



Big Data in Pharmaceuticals

Big Opportunities or
Big Challenges – or Both?



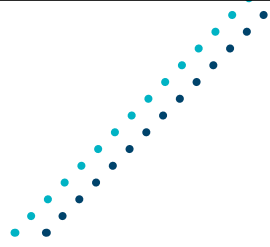
Introduction

Big data covers every facet of our working life. Every aspect of pharmaceutical research and development involves the generation of huge quantities of data, with the expectation that we can turn this information rapidly into useful knowledge, which in turn can be used to make 'data-driven' decisions to better understand and control processes. This derived knowledge can also be used to reduce costs, improve efficiencies, reduce development times and facilitate rapid post-approval changes [1]. It needs to be understood that in addition to having multiple data-users within an organization, these individuals/groups, will have different objectives, i.e. trending, change control, decision making, etc.

In addition, with increased complexity of molecules and processes, there is an increased likelihood of generating larger volumes of supporting data. Biologicals and particularly biosimilars are good examples of this latter scenario. Thus for example, the ability to evaluate and trend significant volumes of analytical data generated from orthogonal methodologies, which often generate conflicting

data, as part of a comparability exercise [2] for biologicals/biosimilars will continue to be challenging. In addition, real time review of data for manufacture and release of biologicals/biosimilars is likely to be a reality in the near future [3]. As such, advanced process control (APC) using process analytical technologies (PAT) will likely become common place, using for example model predicted control (MPC) approaches [4].





However, despite the increased focus on quality in the pharmaceutical industry (quality risk management (ICH Q9) [5], pharmaceutical quality systems (ICH Q10) [6] and lifecycle management (ICH Q12), [1] etc.) there has been a significant increase in the number and severity of quality defects as measured by FDA 483 observations and product recalls. For the pharmaceutical industry to “continue to be successful, drug manufacturing must become agile, rapidly scalable, efficient, reliable-and less costly” [7]. The cost of a poor quality culture can be enormous. For example, it was estimated that the costs incurred by Warner-Lambert over the 10-year period from 1993 to 2002 to effectively manage the quality-related outcomes from an FDA consent decree in 1993; “in terms of product terminations, delays in approvals, and bringing facilities and systems into compliance-

was nearly \$1 billion” [8]. These challenges can only be met by making better use of data and knowledge, allowing us to reduce the enormous cost of poor quality and in parallel improving data integrity.

Adding to this complexity is the likelihood that we will move towards probabilistic understanding of processes, where data will no longer be binary in nature (i.e. yes/no), but will be probability based. A good example of this is release specifications.



How will the industry cope with Big Data?

The industry is also under significant pressure to embrace new manufacturing technologies. FDA recently published its guidance on modernizing the pharmaceutical manufacturing base utilizing emerging technology applications [9] and to facilitate rapid introduction the Agency has introduced an Emerging Technology Team (ETT) [10].

Pharmaceutical companies can submit questions and proposals about the use of any specific emerging technology to the FDA. The ETT assumes a leadership (or co-leadership) role for any cross-functional quality assessments, which can include reviewing or inspection of on-site facility capabilities.

FDA sees this initiative as part of the long-term answer to avoiding drug shortages [11], which have bedevilled the US market in the recent past [12], and it is part of FDA's ongoing promise to streamline and improve current manufacturing practices [13]. FDA have highlighted that this initiative [9] also embraces "testing, quality control, packaging and labelling operations". FDA's guidance on process validation [14] also supports process improvements and ongoing innovation, which include continuous manufacturing. FDA has also highlighted that modernising the manufacturing base needs to be supported by integrated strategies towards product and process understanding. This in turn is underpinned by real-time monitoring of critical process data, which taken together should support better understanding, monitoring and process control [15].

The key concepts underpinning real time process control are multivariate data acquisition using process sensors, e.g. pH, pO₂, flow rates, temperature,

conductivity, etc., and analysers e.g. NIR, Raman, particle size distribution (PSD), etc., allowing process monitoring and subsequent control [16]. Data management tools are also important in that they allow the data to be accessed and correctly interpreted in a reliable fashion.

The regulatory strategy within the EU is comparable, although there has been less explicit emphasis on the role of modernizing the manufacturing base [17]. However, all of these novel approaches generate significant volumes of data, i.e. information rich, which necessitate enhanced data management solutions and enhanced data infrastructure [18], to collect, process and analyse these data rich streams. In turn, these data rich streams should result in "higher quality products, as online analysis systems sample a higher fraction of every batch, sometimes by a factor of 100x or greater, which in turn can be used for real time decision making. Systems that can integrate and control these big data streams will be extremely important" [19].

For analytical PAT systems, intelligent software with full automation of system functionality, which guides end-users of all levels of programming expertise through measurement processes via embedded user-support assistance, will see the greatest uptake. In addition, few PAT-users have expertise in any one particular specialist analytical technology; rather they need to use and be familiar with many different analytical approaches, i. e. they are typically “black box” users. Therefore, the use of intuitive, easy to use software and hardware is not just necessary, but rather it is essential [20].

Indeed, the connectivity of equipment, people, processes, services and supply chains all contribute towards “Pharma 4.0” [21]. The arrival of this new era of “embedded” technologies ensures that information relating to how these attributes are better integrated and are available at all levels of the organization.

Industry 4.0 technologies will enable manufacturers to have better visibility of ongoing operations allowing them to be more responsive to information about changes in raw materials, inventory, assets, quality, waste, output and customer demands, highlighting improvement opportunities and ensuring that actions are taken, saving time, money and resource.



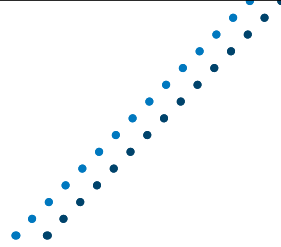
Data Integrity

However, at the same time that both Industry and regulators want to embrace and embed 'big data' into their decision making process there are parallel concerns and significant unease about data integrity in general. Where, data integrity denotes the completeness, consistency and accuracy of all of the established data. These data should in turn be attributable, legible, contemporaneously recorded, original (or a true copy) and accurate, i.e. ALCOA [22]. Another important facet of data integrity is metadata. Data is often meaningless without additional understanding about the context in which the data were generated. Metadata is therefore typically described "as data about data".

Metadata "describes, explains, or otherwise makes it easier to retrieve, use, or manage data". Thus, metadata could include a date/time stamp (electronically or paper-generated) for when the data were generated, a user identification of the analyst/person who performed the test or generated the data, the unique instrument identification that was used to acquire the data, audit trails allowing the data to be reconstructed, etc. Data contextualization brings the 'data to life' and allows data to be transformed into knowledge. Data contextualization should be performed automatically and in advance of the planned activity. This will allow data to be utilized by different persons (or groups) within the organization with different objectives, i.e. quality, KPIs (Key Performance Indicators), engineering. For example, pH data can be reviewed as a time-based measure (trending the data), or comparing the output of different probes or versus the last successful validation exercise. In fermentation processes data can be generated dynamically (time-based data) or statically, e.g. calibration, clean in place

(CIP), etc. In addition, discreet data, e.g. in-process control (IPC) as well as data on critical material attributes (CMA) e.g. microbiological organism strain, media information, etc., can also be simultaneously generated. Data should be conserved throughout the record's retention period as well as all of the associated metadata which is required to recreate the specific cGMP activity, e.g., 21 CFR §§ 211.188 and 211.194. In addition, any unique relationships between the data and their supporting metadata should be retained in a safe, secure and ultimately traceable manner. Data integrity typically manifests itself as either genuine errors, i.e.; "omission of data, errors in data recording" or fraud, i.e. "changing data, deleting data or destroying data" [23].

There is also a need to manage audit trails for the data, meta-data and data contextualization.



FDA has issued draft guidance on the subject of data integrity [22]. The primary reason is that FDA (and other agencies) has observed increasing numbers of cGMP violations involving data integrity during their routine cGMP inspections [24]. FDA and other agencies are increasingly concerned about these trends because data integrity is a critically important element of industry's responsibility "to ensure the safety, efficacy, and quality of drugs, and of FDA's ability to protect the public health". These data integrity related cGMP violations have led to numerous regulatory actions, "including warning letters, import alerts, and consent decrees".

The European Medicines Agency (EMA) has also recently issued a new good manufacturing practice (GMP) guidance to provide assurance of data integrity [25]. These data are generated during the process of "testing, manufacturing, packaging, distribution and monitoring of medicines".

The concern resides around the fact that regulatory agencies rely on these data to evaluate "the quality, safety and efficacy of medicines and to monitor their benefit-risk profile throughout their life span". In addition, better control of data records guarantees that the generated data are accurate and consistent, facilitating "good decision-making by both pharmaceutical manufacturers and regulatory authorities".

In parallel, EMA's GMP/GDP (Good Distribution Practice) IWG (Inspectors Working Group) have developed a set of Q&As (questions & answers) with guidance for stakeholders on actions that assure data integrity and also minimizes data integrity risks at

all stages of the data lifecycle in all pharmaceutical quality systems (PQS) [26]. This advice is applicable to both paper-based and electronic systems.

It specifically addresses :

①

An assessment of those risks to data integrity during "the collection, processing and storage of data"

②

Those risk management activities and measures at different stages in the lifecycle of the data

③

The design and subsequent control of both electronic and paper-based data systems; as well as measures to safeguard data integrity for those activities contracted out to a third party, i.e. contract manufacturing organization (CMO), contract research organization (CRO), etc.

Conclusion

The significant cost of a poor quality culture is often underestimated by the pharmaceutical industry. For example, consider what would be the organizational outcome if the costs for non-conformity (NC) or Out of Specification (OOS) outcomes were presented in Annual Reports? Undoubtedly this would fuel an appetite for significant change.

However, these quality challenges can only be addressed by making better use of data and knowledge to facilitate better and quicker decision making. Although, there are ongoing concerns about data integrity, compliance could actually improve if better 'data-driven' decisions could be made and the human element is removed from the equation.

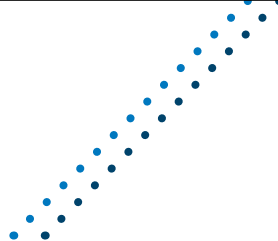
OSIsoft, January 2018



References

- 1.** ICH Q12, 2014. Final Concept Paper Q12: Technical and Regulatory Considerations for Pharmaceutical Product Lifecycle Management dated 28 July 2014. Endorsed by the ICH Steering Committee on 9 September 2014.
- 2.** ICH Q5E, 2004. Comparability of biotechnological/biological products subject to changes in their manufacturing process. Current Step 4 version dated 18 November 2004.
- 3.** T. Alosi. Next Generation Biologics Manufacturing: Delivering the Vision. <https://cdn.osisoft.com/osi/presentations/2017-uc-emea-london/UC17EU-D2LS04-Biogen-Alosi-Next-Generation-Biologics-Manufacturing.pdf>. Accessed on 12 November, 2017.
- 4.** J. Rantanen, J. Khinast. 2015. The Future of Pharmaceutical Manufacturing Sciences. *J. Pharm. Sci.*, 104(11), 3612-3638.
- 5.** ICH Q9, 2005. Quality risk management. Current step 4 version dated 09 November 2005.
- 6.** ICH Q10, 2008. Pharmaceutical quality system. Current step 4 version dated 04 June 2008.
- 7.** J. Woodcock. 2014. Introduction to Flawless: From Measuring Failure to Building Quality Robustness in Pharma. Eds: T. Fuhr, A. Gonce, L.o Positano, P.I Rutten, V. Telpis, McKinsey & Co.
- 8.** S. Ayd. Managing the Cost of Non-Compliance: Creating a quality culture can prevent the costs and challenges associated with receiving a consent decree. *Pharm. Technol.*, 41(11), 54-57.





9. FDA, 2015. Advancement of emerging technology applications to modernize the pharmaceutical manufacturing base. Guidance for industry. December 2015.

10. FDA, 2017. Advancement of Emerging Technology Applications for Pharmaceutical Innovation and Modernization. Guidance for Industry. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), September 2017.

11. FDA, 2013. Strategic Plan for Preventing and Mitigating Drug Shortages. Food and Drug Administration, October 2013.

12. FDA, 2016. Drug Shortages. <http://www.fda.gov/Drugs/DrugSafety/DrugShortages/ucm441579.htm>. Accessed on 29 October 2017.

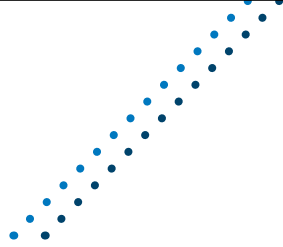
13. FDA, 2004. Pharmaceutical GMP's for the 21st century: A risk based approach. Food and Drug Administration, September 2004. <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/Manufacturing/QuestionsandAnsweronCurrentGoodManufacturingPracticescGMPforDrugs/UCM176374.pdf>.

14. FDA, 2011. Process Validation: General Principles and Practices. Guidance for Industry. January 2011.

15. FDA, 2004. PAT – A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance. Guidance for industry, September 2004.

16. D.P. Elder, 2015. An overview of process analytical control. Eur. Pharm. Rev. 20(6), 44-47.





17. 17. D.P.Elder, 2016. Modernising Pharmaceutical Manufacturing. Eur. Pharm. Rev, 21(5), 7.

18. 18. A. V. Venkatasubramanian, C. Zhao, G. Jogelakar, A. Jain, L. Hailemariam, P. Suresh, P. Akkisetty, K. Morris, G.V. Reklatis. 2006. Ontological informatics infrastructure for pharmaceutical product development and manufacturing. Comput. Chem. Engin., 30(10-12), 1482-1496.

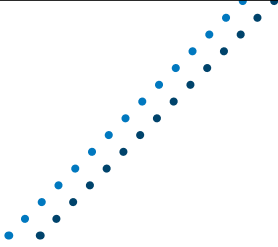
19. 19. S. Byrn, M. Futran, H. Thomas, E. Jayjock, N. Maron, R.F. Mayer, A.S. Mayerson, M. P. Thien, B.L. Trout. 2014. Achieving continuous manufacturing for final dosage formation: Challenges and how to meet them. May 20-21, 2014, Continuous Manufacturing Symposium. J. Pharm, Sci. 104, 792-802.

20. T. Page, Dubina H, G. Fillipi et al. 2015. Equipment and analytical companies meeting continuous challenges May 20-21, 2014 continuous manufacturing symposium. J. Pharm. Sci. 104(3), 821-831.

21. Anon. 2016. IoT, Industry 4.0 and the pharmaceutical manufacturing sector. Manuf. Chem. Pharma., https://www.manufacturingchemist.com/technical/article_page/IoT_Industry_40_and_the_pharmaceutical_manufacturing_sector/123678. Accessed on 15 December 2015.

22. FDA, 2016. Data Integrity and Compliance With cGMP. Guidance for Industry. <https://www.fda.gov/downloads/drugs/guidances/ucm495891.pdf>. April 2016. Accessed on 24 October 2017.





23. C. Platco, T. Cundell. 2017. Data integrity issues in microbial testing. *Amer. Pharm. Rev.*, 20(6), 36-43.

24. B. Unger. 2017. An analysis of FDA warning letters on data governance and data integrity. *PharmaceuticalOnline*. Guest column, 14 July 2017. <https://www.pharmaceuticalonline.com/doc/an-analysis-of-fda-warning-letters-on-data-governance-data-integrity-0001>. Accessed on 24 October 2017.

25. EMA, 2016. Data integrity: key to public health protection. http://www.ema.europa.eu/ema/index.jsp?curl=pages/news_and_events/news/2016/08/news_detail_002589.jsp&mid=WC0b01ac058004d5c1. Accessed on 24 October 2017.

26. EMA, 2015. GMP/GDP Inspectors Working Group. http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/document_listing/document_listing_000161.jsp&mid=WC0b01ac05800296c9. Accessed on 24 October 2016.

