-breed or internal development approach to informatics that enabled them to optimize their departmental efficiency and results. However, this hampered technology transfer between groups, which was often based on documents, presentations, or high-level summary data with limited ability to share the context of data and higher-level “corporate knowledge.” Data management was primarily designed to support IP compliance and regulatory filing, with results reuse and collaboration a secondary task

handled by adjunct knowledge management groups. Some external specialists, biotech partnering, and contract researchers were used, but the drug portfolio was essentially internally driven and owned [2].

Past days



Present days



Well-documented pressures on the life science industry have forced a major rethinking of the pharmaceutical model. The pace of change toward pharmaceutical outsourcing has been startling in the past few years, and the business is expected to grow to \$65 billion by 2018, fueled by a compound annual growth rate of nearly 15 percent.¹ This has transformed the internal focus of these organizations to be much more development-, clinical- and marketing-centric and has triggered significant reduction in discovery and research departments across the board. Many noncore capabilities have been outsourced, ranging from individual groups such as bioanalysis, pharmacokinetics, synthetic chemistry and pharmacology right up to entire functions such as “basic research” and preclinical development. The drive for innovation is increasingly coming from partnerships and shared risk models for new medically active entities (biologic, chemical, technology). Furthermore, the internal IT groups of life science organizations have also been hit by budget constraints and are having to support a very different environment in which data is shared between external partners as part of this outsourcing and externalization drive. This creates significant problems in how to manage and maintain different levels of compliance, audit, and security to support varying levels

of interaction with third parties—all partners are not created equal [3].

The Future is Distributed and Data Centric

Opinions on how the life science ecosystem will look in the future are everywhere, but whether it is smaller pharmaceutical companies with a mainly clinical trial and marketing bias, a major growth in the impact of CROs, or a market dominated by biologics-based drugs, the details of the network topology are really irrelevant. Taking a holistic view of what is already happening, a more fluid and dynamic environment is evolving with collaborations being created, executed, reported, and stopped as a matter of course. The wider use of hosted and cloud technology and distributed data is also providing appropriate business- focused and beneficial solutions, and now companies are finding ways to deal with compliance across geographies while maintaining that all-important data security [4].

Types of Collaboration

The types of collaboration are also evolving, but some real-world examples are given below:

1. **Fee for Service**—where compound pharmacology is assessed by an external lab or academic center and supplied back as simple files, but there is no flow of data from the company to the partner. These types of interactions have been commonplace for many years
2. **Virtualized R&D**— where minimal in-house labs exist and extensive collaboration is carried out with CROs and partners to provide the full spectrum of research, development, clinical, and manufacturing services. There are many examples where elements of the R&D process are outsourced, but a good example of a more “virtualized organization” is Shire plc.²
3. **Hospital Collaborations**—clinical, observational, and pharmacovigilance studies conducted via close hospital collaborations. These collaborations require more advanced systems for near real-time data sharing and to protect patient privacy and follow ethical review board procedures. Various large pharmaceutical companies are starting to “embed” themselves into frontline clinically driven organizations (hospitals, health institutes, etc.), such as Roche with its Translational Medicine Research Collaboration (TMRC) in New York³ and Pfizer with its centers for therapeutic innovation in various US cities.⁴ Here, the closeness of the R&D organization with the health care provider organizations is expected to provide much better “real-world exposure” and therefore the ability to innovate and develop new medicines faster.
4. **Pre-competitive**—where the sharing of data is done on a very large scale for the good of all potential interested parties. Typical examples are the sharing of clinical trials data across broad disease areas as supported by the Innovative Medicines Initiative (IMI).

5. These aggregated studies include thousands of subjects and require industrialized data sharing and access addressed in projects such as eTRIKS [5].

In addition to new ways of collaborating, the move toward a more open R&D environment is creating new businesses and business models that are tapping into the opportunities created by these changes. For example:

1. **Research Service Brokering**—Assay Depot (www.assaydepot.com) is pioneering a service brokerage platform for available screening assays, techniques, and providers. This model also reflects the potential power of aggregation of useful data.
2. **Fully Electronic CROs**—Companies such as AIT are becoming fully electronic and capturing full experimental context for bioanalysis, enabling them to provide detailed data to their customers.
3. **Open Screening**—Eli Lilly's Open Innovation Drug Discovery offering allows anyone to use Lilly's industrialized screening process to test their libraries; in doing so, they also have the opportunity to partner with Lilly if an interesting hit is found.
4. **Patient Data Brokering**—The UK NHS's Clinical Practice Research Datalink⁶ offering aggregates highlevel, population-based, anonymized patient data from across the NHS that is of potential interest and then sells the data to research and development organizations.

Data Systems

1. Simple measurement such as IC/EC50 is not meaningful until the experimental conditions or context is specified and the networks of other data it touches are "explorable."
2. IDBS E-WorkBook allows multiple secure projects to be managed from the same system and tracks all data, metadata, IP, and decisions.
3. Many groups successfully use a LIMS to track samples from receipt to end of analysis and deliver results to end-point tests. While this gives a direct view of a study as would be placed in a final report, some critical information remains in paper format. For example, important data relating to the validation of instruments and software, staff training records, QA audits, metrology data, and information surrounding reagents also need to be captured to provide context. In a virtual environment it is vital for all this LIMS data and additional context to be captured.

Technology Considerations

To effectively support this diversity of collaboration types, life science companies of all types need to have ways of sharing and analyzing data that lend themselves to a dynamic, but still validated, data environment. Some key considerations include:

1. **Hosting and Cloud**with the growth in availability of external storage of data, the cost savings and convenience

of cloud technologies are compelling. For most IP-based organizations, the use of "private cloud" via hosted servers is a more likely option as they know where the data is stored, often important with regional variation in privacy rules, and have greater control of cloud provider data access rights.

2. **Software Infrastructure Management**the systems used must be easy to deploy and update and have the ability to support tens of thousands of users at one time. Also important is the ability to manage the addition and removal of users and privileges easily against a core set of "business rules" without overloading the IT groups.

3. **Security and Audit**—Multiple types and durations of collaboration must be supported, with individual CROs and academic collaborators having their own secure areas to enter data and share project data with the consortia. Security must be linked to types of data and the context of data—it must be possible to control both and have a security model flexible enough to be changed quickly. It must also be possible to do "who, what, when, and where" analysis on the system and to be able to track users' usage and interaction with the system at the data object level.

4. **Domain Flexibility**—Systems need to be able to support structured and unstructured data capture and collection across many domains and disciplines, including both large and small molecule, research and development, omics, imaging, and other molecular techniques, along with patient and market data.

5. **Data Capture and Signing**—As with internal systems, there needs to be capture of the context/metadata as well as the experimental results to ensure the same level of auditability and traceability is maintained when compared with internal systems. Coupled with that there needs to be support for digital signatures/ identity stamping of data so that regulatory requirements such as 21 CFR Part 11, GLP, and GMP are met.

6. **Data Quality**—In a diverging ecosystem, data quality will be the key to competitive advantage, so full data context, checking, and validation are essential.

7. **Data Analysis**—In complex scientific domains, there are many tools and technologies for data analysis and visualization. Systems need to be able to integrate existing and new algorithms flexibly and allow SAS, R, and Matlab scripts, for example, to be run from a common environment. Visualization tools such as Spotfire, QlikView, and Tableau also need to be accessed easily and data exported.

8. **Dashboards**—Oversight of multiple projects and collaborators is essential, so dashboards on project status, the ability to surface data instead of search for it, and extended analytics to show trends and risk analysis are all important as the trend evolves [6].

CONCLUSION

While the life science industry is changing at an unheard-of pace, there are still many opportunities for

R&D organizations, both large and small, to work together effectively to develop better treatments. New business models are here, and technology is now available to support data-driven collaboration. Therefore, despite the obvious hurdles, there is no reason why “virtual R&D” shouldn’t improve significantly on the internally focused processes of the past. We are seeing more advanced organizations implementing structured metadata tagging, controlled at the enterprise level, that delivers the security control and auditability required for full virtual

collaboration. We also see innovative “open” identity management and “trusted status” sharing between collaborators and real interest in integrating these into advanced data management platforms. But for everyone there remains a fundamental need to ensure the context and the provenance is captured along with the data. By capturing this information by default, organizations reduce the risk of losing the value of the collaborative data they have invested so much in generating.

REFERENCES

1. Healthcare Contract Research Outsourcing Market (Contract Research Organizations, Clinical Data Management, Clinical Trial Laboratories, Regulatory Medical Writing, Academic Medical Center) - Global Industry Analysis, Size, Share, Trends and Forecast, 2012 – 2018. Transparency Market Research, January 2013.
2. IDBS press release, Shire Selects IDBS for Virtualized Drug Research Collaboration Platform. April 14 2010: <http://www.idbs.com/Data-Management- News/Press-Release/10APR14.asp>
3. http://www.nytimes.com/2012/09/20/nyregion/roche-swiss-pharmaceutical-giant-to-open-researchcenter-in-manhattan.html?_r=2
4. <http://www.bloomberg.com/news/2012-06-15/pfizersees-harvard-collaboration-as-spark-for-new-drugs.html>
5. <http://www.imi.europa.eu>
6. <http://www.cprd.com/intro.asp>