Xbiom™ - Accelerating innovation in BioPharma

PointCross Life Sciences Inc, a wholly owned subsidiary of PointCross Inc.
1291 E. Hillsdale Blvd., Suite 304, Foster City, CA, 94404, USA
Tel: +1 844-382-7257 Websites: www.xbiom.com or www.pointcross.com
Email: insights@pointcross.com

Date: July 8, 2019

Contents

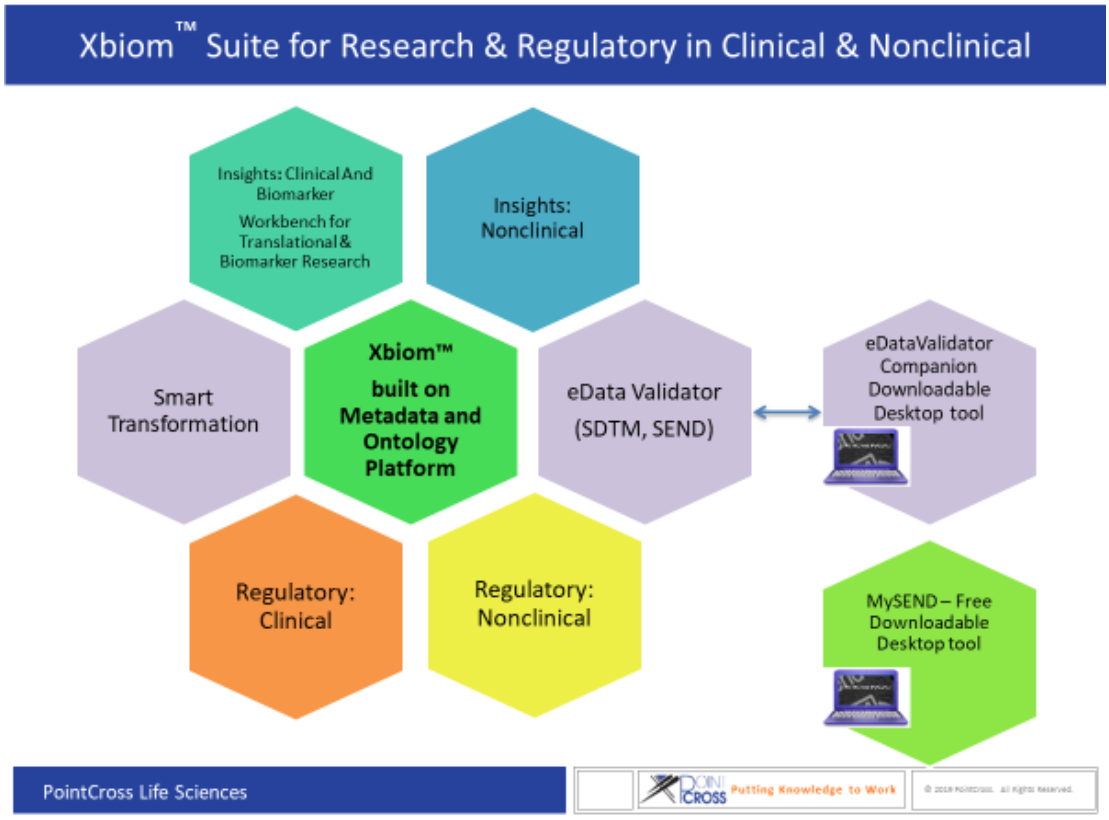
Summary

BioPharma companies have multiple streams of data collection to support their research, development, and regulatory submissions. These streams serve disparate parts of the R&D organizations and separate departments. Therefore, these data rarely gets integrated or unified to serve all the stakeholders. Despite advances in laboratory technology and bioinformatics, most organizations have barriers that prevent the immediate and complete access to an integrated view into clinical trial patient data and their biomarkers in disease or treatment. This is partly due to the lack of unification of the GxP data from clinical trials, that was originally intended for submission to the regulatory agencies; and the various translational assays that may be done by biomarker and translational scientists who are researching alternate targets, understanding pathways, biological mechanisms, and diagnostics. And, partly due to the difficulty of getting access to clinical trail data from the controlled systems of clinical operations. Organizational barriers in a global Biopharma are caused by departmental autonomy, regional regulations, and data parochialism. Xbiom™ creates a virtual global hub for all Biopharma R&D in a globally dispersed organization while maintaining the integrity of the source data.

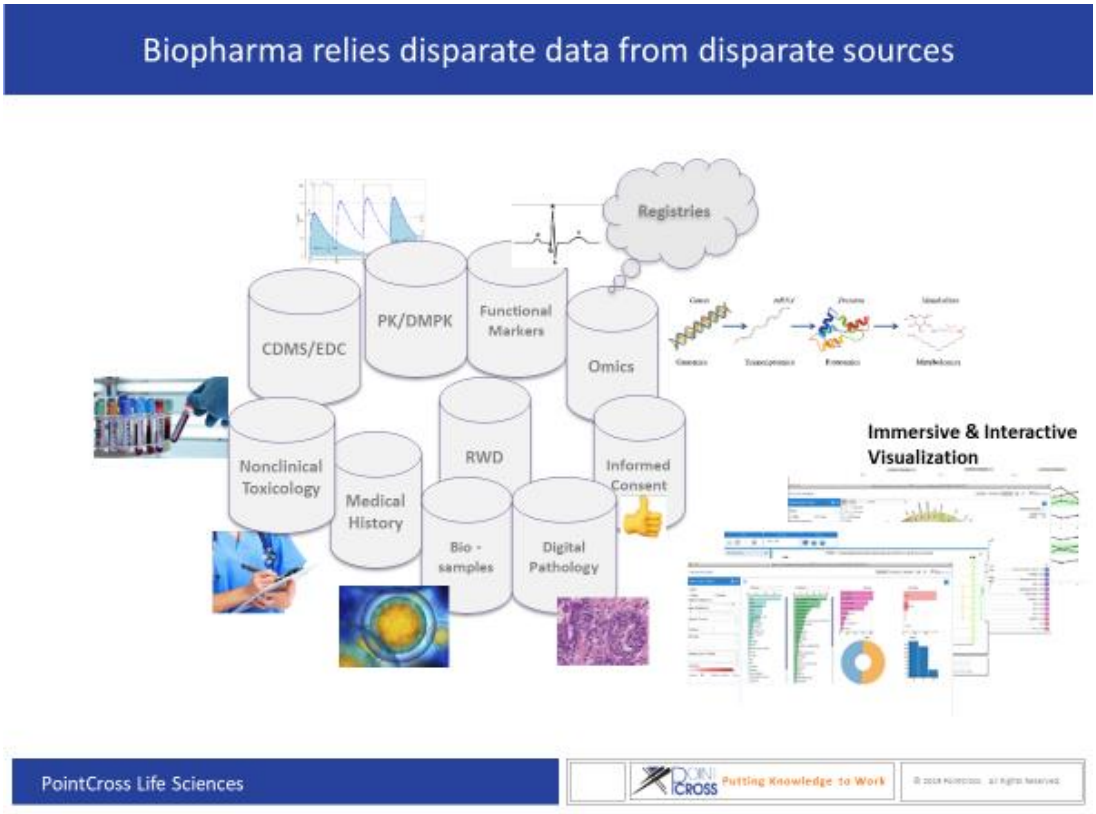
This paper discusses best practices to optimize cross-functional processes in Biopharma R&D not just to create better efficiencies but to do it by accelerating innovation, and anticipating the disruptive changes in how clinical trials will be conducted in the coming years. These best practices are good for today and it prepares for the future and its accompanying uncertainties. We will describe the role of integrating digital technologies that facilitate data transformation into a unified model; holding it for R&D purpose or re-purpose; making it searchable, findable; making results available for analysis such as differential gene expression across cohorts or ex-vivo samples or relating it to clinical end points. We will demonstrate practical implementation through our Xbiom™ solutions platform under a F.A.I.R. principle.

The paper will contemplate the nature of disruptive changes to come in the way clinical trials are planned and conducted. Ideas such as those fostered by Verily and Google, and ideas on “Deep Medicine” (by Dr. Eric Topol) and how AI will inject itself in various aspect of trial designs will change the way organization will need to think of data and how it should be reposed, and re-purposed. We will show how the Smart Transformation capabilities of Xbiom serve these needs.

Translational science is now at the center of most biologics based Biopharmas. These touch data from in-vitro, to ex-vivo to in-vivo animal studies and human clinical trials and they touch drug safety, immunology, molecular and digital biomarkers from proteomics, transcriptomics and metabolomics; and markers for CAR-T for the development of therapies, alternate targets, stratified cohort identification for extra-ordinary responders, and companion diagnostics. Xbiom™ from PointCross serves this need.



Xbiom serves the industry to manage, and exploit the following data.



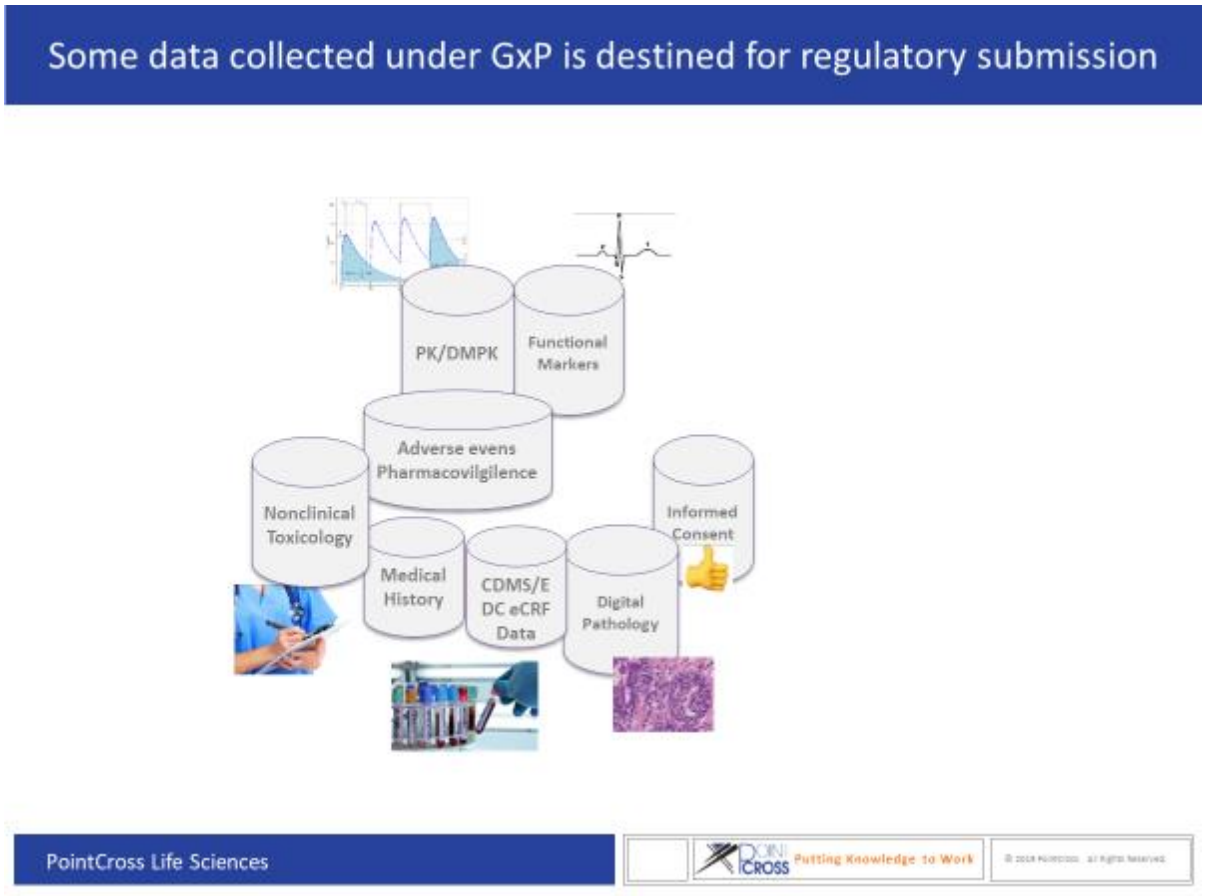
Version 2.0

Regulatory versus translational research data

Regulatory barriers, organizational divisions, data parochialism all conspire to prevent natural confluence of these streams. Sponsor companies such as Biopharma have to avoid the risk of corrupting as-collected submittable GxP data. The unintended consequence is the rise of divisions and data ownership. The implications, as well the benefits, of an integrated approach that preserves the relative differences of these two development organization while raising the possibilities of better synergy and insights leading to better therapies quicker, is discussed and proposed in this paper.

Regulatory data

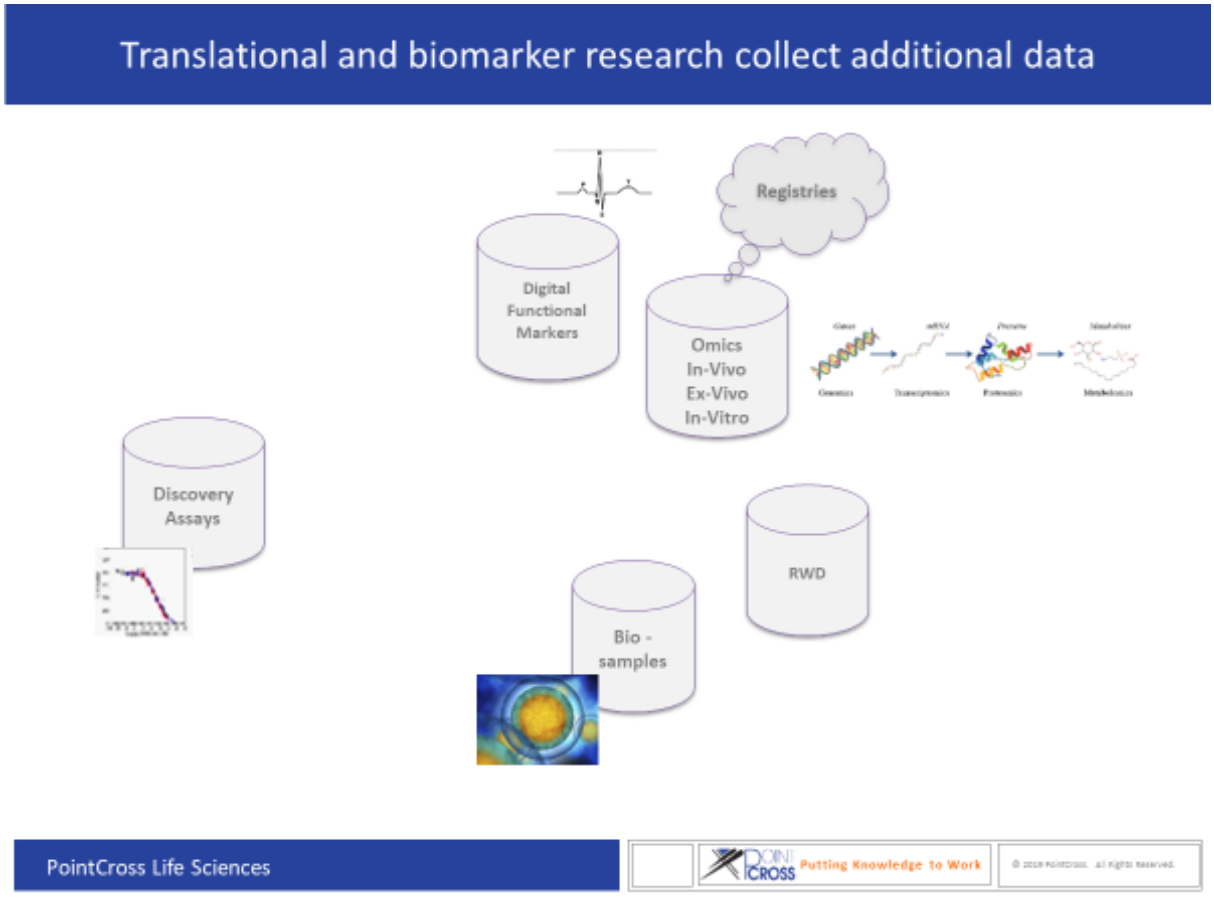
BioPharma companies develop therapeutics and diagnostics that are biologic in nature. All the data collected for regulatory submissions and approval of such as IND or NDA at FDA must be under GxP and the data held in 21 CFR Part 11 records management system where the data cannot be corrupted or altered. This data is governed and managed by the Clinical Operations departments under strict auditable regulatory processes and practices. Data is not accessed freely, but a process and data managers provide access with traceability. Analysis of this data is also not done in a freewheeling manner. Rather, a statistical analysis plan that takes the study design into account is published to the agencies even prior to the actual conduct of the studies. The analysis data sets are generated and prepared for analysis. Today the collected data is usually prepared in SDTM (CDISC) format, and the analysis data sets, and the analysis outputs are in ADaM (CDISC) format. These results are presented in the study reports along with various tables, figures, and listings as well the additional Adverse Events (Safety) data for publication.



Please send all inquiries to insights@pointcross.com

Translational data

Biologics development is frequently supported by translational and biomarker research where proteomic assays such as immunoassays (WB, IF, IHC, ICC, ELISA, and others) are conducted on patient's bio-samples for proteins, and genomic assays including sequencing for transcriptomics assays for DNA, mRNA, miRNA, and metabolomics. Many of these are assays may be ordered by individual researchers. The likelihood that these large datasets will be submitted is low. When the research shows promise, additional studies may be done under GxP that will validate these research findings, and they will be merged into a submission. The types of data collected are shown below.



Disruptive changes in clinical trials

Translational research is often thought of as "bench to bedside" to represent data collected in in-vitro assays, ex-vivo assays from human or animal tissues, and in-vivo studies conducted on animals, human trials; as well as the in-vitro assays on biologic samples.

"Bedside" data includes real-world patient data whose data can be very important in research. A virtual clinical trial can incorporate RWD as they may represent patients at various disease stages, or who are being treated by alternate standards of care.

The ideas evoked by RWD is not sparking many new concepts that integrate advanced sensors with remote reporting facilities, wearable, and other methods of sensing using other than primary transducers. Many techniques are now being considered, such as blood glucose measurements from contact lens type sensors on the eye, no-contact laser interferometry and sonar for mobility, ECG,

and many others. "Deep Medicine" by Dr. Eric Topol provides a clear view of a knowledgeable futurist about the trajectory of these technologies. His work with Verily is further evidence of how his futuristic view will likely come to pass with the disruptive changes in clinical trials. And with that the nature of data and how it will serve regulatory approvals. Recent FDA commissioners such as Dr Califf and Dr. Gottlieb have shown considerable support for these innovations. FDA's CBER is also speaking (such as at Bio-IT 2019) about the ways in which biomarker and genomic analysis data can be shared securely for regulatory review.

GxP clinical processes and practices are well established and understood. The nature of translational and biomarker research is also well known. However, the industry is already imagining the kind of disruption that the merging of biomarker research, bio-informatics, and smart sensor technology will bring to the industry. That may include new types of clinical trials that require fewer or no visits to the clinic, where genomic and deep assays can be done on participating subjects remotely while generating a tsunami of data that blur the difference between traditional eCRF and lab data.

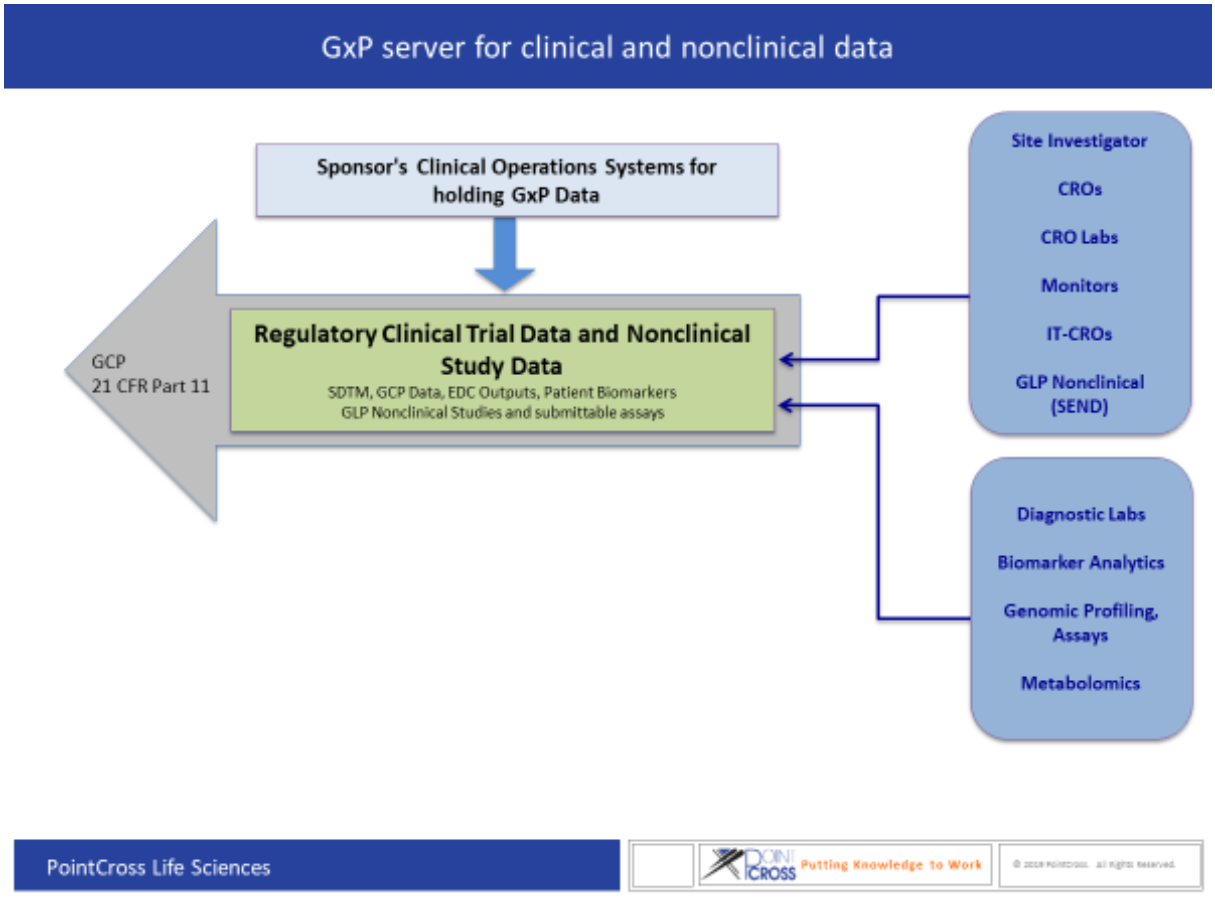
Goal - data unification without disruption

The need for unifying these disparate data is more than something "nice to have," but it is strategic and essential to remain innovative and maintain an agile translational and biomarker-driven drug program. In this paper, we will limit the discussion to unifying contemporary data types. We will consider this taking into account the following goals:

1. Do not require changes to established Clinical Operations processes or systems, retain the integrity of GxP data and the pathway to submissions and auditability
2. Conduct SAP, generation of analysis data sets and the final TFL in established processes for regulatory submission
3. Hold a copy of medically important data from clinical trials in a GxP qualified 21CFR Part 11 server so it can be queried and data needed for research can be "Read-Only" accessed
4. Provide researchers and data scientists the means to readily search, access and analyze data across data streams of patient clinical end-points and bio-sample molecular and digital biomarkers to generate insights that can be tested; validate new diagnostics, and provide study designs to confirm new therapeutic targets
5. Maximum flexibility for searching and data access with direct user interactivity or remote access through RESTful APIs.

Architectural approach

A GCP qualified validated server for staging the as collected clinical data (e.g., SDTM), the analysis data sets (ADaM) and the results (TFL, ISS, ISE) ensures the continued preservation and audit-tracked access of this data. None of the data in this server will be submitted or participate in the clinical operations processes during or after the trial is conducted. Pre-defined, selected, data from the main clinical operation process is read and copied over to this GCP server after the database lock in the main clinical operations system. It is then made accessible by the integrated, unified environment where the researchers can do their data science. Data access is continued to be tracked, and audit tracked. No changes can be made to this data in this 21 CFR Part 11 records management system.

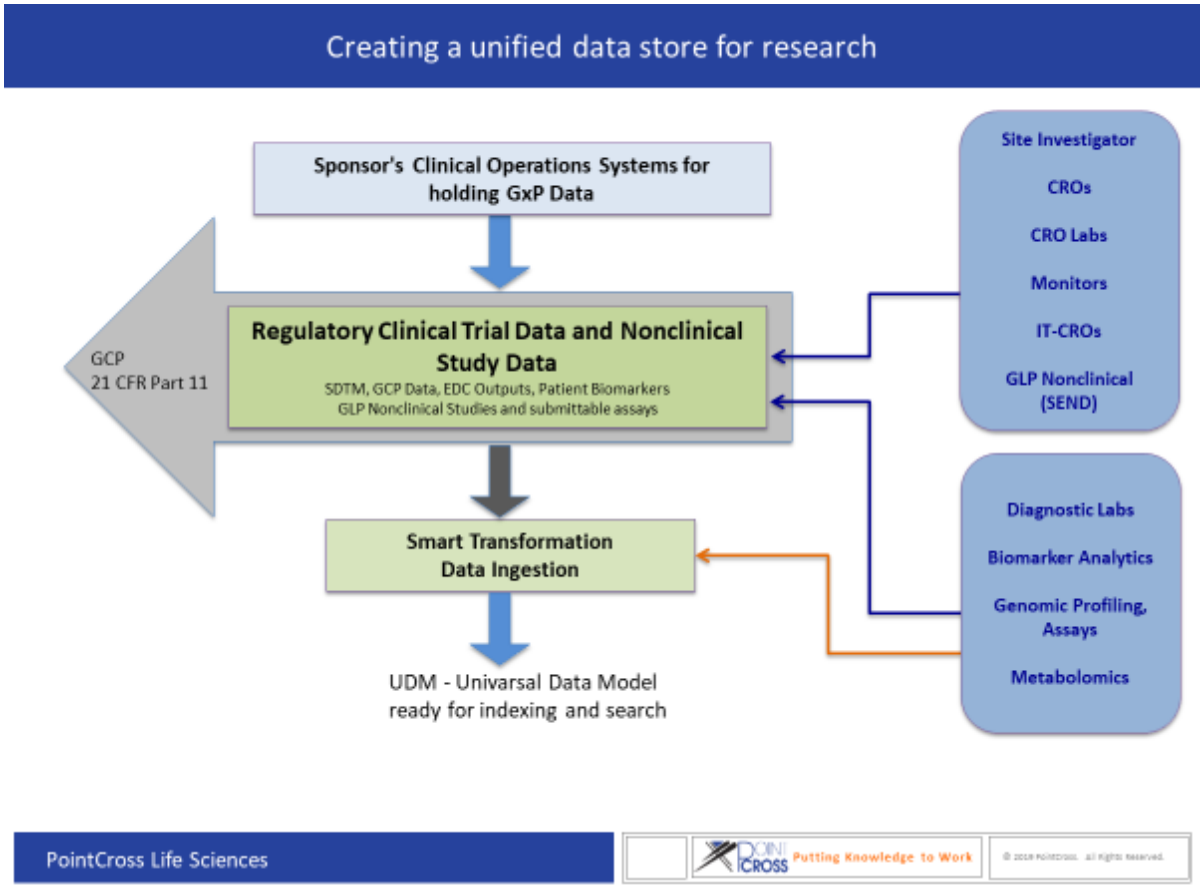


Smart data transformation to a repository model

Clinical data, even when standardized to an exchange model such as SDTM, ADaM does not mean that its data model is standard or uniform. It only means that for a system that understands that standard the data is readable. To find data sets that meet specific criteria across studies, these studies must have the same data model, and that requires a model transformation for each of the studies into a flat or graph model that can be indexed and made searchable. Contemporaneously, the terminology and units used in the study data sets must be harmonized and normalized to an invariant terminology set (an anchor term for each set of synonyms that is permanently indexed) and units. See the following paper on Smart Transformation for more on this topic.

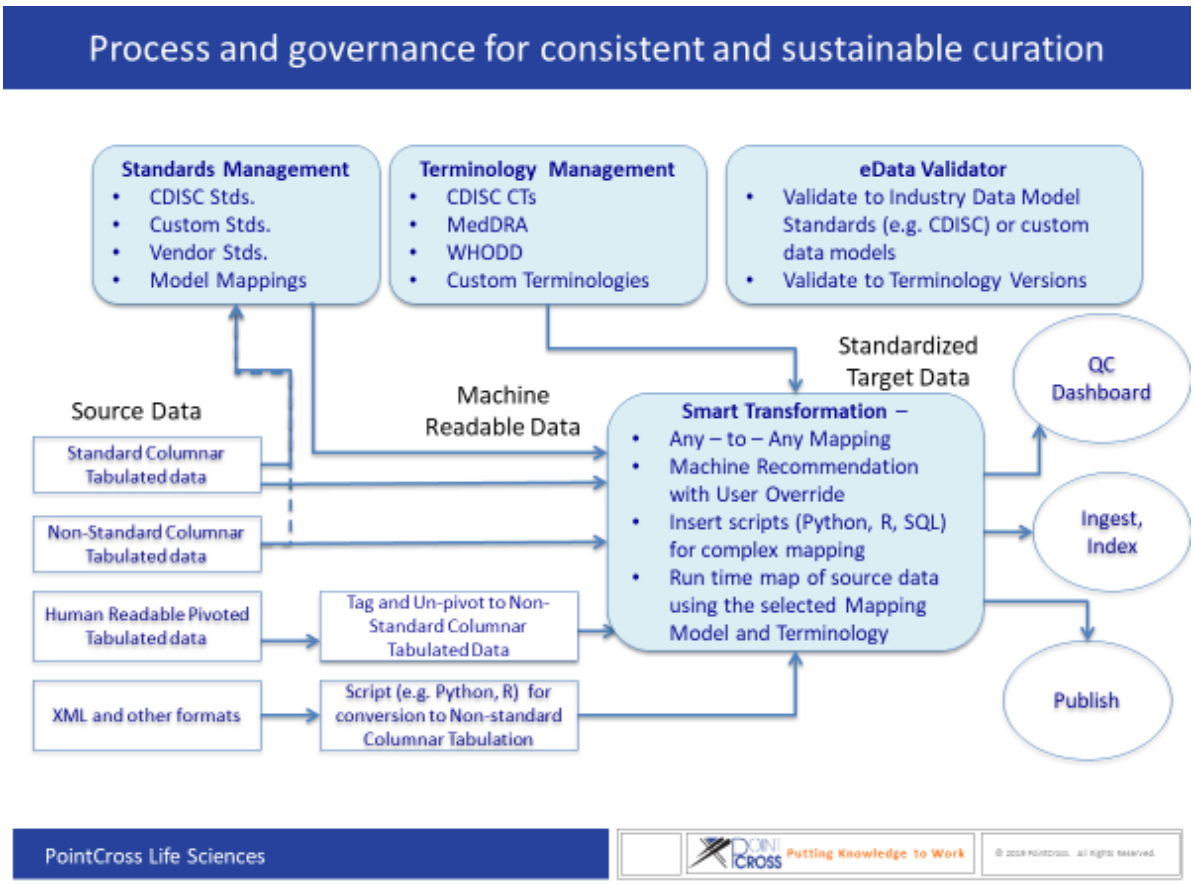
We refer to this as "smart" transformation to reflect the use of AI and machine learning to automate the identification of the source model and the recognition and recommendation of the right mapping models using machine learning. For sponsors who conduct many studies, this reduces the time and improves the quality of the curation and transformation very significantly. Once transformed, the data is available in an SDM (Standard Data Model) - an "exchange" standard such as CDISC SDTM or SEND to a specific IG version and controlled terminology release. The end objective is to generate a repository model, or UDM (Universal Data Model).

This same smart transformation module is also used to normalize the data received directly from various laboratories and assays, where the large variety of biomarker data is reported in custom models with lab specific data formats and terminology. In this case, the source data is not standardized, but they may still have some repeatable characteristics as the same vendor reports similar data from time to time.



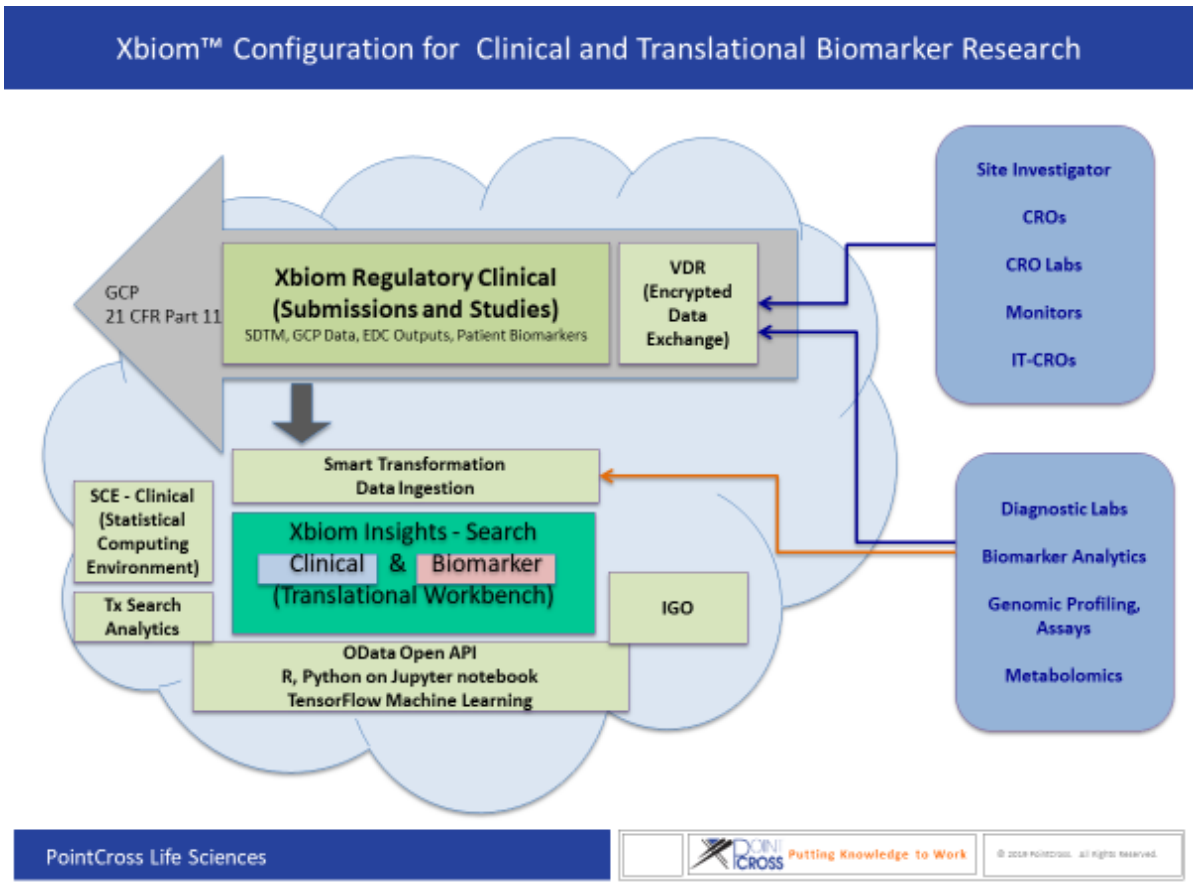
Metadata, Standards and Terminology Governance and Management

For a Bipharma committed to building a re-useable and evergreen, growing knowledge base from all their clinical and biomarker data there are multiple disciplines that must govern and manage certain internal processes with sustained diligence. These are shown in the figure below here a “Standard and Metadata Management” module, a Terminology and its ontology management, and a validation facility are described. For many large Pharma companies these may already be set up with third party solutions. With Xbiom – all these are included. Smart Transformation is supported by these three modules. More details are provided in the “Smart Transformation” paper.



Ancillary components for translational and biomarker research

With the clinical and biomarker data in a searchable indexed store or a queriable graph database, several additional components are needed to support the needs of innovation, research, analysis, and identifying insights. These are tools, and facilities to access data that has been unified so that they can be analyzed and their results visualized. Additional facilities are provided to allow researchers and data scientists with flexible programming tools and environments such as Jupyter with Python to set up complex or Bayesian queries. Many BioPharma companies have teams of bio-statisticians who are trained in working with SAS, R, and other such tools. They are supported with a statistical computing environment (SCE) so that they can query using SDTM terminology and apply their previously validated scripts to run off standard analysis.



Data Visualization

Researchers understand their data when it can be visualized in the ways which they are most familiar with. The Jupyter notebook and programming features such as Python and R provide access to the results of an analysis that can be visualized. The benefits of an OData interface is that this data can be readily accessed directly from third-party data visualization or BI (Business Intelligence) tools such as SpotFire, Tableau, GraphPad, and many others.

Xbiom provides IGO (Interactive Graphics Object) an onboard visualization facility where most of the products of analysis using the onboard analysis or search can be directly charted or plotted using immersive interactive graphics. For data that is generated using the open APIs and external scripts, the IGO can be configured for presenting specialized plots and visualization.

Conclusion

Xbiom provides a commercial off the shelf, solution platform that is available on the AWS cloud as "Software as a Service." It provides for a non-intrusive way to stage the GCP clinical and GLP nonclinical data for ready use by a smart transformation facility to map all the data into a repository model with normalized terminology and harmonized units. This resulting data can be indexed and made available for a search or a graph query. A battery of tools and facilities to interact, query, and analyze this unified data along with a way to visualize results are available. Access from an external system such as Jupyter through an OData interface and Xbiom's open APIs combine to create a powerful ready to install solution that will accelerate innovation at any biologics focused company.

Please send all inquiries to insights@pointcross.com

More importantly, Xbiom integrates data and organizations to work from a common hub that is accessible, findable and repurposable by dispersed organizations within a Biopharma while maintaining the sanctity of data security and access controls in a F.A.I.R environment.