LSCP Life Sciences Consulting Partners

January 2019

***Our Current Thinking*: EDG & RWD/RWE …**

Understanding Data first to reach Quality Data is Cirtical

Go deep to go fast

Govern your data holistically – including those from digitalization – to grow your business while staying compliant (e.g. IMDP+) & reducing costs, mitigate risks

Align and tune before accelerating

Executive Summary

Authors: Ruedi F. Blattmann (lead), **LSCP**

Contributors:

Prof. Peter Aiken (data blueprint), Dr. Beat Widler, Steve Scribner, **LSCP**;
Paul Gunstone (Smartlogic's Semaphore); Stephen Owens (Author-it's Docuvera)

A Business, not a tech decision: *Holistic view, enriched by FDA's RWD/RWE Definitions*

Poor Data Quality will not generate meaningful analytics and insight: Following "Our Current Thinking" November 2018 edition (for convenience added below as to context), we are pleased to see FDA moving forward with their Big Data and regulatory relevance initiative. Here is some high level information of general interest:

FDA RWD & RWE Definitions of December 2018

"Section 505F(b) of the FD&C Act defines RWE as "data regarding the usage, or the potential benefits or risks, of a drug derived from sources other than traditional clinical trials" (21 U.S.C. 355g(b)).5

In developing its RWE program, FDA believes it is helpful to distinguish between the sources of RWD and the evidence derived from that data. Evaluating RWE in the context of regulatory decision-making depends not only on the evaluation of the methodologies used to generate the evidence but also on the reliability and relevance of the underlying RWD; these constructs may raise different types of considerations.

For the purposes of this framework, FDA defines RWD and RWE as follows:

- Real-World Data (RWD) are data relating to patient health status and/or the delivery of health care *routinely collected* from a variety of sources
- Real-World Evidence (RWE) is the clinical evidence about the usage and potential benefits or risks of a medical product *derived from analysis of RWD*

***Our Current Thinking*: EDG & RWD/RWE …**

Understanding Data first to reach Quality Data is Cirtical

Examples of RWD include *data* derived from sources like:

- electronic health records (EHRs);
- medical claims and billing data;
- data from product and disease registries;
- patient-generated data, including from in-home-use settings; and
- data gathered from other sources that can inform on health status, such as mobile devices
- RWD sources (e.g., registries, collections of EHRs, administrative and medical claims databases) can be used for data collection and,
- in certain cases, to develop analysis infrastructure to support many types of study designs to develop RWE, including, but not limited to, randomized trials (e.g., large simple trials, pragmatic clinical trials) and observational studies (prospective or retrospective).

- FDA will continue its efforts to evaluate and explore ways and methods to optimize the utility of RWE. FDA has developed this Framework to guide its RWE Program.
- FDA already has taken the initiative to pilot projects that further the *understanding of potential uses of RWD and RWE* and will continue these efforts.
- Throughout this process, FDA is making stakeholder engagement a key aspect of its RWE Program.

Source: Framework for FDA's RWE Program, Dec 2018

LSCP *Current Opinion*

If industry feels that RWD/RWE can really provide value then they should (with us, of course☺) become PROACTIVE and provide input to FDA to get earlier approval as those "wait and see"!

Our Dr. Beat Widler (Core Team Member & CT SME) *further comments:*

- While some RWD data such as prescribing data (via IMS) has for very long been widely used by pharma companies to make decisions about marketing strategies their use for clinical development purposes has been and still is limited. For instance IMS had published years ago an article about the use of their data sets to optimize protocol design but the uptake was nil. Only recently a company like Clinerion offers access to RWD sets as a service to optimize recruitment strategies & protocol design.

- To use EHR data as a source for in silico development the biggest stumbling blocks are access to such data (often an EHR is a simple PDF of a paper record) and data privacy rules. While FDA also makes reference to ePRO data as RWD this may apply when we think of data collected through applications like Google Health but not when a sponsor collects such data in the frame of a CT - this is nothing new. Data collected through Apps could indeed be used as part of patient monitoring but then in most countries the patient would need to give his consent.

In my opinion companies should *first develop a data quality strategy* of what they want to accomplish in context of EDG before starting to talk about technology.

… and here is what we stated last November:

We have noted that participants and collaborators in Clinical Trials speak another professional language, different from the manufacturing floor and/or the Patients/Providers (RealWorld Data). Authors need to be understood by reviewers/approvers, users and recipients. Business data need to have value within and across all functions, usable/re-usable everywhere - globally.

***Our Current Thinking*: EDG & RWD/RWE …**

Understanding Data first to reach Quality Data is Cirtical

This paper recommends to consider the management of data – your assets - in a "*format-free"* or neutral form in Enterprise Data Governance (EDG). Your systems can then fully leverage the value of your data, whether structured (e.g. from your ERP) or unstructured (e.g. scientific content in a document – understood and interpretable).

A pragmatic approach to consider this *new way* of managing data is outlined below. Integrity, compliance and usability will be increased (data quality) and redundancies reduced. Improved *effectiveness* in context of product lifecycles will be the end result.

This approach can be better understood by first being aware of some history

In the early 1990s, the topics of discussion centered on how Word, WordPerfect and PDF were the tools that could contribute to accelerate time to market. In those times, electronic document management (EDM) was proposed. XEROX' Palo Alto Research Center (PARC) was where the kernel of an EDM concept was born. Documentum was launched and spun off as an independent vendor. Adobe created the concept of universal and unalterable output, to become known as PDF. The industry did not know of a standard yet for the exchange of information with external entities such as Health Authorities.

In Zurich, Switzerland, a thinktank was launched within Xerox to create/design a concept for electronic publishing of a dossier for a New Drug Application (NDA). It was an early step into the concept of digitalization. When the FDA showed interest in the concept, Ciba and Sandoz Pharma (today Novartis) jumped on the bandwagon and assisted in capturing business requirements for a first test. Based on the early successes, a Consortium was created including Sanofi and Weyth (Pfizer today) and was soon joined by Merck KGaA to provide input on a broader scale: XEROX Document Assembler (XDA) was born.

On the Health Authority side, it was suggested to use the paper based Common Technical Document (CTD) as a starter into the e-world. The International Conference on Harmonization (ICH) organized industry and regulators to work together. It started to become the global platform for a Standard, to be called the electronic CTD or eCTD.

This standard became accepted by industry and health authorities to exchange information in a uniform, technical and comprehensible way.

Modules 2-5 were structured in an approach to enable reviewers to fulfil their job in a more efficient and effective way. Experience has shown, however that there is much more room for improvement. Higher granularity of information could provide better understanding.
There was, and still is, a need to exchange information earlier in development than at fulfillment of an NDA or MAA. The IND, CTA and other early applications became new types of submissions that would precede the full Dossier.

At the FDA, Structured Product Labeling (SPL) was launched, based on the eCTD concept and at EMEA (today called EMA) a similar initiative was born called Product Information Management (PIM). PIM took into consideration a much different regulatory structure for Europe that included a wide variety of EU language requirements and still maintain consistency with the English version.

A major event occurred that spotlighted the need to provide a higher focus on quality data. It came with the scandal around Vioxx. Adverse events reported from the manufacturer were recognized as not being compatible nor combinable with the FDA, EMA and others, and therefore was not interpretable.

***Our Current Thinking*: EDG & RWD/RWE …**

Understanding Data first to reach Quality Data is Cirtical

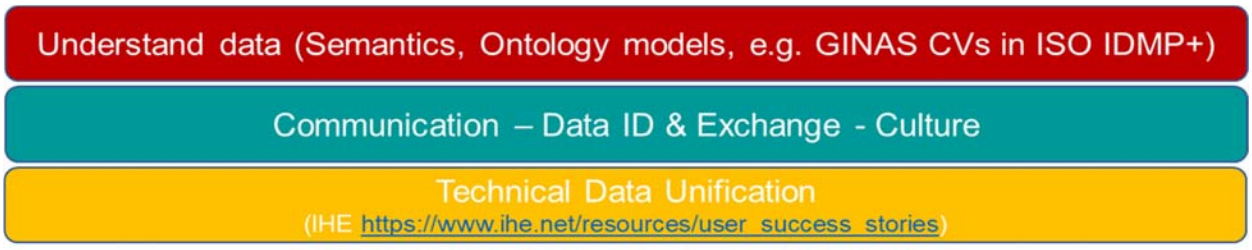
The regulators and the markets both started to recognize the importance of underlying data and the need to harmonize across all platforms. The ICH started an initiative under the term Identification of Medicinal Products (IDMP). This endeavor was transferred to ISO for finalization under the Standards umbrella of ISO Health Level 7 (HL7). This is where Individual Case Safety Reports (ICSR) and SPL were grouped as a collection and maintained. Global implementation guides have been updated and published, ready for implementation guides on the regional level.

Looking at content, granularity today can help achieve agility in the organization, and will extend the benefit to many

Many have come to realize that IDMP should be regarded as the kernel of product data. We have to look at all of those data elements across all industry functions. These elements have been structured towards the consistent re-use of data.

As eCTD release 4 is about to become mandatory over the next few years, with its extended granularity requirements, many business functions beyond regulatory affairs need to have access and use regulated content/components for their communication with the world. The information must be scalable from global to local and support development, government, markets and real world Medical Information. Real World Data (RWD) analytics and Real World Evidence (RWE) in data collection will be critical for validation against standards before ingesting to mission-critical systems.

All these considerations should lead stakeholders into creating a Vision & Strategy (V&S) effort to help understand two very important elements of change: *Structured Component Authoring* (SCA) and *advanced Semantic Technology* (ST). These two components will support agility in how to capture, analyze and merge any type of content (data) and how Enterprise Data Governance (EDG) can provide the better framework.



***No Innovation without Information** (integrated DATA)*

- Standards, combined with harmonized Business Glossaries enable *interoperability*
- Content and data components, used as a basis for innovation, must be based on an evolving EDG data strategy – from capturing, creating and managing to data consumers
- Data must be “neutral” as to format. This needs to be understood to be across all borders, meaning that semantically enriched content and metadata will enable mutual comprehension. Platform technology changes will always happen in the future and this concept will prevent confusion
- Common, or easily convertible units of measure will bridge the sources of data

Data and content can be created/maintained/re-used/merged or integrated through an holistic approach, guided by the concepts of EDG with strongly influenced measures needed to handle its massive extent.

***Our Current Thinking*: EDG & RWD/RWE …**

Understanding Data first to reach Quality Data is Cirtical

A Semantic Technology layer (ST) will deliver a number of primary capabilities and will support a strong Proof-of-Value. Decades of Clinical Trial (CT) data ***can*** be made available to assist the work for today's scientists.

1. The first is the automated extraction from previous CTs of facts, entities, concepts and relationships. Such datasets will be ingested into a "format neutral" repository (e.g. Master Data Management [MDM] or Data Lake, with a preference for a triple store technology). Upon ingestion, the extraction will undergo semantic processing. This will create and surface a rich layer of metadata which is underpinned by a domain model (organized by therapeutic area), thereby ensuring *consistency of terminology*, a foundation for interoperability. This technology is already used in production systems in many Life Sciences organizations today.
2. The second key capability of ST is to provide, within a *semantic metadata hub*, a metamodel of CTs and related data which enables *unification of disparate datasets*, without major Extract, Transform & Load (ETL) or MDM programs. This technology is already in use delivering significant results worldwide.
3. The metamodel adds *meaning to data*, defining for example that data from one source (e.g. a French blood pressure measurement expressed as 16/9) is related to data from another source (e.g. a measurement from the UK expressed as 140/80), and is also related to a concept from a Standard Vocabulary (e.g. ISO, IDMP, GINAS, MeSH). It will provide the neutral format, normalization across units of measure and how that data can be merged/converted/translated or unified to be commonly understood. This will result in data (and metadata) that can be associated with other data ingested or created within the lifecycle. We have created the term "IDMP+" to signify an expanded or more comprehensive set of IDMP data to meet future global business needs.
4. With the data, metadata and metamodel all residing together, *"integrated" (or unified) data* is made available to data consumers, providing agility and self-service. Furthermore, the models can be used to guide the data consumers in the terminology they use to ensure they are analysing or applying all of the data in the repository that is germane to their task.
5. The "neutral format" data can be managed within a *Structured Component Authoring* (SCA) data strategy. That is an approach that is also used in IDMP+ for topics such as medical information, already in production in global LS companies.

AND

If we closely look at SCA with its inherit opportunities and benefits in order to develop a comprehensive and holistic *Enterprise Data Strategy*, we urge you to assume the highest level of content granularity and to manage these newly-defined data elements through EDG on a platform that is focused on the business needs of tomorrow, enabling increased agility!

Let's discuss and elaborate options

ruedi.blattmann@lsconsultingpartners.com

